

Frozen localized molecular orbitals in electron correlation calculations – Exploiting the Hartree–Fock density matrix

Péter R. Surján *, Dóra Kóhalmi, Zoltán Rolik, Ágnes Szabados

Laboratory of Theoretical Chemistry, Institute of Chemistry, Eötvös University, Budapest P.O. Box 32, H1518 Budapest, Hungary

Received 17 July 2007; in final form 7 November 2007

Available online 17 November 2007

Abstract

The Hartree–Fock density matrix is used to generate occupied and virtual molecular orbitals localized on a selected (active) region within a molecule. The orbitals are well suited for high level description of electron correlation in the active site. Orbitals outside the active site are not constructed explicitly, they provide a frozen core for the correlation calculation. Standard correlation methods are straightforward to apply and result local correlation energies. Transforming to locally canonical orbitals facilitates an iteration-free evaluation of local Møller–Plesset(MP n) energies. Selection of active orbitals does not produce dangling bonds since no chemical bonds are cut at the boundary.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Energetics of chemical reactions affected by long range interactions can often be well described by simplified quantum chemical models. On the other hand, electron correlation, that is also important to describe most reactions, may have short radius of action [1,2]. Intention to exploit this duality are present in several efficient algorithms elaborated for large molecules [3–6]. In the present study we describe long range interactions by performing a Hartree–Fock (HF) calculation for the entire system. This is followed by constructing a set of orbitals localized on an active site. The correlation calculation is performed within the selected set of active localized molecular orbitals (LMOs), leaving remote LMOs uncorrelated (frozen). We term this philosophy the frozen LMO (FLMO) approach.

Previously, Maynau et al. followed an FLMO type strategy [1,2,7] with a special scheme for orbital optimization within a local complete active space (CAS) approach. A closely related method termed effective group potential

(EGP) was developed by Daudey, Heully, and coworkers [8–10]. They describe the inactive part of the system by a suitably chosen pseudo-potential, using the formalism of core potentials. The divide-and-conquer-type method of Li et al. [11] works with several overlapping MO regions which are treated on the same footing. A reduction of their method to consider just one MO region localized on the active site would correspond to the FLMO philosophy. The divide-and-conquer philosophy was also used in combination with the MP2 method [12]. The present FLMO scheme differs from this in that we do not introduce any approximation (fragmentation) at the Hartree–Fock level.

The advantages of using local orbitals in the correlation problem are well known [13–19]. Most local correlation methods aim to reproduce the correlation energy for the total system. In contrast to this we shall focus here on the total HF energy plus the correlation energy of an active site.

In the present communication we consider the situation where, having solved the HF equations by some linear scaling algorithm [20–24], individual MOs are lacking, only the density matrix, P is provided. To compute correlation effects in such a case, one may either turn to formulae which give the correlation energy as an explicit functional

* Corresponding author. Fax: +36 1 372 2592.

E-mail address: surjan@chem.elte.hu (P.R. Surján).

of the density matrix [25,26], or one should develop algorithms to get LMOs from matrix P . An iteration-free algorithm of the latter type is outlined below in combination with the FLMO approach.

2. Theory

Assume that the HF density matrix P for a large molecular system is known, but coefficients of individual MOs $C_{i\mu}$ (i being the MO and μ the AO index) are not available. In local correlation methods, following the suggestion of Pulay [27], virtual LMOs are routinely constructed by projecting atomic orbitals out of the occupied space leading to coefficients

$$C_{i\mu} = \delta_{i\mu} - (PS)_{\mu i}$$

where S is the overlap matrix of AOs. The resulting LMOs are not orthonormal.

By analogy, occupied LMOs φ_i can also be obtained by projecting AOs χ_i into the occupied space:

$$|\varphi_i\rangle = \sum_{\mu} (PS)_{\mu i} |\chi_{\mu}\rangle$$

This definition yields the occupied LMO coefficients

$$C_{i\mu} = (PS)_{\mu i}. \quad (1)$$

In other words, the (nonorthogonal) occupied LMOs are simply the columns of matrix PS . Since localizability of the system is reflected in the sparsity of matrix P , coefficients $C_{i\mu}$ in Eq. (1) define localized orbitals, taken into account that the overlap matrix is also considerably sparse.

