

Molecule-Adapted Gaussian Basis Sets Generated in Multi-Parallel Hartree-Fock Scheme

Tong-Hyok Cha^{1,*}, Kwang-Jin Ri², Gum-Chol Kim¹, Yong-Su Pak¹

¹Department of Structure Analysis, Central Institute of Experiment and Analysis, Pyongyang, Democratic People's Republic of Korea

²Department of Physics, University of Sciences, Pyongyang, Democratic People's Republic of Korea

Email address:

cth1965@star-co.net.kp (Tong-Hyok Cha)

*Corresponding author

To cite this article:

Tong-Hyok Cha, Kwang-Jin Ri, Gum-Chol Kim, Yong-Su Pak. (2023). Molecule-Adapted Gaussian Basis Sets Generated in Multi-Parallel Hartree-Fock Scheme. *International Journal of Computational and Theoretical Chemistry*, 11(1), 26-33.

<https://doi.org/10.11648/j.ijctc.20231101.13>

Received: October 28, 2023; **Accepted:** November 16, 2023; **Published:** November 29, 2023

Abstract: For a molecule, all atoms are equally shared in several parallel-computers. Firstly, standard atomic Gaussian parameters (centers and exponents) for each atom are atom-independently and -simultaneously re-optimized by minimizing of molecular energy, while its surrounding atoms are kept in standard ones. For all atoms, by replacing re-optimized Gaussian parameters instead of standard ones, the molecular properties are improved, and such procedure is repeated N -times until a minimum energy is gained. Within the multi-parallel Hartree-Fock (MPHF) scheme, the McLean's 6-311G set for atoms (H, C, N, O) are re-optimized in various molecular environments, and used to calculate some HF-properties such as molecular energy and its components, and dipole moments. In the MPHF scheme, the eigen energy-related properties (E , ϵ_{HOMO} , ϵ_{LUMO}) are soon converged at $N=1\sim 2$, but the wave function-related properties ($\langle K \rangle$, $\langle V \rangle$, μ) are slowly converged to $N=5\sim 6$. Their limit values are closer to those obtained from the highest quality sets (6-311++G(3df,3pd), aug-cc-pVQZ). By introducing the multi-parallel calculus algorithm in which several computers are simultaneously operated, the time-consuming in the MPHF/6-311G scheme might become smaller than one in the HF scheme based on the 6-311++G(3df,3pd) set. The generation technique of molecule-adapted Gaussian basis sets with the multi-parallel calculus algorithm can be also imitated in the high level schemes such as CASSCF, MP2, DFT and HF over ab initio molecular dynamics.

Keywords: Gaussian Basis Set, Hartree-Fock Method, Parallel Computer, Floating Shift, Scale Factor

1. Introduction

In the variational Hartree-Fock (HF) scheme, a trial wave function is constructed as a linear combination of one-electron basis functions, where one minimizes the molecular energy by adjusting the linear expansion coefficients. With the plane-wave functions [1], the Gaussian basis functions (GBFs) are popular because of the well-established analytical forms of one- and two-electron integrals [2]. A Cartesian GBF is defined by four kinds of parameters: (i) angular momentum $\{m, n, l\}$, (ii) center-coordinates $\{A_x, A_y, A_z\}$, (iii) exponents $\{a\}$ and (iv) contraction coefficients $\{c\}$. In the Gaussian basis set limit, the more appending of GBFs with higher angular momentum can yield the lower molecular energy.

In a fully variational HF scheme [3-5], three kinds of parameters (ii-iv) are simultaneously optimized by minimizing of HF energy. However, usually, the GBF-centers are regarded as to be fixed on their atoms, and the exponents and contraction coefficients do as to be pre-determined for free atoms or small sample molecules. For a long day, sorts of atomic Gaussian parameters were pre-determined by several research groups [6-10], and stored in well-known commercial quantum chemistry programs. Those sets so-called as atomic orbital basis sets were fairly summarized by Jensen [11].

On the other hand, the molecule-adapted GBFs were occasionally studied by several research groups. A molecular selection rule of GBF-angular momentum was suggested by Zhang, and examined for a large number of molecules such as hydrogen-bonds, hydroxylates, carboxylic acids, amines,

alkyl-alkene hydrocarbons, oxides and metal-compounds [12]. The molecular floating of GBF-centers was discussed by Frost, Christoffersen, Huber, Pakiari and et al [13-16]. The GBF-exponents adapted to small molecules containing atoms from different rows of the periodic table, were determined within the molecular improved generator coordinate Hartree-Fock (MIGCHF) scheme developed by Jorge and co-workers [17, 18]. The GBF-exponents were partially re-optimized for valence shell or included extra orbitals around the region of chemical bond [19, 20].

In recent, with rapidly increasing of computer-performance, there appeared a few studies for generating of system-specific GBFs in the Python program package for differentiable quantum chemistry (DQC) [21-23]. By taking the advantage of automatic differentiation techniques, Alán and co-workers could compute numerically exact gradients with respect to any parameter without the need for implementing analytical gradients explicitly, and re-optimize the Gaussian parameters of STO-2G and -3G minimal basis sets for small molecules (H_2O , NH_3 , CH_4 , HF , CH_2O) [23]. By Kasim and co-workers [21], the Dunning's smallest cc-pVDZ sets for a training set (CH , CH_3 , CH_4 , C_2H_2 , C_2H_4) were re-optimized and applied to lower the DFT(PW92) energies for several hydrocarbons, and compared with those based on the cc-pVTZ sets.

