

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/221817756>

Quantum Molecular Mechanics–A Noniterative Procedure for the Fast Ab Initio Calculation of Closed Shell Systems

ARTICLE *in* JOURNAL OF COMPUTATIONAL CHEMISTRY · APRIL 2012

Impact Factor: 3.59 · DOI: 10.1002/jcc.22921 · Source: PubMed

READS

20

2 AUTHORS, INCLUDING:



Gustavo L.C. Moura

Federal University of Pernambuco

10 PUBLICATIONS 139 CITATIONS

SEE PROFILE

Quantum Molecular Mechanics—A Noniterative Procedure for the Fast *Ab Initio* Calculation of Closed Shell Systems

Gustavo L. C. Moura^[a] and Alfredo M. Simas^{*[a]}

In this article, we advance the foundations of a strategy to develop a molecular mechanics method based not on classical mechanics and force fields but entirely on quantum mechanics and localized electron-pair orbitals, which we call quantum molecular mechanics (QMM). Accordingly, we introduce a new manner of calculating Hartree–Fock *ab initio* wavefunctions of closed shell systems based on variationally preoptimized nonorthogonal electron pair orbitals constructed by linear combinations of basis functions centered on the atoms. QMM is noniterative and requires only one extremely fast inversion of a single sparse matrix to arrive to the one-particle density matrix, to the electron density, and consequently, to the *ab initio* electrostatic potential around the molecular system, or cluster of molecules. Although QMM neglects the smaller polarization effects due to intermolecular interactions, it fully takes into consideration polarization effects due to the much

stronger intramolecular geometry distortions. For the case of methane, we show that QMM was able to reproduce satisfactorily the energetics and polarization effects of all distortions of the molecule along the nine normal modes of vibration, well beyond the harmonic region. We present the first practical applications of the QMM method by examining, in detail, the cases of clusters of helium atoms, hydrogen molecules, methane molecules, as well as one molecule of HeH^+ surrounded by several methane molecules. We finally advance and discuss the potentialities of an exact formula to compute the QMM total energy, in which only two center integrals are involved, provided that the fully optimized electron-pair orbitals are known. © 2012 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.22921

Introduction

Molecular mechanics (MM) has been playing an important role in the modeling of large molecular systems and assemblies. Since its inception, almost 70 years ago,^[1,2] increasingly better, more accurate, and diverse force fields^[3–6] have been helping research in areas such as biochemistry^[7,8] and drug design.^[9,10]

The main difficulty with MM lies in the availability of force fields for the particular system of interest—the need to parameterize them always being a hindrance to a more generic application. Nevertheless, all-atom force fields, both in the commercial and in the public domain, are available, such as CHARMM22,^[11] AMBER03,^[12] OPLS-AA,^[13] and GROMOS-45a3.^[14]

With the increasing speed of computers, larger and larger systems can be now treated by these methods, such as a whole virus (~1 million atoms)^[15] and a 1.5 billion atoms materials simulation.^[16]

Although there have been encouraging attempts to model the electron density in a MM environment by fitting to Gaussian functions,^[17,18] MM does not provide users with a wave function and therefore is limited with respect to its ability in predicting electronic properties.

To address this inherent limitation of MM, methodologies mixing quantum mechanics (QM) with MM,^[19] known as QM/MM,^[20,21] are increasingly popular due to their ability to correctly address many problems of interest,^[22] mainly in biochemistry and pharmacology applications.^[23] In QM/MM, the system of interest is divided into two regions: the smaller, where the chemical interaction takes place, is treated by QM,

whereas the remaining region is treated by MM—the interface being the subject of coalescence to integrate one region into the other in a seamless manner. Within QM/MM, the idea of using QM as a framework for molecular dynamics was first proposed by Gao in 1997^[24] and demonstrated in 1998.^[25] Moreover, the use of QM as a force field has been implemented in the electronic structure based force field called the explicit polarization potential, X-Pol,^[6] and used to carry out molecular dynamics simulations on a protein.^[26]

Nevertheless, for systems like a single protein, conditions are now set for the emergence of another type of MM, based on neither classical nor semiclassical force fields, but now based exclusively on QM. That is for the advent of a quantum MM (QMM) method.

The purpose of QMM, advanced in this article, is also to replace, in the course of time, the role of MM in QM/MM, thus eventually giving rise to QM/QMM methodologies. As such, the more delicate intramolecular and intermolecular interactions in the small part of the system that is of major interest can be modeled by high-level *ab initio* methods, leaving QMM to model the remaining system.

Thus, QMM should be based on the Born–Oppenheimer separation, should be noniterative, should be linearly scalable,

[a] G. L. C. Moura, A. M. Simas

Departamento de Química Fundamental, CCEN, Universidade Federal de Pernambuco, Recife, Pernambuco 50590-470, Brazil
E-mail: simas@ufpe.br

Contract/grant sponsor: CNPq; Contract/grant sponsor: FACEPE (Pronex); Contract/grant sponsor: INCT INAMI; Contract/grant sponsor: RENAMI

© 2012 Wiley Periodicals, Inc.

and should reproduce well the results and trends of regular quantum chemistry calculations.

The achievement of such an attractive set of properties comes at a price, though: the orbitals must be known in advance. Accordingly, like its MM classical counterpart, QMM must be parameterized for the systems of interest. However, the success of classical MM seemingly substantiates the expectation that the QMM orbitals, parameterized in the specific manner described in this article, will indeed be transferrable from molecule to molecule.

In subsequent articles, we will explore the introduction of electron correlation and dispersion interactions into QMM. But at first, to establish its foundations, we will build QMM as an *ab initio* method equivalent to a Hartree–Fock (HF) calculation for the purpose of introducing the basic methodology for the calculation of molecular closed shell systems. Finally, we exemplify its potential, by reporting sample calculations on systems containing several helium atoms, hydrogen molecules, methane molecules, and a single molecule of HeH^+ embedded in a cluster of methane molecules.

As will be shown, to obtain the complete QMM wave function from the parameterized orbitals, it suffices to invert a single sparse matrix.

Method

Of central importance to the QMM model is matrix \mathbf{T} , defined as

$$\mathbf{T} = \mathbf{S}^{-1}, \quad (1)$$

where \mathbf{S} is the overlap matrix between normalized nonorthogonal doubly occupied spatial orbitals ϕ ,

$$S_{\mu\nu} = \int \phi_\mu^*(r) \phi_\nu(r) dr, \quad (2)$$

where r represents the spatial coordinates of the electron. For large systems, \mathbf{S} is a sparse matrix, and the algorithm for inverting a sparse matrix is linearly scalable.^[27]

Equation (1) follows naturally when we relate the nonorthogonal spatial orbitals ϕ to a set of orthogonal spatial orbitals χ obtained using Lowdin's symmetric orthogonalization^[28]

$$\chi = \phi \mathbf{S}^{-\frac{1}{2}} = \phi \mathbf{T}^{\frac{1}{2}} \quad (3)$$

For a monodeterminantal N -electron wave function, constructed from orthogonal doubly occupied spatial orbitals χ , it is well known that the electron density is given by

$$\gamma(r) = 2 \sum_{\mu}^{N/2} \chi_\mu^*(r) \chi_\mu(r) \quad (4)$$

Combining eqs. (3) and (4), it can be easily shown that the electron density can be written in terms of ϕ as^[29]

$$\gamma(r) = 2 \sum_{\mu,\nu}^{N/2} \phi_\mu^*(r) \phi_\nu(r) T_{\mu\nu}. \quad (5)$$

Therefore, we recognize that matrix $2\mathbf{T}$ is simply an electronic density matrix in a nonorthogonal basis set, which can be easily calculated provided that the orbitals ϕ are known.

The energy of a single Slater determinant constructed from nonorthogonal spin orbitals was studied by Lowdin.^[28] Within the Born–Oppenheimer separation, for a system with M nuclei, described by a set of N nonorthogonal normalized electronic spin orbitals ϕ , the total energy is given, in atomic units, by

$$E_{\text{tot}} = \sum_{\mu,\nu}^N \langle \mu | H_{\text{core}} | \nu \rangle T_{\mu\nu} + \sum_{\mu,\nu,\kappa,\lambda}^N \langle \mu\nu | G | \kappa\lambda \rangle [T_{\mu\nu} T_{\kappa\lambda} - T_{\mu\lambda} T_{\nu\kappa}] + \sum_{g < h}^M \frac{Z_g Z_h}{|R_g - R_h|}, \quad (6)$$

where $\langle \mu | H_{\text{core}} | \nu \rangle$ is a one-electron integral and $\langle \mu\nu | G | \kappa\lambda \rangle$ is a two-electron integral. The one-electron integrals are given by

$$\langle \mu | H_{\text{core}} | \nu \rangle = \int \phi_\mu^* H_{\text{core}} \phi_\nu dx, \quad (7)$$

with H_{core} in atomic units, being

$$H_{\text{core}} = \frac{p^2}{2} - \sum_g^M \frac{Z_g}{|R_g - r|}, \quad (8)$$

where x are the spatial and spin coordinates of the electron, p is the linear momentum operator, Z_g is the atomic number of nucleus g and R_g is the nuclear position. The two-electron repulsion integrals, also in atomic units, are given by

$$\langle \mu\nu | G | \kappa\lambda \rangle = \int \phi_\mu^*(1) \phi_\nu(1) r_{12}^{-1} \phi_\kappa^*(2) \phi_\lambda(2) dx_1 dx_2 \quad (9)$$

where r_{12} is the interelectronic distance.

