The variational explicit polarization potential and analytical first derivative of energy: Towards a next generation force field

Wangshen Xie, Lingchun Song, Donald G. Truhlar,[a] and Jiali Gao[a]
Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, USA

(Received 6 March 2008; accepted 1 May 2008; published online 20 June 2008)

A previous article proposed an electronic structure-based polarizable potential, called the explicit polarization (X-POL) potential, to treat many-body polarization and charge delocalization effects in polypeptides. Here, we present a variational version of the X-POL potential, in which the wave function of the entire molecular system is variationally optimized to yield the minimum total electronic energy. This allows the calculation of analytic gradients, a necessity for efficient molecular dynamics simulations. In this paper, the detailed derivations of the Fock matrix and analytic force are presented and discussed. The calculations involve a double self-consistent-field procedure in which the wave function of each fragment is self-consistently optimized in the presence of other fragments, and in addition the polarization of the entire system is self-consistently optimized. The variational X-POL potential has been implemented in the Chemistry at Harvard Molecular Mechanics (CHARMM) package and tested successfully for small model compounds. [DOI: 10.1063/1.2936122]

I. INTRODUCTION

The explicit polarization (X-POL) potential has been proposed for use as a force field for simulations of large molecules in condensed phases (note that the acronym X-POL is also synonymous for Xie's polarization).[1] The X-POL potential is based on division of a system into fragments within which the interactions are treated using molecular-orbital (MO) theory while interactions between fragments are treated using a combined quantum mechanical and molecular mechanical[2–5] (QM/MM) calculation. In the QM/MM calculation, the fragments other than the one currently under consideration are represented by MM point charges that are computed by Mulliken[6] population analysis. Polarization of the wave function of each fragment is included in the context of this QM/MM framework. Electronic structure calculations are sequentially performed for each fragment in the presence of Mulliken charges that are obtained from the wave functions of all other fragments. We iteratively cycle through all the fragments until the total energy or electron density is converged to a given tolerance.[1] One of the advantages of dividing the system into fragments is that we need to construct and diagonalize Fock matrices only for small fragments, so that the computational bottleneck of diagonalizing the Fock matrix of the entire system is avoided. Since the interfragment interactions are treated using a QM/MM method, the number of integrals to be evaluated is significantly decreased compared to a full QM calculation of the entire system. Another advantage of the X-POL potential is that it can be easily parallelized by assigning a certain number of fragments to each processor. Communication between processors will be mainly for Mulliken charges.

The X-POL potential was originally called the MO derived empirical potential (MODEL) for liquids;[7] it has been applied to the Monte Carlo simulation of liquid water[8] and liquid hydrogen fluoride,[9] and it was recently extended to the treatment of polypeptides in which the fragments are covalently bonded.[1] The wave function for each fragment is updated iteratively in the electric field from all other fragments until self-consistency is achieved for the entire system. Thus, each part of the system is treated at the same level. However, the Fock matrix used previously[1,7–9] was not derived variationally, so it is not guaranteed that the converged wave function yields the minimum electronic energy of the entire system. In fact, tests showed that the previous converged X-POL potential does not have the lowest energy for a given molecular geometry, although the difference from the true minimum is very small. The nonvariational behavior of the previous formulation makes it difficult to calculate an analytic gradient. Although it is possible to use the coupled-perturbed Hartree–Fock method[10] to determine forces iteratively, this would add additional cost to the computational scheme. Approximate gradients obtained previously are not consistent with the energies, as found in the fragment MO method[11] (FMO) because of the neglect of the derivative of the wave function (Pulay force).[12]

In this work, we derive a set of variational equations for the X-POL potential to yield an analytic gradient. The method is similar to that used by Roothaan in deriving the Hartree–Fock equations.[13] We show that the Fock matrix for the variational X-POL potential has three extra terms due to the variation of Mulliken charges. We present a method for calculating analytic forces for the entire system in which the wave function of each fragment is self-consistently optimized in the presence of other fragments, and in addition the

[a]Electronic addresses: truhlar@umn.edu or jgao@umn.edu.

polarization of the entire system is self-consistently optimized, which is called the double self-consistent-field (DSCF) method.