Without the analytical [3-5, 24] and numerical [21-23] gradients, by an *iterative perturbation* technique, the atom-derived standard Gaussian basis sets were re-optimized in various molecular environments [25-27]. It was introduced the structure unit group (SUG) consisted of the atom (called as core-atom) and the neighbor ones primarily combined with it, in which the GBFs only for core- and its hydrogen atoms were perturbed in that SUG-environment, while those for other atoms were kept in atom-derived standard ones. The SUG-perturbed GBFs were compared with those re-optimized in molecular environments [25, 26], in which, to accelerate the molecular re-optimization of GBFs, Cha suggested the following multi-parallel Hartree-Fock (MPHF) scheme:

- (i) The Gaussian parameters for each atom are re-optimized by minimizing of molecular energy, while those for other atoms are kept in atom-derived standard ones. For all the atoms, the re-optimized Gaussian parameters (we call it as the *first* molecule-adapted Gaussian parameters (MAGPs)) are replaced the atom-derived standard ones.
- (ii) The first MAGPs for each atom are again re-optimized by minimizing of molecular energy, while those for other atoms are kept in the first MAGPs. For all the atoms, the second MAGPs are replaced the first ones, and the molecular energy is again calculated.
- (iii) The procedure (ii) is repeated N -times until the molecular energy is no more lowered.
- (iv) To save a computational time, all the atoms of a molecule are equally shared in several parallel-computers, in which the above re-optimizations begin at the same time with one another.

The aim of present article is to study, within the MPHF scheme, how way the standard atomic GBFs become adapted to molecular environments, and how much it will improve the accuracy of HF calculations. The article consists of three Sections. In Section 2, the MPHF scheme using parallel-computers is described. The N -th molecule-adapted 6-311G [28] is briefly denoted as Mol(N)/6-311G. In Section 3, for HCONH_2 , HCOOH and $\text{H}_2\text{C}_2\text{HOH}$ molecules, the HF properties obtained from the Mol(N)/6-311G ($N=1, 2, \dots$) are showed in detail. The molecular energy and its kinetic and potential components, virial ratio, frontier (HOMO, LUMO) energies and dipole moment are calculated, and compared with the HF/6-311++(3df,3pd) ones. The calculus time of four-center-two-electron repulsion integrals for the Mol(N)/6-311G is compared with one for the high quality sets such as the 6-311++(3df,3dp) and aug-cc-pVQZ.

2. Methodology

For an atom A , the valence-shell Gaussian type orbitals (GTOs) are float-scaled in molecular environment, and replaced as the following form;

$$\chi_{Mol}(\mathbf{r}, \mathbf{A}) = N_{norm} (x - A_x - \delta_x)^m (y - A_y - \delta_y)^n (z - A_z - \delta_z)^l \times \exp\left[-(r - A - \delta)^2 \zeta^2 \alpha\right], \quad (1)$$

where N_{norm} is the normalization constant, α is the exponent, (A_x, A_y, A_z) is the Cartesian coordinates of atom A . The scaling factor ζ and floating shift δ are varied to minimize the molecular energy within the HF scheme, while the GTOs belong in other atoms are kept in atom-derived standard ones. At the first, the valence-shell GTOs of atom A are floated as $(\delta_x, \delta_y, \delta_z)$ around its neighbor atoms. To find optimized-floating shifts, we repeated the HF energy-minimizing by shifting GTO-centers as ± 0.1 bohr from that atom's coordinates (A_x, A_y, A_z) . In next, to find optimized-scale factors, we repeated the HF energy-minimizing by changing the scale factors as 0.02. Similar to the float-scaling (FS) of GTOs in SUG-environment [27], the GTOs for a hydrogen atom are float-scaled together with those for its mother-atom. For all the atoms, the *first* MAGPs are replaced the atom-derived standard ones, and used to calculate the lower molecular energy. At the second, the first MAGPs for each atom are again re-optimized by minimizing of molecular energy, while those for other atoms are kept in the first MAGPs. For all the atoms, after the second MAGPs are replaced the first MAGPs, the molecular energy is calculated, and compared with the prior one. This procedure is repeated N -times until the molecular energy is no more lowered.

3. Results and Discussions

3.1. CHOOH Molecule

In the Mclean's 6-311G sets for heavy atoms (C, N, O, F), six kinds of valence shell-GBFs, $\{\text{S}_{3G}, \text{S}_{1G}, \text{S}'_{1G}, \text{P}_{(x,y,z)3G}$,

$P_{(x,y,z)1G}$, $P'_{(x,y,z)1G}$, are float-scaled within the MPHf scheme, and for hydrogen atoms, three kind of those, $\{S_{3G}, S_{1G}, S'_{1G}\}$. The MPHf calculation was repeated as N times until the value of E was not more lowered than ~ 1.0 mh. The floating shifts and scale factors generated from $N=1$ to $N=8$ are presented in Supplementary material-Part A1. By using serial CHOOH-adapted Mol(N)/6-311G sets, some properties are calculated and comparatively showed in Table 1. We can know the following facts:

- (i) The value of E ($N=1$) becomes lowered as ~ 60 mh than the E ($N=0$) which is calculated by the original 6-311G. Except for $N=1$, with increasing of $N>1$, the value of E is slightly lowered as ~ 1 mh (Figure 1), and converges to a certain value (-188.79026 hartree).
- (ii) The value of dipole moment is vibrante-increased from 1.5562 to 1.8137 Debye (Figure 2).