When we use the constraint of same spatial nonorthogonal orbitals for electrons with different spins (doubly occupied orbitals), eq. (6) becomes^[30]

$$E_{\text{tot}} = 2 \sum_{\mu,\nu}^{N/2} \langle \mu | H_{\text{core}} | \nu \rangle T_{\mu\nu} + \sum_{\mu,\nu,\kappa,\lambda}^{N/2} \langle \mu\nu | G | \kappa\lambda \rangle [2T_{\mu\nu} T_{\kappa\lambda} - T_{\mu\lambda} T_{\nu\kappa}] + \sum_{g < h}^M \frac{Z_g Z_h}{|R_g - R_h|}, \quad (10)$$

where μ , ν , κ , and λ , instead of indexing spin-orbitals, as in eq. (6), are now indexing the nonorthogonal spatial orbitals ϕ .

Lowdin^[31] showed that, for a Hamiltonian of the form

$$H = \sum_n^N \left(\frac{p_n^2}{2} - \sum_g^M \frac{Z_g}{|R_g - r_n|} \right) + \sum_{n < m}^N \frac{1}{r_{nm}} + \sum_{g < h}^M \frac{Z_g Z_h}{|R_g - R_h|}, \quad (11)$$

the total energy of this system is given by

$$E_{\text{tot}} = \int \left(\frac{p_1^2}{2} - \sum_g^M \frac{Z_g}{|R_g - r_1|} \right) \gamma^s(1'|1) dx_1 + \int \frac{\Gamma^s(1', 2'|1, 2)}{r_{12}} dx_1 dx_2 + \sum_{g < h}^M \frac{Z_g Z_h}{|R_g - R_h|}, \quad (12)$$

where $\gamma^s(1'|1)$ and $\Gamma^s(1', 2'|1, 2)$ are, respectively, the first- and second-order reduced density matrices in terms of spin-orbitals. Comparing eqs. (12) and (10), for the special case of

doubly occupied nonorthogonal spatial orbitals, the spinless reduced density matrices are given by

$$\gamma(1'|1) = 2 \sum_{\mu,\nu}^{N/2} \phi_{\mu}^*(r'_1) \phi_{\nu}(r_1) T_{\mu\nu} \quad (13)$$

and

$$\Gamma(1', 2'|1, 2) = \sum_{\mu,\nu,\kappa,\lambda}^{N/2} \phi_{\mu}^*(r'_1) \phi_{\nu}(r_1) \phi_{\kappa}^*(r'_2) \phi_{\lambda}(r_2) \times [2T_{\mu\nu}T_{\kappa\lambda} - T_{\mu\lambda}T_{\nu\kappa}] \quad (14)$$

Let us, for a moment, assume that we already have a set of optimized doubly occupied orbitals ϕ . This fact alone provides us with matrix **T**, easily obtained as the inverse of the overlap matrix **S** for these orbitals. Therefore, we have already the molecular electronic density and from that we can quickly obtain the electrostatic potential around the molecule, as well as its multipole moments. This calculation can be performed without the need to calculate the very costly one- and two-electron integrals that, in the case of the HF method, are needed to obtain the molecular orbitals.

So, if we already have a set of preoptimized spatial orbitals, it is a trivial task to construct the electronic density and obtain several other dependent molecular properties. All we need to do is to calculate the inverse **T** of the overlap matrix **S**. When needed, the energy can be calculated using eq. (10).

One method that uses floating basis sets and that also uses eq. (10) to calculate total energies is the floating spherical Gaussian orbital (FSGO) model.^[30] In the FSGO model, each doubly occupied spatial orbital ϕ is represented by a single spherical Gaussian function

$$G_i(\alpha_i, X_i) = \left(\frac{2\alpha_i}{\pi}\right)^{\frac{3}{4}} \exp[-\alpha_i(r - X_i)^2], \quad (15)$$

and both the exponent α_i and the spatial position X_i of each spherical Gaussian are fully optimized to minimize the energy of the system. However, the simultaneous nonlinear optimization of all α_i and X_i is a very complex and costly procedure. Through the optimization of X_i , the orbitals of the FSGO model are allowed to float around the system, uncoupled to the positions of the nuclei (R_g). A double-Gaussian modification of the FSGO model was also proposed.^[32] In this model, instead of a single spherical Gaussian function, each spatial orbital is expanded as a linear combination of two concentric spherical Gaussian functions.

A technique which also uses localized orbitals to calculate large systems is Stewart's semiempirical local molecular orbital (LMO)^[33,34] approach. However, there are three major differences between our QMM method and Stewart's LMO approach for proteins. While our QMM method uses only the strictly necessary number of occupied orbitals, the LMO approach also includes the unoccupied orbitals necessary to complete the number of valence orbitals of the molecule. Furthermore,

Stewart uses the usual ZDO approximation of semiempirical methods, that is, the LMOs are assumed orthogonal to each other, whereas in QMM the localized orbitals are nonorthogonal by construction. Finally, in QMM there is no iteration whatsoever to obtain the wave function (only a matrix inversion) and the preoptimized occupied orbitals are strictly localized, while in Stewart's LMO approach there is an iterative process to annihilate the interaction between the occupied and unoccupied LMOs. During this iterative process, the initially monoatomic and diatomic LMOs become somewhat delocalized, though not as delocalized as the conventional molecular orbitals resulting from the full diagonalization of the Fock matrix.

By contrast, in QMM we strive to develop a method where a reasonably accurate wave function is known *a priori* in such a way that we can obtain the properties of a molecule without resorting to a full optimization of the wave function each time we perform a calculation. To that end, we need to consider the conditions in which we could have a preoptimized wave function for molecular calculations.

Because the exact wave function will have to obey the electron-nuclear Kato cusp condition,^[35] it is reasonable to assume that, considering a molecular orbital made up of a linear combination of either Slater or Gaussian functions, as the basis set increases in size, the functions in the fully optimized linear combination, originally located in between the bond, will move apart toward the nuclei. Despite being a reasonable assumption, we are not aware of any previous work that has determined this conjecture to be a fact. In the Supporting Information of this article, we show that to be true for the cases of H_2 and HeH^+ molecules.

Based on the results described above, in QMM we adopt the strategy of preoptimizing the exponents and the linear combination coefficients of the primitives of each of the non-orthogonal doubly occupied orbitals describing the bonding on the ground state of closed-shell molecules. We also address the problem of optimizing the spatial positions of the basis functions. Because the primitives converge to the positions of the nuclei, instead of using floating basis functions, we use a linear combination of s and, when needed, p Gaussian primitives centered on the nuclei to describe the doubly occupied orbitals ϕ . To truly have accurate wave functions, it is very important that the doubly occupied spatial orbitals are fully optimized to describe the bonding for all nuclear configurations, otherwise the response of the electron density to variations in geometry would be completely neglected. With that in mind, the philosophy behind the QMM model is very similar to that involved in a typical MM calculation. As in MM, where we need to have force fields for all situations, in QMM we also need to have preoptimized pair orbitals for all types of situations.

In QMM, depending on their nature, each core orbital is expanded in terms of only s or only p primitives centered on a single nucleus. In the case of saturated molecules, each sigma bond orbital is constrained to be composed of primitives centered on only two nuclei. In this case, the p primitives used to describe the orbitals are aligned parallel to the internuclear axis. The strategy to describe nonbonded lone pairs is similar

to the one used to describe sigma bonds. We use a dummy atom that does not possess any basis functions with the sole purpose of giving directionality to the lone pair. In the case of nonconjugated pi bonds, the bond orbitals are constrained to be composed of only p primitives centered on only two atoms and placed orthogonally to the internuclear axis.

Hydrogen bonding is very important in biomolecules. In QMM, we have the option of adopting no special measures to describe such interactions and let them be described as a simple electrostatic interaction between a lone pair and a nearby hydrogen atom. Or, perhaps, we may need to parameterize a special type of sigma bonding orbital to describe such interactions. The description of such bonds within the framework of the QMM model is still a matter of research. The description of molecules containing conjugated delocalized bonds in the QMM model, such as in the case of aromatic rings, is still under evaluation. We may need to allow the pi bonding orbitals to spread over several atoms, or we may need to adopt a resonating picture, where such systems are described as a superposition of several resonance structures. This is also a matter of research.