II. THEORY

II.A. Noncovalently bound clusters

We first consider a system consisting of M molecules that are not covalently bonded to each other (Sec. II.B extends the treatment to a system in which fragments are chemically connected). The electronic interactions within each molecule are described by a QM electronic structure method, and the interactions between molecule A ($A=1,2,\ldots,M$) and the other $M-1$ molecules are treated by using an electronic wave function for A and partial atomic charges (point charges at the nuclei) for the $M-1$ other molecules; the partial atomic charges for those molecules are computed by Mulliken[6] population analysis of their wave functions. [This assumption for the interfragment treatment is called a QM/MM (Refs. 2–5) treatment]. The N-electron wave function of each QM fragment is constructed as a Slater determinant of the occupied molecular spin orbitals ψ_k as follows:

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1^1 & \psi_2^1 & \cdots & \psi_N^1 \\ \psi_1^2 & \psi_2^2 & \cdots & \psi_N^2 \\ \cdots & \cdots & \cdots & \cdots \\ \psi_1^N & \psi_2^N & \cdots & \psi_N^N \end{vmatrix}, \tag{1}$$

where ψ_k^j denotes electron i in spin orbital k, and N is the number of explicitly treated electrons in molecule m. We assume a closed-shell wave function for each fragment. Then, each spin orbital ψ_k is the product of an orbital φ_i and a spin function (α or β), and the total electronic energy of the system is written as

$$E_{\text{elec}} = \sum_{m=1}^{M} \left\{ \left[2\sum_{i\in m} H_i + \sum_{ij\in m} (2J_{ij} - K_{ij}) + E_{\text{nucl}}^m \right] + \frac{1}{2}\left(2\sum_{i\in m} I_i + \sum_{B\in m} L_B \right) \right\}, \tag{2}$$

where i and j represent orthonormal MOs, m denotes a fragment, and B represents an atom in a fragment. The terms in the square brackets are the Hartree–Fock electronic energy and the nuclear-nuclear interaction energy within fragment m. The terms in the final parentheses are QM/MM interactions between fragment m and all other fragments; these are multiplied by $1/2$ to avoid double counting. The symbols in Eq. (2) are defined below.

In Eq. (2), H_i is the expectation value of the one-electron Hamiltonian H as follows:

$$H_i = H_i^* = \int \varphi_i^* H \varphi_i \, d\nu, \tag{3}$$

where an asterisk denotes complex conjugation. H_i is the sum of the electronic kinetic energy and the electron-nuclei Coulomb attraction (only for the nuclei in the same molecule), and J_{ij} and K_{ij} are Coulomb and exchange integrals defined by

$$J_{ij} = J_{ji} = J_{ij}^* = J_{ji}^* = e^2 \int dr^\mu dr^\nu \frac{\varphi_i^{\mu *} \varphi_j^{\nu *} \varphi_i^\mu \varphi_j^\nu}{r^{\mu\nu}}, \tag{4}$$

$$K_{ij} = K_{ji} = K_{ij}^* = K_{ji}^* = e^2 \int dr^\mu dr^\nu \frac{\varphi_i^{\mu *} \varphi_j^{\nu *} \varphi_j^\mu \varphi_i^\nu}{r^{\mu\nu}}, \tag{5}$$

where e is the atomic unit of charge, μ and ν label electrons, and $r^{\mu\nu}$ is the distance between two electrons. It will be useful to define the Coulomb operator J_i and the exchange operator K_i by

$$J_i^\mu \varphi^\mu = e^2 \int dr^\nu \frac{\varphi_i^{\nu *} \varphi_i^\nu}{r^{\mu\nu}} \varphi^\mu, \tag{6}$$

$$K_i^\mu \varphi^\mu = e^2 \int dr^\nu \frac{\varphi_i^{\nu *} \varphi^\nu}{r^{\mu\nu}} \varphi_i^\mu. \tag{7}$$

The quantity I_i in Eq. (2) is the energy of interaction of orbital i with the electric potential due to the MM charges of the other fragments, that is,

$$I_i = \sum_{A\notin m} q_A e \int dr^\mu \frac{\varphi_i^{\mu *} \varphi_i^\mu}{r^{\mu A}}, \tag{8}$$

where $A \notin m$ denotes that the sum is over nuclei that are not in the same molecule m in which the orbital φ_i is centered, q_A is the partial atomic charge (in a.u.) on MM atom A, and $r^{\mu A}$ is the distance between atom A and electron μ. Finally, L_B represents the interaction of the nucleus of QM atom B in fragment m and the partial atomic charges q_A in atoms A in other fragments as follows

$$L_B = \sum_{A\in n, n\neq m} \frac{Z_B q_A}{r^{AB}}. \tag{9}$$

Note that the above formalism is written for theoretical methods that treat all electrons explicitly. In semiempirical MO theory, only valence electrons are explicit, and the total number of electrons and the effective nuclear charges should be adjusted accordingly.