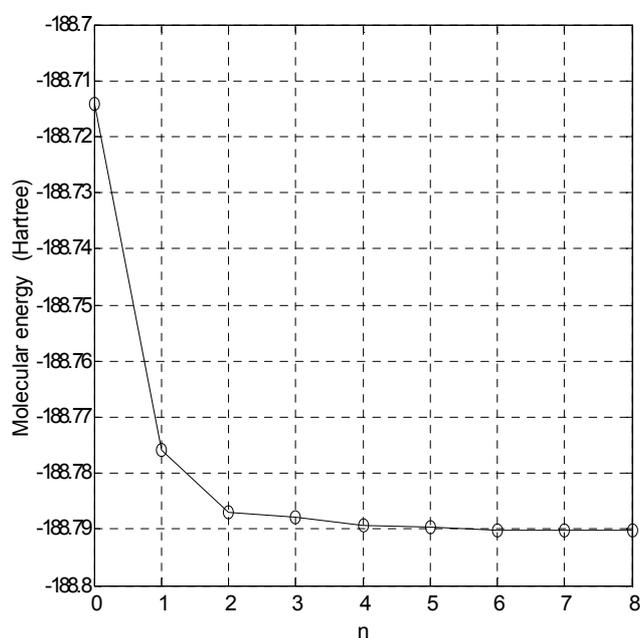


Figure 1. The plot of molecular energy.

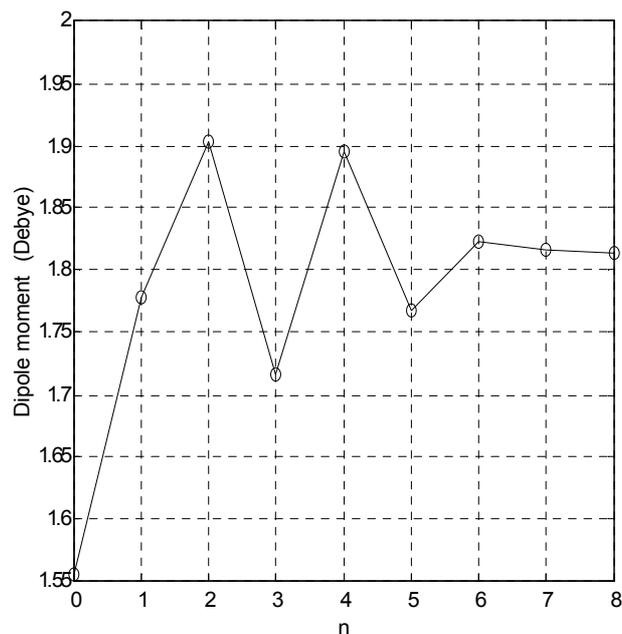


Figure 2. The plot of dipole moment.

Under the adiabatic approximation, since two energy-components $\langle K \rangle$ and $\langle V \rangle$ are closely related to the quantum-mechanical theorems such as the Virial and Hellmann-Feynman theorems, two energy components ($\langle K \rangle$, $\langle V \rangle$) and dipole moment μ are estimated as not “eigenvalues” but expectation values of “eigen-function” [26]. Therefore, those values depend much more sharply on the accuracy of trial wave function Ψ than the total energy E . Table 1 and Figures 1-2 show that, in the MPHf scheme, the eigen energy-related quantities (E , ϵ_{HOMO} , ϵ_{LUMO}) are soon converged at $N=1\sim 2$, but the eigen function-related properties ($\langle K \rangle$, $\langle V \rangle$, μ) are slowly converged to $N=5\sim 6$. Table 2 shows the floating shifts (δ_X , δ_Y) and scale factors ζ for the CHOOH-adapted Mol(8)/6-311G. Table 3 shows the comparison with the results obtained from the highest sets (6-311++G(3df,3pd), aug-cc-pVQZ) stored in “Gaussian 09” [29].

Table 1. The MPHf property-series for CHOOH-adapted Mol/6-311G ($N=1-8$).

n	E	$\langle K \rangle$	$\langle V \rangle$	ϵ_{HOMO}	ϵ_{LUMO}	μ
0	-188.71417	188.60790	-377.32207	-0.4767	0.1354	1.5562
1	-188.77599	188.49372	-377.26937	-0.4559	0.1881	1.7780
2	-188.78693	188.48978	-377.27626	-0.4641	0.1713	1.9031
3	-188.78797	188.47211	-377.25948	-0.4598	0.1781	1.7160
4	-188.78933	188.49287	-377.28162	-0.4636	0.1676	1.8950
5	-188.78972	188.46719	-377.25634	-0.4614	0.1727	1.7684
6	-188.79010	188.46376	-377.25331	-0.4629	0.1694	1.8236
7	-188.79016	188.46049	-377.25008	-0.4625	0.1672	1.8158
8	-188.79026	188.46262	-377.25229	-0.4627	0.1670	1.8137

Table 2. The floating shifts and scale factors for the CHOOH-adapted Mol(8)/6-311G.

Atom	Floating shifts (bore)						Scale factors					
	S_{3G}	S_{1G}	S'_{1G}	P_{3G}	P_{1G}	P'_{1G}	S_{3G}	S_{1G}	S'_{1G}	P_{3G}	P_{1G}	P'_{1G}
C	(0.0, 0.0)	(0.4,-0.1)	(0.1, 0.5)	(0.0, 0.0)	(0.0, -0.1)	(-0.2, 0.1)	1.00	1.16	1.34	1.10	1.12	1.14
O ¹	(0.0, 0.0)	(-0.1,-0.1)	(0.0, 0.0)	(0.0, 0.0)	(0.1, 0.1)	(0.1, 0.0)	1.00	1.02	1.06	0.96	0.94	0.94
O ²	(0.0, 0.0)	(0.1, -0.2)	(-0.2,-0.3)	(0.0, 0.0)	(-0.1, 0.0)	(-0.1, 0.0)	1.00	1.04	1.12	0.98	0.96	0.96
H ¹	0.0	1.3	0.2				0.94	1.48	1.42			
H ²	0.1	0.6	-0.3				0.96	1.14	1.30			

Table 3. The HF results obtained from CHOOH-adapted Mol(8)/6-311G.