In short, the steps involved in a QMM calculation are:

1. Given the geometry of the molecule, assign to it the fully preoptimized orbitals ϕ , and calculate the overlap matrix **S**.
2. Obtain **T**, the inverse of the overlap matrix.
3. If you want the electrostatic potential, simply compute it using **T** via eq. (5), and stop.
4. However, if you want the energy, calculate the one- and two-electron integrals as needed, and use them only once in eq. (10).

As we can see, during a QMM calculation we need storage space only for matrices **S** and **T**. For large molecules, the overlap matrix is sparse, and we can use this fact to reduce the storage requirements of this matrix. The matrix **T**, however, is not sparse, and we need to store all of its elements.

There are two important advantages of the QMM method over the HF method. First, steps 1 and 2 of the QMM algorithm involve quick and easy calculations of **S** and **T**. Therefore, unlike the HF method, which often encounters convergence problems to calculate the molecular orbitals, the QMM method has no convergence issues (there is no SCF condition to be achieved). The second advantage is that the QMM method has a lower cost, in terms of both CPU and memory usage. That is because, for a QMM calculation of the energy, each one- and two-electron integral needs to be evaluated only once, and there is no need to store them for later use. Therefore, the QMM method allows the calculation of much larger systems for a given hardware configuration.

It is important to stress that, in its present form, QMM does not include electron correlation in the calculation and therefore, provides results similar to HF. One approach that could allow the inclusion of correlation effects in QMM would be to replace the preoptimized HF bonding orbitals by preoptimized Kohn–Sham bonding orbitals within the framework of density functional theory. An appropriate functional would need to be selected, and the preoptimization of the bonding orbitals

should then be performed using such functional. Another effect that is still absent in this implementation of QMM is the inclusion of dispersion. Such effects could, for example, be included in a QMM calculation using corrections, for example, through the effective fragment potential method.^[36,37]

In its present form, QMM does not yet have analytical first and second derivatives with respect to the nuclear positions. Once such derivatives are implemented, it will be possible to perform easier geometry optimizations, normal mode analyses, and obtain thermodynamic properties in an efficient manner.

Results and Discussion: First Practical Implementation of QMM

A key step in the development of the QMM method is the parameterization of the nonorthogonal doubly occupied orbitals describing each bond in the molecule we want to study. This parameterization involves the preoptimization of the exponents, and the linear combination coefficients of the Gaussian primitives used to expand each bonding orbital. Therefore, we need to select small representative molecules containing each type of bond we wish to study and optimize the parameters of the orbitals describing such bonds. In this optimization, we use eq. (10) to calculate the total energy and variationally minimize this energy as a function of the orbital parameters. In our program, the molecular integrals were obtained using the McMurchie–Davidson scheme.^[38]

As such, we report the first implementation of the QMM method: initially for the helium atom; subsequently for the hydrogen molecule; and, finally, we extend QMM to treat the methane molecule.

Accordingly, we now describe the case of the helium atom. Naturally, we optimized orbitals for this atom containing only

Table 1. Exponents (α) and linear coefficients (c) for the occupied orbital of the helium atom expanded using three, four, five, and six s primitives.

Parameter	Number of Gaussian primitives			
	He(3s)	He(4s)	He(5s)	He(6s)
α_1	13.623099	38.354933	98.078311	234.063402
α_2	1.998943	5.768907	14.764405	35.174020
α_3	0.382943	1.239941	3.318583	7.991109
α_4	–	0.297578	0.874139	2.212422
α_5	–	–	0.244599	0.667069
α_6	–	–	–	0.208947
c_1	0.122125	0.046420	0.019070	0.008401
c_2	0.622604	0.301958	0.137981	0.063421
c_3	1.000000	0.916115	0.555321	0.295542
c_4	–	1.000000	1.226207	0.883574
c_5	–	–	1.000000	1.552789
c_6	–	–	–	1.000000
E_{tot} (a.u.)	–2.835680	–2.855160	–2.859895	–2.861153
Virial ratio	1.99999999	2.00000003	1.99999999	2.00000001

Due to the normalization constraint, the linear coefficients are given in relation to the most diffuse primitive of the orbital. They need to be normalized before being used in a calculation. Also given are the total energies and the virial ratios obtained using each of the orbital expansions. For comparison, the total energy obtained from an LCAO aug-cc-pV5Z calculation is equal to –2.861543 a.u.

s primitives. In Table 1, we have the exponents and linear coefficients for the orbitals expanded using three, four, five, and six primitives. Taking into account that, due to the normalization constraint, the orbital parameters are not all independent, the linear coefficients reported in Table 1 are given in relation to the most diffuse s primitive of the orbital. The first step of a calculation that uses these orbitals is to normalize them. As expected, in Table 1, we have that the total energy of the helium atom decreases, as we increase the number of variational parameters (for six s primitives $E_{\text{tot}} = -2.861153$ a.u.). We also note that the virial ratios obtained using these orbitals are very close to 2. The total energies also compare really well with the one from a large basis set HF LCAO calculation performed with the aug-cc-pV5Z basis set ($E_{\text{tot}} = -2.861543$ a.u.) that contains 9s, 5p, 4d, 3f, and 2g Gaussian primitives.

An important feature of the QMM method is that it needs to be able to give reasonable descriptions of noncovalent interactions. Therefore, we have prepared 27 clusters of helium containing 27 atoms each. These clusters were prepared by first placing the atoms on a regular cubic grid with a separation of 2 Å, and then applying a random Gaussian perturbation, with zero mean and standard deviation of 0.1 Å in each of the three directions, to the position of each atom in such a way as to break the cubic symmetry of the cluster. In some of the clusters, the minimum distance between the helium atoms was as small as 1.565 Å, resulting in strong interactions between nearby atoms. As a reference for the interaction energies, in the absence of electronic correlation effects, we performed HF LCAO calculations with the 6-31G basis set for each cluster. The total QMM energy was calculated using the preoptimized orbital with four s primitives shown in Table 1. For comparison, we also performed QMM-like calculations (ones that do not preoptimize the occupied orbitals) using doubly occupied orbitals with the exponents 34.519439, 6.345798, 1.115947, and 0.327336, and the same linear coefficients used in the preoptimized orbital.

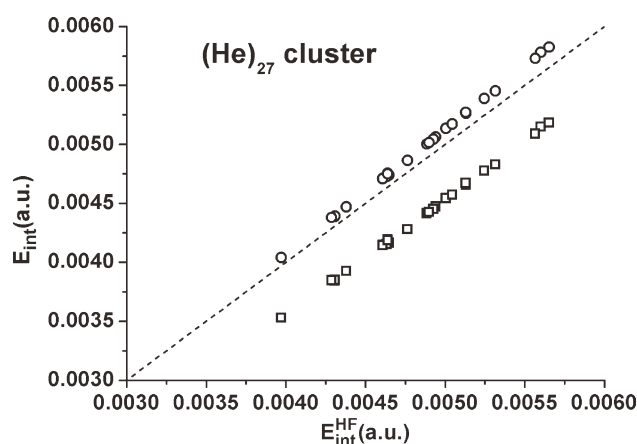


Figure 1. Comparison between the interaction energies per atom, for clusters of 27 helium atoms, calculated with the HF/6-31G method with the ones calculated using the QMM method (circles) and a QMM-like method with not fully optimized orbitals (squares). The dashed line has slope one.

Table 2. Exponents (α) and linear coefficients (c) of the primitives, centered on each hydrogen atom, for the symmetric occupied orbital of H_2 optimized at the equilibrium bond length.

Primitive	α	C
p	1.026825	0.079555
1s	15.983453	0.053171
2s	2.404039	0.362192
3s	0.550756	1.144306
4s	0.150582	1.000000

Due to the normalization constraint, the linear coefficients are given in relation to the most diffuse s primitive of the orbital. They need to be normalized before being used in a calculation.

In Figure 1, we compare the interaction energies per atom calculated with the HF/6-31G method with the ones calculated using the QMM and QMM-like methods. We can easily see that the interaction energies obtained with the QMM method are much closer to the LCAO results than the ones obtained with a QMM-like method: a method that does not preoptimize the occupied orbitals. The interaction energies calculated using the QMM method are slightly higher by 0.00012 a.u. per atom, equivalent to 0.075 kcal mol⁻¹, when compared to the ones obtained using the HF/6-31G method. But this is not surprising, considering that in the HF calculations the clusters are stabilized due to the basis set superposition error (BSSE). The QMM method is free from this sort of error.