In practice, MOs are constructed as linear combinations of atomic basis functions

$$\varphi_i = \sum_p \chi_p C_{pi}, \tag{10}$$

where χ_p are normalized nuclear-centered basis functions, and the normalization condition is

$$\int dr^{\mu} \chi_p^{\mu^*} \chi_p^{\mu} = 1. \tag{11}$$

Although there is actually a different set of MOs and a different set of coefficients in each molecule m and we could write φ_i^m and C_{pi}^m, we omit these superscripts to keep the notation manageable. We have

$$\varphi_i = \boldsymbol{\chi} \mathbf{c}_i, \tag{12}$$

where $\boldsymbol{\chi}$ is a row vector with elements χ_p, and $\mathbf{c}_i$ is a column vector with elements C_{pi}. Let $\boldsymbol{\Phi}$ denote a row vector with elements φ_i; then,

$$\boldsymbol{\Phi} = \boldsymbol{\chi} \mathbf{C}. \tag{13}$$

Following Roothaan's derivation[13] of the Hartree–Fock equations, we define, for any one-electron operator M,

$$M_{pq} = \int dr^{\mu} \chi_p^{\mu^*} M \chi_q^{\mu}, \tag{14}$$

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1N} \\ M_{21} & M_{22} & \cdots & M_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ M_{N1} & M_{N2} & \cdots & M_{NN} \end{bmatrix}. \tag{15}$$

Then, the matrix elements of the one-electron operators can be written as

$$\begin{aligned} H_i &= \mathbf{c}_i^{\dagger} \mathbf{H} \mathbf{c}_i, \\ I_i &= \mathbf{c}_i^{\dagger} \mathbf{I}_A \mathbf{c}_i, \\ J_{ij} &= \mathbf{c}_i^{\dagger} \mathbf{J}_j \mathbf{c}_j = \mathbf{c}_j^{\dagger} \mathbf{J}_i \mathbf{c}_j, \\ K_{ij} &= \mathbf{c}_i^{\dagger} \mathbf{K}_j \mathbf{c}_i = \mathbf{c}_j^{\dagger} \mathbf{K}_i \mathbf{c}_j, \end{aligned} \tag{16}$$

where the obelisk denotes a Hermitian conjugate. With the above definition, the total electronic energy of Eq. (2) can be rewritten as

$$E_{\text{elec}} = \sum_{m=1}^{M} \left\{ \left[2 \sum_{i \in m} \mathbf{c}_i^{\dagger} \mathbf{H} \mathbf{c}_i + \sum_{ij \in m} (2 \mathbf{c}_i^{\dagger} \mathbf{J}_j \mathbf{c}_i - \mathbf{c}_i^{\dagger} \mathbf{K}_j \mathbf{c}_i) + E_{\text{nucl}}^m \right] + \frac{1}{2} \left(2 \sum_{i \in m} \mathbf{c}_i^{\dagger} \mathbf{I} \mathbf{c}_i + \sum_{B \in m} L_B \right) \right\}. \tag{17}$$

To find the linear combinations of atomic orbitals (LCAOs) that give the minimum of Eq. (17), we vary the vectors $\mathbf{c}_i$ by infinitesimal amounts $\delta \mathbf{c}_i$ and find the variation of the energy. In Eq. (17), matrices $\mathbf{H}$, $\mathbf{J}_j$, and $\mathbf{K}_j$ are independent of variations of $\mathbf{c}_i$; however, the QM/MM interaction matrix $\mathbf{I}$ in the terms of Eq. (17) referring to molecule m depends on the $\mathbf{c}_i$ of other molecules through the Mulliken charges on the MM atoms.