The advantage of constructing LMOs by Eq. (1) is that no iterative procedure is needed to maximize a localization functional. The disadvantage is that the resulting LMOs are not only overlapping, but also linearly dependent, since the number of AOs available for projection is much larger than the number of occupied orbitals.

A natural way of getting rid of linear dependence is canonical orthogonalization [28], resulting a smaller but linearly independent set of orthogonal vectors corresponding to nonzero eigenvalues of the overlap matrix. Though canonical orthogonalization produces delocalized MOs, this poses no problem in the FLMO approach, since the MOs become delocalized only within the selected subspace. In other words, we produce *locally delocalized* MOs on the active site.

In the orthonormal LMO basis the effective Hamiltonian for the active site can be written as

$$\hat{H}^{\text{eff}} = \sum_{ij} h_{ij}^{\text{eff}} a_i^{\dagger} a_j + \frac{1}{2} \sum_{ijkl} [ij|kl] a_i^{\dagger} a_j^{\dagger} a_l a_k, \quad (2)$$

where $[ij|kl]$ is a standard two-electron integral in the [12|12] notation, and the effective core that accounts for the field of uncorrelated (frozen) orbitals can be given as

$$h_{ij}^{\text{eff}} = \sum_{\mu\nu} C_{i\mu} h_{\mu\nu}^{\text{eff}} C_{\nu j},$$

with the effective core in the AO basis

$$h_{\mu\nu}^{\text{eff}} = h_{\mu\nu} + \frac{1}{2} \sum_{\lambda\sigma} P_{\lambda\sigma}^{\text{frozen}} [\mu\lambda||\nu\sigma],$$

where $h_{\mu\nu}$ is the standard one-electron integral over AOs, and the double-bar notation is used for the antisymmetrized two-electron AO integral. To avoid explicit construction of frozen LMOs, the frozen core density matrix is computed as the difference $P^{\text{frozen}} = P - P^{\text{active}}$, where the active part of the density matrix is given by

$$P_{\lambda\sigma}^{\text{active}} = \sum_{i \in \text{active}}^{\text{occ}} C_{i\lambda} C_{i\sigma}.$$

This leads to the following expression of the effective core

$$h_{\mu\nu}^{\text{eff}} = F_{\mu\nu} - \frac{1}{2} \sum_{\lambda\sigma} P_{\lambda\sigma}^{\text{active}} [\mu\lambda||\nu\sigma] \quad (3)$$

where F denotes the entire Fockian in the SCF calculation.

Once the effective Hamiltonian is formed, any of the standard methods to compute the correlation energy can be used for the active site. If the method chosen requires canonical MOs (like a standard MP n calculation with a diagonal Fockian), diagonalization of the limited-size active Fockian may be performed to yield a set of *locally canonical* MOs. In this manner local MP n correlation energies can be evaluated without the need of amplitude iterations [13,29] or approximations that change the partitioning [30–33].

In brief, the steps constituting the proposed FLMO procedure are the following:

- (1) Determine P for the whole system in the HF approximation.
- (2) Select the atoms which form the active site.
- (3) Project all AOs of the selected atoms to the occupied and virtual spaces.
- (4) Remove linear dependence of the resulting LMOs by canonical orthogonalization.
- (5) Form the effective Hamiltonian for the active LMOs.
- (6) Perform the electron correlation calculation with this effective Hamiltonian in the subspace of the orthogonalized active LMOs.

The FLMO approach is useful for molecular systems in which a local contribution of the correlation energy is of our interest. The approach is ‘infinitely refinable’ by systematic enlargement of the active site. Selection of active LMOs does not produce dangling bonds since no chemical bonds are cut at the site boundary. The approach is size-consistent and size-extensive if the method used to evaluate the correlation energy within the active site is such. Systems with significant charge transfer can also be treated if the Hartree–Fock description gives a sufficiently accurate account of the charge distribution. Note that dividing the whole system into two disjunct pieces and evaluating the FLMO energy for both of the fragments, the sum of these

two will not be the total correlation energy, since the inter-fragment correlation will still be lacking.