Our next implementation of the QMM method was carried out for the study of the hydrogen molecule. Here, we have optimized the bonding orbital of this molecule using, for each atom, four s primitives and one p primitive parallel to the internuclear axis (the positive lobe of the p primitive, centered in one nucleus, points to the direction of the other nucleus). We have first performed the simultaneous optimization of the bonding orbital parameters (exponents and linear coefficients) and the internuclear distance to obtain an optimized bond length. In Table 2, we have the exponents and linear coefficients of the bonding orbital optimized for the optimized bond length. In Table 3, we compare the optimized properties of the hydrogen molecule obtained using the QMM method with the ones obtained from HF LCAO calculations. We can observe that the QMM method gives an optimized bond length (1.3870 a.u.) that is very similar to the one obtained with the large LCAO basis set aug-cc-pV5Z (1.3863 a.u.). The

Table 3. Optimized bond lengths (R_{opt}) and total energies of the H_2 molecule calculated using the QMM method and other HF LCAO calculations.

Method—basis set	R_{opt} (a.u.)	E_{tot} (a.u.)
QMM— H_2 (4s1p, 4s1p)	1.3870	-1.131590
LCAO—6-31G ^[a]	1.3795	-1.126828
LCAO—6-31G(d,p) ^[a]	1.3844	-1.131334
LCAO—aug-cc-pVDZ ^[a]	1.4137	-1.128826
LCAO—aug-cc-pVQZ ^[a]	1.3865	-1.133509
LCAO—aug-cc-pV5Z	1.3863	-1.133648

In QMM, the orbital was expanded using four s and one p primitives centered on each hydrogen atom.

[a] From <http://cccbdb.nist.gov/>.

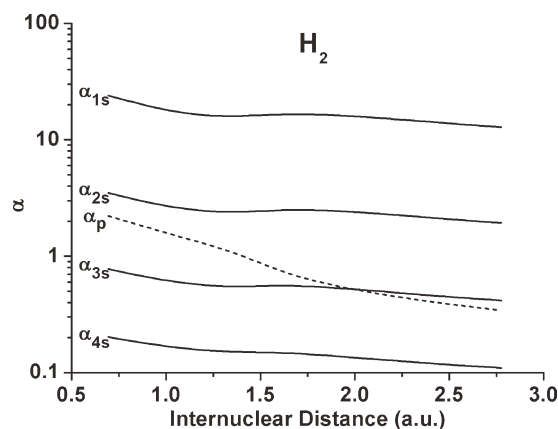


Figure 2. Exponents, α , of the preoptimized occupied orbital of the hydrogen molecule plotted as a function of the internuclear distance. The solid lines are for the s primitives and the dashed line for the p primitive. The raw values are present in Table S5 of the Supporting Information.

bond length comparison with an LCAO basis set of similar size (6-31G(d,p), which for the hydrogen atom does not have the d primitives), is also favorable (1.3844 a.u.). The total energy obtained using the QMM method is -1.131590 a.u., lower than the one obtained from a HF LCAO calculation with the 6-31G(d,p) basis set of -1.131334 a.u.

Following our MM-like approach for the QMM model, where we need to have preoptimized bonding orbitals for all nuclear configurations, we next optimized the exponents and linear coefficients of the bonding orbital of H_2 for internuclear distances in the interval between half and twice the optimized bond length. In Table S5 of the Supporting Information, we have a list of these parameters as a function of the internuclear distance. In Figures 2 and 3, we have, respectively, the values of the exponents and the linear coefficients plotted as a function of the internuclear distance.

Someone may ask if the investment in preoptimizing the bonding orbital in the QMM method is really worthwhile. Now, we show that this step is important to allow the system to

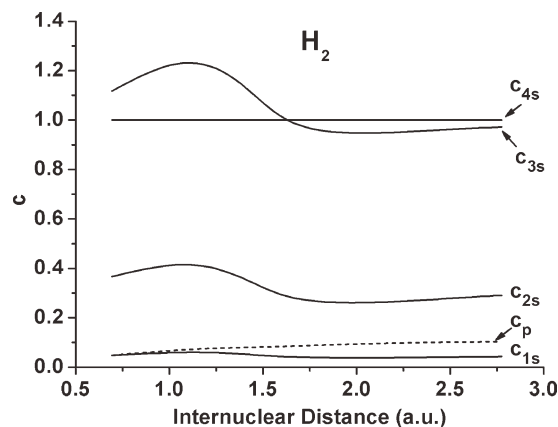


Figure 3. Linear coefficients, c , of the preoptimized occupied orbital of the hydrogen molecule plotted as a function of the internuclear distance. The linear coefficients are given in relation to the most diffuse s primitive of the orbital, that is, $c_{4s} = 1$. The solid lines are for the s primitives and the dashed line for the p primitive. The raw values are present in Table S5 of the Supporting Information.

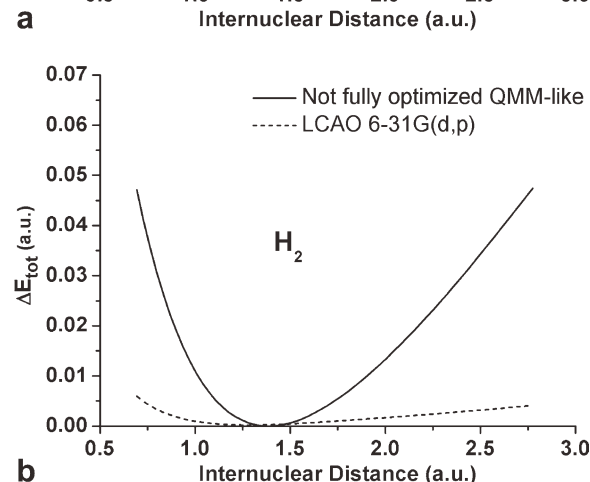
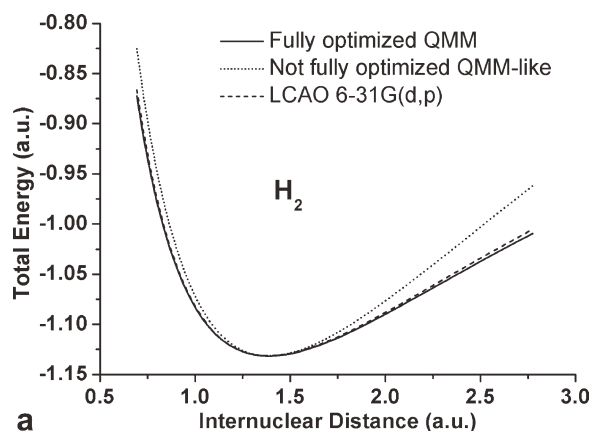


Figure 4. a) Total energy for the hydrogen molecule, as a function of the internuclear distance, obtained using three approaches: QMM calculations using fully optimized orbitals (solid line), QMM-like calculations using not fully optimized orbitals (dotted line) and 6-31G(d,p) HF LCAO calculations (dashed line). b) Difference in total energy between the QMM calculation with a fully optimized orbital and (solid line) QMM-like calculation without full optimization of the orbital and (dashed line) 6-31G(d,p) HF LCAO calculation. Data for generating this figure is present in Table S6 of the Supporting Information.

properly polarize, when it distorts from the equilibrium geometry. In Figure 4a, we compare the total energy for the hydrogen molecule, as a function of the internuclear distance, obtained using two approaches: the QMM method where the bonding orbital is fully optimized and a QMM-like method where we use the orbital optimized only for the equilibrium bond length (these data are shown in Table S6 of the Supporting Information). In Figure 4b, we have the difference in energy obtained using these two sets of orbitals. We can clearly see that, far from the equilibrium bond length, the energy obtained using the orbital that was not optimized as a function of the internuclear distance is considerably higher than the energy obtained using the fully optimized orbital. Therefore, we have that in the QMM method, it is very important to provide fully optimized orbitals. For comparison, in Figures 4a and 4b, we also have HF LCAO results obtained using the 6-31G(d,p) basis set. The energy difference between the fully optimized QMM calculation and the HF calculation is one order of magnitude smaller than the one for the case of using a not fully optimized QMM-like method. Although we are

using a fixed basis set in the LCAO calculation, the split valence nature of the 6-31G(d,p) basis set allows the HF orbital to better adjust to the occupied orbital. Finally, we may conclude that the QMM method is capable of yielding excellent results for the hydrogen molecule for all internuclear distances studied.