The Mulliken charge on atom A in molecule m is given by

$$q_A = Z_A - \sum_{p \in A} (\mathbf{PS})_{pp} = Z_A - \sum_{p \in A} (\mathbf{SP})_{pp}, \tag{18}$$

where Z_A is the atomic number of atom A in an all-electron treatment or the effective nuclear charge equal to the number of valence electrons on neutral atom A in a valence-electron treatment, $\mathbf{S}$ is the overlap matrix defined as

$$S_{pq} = \int dr^{\mu} \chi_p^{\mu^*} \chi_q^{\mu}, \tag{19}$$

and the elements of the density matrix are

$$P_{pq} = 2 \sum_i C_{qi}^* C_{pi}. \tag{20}$$

It is useful to introduce, for each molecule m, a $N^{Bm} \times N^{Bm}$ (where N^{Bm} is the number of basis functions on molecule m) matrix with binary elements

$$\Delta_{pq}^{(\lambda)} = (\boldsymbol{\Delta}^{(\lambda)})_{pq} = \delta_{pq} \delta_{p\lambda}, \tag{21}$$

where δ_{pq} and $\delta_{p\lambda}$ are Kronecker deltas. The Mulliken charge on atom A can be expressed as

$$\begin{aligned} q_A &= Z_A - \sum_{p \in A} \sum_{\lambda \in m(A)} S_{p\lambda} P_{\lambda p} \\ &= Z_A - \sum_{p \in A} \sum_{\lambda \in m(A)} S_{p\lambda} \left(2 \sum_i \sum_{\sigma} \sum_{\rho} C_{\sigma i}^* \delta_{\sigma\lambda} \delta_{p\rho} C_{\rho i} \right) \\ &= Z_A - 2 \sum_i \sum_{p \in A} \sum_{\lambda \in m(A)} S_{p\lambda} \mathbf{c}_i^{\dagger} \mathbf{T}^{(\lambda,p)} \mathbf{c}_i, \end{aligned} \tag{22}$$

where we have defined $T_{\sigma\rho}^{(\lambda,p)} = \Delta_{\sigma\lambda}^{(\lambda)} \Delta_{p\rho}^{(p)}$, and where $\lambda \in m(A)$ means that λ is in the same molecule m as atom A. Thus, the variation of q_A is given by

$$\begin{aligned} \delta_{qA} = &- 2 \sum_i \sum_{p \in A} \sum_{\lambda \in m(A)} S_{p\lambda} (\delta \mathbf{c}_i^{\dagger}) \mathbf{I}^{(\lambda,p)} \mathbf{c}_i \\ &- 2 \sum_i \sum_{p \in A} \sum_{\lambda \in m(A)} S_{p\lambda} \mathbf{c}_i^{\dagger} \mathbf{T}^{(\lambda,p)} (\delta \mathbf{c}_i). \end{aligned} \tag{23}$$

We define the QM/MM energy as

$$E_{\text{QM/MM}} = \frac{1}{2} \sum_{m=1}^{M} \left(2 \sum_{i \in m} \mathbf{c}_i^{\dagger} \mathbf{I} \mathbf{c}_i + \sum_{B \in m} L_B \right). \tag{24}$$

One can show that variation of $E_{\text{QM/MM}}$ due to variation of the Mulliken charges on atoms in fragment m is

$$\delta E_{\text{QM/MM}} = \frac{1}{2} \sum_{n \neq m}^{M} \sum_{A \in m} \left[\sum_{pq \in n} P_{pq} I_{pq}^0 + \sum_{B \in n} L_B^0 \right] \delta_{qA}, \tag{25}$$

where the superscript 0 denotes that the matrix element is calculated by setting the MM charges on atom A to $+e$.

The variation of total electronic energy is then given by

$$\delta E_{\mathrm{elec}} = 2\sum_{m=1}^{M}\sum_{i\in m}(\delta\mathbf{c}_i^{\dagger})\left[\mathbf{H} + \sum_{j\in m}(2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\mathbf{I} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I_{pq}^{0} + \sum_{B\in n} L_B^0\right)\sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)}\right]\mathbf{c}_i$$
$$+ 2\sum_{m=1}^{M}\sum_{i\in m}\mathbf{c}_i^{\dagger}\left[\mathbf{H} + \sum_{j\in m}(2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\mathbf{I} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I_{pq}^{0} + \sum_{B\in n} L_B^0\right)\sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S'_{p\lambda}\mathbf{T}^{(\lambda,p)}\right](\delta\mathbf{c}_i). \quad (26)$$