Apart from stressing the usefulness of considering an active site within a large molecule, the message we wish to put through presently is the application of density matrix projected LMOs as defined by Eq. (1) along with their canonical orthogonalization. This concept is extremely simple and fitted for the FLMO framework.

3. Numerical illustrations

The role of the preliminary calculations reported below is merely to illustrate the feasibility and accuracy of the FLMO scheme based on the density-matrix-projected LMOs.

Sample MP2/6-31G calculations have been performed for a zig-zag chain built of 16 HF molecules. The geometry is described by the following parameters: $r(\text{HF}) = 0.97 \text{ \AA}$, $r(\text{FF}) = 2.5 \text{ \AA}$, $\angle(\text{FHF}) = 180^\circ$, $\angle(\text{HFH}) = 116^\circ$. These data are close to the experimental geometry in a crystalline structure [34].

Fig. 1. shows potential curves for the process when the H-F bond length of the terminal HF molecule is varied. The same process is computed for a single HF molecule, denoted ‘one HF in vacuo’. For reference, the full MP2 energy was also obtained for the whole chain denoted by ‘exact MP2’. These curves are to be compared to the FLMO calculation with the elongated HF molecule as the active site ‘FLMO, one active HF’.

The potential curve for the isolated HF molecule is quite different from that of the in-chain curve, while the FLMO method provides a good approximation to the exact result. The latter predicts practically the same equilibrium bond length while in the former case the minimum is shifted by some 0.1 Å.

To get a better view into the accuracy of the FLMO approach, we plotted energy errors with respect to the full MP2 calculation in Fig. 2. MP2 results for two terminal

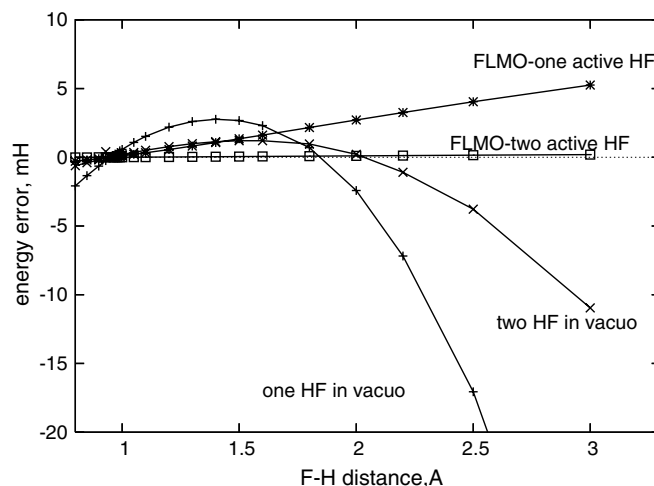


Fig. 2. Energy errors with respect to the MP2 calculation for the full chain of 16 HF units.

HF molecules are also shown (‘two HF in vacuo’). Similarly, the FLMO results with the two terminal HF molecules constituting the active site (‘FLMO – two active HF’) are also plotted. The Figure shows that the FLMO approach improves both (i) the error of the isolated molecule treatment, on the mH scale and (ii) the curvature around equilibrium.

In Table 1 we report dissociation energies corresponding to the process when the terminal H atom is removed from the HF chain. Since MP2 is inapplicable to follow this process, we used a multi-reference treatment. Lengths of non-dissociating bonds were relaxed at the Hartree–Fock level, and a generalized valence bond (GVB) calculation was performed for the active site. To initiate the GVB orbital optimization, the active MOs were subjected to a Boys’ localization procedure. The optimal GVB wave function was used as the reference function for a subsequent multi-configuration perturbation theory (MCPT) calculation [35]. Results are tabulated for one and two isolated HF molecules, respectively, as well as for the whole chain with one, two and three HF molecules constituting the active site. There are two interesting aspects to observe in the table. One is the apparent convergence of FLMO computed approximations as the size of the active site gets

Table 1
Dissociation energies corresponding to the removal of a terminating hydrogen atom in a HF chain (kJ/mol)

System	Method		
	GVB	MCPT2	MCPT3
One HF molecule in vacuo	478	473	461
HF dimer in vacuo	650	540	522
16 HF molecules, FLMO (one active HF molecule)	508	508	492
16 HF molecules, FLMO (two active HF molecules)	417	507	418
16 HF molecules, FLMO (three active HF molecules)	408	506	421