In the case of systems away from their equilibrium geometries, we cannot use the traditional virial ratio. However, please consider the virial theorem demonstrated by Slater,^[39] originally for diatomic molecules, but generalizable to a polyatomic system within the Born–Oppenheimer separation as:

$$E_{\text{pot}} = -2E_{\text{kin}} - \sum_g^M \sum_{i=1}^3 R_{gi} \frac{\partial E_{\text{tot}}}{\partial R_{gi}}, \quad (16)$$

where R_{gi} is the i th Cartesian coordinate of nucleus g . In eq. (16), E_{pot} is the potential energy contribution to the total energy, which contains the electron–nuclei attraction (a one-electron contribution), electron–electron repulsion (a two-electron contribution), and nuclear repulsion (a classical contribution). The kinetic energy contribution (E_{kin}) is a one-electron contribution that can be obtained without much effort (when compared with the two-electron integrals). The term $\frac{\partial E_{\text{tot}}}{\partial R_{gi}}$ is the negative of the i th coordinate of the force acting on nucleus g .

Thus, an extended virial ratio can be defined as:

$$\text{virial ratio} = - \frac{E_{\text{pot}} + \sum_g^M \sum_{i=1}^3 R_{gi} \frac{\partial E_{\text{tot}}}{\partial R_{gi}}}{E_{\text{kin}}}. \quad (17)$$

In Table S7 of the Supporting Information, we present the extended virial ratio, as a function of the internuclear distance, for the hydrogen molecule, and in Figure 5, we have a plot of these data. We can observe that, in the case of QMM calculations, all the extended virial ratios are very close to 2. However, in the case of the QMM-like calculations using not fully optimized orbitals, there were sizeable deviations from the

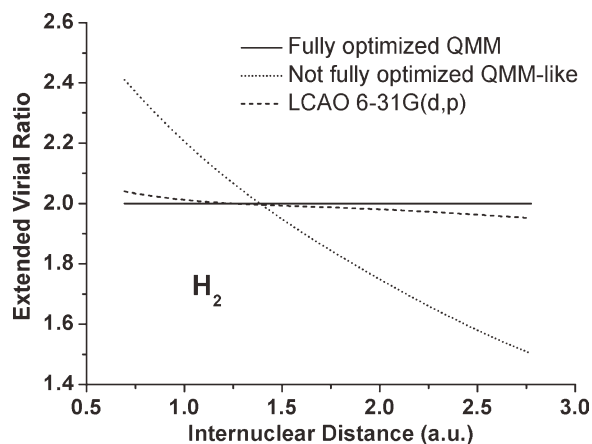


Figure 5. Extended virial ratio for the hydrogen molecule, as a function of the internuclear distance, obtained using three approaches: (solid line) QMM calculations; (dotted line) QMM-like calculations using not fully optimized orbitals; and (dashed line) 6-31G(d,p) HF LCAO calculations. Data for generating this figure are present in Table S7 of the Supporting Information.

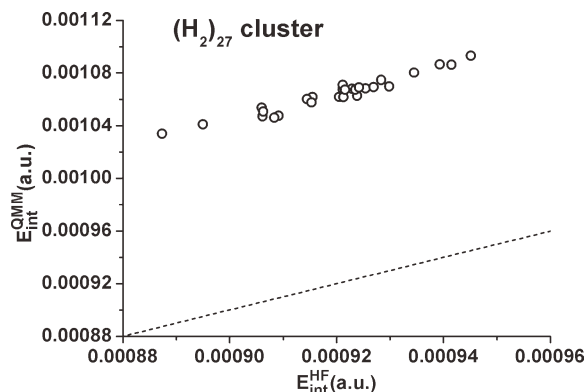


Figure 6. Comparison between the interaction energies per molecule, for clusters of 27 molecules of H_2 , calculated with the HF/6-31G(d,p) method with the ones calculated using the QMM method. The dashed line has slope one.

value of 2 far from the equilibrium bond length. The results obtained from HF LCAO calculations using the 6-31G(d,p) basis set are also not fully satisfactory from the perspective of the virial ratio.

The next step was to assess the quality of the interaction energies of the QMM method using 27 clusters with 27 H_2 molecules each. Unlike the clusters of helium atoms, where we had only translational degrees of freedom, in the case of the H_2 molecule we also had to consider the rotational degrees of freedom of each molecule. We chose to disregard the translational degrees of freedom when preparing the clusters of H_2 . The internuclear distance of each molecule of the clusters was fixed at the equilibrium bond length. Each molecule was then randomly rotated in space, and its center of mass placed on a regular cubic grid with a separation of 3.0 Å.

In Figure 6, we compare the interaction energies per molecule calculated with the HF/6-31G(d,p) method with the ones calculated using the QMM method. We can observe that the QMM method yields interaction energies that are higher when compared to the HF/6-31G(d,p) method by only 0.00014 a.u. per molecule, equivalent to $0.088 \text{ kcal mol}^{-1}$. This slight displacement may be attributed to a combination of BSSE with the fact that the HF/6-31G(d,p) calculation includes p primitives that are orthogonal to the axis of the H_2 molecule. These nonaxial p primitives, that are absent in the QMM method, contribute to the stabilization of the clusters in the HF calculations.

Finally, the QMM method was implemented for the study of the methane molecule. This molecule has two types of occupied orbitals: one core orbital in the carbon and four symmetry equivalent CH bonding orbitals. We chose to expand the core orbital of the carbon atom using six s primitives and each CH bonding orbital using four s primitives and four p primitives centered on the carbon, and four s primitives centered on the hydrogen. First, we performed the simultaneous optimization of the orbital parameters (exponents and linear coefficients) and the CH internuclear distance to obtain an equilibrium bond length. In Table 4, we have the exponents and linear coefficients of the orbitals of the methane molecule for

Table 4. Exponents (α) and linear coefficients (c) for the occupied orbitals of the methane molecule optimized at the CH optimized bond length.

Orbital	Parameter	α	c
Carbon core	1s	2720.123585	0.006802
	2s	408.559744	0.051854
	3s	92.962048	0.251277
	4s	26.104253	0.824608
	5s	8.255769	1.574281
	6s	2.755212	1.000000
CH bonding	C-1s	63.838963	-0.003289
	C-2s	15.524491	-0.033791
	C-3s	4.629047	-0.177726
	C-4s	0.395453	1.421425
	C-1p	11.645746	0.058286
	C-2p	2.495355	0.327410
	C-3p	0.650327	0.975977
	C-4p	0.189676	0.792381
	H-1s	22.362545	0.032393
	H-2s	3.382619	0.227930
	H-3s	0.705203	1.025966
	H-4s	0.158927	1.000000

Due to the normalization constraint, the linear coefficients are given in relation to the most diffuse s primitive of the orbital. They need to be normalized before being used in a calculation.

the optimized bond length. In Table 5, we compare the optimized properties of the methane molecule, obtained using the QMM method, with the ones from HF LCAO calculations. We can observe in Table 5 that, except when using the aug-cc-pVDZ basis set in the LCAO calculation, the QMM optimized CH bond length compares really well with the values obtained with the HF method.

The next step of our study of methane was to optimize the occupied orbitals as a function of the molecular geometry. To strictly obey our MM-like approach for the QMM model, where we would need to have preoptimized bonding orbitals for all nuclear configurations, we would need, in the case of polyatomic molecules, to perform a very complicated multidimensional optimization of the bonding orbitals as a function of all bond stretches, bending and torsions. However, for practicality, as a first attempt, we will concentrate only on the effects that bond stretches have on the bonding orbitals. This first

Table 5. Comparison between the CH optimized bond length (R_{opt}) and total energy (E_{tot}) of the CH_4 molecule calculated using the QMM method and the results of HF LCAO calculations.

Method—basis set	R_{opt} (a.u.)	E_{tot} (a.u.)
MQM—C(6s), CH(4s4p, 4s)	2.0434	-40.181669
LCAO—4-31G	2.0430	-40.139767
LCAO—6-31G ^[a]	2.0449	-40.180554
LCAO—6-31G(d,p) ^[a]	2.0475	-40.201705
LCAO—aug-cc-pVDZ ^[a]	2.0587	-40.199633
LCAO—aug-cc-pVQZ ^[a]	2.0449	-40.216345

In the QMM calculation, the core orbital of the carbon atom was expanded using six s primitives. Each of the four symmetry equivalent CH bonding orbitals were expanded using four s primitives and four p primitives centered on the carbon, and four s primitives centered on the hydrogen.

[a] From <http://cccbdb.nist.gov/>.