Using the Hermitian property of $\mathbf{H}$, $\mathbf{I}$, $\mathbf{J}_j$, $\mathbf{K}_j$, and $\mathbf{T}^{(\lambda,p)}$, we may rewrite Eq. (26) as

$$\delta E_{\mathrm{elec}} = 2\sum_{m=1}^{M}\sum_{i\in m}(\delta\mathbf{c}_i^{\dagger})\left[\mathbf{H} + \sum_{j\in m}(2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\mathbf{I} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I_{pq}^{0} + \sum_{B\in n} L_B^0\right)\sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)}\right]\mathbf{c}_i$$
$$+ 2\sum_{m=1}^{M}\sum_{i\in m}(\delta\mathbf{c}_i^{T})\left[\mathbf{H}^{*} + \sum_{j\in m}(2\mathbf{J}_j^{*} - \mathbf{K}_j^{*}) + (1/2)\mathbf{I}^{*} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I_{pq}^{0} + \sum_{B\in n} L_B^0\right)\sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)*}\right]\mathbf{c}_i^{*}, \quad (27)$$

where superscript T denotes a transpose. We define the Fock matrix for fragment m as follows:

$$\mathbf{F} = \mathbf{H} + \sum_{j\in m}(2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\mathbf{I} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I_{pq}^{0} + \sum_{B\in n} L_B^0\right)\sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)}. \quad (28)$$

From Eq. (28), we immediately see that each fragment m is fully polarized by the rest of the system such that half of the polarization comes from the Mulliken charges of the rest of the system and the other half from charge density plus nuclei charges of the rest of the system. The variation of total electronic energy is then readily expressed as

$$\delta E_{\mathrm{elec}} = \sum_{m=1}^{M}\left[\sum_{i\in m} 2(\delta\mathbf{c}_i^{\dagger})\mathbf{F}_m\mathbf{c}_i + 2(\delta\mathbf{c}_i^{T})\mathbf{F}_m^{*}\mathbf{c}_i^{*}\right]. \quad (29)$$

We assume that the LCAOs satisfy the orthonormal condition

$$\mathbf{c}_i^{\dagger}\mathbf{S}\mathbf{c}_j = \delta_{ij}. \quad (30)$$

The restriction on $\delta\mathbf{c}_i$ resulting from the orthonormality of the MOs is obtained by varying Eq. (30), which yields

$$(\delta\mathbf{c}_i^{\dagger})\mathbf{S}\mathbf{c}_j + \mathbf{c}_i^{\dagger}\mathbf{S}(\delta\mathbf{c}_j) = 0 \quad (31)$$

or

$$(\delta\mathbf{c}_i^{\dagger})\mathbf{S}\mathbf{c}_j + (\delta\mathbf{c}_i^{T})\mathbf{S}^{*}\mathbf{c}_i^{*} = 0. \quad (32)$$

We multiply these restricting conditions by the Lagrangian multipliers $-2\varepsilon_{ij}$ and add them together as follows:

$$-2\sum_{ij}(\delta\mathbf{c}_i^{\dagger})\mathbf{S}\mathbf{c}_j\varepsilon_{ji} + \mathbf{c}_i^{\dagger}\mathbf{S}(\delta\mathbf{c}_j)\varepsilon_{ji} = 0 \quad (33)$$

or

$$-2\sum_{ij}(\delta\mathbf{c}_i^{\dagger})\mathbf{S}\mathbf{c}_j\varepsilon_{ji} + (\delta\mathbf{c}_i^{T})\mathbf{S}^{*}\mathbf{c}_i^{*}\varepsilon_{ji} = 0. \quad (34)$$

Adding Eq. (34) to the variation of total electronic energy given by Eq. (29), we obtain

$$\delta E_{\mathrm{elec}} = \sum_{m=1}^{M}\left[\sum_{i\in m} 2(\delta\mathbf{c}_i^{\dagger})\left(\mathbf{F}_m\mathbf{c}_i - \sum_{j\in m}\mathbf{S}\mathbf{c}_j\varepsilon_{ji}\right) + 2(\delta\mathbf{c}_i^{T})\times\left(\mathbf{F}_m^{*}\mathbf{c}_i^{*} - \sum_{j\in m}\mathbf{S}^{*}\mathbf{c}_j^{*}\varepsilon_{ji}\right)\right]. \quad (35)$$