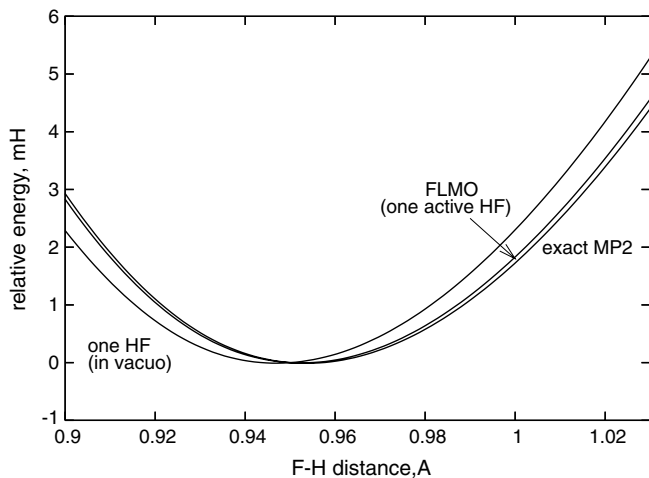


Fig. 1. Potential curves at around equilibrium for the F...H dissociation in a HF chain. The curves are shifted to a common energy minimum.

Table 2

Isomerization in C₁₀H₂₂ with standard geometry $R_{CC} = 1.54 \text{ \AA}$, $R_{CH} = 1.09 \text{ \AA}$. The standard 6-31G basis set is used

Method	E_{tot} (a.u.)		Difference (kJ/mol)
	Decane	2-Methyl nonane	
Hartree–Fock	–391.336359	–391.276447	157
MP2 FLMO [5] ^a	–391.880686	–391.821979	154
MP3 FLMO [5] ^a	–391.933759	–391.874040	157
MP2 FLMO [6] ^b	–391.973428	–391.918171	145
MP3 FLMO [6] ^b	–392.034979	–391.978668	148
MP2	–392.268134	–392.214923	139
MP3	–392.358575	–392.304835	141

^a Active site consists of the first 5 carbon atoms and the hydrogens attached.

^b Active site consists of the first 6 carbon atoms and the hydrogens attached.

enlarged: there is less than 10 kJ/mol difference between 2 and 3 HF molecules forming the active site. The false trend of ‘in vacuo’ calculations with increasing the molecular fragment is also noteworthy, i.e., dissociation energies increase upon stepping from one HF unit to two. This is opposed to the decrease observed for the same step in the FLMO results.

The above demonstration – a quasi one-dimensional H-bonded chain – admittedly represents an ‘easy’ case for the FLMO philosophy. In 3-D structures and if the system is covalently bound much larger active sites may need to be selected to achieve similar accuracy. An example is given by Table 2 where an isomerization energy of an alkane chain is shown with two active spaces of different sizes. The FLMO method is quite successful also in this case, if a sufficiently large active site is used. Though the isomerization energy is not yet significantly improved from its Hartree–Fock value if only a C₅H₁₁ unit is selected, it gives values close to the exact MP2 and MP3 ones if C₆H₁₃ is used as the active site. Note that even for the larger active space and even for this short chain there is a factor of 6.8 in favor of FLMO in terms of computational time for MP3 (the same factor is 4.9 for MP2). For a metallic system the situation is clearly much more difficult, and it is evident that the FLMO approach is not one’s choice when describing collective effects.

The goal of the present applications is simply to show the efficacy of the FLMO approach, utilizing neither any sophisticated computational techniques nor requiring extreme computational resources. More voluminous calculations are in progress and planned to be reported subsequently.

Acknowledgments

This work was supported by the funds OTKA T-04971-NI-67702 and OMFB-01445/2006.