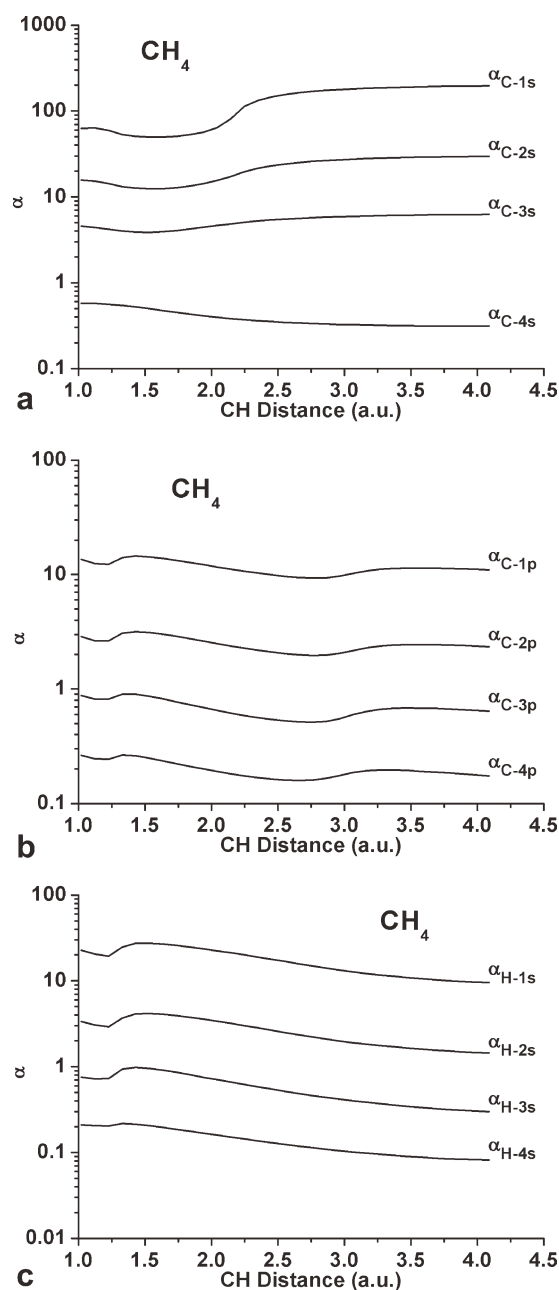


Figure 7. Exponents, α , of the preoptimized bonding orbital of the methane molecule plotted as a function of the CH internuclear distance for a) s primitives of the carbon, b) p primitives of the carbon, and c) s primitives of the hydrogen. Data for generating this figure are present in Table S8 of the Supporting Information.

approach is grounded on the fact that the molecular energies are more sensitive to changes in internuclear distances than to changes in bond and dihedral angles, as evidenced by larger force constants for stretching modes than for bending or torsional modes (0.379 a.u. for stretching and 0.128 a.u. for bending, from a methane HF-LCAO 6-31G calculation). We also kept the parameters of the core orbital of the carbon atom fixed at their values for the optimized bond length.

We therefore optimized the exponents and linear coefficients of the bonding orbitals of CH_4 , for internuclear distances between half and twice the CH optimized bond length,

following the symmetric stretching normal mode of the molecule. In Table S8 of the Supporting Information, we have a list of these parameters as a function of the CH internuclear distance. In Figures 7 and 8, we have, respectively, the values of the exponents and the linear coefficients plotted as a function of the internuclear distance.

In Figure 9a, we have the QMM total energy for the methane molecule, as a function of the CH internuclear distance, obtained when we distort the molecule following the symmetric stretching normal mode (these data are also shown in

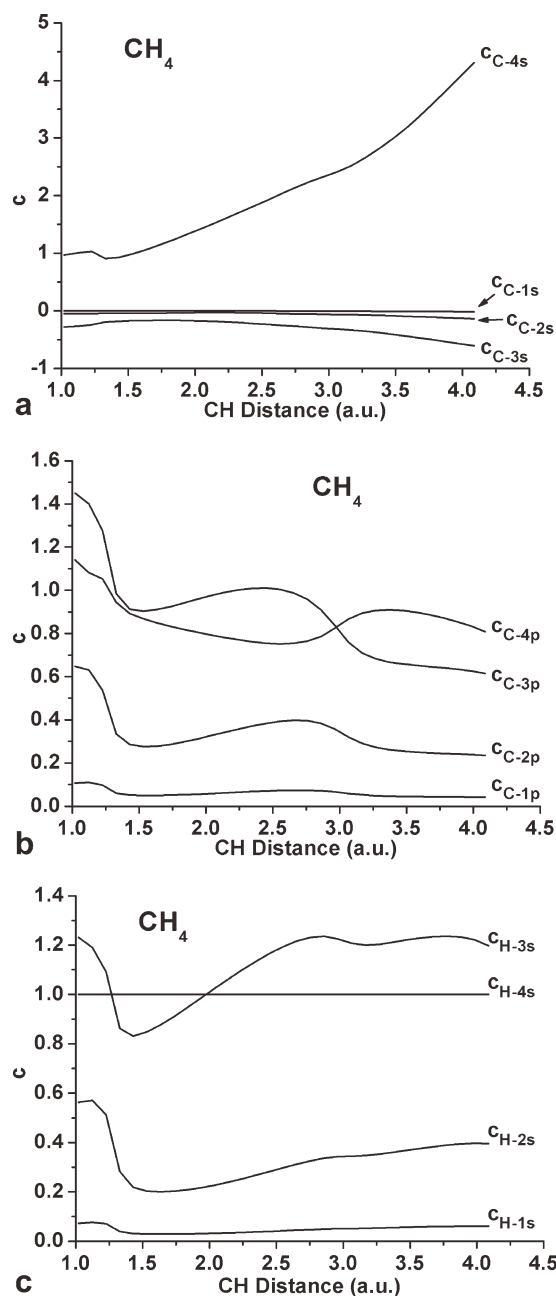


Figure 8. Linear coefficients, c , of the preoptimized bonding orbital of the methane molecule plotted as a function of the CH internuclear distance for a) s primitives of the carbon, b) p primitives of the carbon, and c) s primitives of the hydrogen. The linear coefficients are given in relation to the most diffuse s primitive of the orbital, that is, $c_{H-4s} = 1$. Data for generating this figure are present in Table S8 of the Supporting Information.

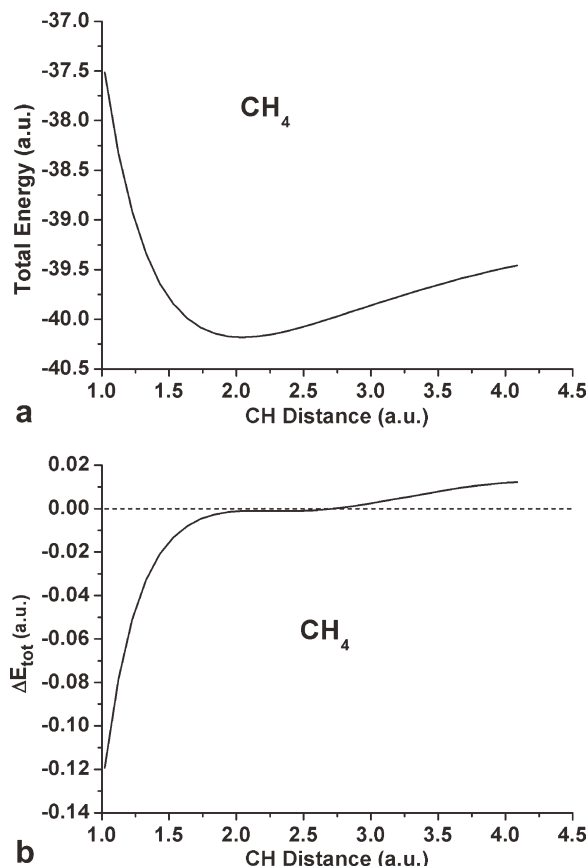


Figure 9. a) QMM total energy for the methane molecule, as a function of the CH internuclear distance, obtained when we distort the molecule following the symmetric stretching normal mode. b) Difference between the total energies calculated using QMM and the ones calculated using the HF/6-31G method ($E^{\text{QMM}} - E^{\text{HF/6-31G}}$). Data for generating this figure are present in Table S9 of the Supporting Information.

Table S9 of the Supporting Information). In Figure 9b, we have the difference between the energies obtained using the QMM method and the ones from HF/6-31G calculations. As we can observe, for shorter CH distances, the QMM method yields energies that are considerably lower than the HF ones. In the region around the equilibrium CH distance, the energies of the two methods are comparable, with a small advantage for the QMM method. For longer distances, however, the HF energies are slightly lower than the QMM ones. This effect is probably due to the way the QMM method polarizes the molecule, keeping a strictly tetrahedral symmetry in a region where the CH bonds are starting to break, and the resulting atoms are trying to acquire spherical symmetry.