For each fragment, the conditions for $\delta E_{\mathrm{elec}}=0$ for any choice of the vectors $\delta\mathbf{c}_i$ and $\delta\mathbf{c}_i^{*}$, or $\delta\mathbf{c}_i^{\dagger}$ and $\delta\mathbf{c}_i^{T}$, are given by

$$\mathbf{F}\mathbf{c}_i = \sum_j \mathbf{S}\mathbf{c}_j\varepsilon_{ji},$$
$$\mathbf{F}^{*}\mathbf{c}_i^{*} = \sum_j \mathbf{S}^{*}\mathbf{c}_j^{*}\varepsilon_{ji}. \quad (36)$$

Since ε_{ji} are the elements of a Hermitian matrix,[13] the two equations are equivalent:

$$\mathbf{F}\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\varepsilon}. \quad (37)$$

Note that although the derivation above is similar to that for the Roothaan equations,[13] Eq. (37) treats intermolecular interactions among fragments by a combined QM/MM algorithm[2–5] in which the MM potential is derived from the corresponding molecular wave functions of other fragments.[1,7]

II.B. Covalent bonding of fragments

Here, we extend the variational X-POL potential to proteins or other systems in which the QM fragments are covalently connected. The partition of a protein gives rise to one or two boundary atoms on each peptide unit connected to adjacent fragments (in general, two, but only one at the N- and C-termini). As in our previous work,[1] we choose to assign two sets of hybrid orbitals (HOs) (active and auxiliary orbitals) on each boundary atom, and this boundary atom is

effectively partitioned into an "active atom" and an "auxiliary atom," each of which has half of the nuclear charge (active and auxiliary nuclei). For each fragment m, the number of electrons is equal to the number of valence electrons on non-boundary plus 2 for each boundary atom.

In the following derivation, we use the hybrid basis set, which is defined in the previous paper[1] and is convenient for discussion purposes. The total electronic energy of a protein is defined as

$$E_{\text{elec}} = E_{\text{act}} + E_{\text{aux}} + E_{\text{nucl}}, \tag{38}$$

in which the last term is the QM core-core repulsion energy, and E_{act} is the electronic energy involving only active orbitals:

$$E_{\text{act}} = \sum_{m=1}^{M} \left[2\sum_{i\in m}^{N_{\text{act}}} \mathbf{c}_i^{\dagger}\mathbf{H}\mathbf{c}_i + \sum_{ij\in m}^{N_{\text{act}}} (2\mathbf{c}_i^{\dagger}\mathbf{J}_j\mathbf{c}_i - \mathbf{c}_i^{\dagger}\mathbf{K}_j\mathbf{c}_i) + \frac{1}{2}\left(2\sum_{i\in m}^{N_{\text{act}}} \mathbf{c}_i^{\dagger}\mathbf{I}\mathbf{c}_i + \sum_{B\in m} L_B \right) \right], \tag{39}$$

where N_{act} is the number of active orbitals (equal to the number of valence orbitals on non-boundary nuclei plus 2 for each boundary atom), where we use $\mathbf{c}_i$ to denote the coefficients that define the wave function in the hybrid basis for simplicity. To avoid double counting, all QM/MM interactions are multiplied by 1/2. The only exception is that orbitals on a boundary atom experience attraction from the full boundary nuclei charge when they are "active." Consequently, they do not interact with their nuclei when they become "auxiliary" orbitals, which will be denoted a and b. The effective potential due to the auxiliary orbitals on each boundary atom is approximated by a 2×2 auxiliary charge density matrix[1] so that the contribution from auxiliary orbitals can be written in the HO basis as

$$E_{\text{aux}} = \frac{1}{2}\sum_{m=1}^{M}\sum_{ab\in m}^{N_{\text{aux}}} P_{ab}(H_{ab} + G_{ab}). \tag{40}$$

where N_{aux} is the number of auxiliary orbitals (two or four). Note that we have used the notations a and b for auxiliary orbitals. The two-electron integrals are

$$G_{ab} = \sum_{\lambda\sigma\in m}^{N_{\text{act}}} P_{\lambda\sigma}[(ab,\lambda\sigma) - (1/2)(a\sigma,\lambda b)]. \tag{41}$$

Note that G_{ab} does not include the interaction between auxiliary orbitals (no interaction within the MM region). Since a and b are auxiliary orbitals, H_{ab} does not include the attraction from the corresponding boundary nuclei, which is accounted for in the fragment that a and b are treated as active orbitals.