	6-311G	Mol(8)/6-311G	6-311++G(3df,3pd)	aug-cc-pVQZ
E	-188.71417	-188.79026	-188.83185	-188.84798
$\langle K \rangle$	188.60790	188.46262	188.38207	188.43171
$\langle V \rangle$	-377.32207	-377.25229	-377.21392	-377.27969
μ_x	1.3574	1.6449	1.6282	1.6296
μ_y	0.7610	0.7641	0.8555	0.8498
μ_z	0.0000	-0.0001	0.0000	0.0000
μ	1.5562	1.8137	1.8392	1.8379

Table 4. The MPHf/6-311G results for CHONH₂ molecule.

n	E	$\langle K \rangle$	$\langle V \rangle$	ϵ_{HOMO}	ϵ_{LUMO}	μ
0	-168.89661	168.83583	-337.73245	-0.4127	0.1432	4.5941
1	-168.93959	168.66296	-337.60239	-0.4100	0.1503	4.1390
2	-168.94724	168.66542	-337.61260	-0.4087	0.1517	4.4274
3	-168.94982	168.67772	-337.62744	-0.4105	0.1479	4.2214
4	-168.95005	168.68937	-337.63934	-0.4079	0.1468	4.3428
5	-168.95034	168.69336	-337.64015	-0.4098	0.1437	4.1691
6	-168.95074	168.68957	-337.64021	-0.4088	0.1451	4.2803

3.2. CHONH₂ Molecule

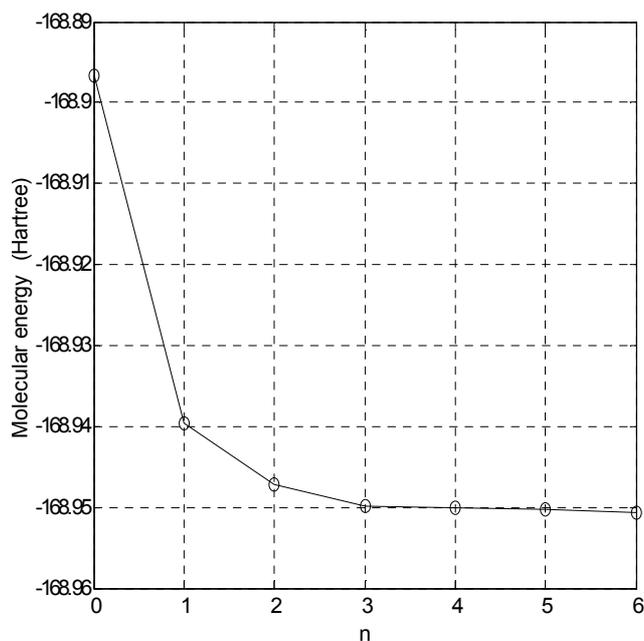


Figure 3. The plot of molecular energy.

For CHONH₂ molecule, the MPHf calculation is repeated as N times until the value of E is not more lowered than ~ 1.0 mh. By using the CHONH₂-adapted Mol(N)/6-311G($N=1-6$), some HF properties are calculated (Table 4). The value of E ($N=1$) becomes lowered as ~ 60 mh than the E ($N=0$) which was calculated by the original 6-311G. With increasing of N (from 2 to 6), the value of E is slightly lowered as ~ 1 mh (Figure 3), and converges to a certain value (-188.95074 hartree). The value of dipole moment is vibrated-decreased from 4.5941 to 4.2803 Debye (Figure 4).

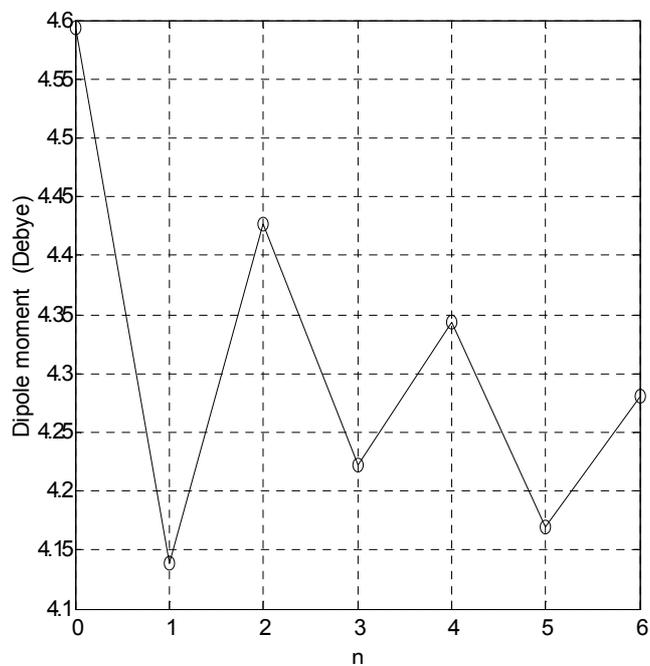


Figure 4. The plot of dipole moment.