References

- [1] D. Maynau, S. Evangelisti, N. Guihery, C.J. Calzado, J.-P. Malrieu, *J. Chem. Phys.* 116 (2002) 10060.
- [2] C. Angeli, S. Evangelisti, R. Cimraglia, D. Maynau, *J. Chem. Phys.* 117 (2002) 10525.
- [3] F. Bessac, S. Hoyau, D. Maynau, *J. Chem. Phys.* 123 (2005) 104105.
- [4] M. Sironi, A. Genoni, M. Vivera, S. Pieraccini, M. Ghitti, *Theor. Chim. Acta* 117 (2007) 685.
- [5] T.-S. Lee, W. Yang, *Int. J. Quantum Chem.* 69 (1998) 397.
- [6] T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, K. Kitaura, *Chem. Phys. Letters* 351 (2002) 475.
- [7] S. Borini, D. Maynau, S. Evangelisti, *J. Comput. Chem.* 26 (2005) 12.
- [8] F. Alary, R. Poteau, J.-L. Heully, J.P. Daudey, *Theor. Chim. Acta* 104 (2000) 174.
- [9] R. Poteau, J. Ortega, F. Alary, A.R. Solis, J.-C. Barthelat, J.-P. Daudey, *J. Phys. Chem. A* 105 (2001) 198.
- [10] R. Poteau, F. Alary, H.A.E. Makrim, J.-L. Heully, J.-C. Barthelat, J.-P. Daudey, *J. Phys. Chem. A* 105 (2001) 206.
- [11] S. Li, J. Shen, W. Li, Y. Jiang, *J. Chem. Phys.* 125 (2006) 074109.
- [12] M. Kobayashi, T. Akama, H. Nakai, *J. Chem. Phys.* 125 (2006) 204106.
- [13] S. Saebø, P. Pulay, *J. Chem. Phys.* 86 (1987) 914.
- [14] M. Häser, J. Almlöf, *J. Chem. Phys.* 96 (1992) 489.
- [15] P.E. Maslen, M. Head-Gordon, *Chem. Phys. Letters* 283 (1998) 102.
- [16] M. Schütz, G. Hetzer, H.-J. Werner, *J. Chem. Phys.* 111 (1999) 5691.
- [17] P. Ayala, G. Scuseria, *J. Chem. Phys.* 110 (1999) 3660.
- [18] G.E. Scuseria, P.Y. Ayala, *J. Chem. Phys.* 111 (1999) 8330.
- [19] R. Ahlrichs, S. Elliot, U. Huniar, *Ab initio Treatment of Large Molecules*, second ed., in: J. Grotendorst (Ed.), *Modern Methods and Algorithms of Quantum Chemistry*, vol. 3, John von Neumann Institute for Computing, NIC Series, Jülich, 2000, p. 7.
- [20] S. Goedecker, *Rev. Mod. Phys.* 71 (1999) 1085.
- [21] K. Németh, G.E. Scuseria, *J. Chem. Phys.* 113 (2000) 6035.
- [22] M. Challacombe, *J. Chem. Phys.* 110 (1999) 2332.
- [23] X.-P. Li, R.W. Nunes, D. Vanderbilt, *Phys. Rev. B* 47 (1993) 10891.
- [24] D. Kóhalmi, Á. Szabados, P.R. Surján, *Phys. Rev. Letters* 95 (2005) 13002.
- [25] P.R. Surján, *Chem. Phys. Letters* 406 (2005) 318.
- [26] M. Kobayashi, H. Nakai, *Chem. Phys. Letters* 420 (2006) 250.
- [27] P. Pulay, *Chem. Phys. Letters* 100 (1983) 151.
- [28] P.-O. Löwdin, *Adv. Phys.* 5 (1966) 1.
- [29] P. Pulay, S. Saebø, *Theor. Chim. Acta* 69 (1986) 357.
- [30] E.R. Davidson, C.F. Bender, *J. Chem. Phys.* 56 (1972) 4334.
- [31] E. Kapuy, F. Bartha, F. Bogár, Z. Csépes, C. Kozmutza, *Int. J. Quantum Chem.* 37 (1990) 139.
- [32] J.E. Subotnik, M. Head-Gordon, *J. Chem. Phys.* 122 (2005) 034109.
- [33] J. Pipek, F. Bogár, *Topics Curr. Chem.* 203 (1999) 43.
- [34] N. Wiberg, A.F. Holleman, E. Wiberg, *Inorg. Chem.*, vol. IV, Academic, New York, 2001.
- [35] Á. Szabados, Z. Rolik, G. Tóth, P.R. Surján, *J. Chem. Phys.* 122 (2005) 114104.