Now, we need to find the variation of the total energy with respect to the perturbation of the MOs. To solve this problem, we first need to find the variation of the Mulliken charges of the MM atoms and the auxiliary density matrix elements. We define the transformation matrix $\mathbf{T}$ that transforms the MOs in the atomic basis set to the hybrid basis set

$$\mathbf{c}^{H} = \mathbf{T}^{-1}\mathbf{c}^{\text{AO}}, \tag{42}$$

where superscript H denotes hybrid basis and superscript AO denotes the corresponding spherical harmonic basis. Then, the variation of Mulliken charges can be written in the hybrid basis set as

$$\delta q_A = -2\sum_{i}^{N_{\text{act}}}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}(\delta\mathbf{c}_i^{\dagger})(\mathbf{T}^{\dagger}\mathbf{T}^{(\lambda,p)}\mathbf{T})\mathbf{c}_i - 2\sum_{i}^{N_{\text{act}}}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{c}_i^{\dagger}(\mathbf{T}^{\dagger}\mathbf{T}^{(\lambda,p)}\mathbf{T})(\delta\mathbf{c}_i). \tag{43}$$

For each auxiliary density matrix element P_{ab}, we define a matrix $\mathbf{T}^{(a,b)}$, which connects MOs to auxiliary charge density

$$P_{ab} = 2\sum_{i}^{N_{\text{act}}} \mathbf{c}_i^{\dagger}\mathbf{T}^{(a,b)}\mathbf{c}_i. \tag{44}$$

We immediately find that the $\mathbf{T}^{(a,b)}$ matrix has only one nonzero element, which is $T_{pp'}^{(a,b)} = \delta_{ap}\delta_{bp'}$. Variation of auxiliary charge density due to variation of MOs is thus

$$\delta P_{ab} = 2(\delta\mathbf{c}_i^{\dagger})\mathbf{T}^{(a,b)}\mathbf{c}_i + 2\mathbf{c}_i^{\dagger}\mathbf{T}^{(a,b)}(\delta\mathbf{c}_i). \tag{45}$$

Consequently, the variation of electronic energy due to the interaction involving auxiliary orbitals can be written explicitly as

$$\delta E_{\text{aux}} = \sum_{m=1}^{M}(\delta\mathbf{c}_i^{\dagger})\left[\sum_{ab\in m\pm 1}^{N_{\text{aux}}} (H_{ab} + G_{ab})\mathbf{T}^{(a,b)} + (1/2)\sum_{k\in m}^{N_{\text{aux}}} (2\mathbf{J}_k - \mathbf{K}_k) \right]\delta\mathbf{c}_i + \sum_{m=1}^{M}\mathbf{c}_i^{\dagger}\left[\sum_{ab\in m\pm 1}^{N_{\text{aux}}} (H_{ab} + G_{ab})\mathbf{T}^{(a,b)} + (1/2)\sum_{k\in m}^{N_{\text{aux}}} (2\mathbf{J}_k - \mathbf{K}_k) \right](\delta\mathbf{c}_i). \tag{46}$$

Note that the variation of active orbitals in fragment m will introduce variation of E_{aux} at its neighboring fragments $m\pm 1$. At this point, the variation of total electronic energy can be readily written in the hybrid basis

$$\delta E_{\text{elec}} = 2\sum_{m=1}^{M}\sum_{i\in m}(\delta\mathbf{c}_i^{\dagger})\left[\mathbf{H} + \sum_{j\in m}^{N_{\text{act}}} (2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\sum_{k\in m}^{N_{\text{aux}}} (2\mathbf{J}_k - \mathbf{K}_k) + 1/2\mathbf{I} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I_{pq}^{0} + \sum_{B\in n} L_B^{0} \right)\sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)} + (1/2)\sum_{ab\in m\pm 1}^{N_{\text{aux}}} (H_{ab} + G_{ab})\mathbf{T}^{(a,b)} \right]\mathbf{c}_i + \text{c.c.}, \tag{47}$$

where c.c. denotes complex conjugate. Following the same line of argument as in the previous derivation, we find that the Fock matrix for each fragment in the hybrid basis is

$$\mathbf{F} = \mathbf{H} + \sum_{j\in m}^{N_{\mathrm{act}}} (2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\sum_{k\in m}^{N_{\mathrm{aux}}} (2\mathbf{J}_k - \mathbf{K}_k) + (1/2)\mathbf{I} - (1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I^0_{pq} + \sum_{B\in n} L^0_B\right) \times \sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)} + (1/2)\sum_{ab\in m\pm 1}^{N_{\mathrm{aux}}} (H_{ab} + G_{ab})\mathbf{T}^{(a,b)}. \quad (48)$$

The optimized MOs are obtained by solving Eq. (37) in the hybrid basis.