Table 5 shows the floating shifts (δ_x , δ_y) and scale factors ζ for the CHONH₂-adapted Mol(6)/6-311G. Table 6 shows the comparison between the MPHf and HF results. The results obtained from the CHONH₂-adapted Mol(6)/6-311G are closer to ones from the 6-311++ G(3df,3pd) and aug-cc-pVQZ sets. The MPHf-dipole moment μ becomes the smallest among four values, and closer to one (4.0816 Debye) obtained from the DFT/B3LYP/6-311G++(3df,3pd).

Table 5. The floating shifts and scale factors for the CHONH_2 -adapted Mol(6)/6-311G.

Atoms	Floating shifts (bore)						Scale factors					
	S_{3G}	S_{1G}	S'_{1G}	P_{3G}	P_{1G}	P'_{1G}	S_{3G}	S_{1G}	S'_{1G}	P_{3G}	P_{1G}	P'_{1G}
C	(0.0, 0.0)	(-0.6, -0.2)	(1.1, -1.2)	(0.0, 0.0)	(0.0, -0.1)	(0.6, -0.2)	1.00	1.12	1.20	1.08	1.08	1.04
O	(0.0, 0.0)	(-0.3, 0.2)	(-0.5, 0.4)	(0.0, 0.0)	(-0.1, 0.1)	(-0.1, 0.0)	1.00	1.08	1.44	0.96	0.94	0.98
H ¹	0.0	1.4	0.7				0.90	1.34	1.34			
N	(0.0, 0.0)	(0.2, 0.1)	(-0.5, 0.0)	(0.0, 0.0)	(0.0, 0.0)	(0.1, 0.1)	1.00	1.00	1.02	1.02	1.00	0.98
H ²	0.1	0.3	-0.5				0.92	0.98	1.02			
H ³	0.1	0.3	-0.5				0.92	0.98	1.02			

Table 6. The HF results obtained from the CHONH_2 -adapted Mol(6)/6-311G.

	6-311G	Mol(6)/6-311G	6-311++G(3df,3pd)	aug-cc-pVQZ
E	-168.89661	-168.95074	-168.99684	-169.01108
$\langle K \rangle$	168.83583	168.68957	168.64098	168.67750
$\langle V \rangle$	-337.73245	-337.64021	-337.63782	-337.68859
μ_x	-4.3384	-4.0695	-4.1945	-4.1756
μ_y	1.5113	1.3268	1.3788	1.3651
μ_z	0.0003	0.0003	0.0002	0.0002
μ	4.5941	4.2803	4.4153	4.3931

Table 7. The MPHF/6-311G results for CH_2CHOH molecule.

N	E	$\langle K \rangle$	$\langle V \rangle$	ϵ_{HOMO}	ϵ_{LUMO}	μ
0	-152.85971	-152.68337	-305.54309	-0.3559	0.1405	2.3033
1	-152.89025	152.58255	-305.47223	-0.3511	0.1795	1.8187
2	-152.89668	152.58629	-305.48243	-0.3501	0.1645	1.7947
3	-152.89748	152.54796	-305.44538	-0.3503	0.1812	1.7392

3.3. CH_2CHOH Molecule

For CH_2CHOH molecule, the MPHF calculation was also repeated as N times until the value of E was not more lowered than ~ 1.0 mh. By using the CH_2CHOH -adapted Mol(N)/6-311G ($N=1-3$), some HF properties are calculated (Table 7). The value of E ($N=1$) becomes lowered as ~ 30 mh than the E ($N=0$) which was calculated by the original 6-311G. With increasing of N (from 2 to 3), the value of E is slightly lowered as ~ 1 mh (Figure 5), and converges to a certain value (-152.89748 hartree). The value of dipole moment is decreased from 2.3033 to 1.7392 Debye (Figure 6).

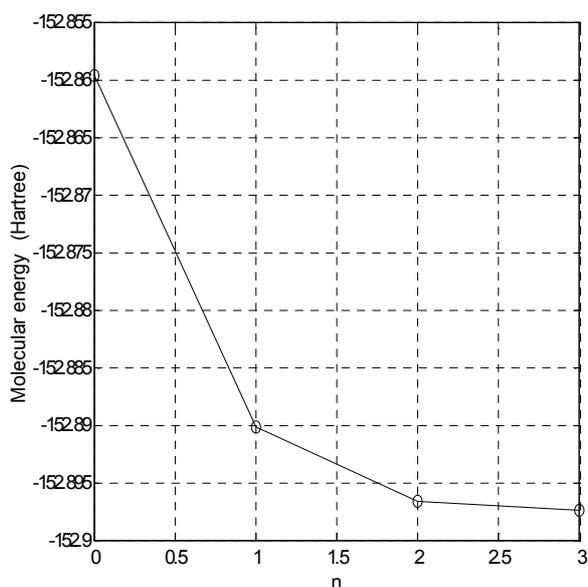


Figure 5. The plot of molecular energy.

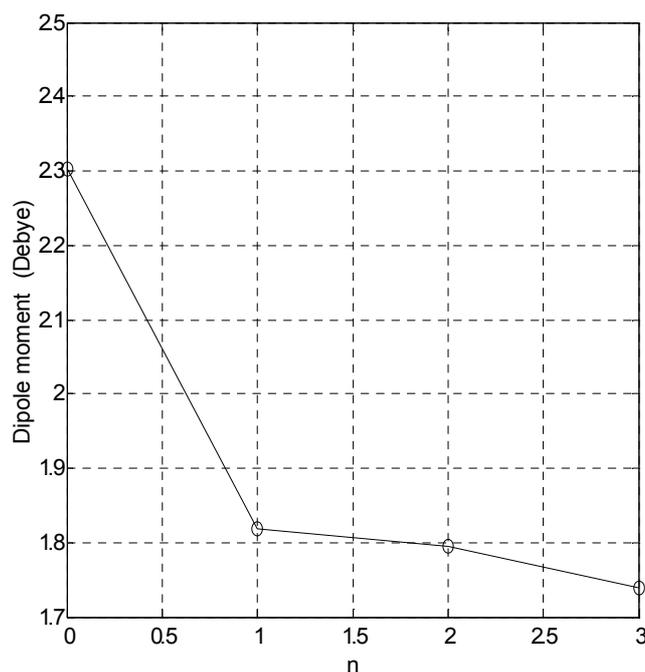


Figure 6. The plot of dipole moment.