Next, we need to show that the bonding orbitals parameters, optimized only for the symmetric stretching normal mode, also give good results, when we distort the methane molecule along other normal modes. We calculated the normal modes of the CH_4 molecule using the HF LCAO method with the 6-31G basis set. We calculated the change in the total energy as the molecule distorts $\pm 5\%$, $\pm 10\%$, and $\pm 15\%$ of the equilibrium CH bond length along each normal mode. In Figure 10, we show a plot of the variation of the total energy calculated using the QMM method versus the values calculated

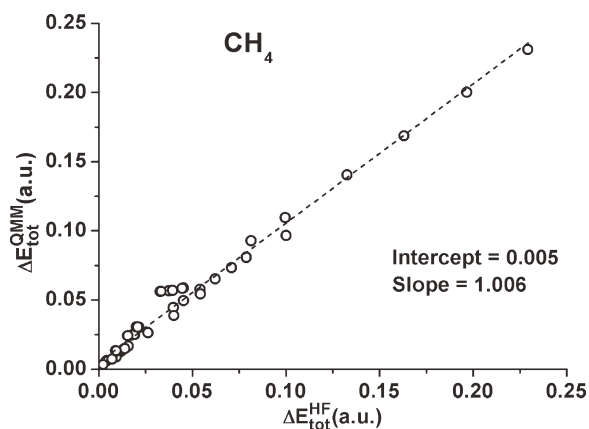


Figure 10. Comparison between the variations in total energy of the methane molecule, along the normal modes, obtained using the QMM method and the values obtained from HF LCAO calculations with the 6-31G basis set.

using the HF LCAO method. As we can see, results from both methods compare really well with each other, showing that our QMM strategy, of optimizing the bonding orbital only for the symmetric stretching normal mode, is appropriate to describe variations of energy along all the normal modes of the methane molecule. Therefore, complete preoptimization is not needed, and we can reasonably expect that some transferability of pair orbitals from one molecule to another with the same type of chemical environment may be warranted.

These results indicate that QMM appears to correctly describe the response of the electron density to variations in geometry and thus seemingly takes into account polarization and charge transfer effects between atoms involved in a chemical bond, which are much more significant than the same effects between nonbonded atoms. To put this assertion to a test, we, once more, decided to investigate the quality of QMM interaction energies. Accordingly, we calculated 27 clusters with 27 CH_4 molecules each. As in the case of the hydrogen molecule, we chose to disregard the translational degrees of freedom of the individual molecules when preparing the clusters of CH_4 , and considered only the rotational degrees of freedom. The geometry of each molecule of the cluster was fixed at the equilibrium CH bond length. Each molecule was then randomly rotated in space and its center of mass placed on a regular cubic grid with a separation of 4.0 Å. In Figure 11, we compare the interaction energies per molecule, calculated with the HF/6-31G method, with the ones calculated using the QMM method. We can observe that the QMM method yields interaction energies that are, again, higher when compared to the HF/6-31G method by only 0.00031 a.u. per molecule, equivalent to 0.19 kcal mol⁻¹. This displacement is probably due to the presence of BSSE in the HF calculation that is absent in the QMM calculation.

To assess how the QMM method performs in the presence of a polarizing group, we performed 27 calculations of a helium hydride cation surrounded by 26 methane molecules. We chose the helium hydride cation immersed in a cluster of methane molecules, to investigate how would QMM behave in

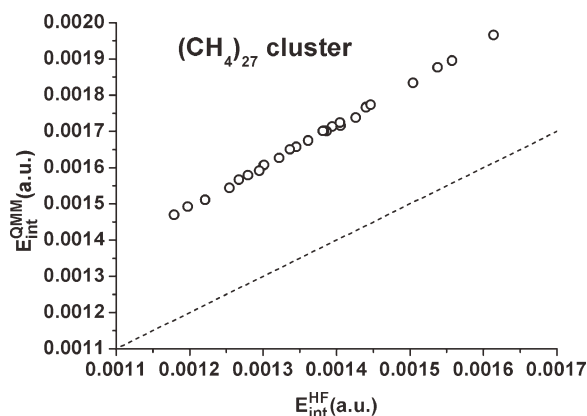


Figure 11. Comparison between the interaction energies per molecule, for clusters of 27 CH_4 molecules, calculated with the HF/6-31G method with the ones calculated using the QMM method. The dashed line has slope one.

the case of a highly polarized and charged molecule immersed in a nonpolar medium. The procedure used was similar to the one used above to study the methane clusters. As such, we generated new clusters of methane molecules and replaced the central methane molecule of the cluster by a randomly oriented helium hydride cation in its QMM equilibrium geometry interatomic distance of 1.4539 a.u. using a preoptimized orbital obtained using four s and one p Gaussian primitives centered on each nucleus. For the methane molecule, we used the same preoptimized orbitals used in the study of the CH_4 clusters. In Figure 12, we compare the interaction energies per molecule calculated with the HF method, with the ones calculated using the QMM method. The HF energies were obtained using the 6-31G basis set for methane and the 6-31G(d,p) basis set for HeH^+ . We can observe that, for the most part, the QMM energies are higher when compared to the HF ones by only 0.00094 a.u. or 0.59 kcal mol⁻¹. Note that the HF results

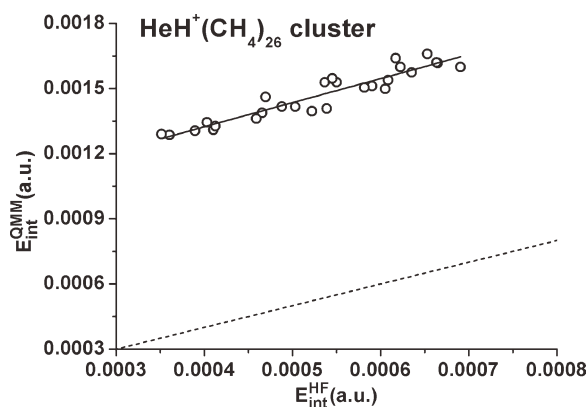


Figure 12. Comparison between the interaction energies per molecule, for arrangements of HeH^+ surrounded by 26 CH_4 molecules, calculated with the HF method with the ones calculated using the QMM method. The dashed line has slope one and intercept zero, and the solid line is a least squares linear fit with slope 1.1 and intercept 8.8×10^{-4} a.u. The HF energies were obtained using the 6-31G basis set for methane and the 6-31G(d,p) basis set for HeH^+ .

include the BSSE effect, not extant in QMM, which may perhaps explain part of this deviation. However, the trends in interactions in the charged clusters are indeed comparable to the HF ones, because the slope of the fitted line is 1.1. Therefore, the preoptimized orbitals in the QMM method are effectively capable of describing the clusters, despite the fact that they do not relax due to intermolecular interactions.

Discussion: Possibilities for Future Work

We now turn to derive an exact formula to compute the QMM total energy in which only two center integrals are involved. For that, we must assume that the fully optimized electron-pair orbitals are known.

To show how that can be done, first let us consider the virial theorem as in eq. (16). To be able to obtain E_{pot} using eq. (16), we need the derivatives of the total energy $E_{\text{tot}} = E_{\text{pot}} + E_{\text{kin}}$ with respect to the nuclear positions. This is apparently a problem, because we would need to already know E_{pot} to calculate E_{tot} . However, because we already have a fully optimized wave function for our molecule, we can use the Hellmann–Feynman theorem^[40,41] in the following form:

$$\frac{\partial E_{\text{tot}}}{\partial R_{gi}} = \left\langle \Psi \left| \frac{\partial H}{\partial R_{gi}} \right| \Psi \right\rangle \quad (18)$$

The only terms of the Hamiltonian [eq. (11)] that depend on the nuclear coordinates are the electron–nuclei attraction and nuclear repulsion. Taking the derivative of these two terms we obtain

$$\frac{\partial H}{\partial R_{gi}} = - \sum_n Z_g \frac{R_{gi} - r_{ni}}{|R_g - r_n|^3} - \sum_{h \neq g} Z_g Z_h \frac{R_{gi} - R_{hi}}{|R_g - R_h|^3} \quad (19)$$

The first term in eq. (19) results in a one-electron integral, and the second is a classical term. Therefore, using eq. (19) in eqs. (18) and (16), we can calculate E_{pot} without resorting to any two-electron integrals. Having E_{pot} , we can calculate E_{tot} , because we no longer need to calculate the two-electron integrals to obtain the total energy.

$$E_{\text{tot}} = - \sum_{\mu, \nu}^{N/2} \langle \mu | p^2 | \nu \rangle T_{\mu\nu} + \sum_g^M \sum_{i=1}^3 Z_g R_{gi} \\ \times \left(2 \sum_{\mu, \nu}^{N/2} \left\langle \mu \left| \frac{R_{gi} - r_i}{|R_g - r_i|^3} \right| \nu \right\rangle T_{\mu\nu} + \sum_{h \neq g}^M Z_h \frac{R_{gi} - R_{hi}}{|R_g - R_h|^3} \right) \quad (20)$$

Thus, we could even perform three-dimensional molecular calculation using Slater type orbitals. An added bonus from eq. (19) is that, because we already know how to calculate the forces acting on the nuclei, we can straightforwardly perform geometry optimizations.