Table 8 shows the floating shifts (δ_x , δ_y) and scale factors ζ for the CH_2CHOH -adapted Mol(6)/6-311G. Table 9 shows the MPHF results. The results obtained from the CH_2CHOH -adapted Mol(3)/6-311G are closer to ones from the 6-311++G (3df,3pd) and aug-cc-pVQZ sets, while the dipole moment μ is also the smallest among four values, and closer to one (1.7181 Debye) obtained from the DFT/B3LYP/6-311G++(3df,3pd).

Table 8. The floating shifts and scale factors for the CH_2CHOH -adapted Mol(3)/6-311G.

Atoms	Floating shifts (bore)						Scale factors					
	S_{3G}	S_{1G}	S'_{1G}	P_{3G}	P_{1G}	P'_{1G}	S_{3G}	S_{1G}	S'_{1G}	P_{3G}	P_{1G}	P'_{1G}
C ¹	(0.0, 0.0)	(0.3, 0.0)	(-0.1, 0.6)	(0.0, 0.0)	(-0.1, 0.1)	(0.0, 0.0)	1.00	1.02	1.02	1.00	1.00	1.00
C ²	(0.0, 0.0)	(-0.4,-0.2)	(1.1,-0.5)	(0.0, 0.0)	(0.0,-0.1)	(0.2, 0.3)	1.00	1.02	1.02	0.98	1.02	1.06
H ¹	0.1	0.4	0.3				0.94	1.02	1.10			
H ²	0.1	0.4	0.3				0.94	1.02	1.10			
H ³	0.1	0.5	-0.5				0.96	1.10	1.22			
O	(0.0, 0.0)	(0.2, 0.1)	(0.1,-0.2)	(0.0, 0.0)	(0.0, 0.1)	(0.0, 0.1)	1.00	0.98	1.02	1.00	0.98	0.98
H ⁴	0.0	0.7	-0.1				0.98	0.94	1.24			

Table 9. The HF results obtained from CH_2CHOH -adapted Mol(3)/6-311G.

	6-311G	Mol(3)/6-311G	6-311++G(3df,3pd)	aug-cc-pVQZ
E	-152.85971	-152.89748	-152.94653	-152.95968
$\langle K \rangle$	152.68337	152.54796	152.53714	152.56598
$\langle V \rangle$	-305.54309	-305.44538	-305.48368	-305.52566
μ_x	-0.2293	-0.4178	-0.3518	-0.3569
μ_y	2.2919	1.6883	1.8037	1.7840
μ_z	0.0037	0.0037	0.0030	0.0029
μ	2.3033	1.7392	1.8377	1.8193

Consequently, the iterative perturbation procedures in the MPHF scheme are repeated as $N=3\sim 10$ times until the molecular energy E is not more lowered than ~ 1.0 mH.

3.4. MPHF Scheme Based on *SUG*-Perturbed GBFs

On the other hand, instead of standard atomic GBFs, the *SUG*-perturbed GBFs might be used as initial ones for the MPHF scheme, and one can load their floating shifts and

scale factors from the database “*SUG*-GP” [25]. If do so, all the MPHF properties are soon converged at $N=1\sim 2$. Table 10 shows some MPHF ($N=2$) results started from the *SUG*-perturbed GBFs. All the dipole moments become very closer to ones obtained from the highest quality sets. Figure 7 shows the comparison of the HF-dipole moments calculated by 6-311G, Mol(2)/6-311G and aug-cc-pVQZ sets.

Table 10. The MPHF ($N=2$) results started from the *SUG*-perturbed GBFs.

Molecule	Quantities	6-311G	Mol(2)/6-311G	6-311++G(3df,3pd)	aug-cc-pVQZ
CH_3OH	E	-115.01917	-115.06235	-115.08861	-115.09875
	$ k_{virial-2} \times 10^{-4}$	11.55161	1.632175	5.91601	4.45443
	$\langle T \rangle$	115.15219	115.08063	115.02056	115.04750
	$\langle V \rangle$	-230.17136	-230.14247	-230.10917	-230.14626
	μ	2.3486	1.8631	1.8648	1.8347
	CH_3NH_2	E	-95.19056	-95.21437	-95.25245
$ k_{virial-2} \times 10^{-4}$		17.80592	5.225134	1.901765	2.27270
$\langle T \rangle$		95.36036	95.26368	95.27057	95.28219
$\langle V \rangle$		-190.55092	-190.47759	-190.52302	-190.54273
μ		1.7819	1.3937	1.4251	1.4168
NH_2OH		E	-130.96352	-131.00017	-131.04105
	$ k_{virial-2} \times 10^{-4}$	17.27780	4.65215	2.94965	2.19966
	$\langle T \rangle$	131.19019	131.06100	131.00241	131.02444
	$\langle V \rangle$	-262.15372	-262.06098	-262.04347	-262.07770
	μ	0.8073	0.6322	0.5728	0.5622
	$CHCOH$	E	-151.63196	-151.66851	-151.71268
$ k_{virial-2} \times 10^{-4}$		17.28615	27.57321	30.77114	29.11440
$\langle T \rangle$		151.37030	151.27452	151.24727	151.28500
$\langle V \rangle$		-303.00226	-303.02345	-302.95995	-303.01046
μ		2.1327	1.6520	1.6199	1.6074
$CH_3CH_2NH_2$		E	-134.21298	-134.24649	-134.29658
	$ k_{virial-2} \times 10^{-4}$	1.84990	11.42336	16.28166	16.057994
	$\langle T \rangle$	134.18816	134.09331	134.07828	134.09244
	$\langle V \rangle$	-268.40114	-268.33980	-268.37486	-268.40022
	μ	1.7613	1.4132	1.4103	1.4057
	CH_3CH_2OH	E	-154.04778	-154.08705	-154.13837
$ k_{virial-2} \times 10^{-4}$		6.29329	17.98797	21.89626	20.75788
$\langle T \rangle$		153.95090	153.81019	153.80160	153.83218
$\langle V \rangle$		-307.99868	-307.89705	-307.93998	-307.98369
M		2.1163	1.5547	1.7029	1.6875

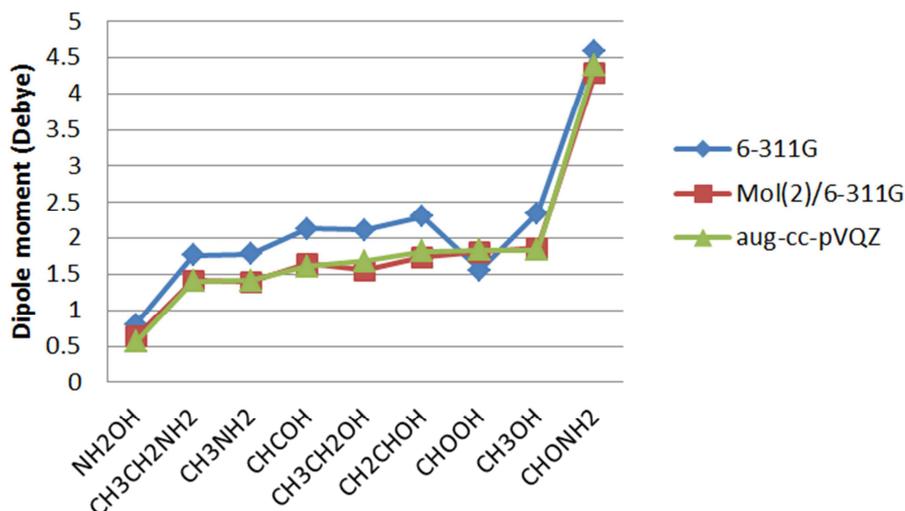


Figure 7. The HF-dipole moments resulted from 6-311G, Mol(2)/6-311G and aug-cc-pVQz sets.

3.5. Four-Center-Two-Electron-Repulsion Integrals in the MPHf Scheme

The two-electron repulsion integral computation is a main time-consuming step in the MPHf scheme. In our parallel-calculus algorithm, the total number of two-center repulsion integrals is equally divided by several computers: Q_i ($i=1-K$) (as described in Supplementary-Part A2). In the case of molecules consisted of only atoms (C, N, O, F) in the first-row of periodic table, the relation between the Q_1 (the total number of two-center-repulsion integrals to be calculated in the 1th-computer) and atom-numbers N_{atom} ($=1\sim 10$) for the MPHf/Mol($N=2$)/6-311G and HF/6311++G(3df,3pd) is showed in Figure 8, where the number K of computers equals with 5 and the number S of parallel-working MATLAB programs in each computer equals with 3.

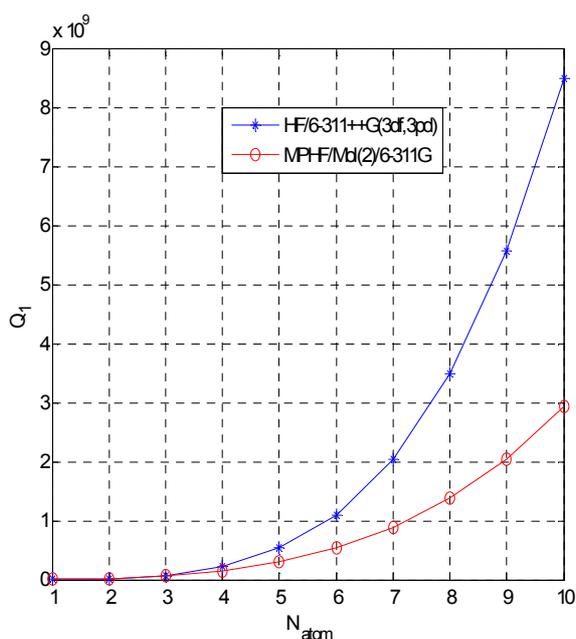


Figure 8. The numbers of two-electron repulsion integrals in the MPHf/Mol(3)/6-311G and HF/6-311++G(3df,3pd) schemes.

As shown in Figure 8, by introducing the multi-parallel calculus algorithm in which several computers are simultaneously operated, the time-consuming in the MPHf scheme might become smaller than one in the standard HF scheme based on the highest quality sets such as 6-311++G(3df,3pd) or aug-cc-pVQZ.