The main obstacle to the calculation of the total energy using eq. (20) is the fact that the forces calculated using the Hellmann–Feynman theorem are very sensitive to even the smallest inaccuracies on the wave function Ψ . Therefore, to use eqs. (19) and (20), we need to have a wave function that

strictly obeys the Hellmann–Feynman theorem, that is, a function which is fully optimized with respect to all adjustable parameters. Besides exact HF, other methods that are known to generate such wave functions that obey the Hellmann–Feynman theorem are the ones which use floating basis sets.^[30,42] These are requirements that are difficult to be guaranteed in an *ab initio* calculation. So, we expect the usefulness of eq. (20) to be severely limited in an *ab initio* context. However, a semiempirical version of QMM may simply accept by definition that the wave function is fully optimized and resort to obtain parameters assuming eq. (20) is the correct one. Research on a semiempirical version of QMM, based on eq. (20), is presently being attempted in our laboratories.

Conclusions

In this article, we advanced a strategy to develop a MM method based, not on classical mechanics and force fields, but entirely on QM and localized electron-pair orbitals, which we call QMM. The ultimate goal of QMM is to replace MM in QM/MM calculations, giving rise to a QM/QMM method.

Accordingly, we introduced a new manner of calculating HF *ab initio* wave functions of closed shell systems based on variationally preoptimized nonorthogonal electron pair orbitals constructed by linear combinations of basis functions centered on the atoms. We showed that these orbitals naturally seek to obey the electron–nucleus cusp conditions and the extended virial theorem—the more so, the larger the basis set used. Once one possesses these preoptimized electron-pair orbitals, QMM requires only one single inversion of a sparse matrix to arrive to the electron density and consequently, to the electrostatic potential around the molecular system, or cluster of molecules. And, as we have shown, the results compare very well with traditional LCAO HF calculations and are considerably faster, because the calculation is noniterative.

Finally, we examined in detail the cases of clusters of helium atoms, hydrogen molecules and methane molecules. In the case of methane, this first implementation of QMM was able to satisfactorily reproduce the energetics of all distortions of the molecules along the nine normal modes of vibration, well beyond the harmonic region.

Indeed, in QMM, the orbitals are not fixed for each chemical bond. They vary according to variations in geometry. Consequently, QMM appears to correctly describe the response of the electron density to variations in geometry and thus takes into account polarization and charge transfer effects between atoms involved in a chemical bond. As an example of a stronger interaction, we considered one HeH^+ molecule embedded in a cluster of 26 methane molecules. Intermolecular polarization should play a more substantial role in this case. Nevertheless, our results show that the interaction energies are correctly described and compare well with HF calculation results for the same system; the small differences found, being attributable to BSSE—an effect nonexistent in QMM.

Moreover, assuming that the QMM spatial orbitals are fully optimized for the system under consideration, we showed that the calculation of the total energy (E_{tot}) can actually be carried


out without resorting to any two-electron integrals. Although this assumption appears impractical for *ab initio* QMM calculations, we expect such an approach to prove important for the development of the semiempirical version of QMM, presently being carried out in our laboratories.

We conclude that QMM is a promising quantum mechanical technique to describe molecules and clusters of molecules of considerable size, with potential to be applied, for example, to Monte Carlo simulations.

Research to extend QMM to introduce electron correlation and dispersion into the method, as well as to extend it to molecules of biological interest, is presently being carried out in our laboratories.

Keywords: quantum molecular mechanics · *ab initio* · nonorthogonal orbitals · atomic and molecular clusters

How to cite this article: G. L. C. Moura, A. M. Simas, *J. Comput. Chem.* **2012**, *33*, 958–969. DOI: 10.1002/jcc.22921

 Additional Supporting Information may be found in the online version of this article.

- [1] T. L. Hill, *J. Chem. Phys.* **1946**, *14*, 465.
- [2] F. H. Westheimer, J. E. Mayer, *J. Chem. Phys.* **1946**, *14*, 733.
- [3] N. Allinger, *J. Am. Chem. Soc.* **1977**, *99*, 8127.
- [4] N. Allinger, Y. Yuh, J. Lii, *J. Am. Chem. Soc.* **1989**, *111*, 8551.
- [5] M. Clark, R. Cramer, N. Vanopdenbosch, *J. Comput. Chem.* **1989**, *10*, 982.
- [6] W. Xie, J. Gao, *J. Chem. Theory Comput.* **2007**, *3*, 1890.
- [7] J. Ponder, D. Case, *Protein Simulations* **2003**, *66*, 27.
- [8] W. F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. van der Vegt, H. B. Yu, *Angew. Chem. Int. Ed.* **2006**, *45*, 4064.
- [9] <http://autodock.scripps.edu/news/autodocks-role-in-developing-the-first-clinically-approved-hiv-integrase-inhibitor> (accessed 2011).
- [10] J. L. Knight, C. L. Brooks, *J. Chem. Theory Comput.* **2011**, *7*, 2728.
- [11] A. MacKerell, D. Bashford, M. Bellott, *J. Phys. Chem. B* **1998**, *102*, 3586.
- [12] Y. Duan, C. Wu, S. Chowdhury, *J. Comput. Chem.* **2003**, *24*, 1999.
- [13] W. Jorgensen, D. Maxwell, J. TiradoRives, *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- [14] L. Schuler, X. Daura, G. W. Van, *J. Comput. Chem.* **2001**, *22*, 1205.
- [15] P. Freddolino, A. Arkhipov, S. Larson, *Structure* **2006**, *14*, 437.
- [16] C. Rountree, R. Kalia, E. Lidorikis, *Annu. Rev. Mater. Res.* **2002**, *32*, 377.
- [17] N. Gresh, G. A. Cisneros, T. A. Darden, J.-P. Piquemal, *J. Chem. Theory Comput.* **2007**, *3*, 1960.
- [18] J.-P. Piquemal, G. A. Cisneros, P. Reinhardt, N. Gresh, T. A. Darden, *J. Chem. Phys.* **2006**, *124*, 104101.
- [19] A. Warshel, A. Bromberg, *J. Chem. Phys.* **1970**, *52*, 1262.
- [20] A. Warshel, M. Karplus, *J. Am. Chem. Soc.* **1972**, *94*, 5612.
- [21] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227.
- [22] J. Gao, X. Xia, *Science* **1992**, *258*, 631.
- [23] G. Seabra, R. Walker, A. Roitberg, *J. Phys. Chem. A* **2009**, *113*, 11938.
- [24] J. Gao, *J. Phys. Chem. B* **1997**, *101*, 657.
- [25] J. Gao, *J. Chem. Phys.* **1998**, *109*, 2346.
- [26] W. Xie, M. Orozco, D. G. Truhlar, J. Gao, *J. Chem. Theory Comput.* **2009**, *5*, 459.
- [27] S. Schweizer, J. Kussmann, B. Doser, C. Ochsenfeld, *J. Comput. Chem.* **2008**, *29*, 1004.
- [28] P. Lowdin, *J. Chem. Phys.* **1950**, *18*, 365.
- [29] P. Lowdin, *Phys. Rev.* **1955**, *97*, 1490.
- [30] A. Frost, *J. Chem. Phys.* **1967**, *47*, 3707.
- [31] P. Lowdin, *Phys. Rev.* **1955**, *97*, 1474.
- [32] R. Rouse, A. Frost, *J. Chem. Phys.* **1969**, *50*, 1705.
- [33] J. J. P. Stewart, *Int. J. Quant. Chem.* **1996**, *58*, 133.
- [34] J. J. P. Stewart, *J. Mol. Model.* **2009**, *15*, 765.
- [35] T. Kato, *Commun. Pure Appl. Math.* **1957**, *10*, 151.
- [36] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, D. Cohen, *J. Chem. Phys.* **1996**, *105*, 1968.
- [37] D. D. Kemp, J. M. Rintelman, M. S. Gordon, J. H. Jensen, *Theor. Chem. Accounts* **2009**, *125*, 481.
- [38] L. McMurchie, E. Davidson, *J. Comput. Phys.* **1978**, *26*, 218.
- [39] J. Slater, *J. Chem. Phys.* **1933**, *1*, 687.
- [40] H. Hellmann, Einführung in die Quantenchemie; Franz Deuticke: Leipzig, **1937**; p. 285.
- [41] R. Feynman, *Phys. Rev.* **1939**, *56*, 340.
- [42] M. Tachikawa, K. Taneda, K. Mori, *Int. J. Quant. Chem.* **1999**, *75*, 497.

Received: 2 April 2011
 Revised: 14 October 2011
 Accepted: 3 December 2011
 Published online on 8 February 2012