On the other hand, the following problems are still open, and will be more discussed in next:

- 1) When the higher-quality sets including diffusion and polarization functions are re-optimized, how much the MPHf accuracy will be improved?
- 2) The comparison of GBF-parameters (centers, exponents, contracted coefficients) re-optimized in the Python program package [21-23] and our MPHf scheme.
- 3) The mathematical completeness of molecule-adapted Gaussian basis sets (as briefly hinted in Supplementary-Part A3).

4. Conclusions

Within the multi-parallel Hartree-Fock (MPHF) scheme, the McLean's 6-311G set for atoms (H, C, N, O) are re-optimized in various molecular environments. All the atoms of a molecule are equally shared in several parallel-computers. Firstly, the valence-shell GBFs belong in each atom are float-scaled by minimizing of molecular energy, while those for other atoms are kept in the standard atomic GBFs. For all atoms, by replacing the float-scaled GBFs instead of standard ones, the HF-properties are improved, and such procedure is repeated until a minimum energy is gained. For CHOOH, CHONH₂, CH₂CHOH molecules, the molecule-adapted Mol(N)/6-311G sets are generated, and used to calculate some HF-properties such as molecular energy and its components, and dipole moments. In the MPHf scheme, the eigen energy-related properties (E , ϵ_{HOMO} , ϵ_{LUMO}) are soon converged at $N=1\sim 2$, but the wave function-related properties ($\langle K \rangle$, $\langle V \rangle$, μ) are slowly converged to $N=5\sim 6$. Their limit values are closer to those obtained from the highest quality sets (6-311++G(3df,3pd), aug-cc-pVQZ).

When the SUG-perturbed GBFs are used as initial ones for the MPHf scheme, all the MPHf properties are soon converged at $N=1\sim 2$. By introducing the multi-parallel calculus algorithm in which several computers are simultaneously operated, the time-consuming in the MPHf/Mol(N)/6311G scheme might become smaller than one in the HF scheme based on the highest quality sets (6-311++G (3df,3pd), aug-cc-pVQZ). The generation technique of molecule-adapted Gaussian basis sets with the multi-parallel calculus algorithm can be also imitated in the high level schemes such as CASSCF, MP2, DFT and HF over ab initio molecular dynamics.

ORCID

Tong-Hyok Cha: <https://orcid.org/0009-0000-4062-8490>

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] G. Lippert, J. Hutter, M. Parrinello, *Theor. Chem. Acc.* 103 (1999) 124.
- [2] F. Jensen. *Introduction to Computational Chemistry*. -2nd, England: John Wiley & Sons Ltd; 2007, pp. 192-231.
- [3] M. Tachikawa, Y. Osamura, *J. Chem. Phys.* 113 (2000) 4942.
- [4] M. Tachikawa, K. Taneda, K. Mori, *Int. J. Quantum Chem.* 75 (1999) 497.
- [5] N. Shimizu, T. Ishimoto, M. Tachikawa, *Theor. Chem. Acc.* 130 (2011) 679.
- [6] T. H. Dunning, *J. Chem. Phys.* 53 (1970) 2823.
- [7] W. J. Hehre, R. Ditchfie, J. A. Pople, *J. Chem. Phys.* 56 (1972) 2257.
- [8] F. J. Jensen, *Chem. Phys.* 115 (2001) 9113.
- [9] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* 7 (2005) 3297.
- [10] A. C. Neto, E. P. Muniz, R. Centoducatte, F. E. Jorge, *Theochem: J. Mol. Struct.* 718 (2005) 219.
- [11] F. Jensen, *Wiley Interdisciplinary Review, Com. Mol. Sci.* 3 (2013) 273. R.
- [12] Q. Zhang, W. J. Fan, *Int. J. Quant. Chem.* 115 (2015) 570.
- [13] A. A. Frost, *J. Chem. Phys.* 47 (1967) 3707.
- [14] R. E. Christoffersen, *Advan. Quant. Chem.* 6 (1972) 333.
- [15] H. Huber, *Chem. Phys. Lett.*, 62 (1979), 95.
- [16] A. H. Pakiari, M. Solimannejad, *J. Mol. Str. (Theochem)* 583 (2002) 249.
- [17] M. T. Barreto, E. P. Muniz, F. E. Jorge, R. Centoducatte, *Brazilian J. of Phys.* 35 (2005) 965.
- [18] F. E. Jorge, A. C. Neto, *J. Mol. Str. (Theochem)* 589–590 (2002) 359.
- [19] D. Neisius, G. Verhaegen, *Chem. Phys. Lett.* 78 (1981) 147.
- [20] D. Neisius, G. Verhaegen, *Chem. Phys. Lett.* 89 (1982) 228.
- [21] M. F. Kasim, S. Lehtola, S. M. Vinko, *J. Chem. Phys.* 156 (2022) 084801.
- [22] R. A. Shaw, J. G. Hill, *Mol. Phys.* 116 (2018) 1460.
- [23] T. Tamayo-Mendoza, C. Kreisbeck, R. Lindh, A. Aspuru-Guzik, *ACS Cent. Sci.* 4 (2018) 559.
- [24] T. Ishimito, M. Tachikawa, *Int. J. Quantum Chem.* 9 (2010) 8.
- [25] T. H. Cha, G. C. Kim, K. J. Ri, *Chem. Phys.* 570 (2023) 111890.
- [26] T. H. Cha, *Chem. Phys. Lett.* 765 (2021) 138285.
- [27] T. H. Cha, *Chem. Phys. Lett.* 773 (2021) 138587.
- [28] A. D. McLean, G. S. Chandler, *J. Chem. Phys.* 72 (1980) 5639.
- [29] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09, Revision D. 01, Gaussian, Inc., Wallingford, CT, 2009.*