III. IMPLEMENTATION

III.A. Energy

We rewrite the Fock matrix in Eq. (48) as follows:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 \quad (49)$$

with

$$\mathbf{F}_1 = \mathbf{H} + \sum_{j\in m}^{N_{\mathrm{act}}} (2\mathbf{J}_j - \mathbf{K}_j) + (1/2)\sum_{k\in m}^{N_{\mathrm{aux}}} (2\mathbf{J}_k - \mathbf{K}_k) + (1/2)\mathbf{I},$$

$$\mathbf{F}_2 = -(1/2)\sum_{n\neq m}^{M}\left(\sum_{pq\in n} P_{pq}I^0_{pq} + \sum_{B\in n} L^0_B\right) \times \sum_{A\in m}\sum_{p\in A}\sum_{\lambda\in m(A)} S_{p\lambda}\mathbf{T}^{(\lambda,p)}, \quad (50)$$

$$\mathbf{F}_3 = (1/2)\sum_{ab\in m\pm 1}^{N_{\mathrm{aux}}} (H_{ab} + G_{ab})\mathbf{T}^{(a,b)}.$$

Starting from an initial guess for the electron density and charges on the MM atoms, we use the DSCF algorithm[14] to self-consistently variationally determine the total energy of the system by optimizing both the wave function of each individual molecular fragment in the presence of the rest of the system and the explicit polarization of the entire system. This DSCF optimization procedure is implemented here in the neglect-of-diatomic-differential-overlap[15] (NDDO) approximation [which underlies such popular models as AM1 (Ref. 16) and PM3], and it involves the following steps.

(a) Scale the MM charges and density matrix elements for auxiliary orbitals by 1/2, then transform the density matrix into the AO basis [Eq. (42)]. Construct $\mathbf{F}_1$ in the AO basis including all basis functions on the two boundary atoms and then transform it into the HO basis using the transformation matrix $\mathbf{T}$. $\mathbf{I}^0$ can be calculated together with $\mathbf{I}$. Transform $\mathbf{I}^0$ in the AO basis to the HO basis. Then, $\mathbf{F}_2$ can be calculated by adding QM core-unit charge interaction to $\mathbf{L}^0$ and putting the elements in the appropriate positions in the matrix. Set the auxiliary density matrix element in the HO basis to zero and transform it to the AO basis. Construct the one-electron $\mathbf{H}$ matrix for boundary atoms by neglecting kinetic energy and attraction of orbitals to their own nuclei and construct the two-electron $\mathbf{G}$ matrix using the AO basis and then transform to the hybrid basis [see Eq. (50)]. Then, the $\mathbf{F}_3$ matrix is constructed using the elements in the transformed $\mathbf{H}$ and $\mathbf{G}$ matrices. The elements of $\mathbf{F}$ are then constructed by Eq. (49).

(b) Diagonalize the Fock matrix $\mathbf{F}$ and calculate the electron density matrix. Calculate the total electronic energy associated with each fragment and check for convergence. If not converged, repeat steps (a) and (b). This represents the "inner" SCF optimization of the individual molecular wave functions.

(c) When convergence is achieved, calculate the Mulliken charges for the QM fragment and build $\mathbf{F}_2$ and $\mathbf{F}_3$ for calculation on the adjacent fragments. Move to the next QM fragment and repeat the above steps.

(d) After looping over all fragments, calculate the total electronic energy and test for convergence. Repeat (a), (b), and (c) until the change of the total electronic energy of the entire system satisfies a given convergence criterion. This constitutes the "system" SCF optimization of the explicit polarization of the fully interacting system.

III.B. Analytical first derivative of energy

The total energy of the system consists of the electronic energy (including, as usual, nuclear repulsion) of the QM subsystem plus the van der Waals terms (vdW) as follows

$$E_{\mathrm{tot}} = E_{\mathrm{elec}} + E_{\mathrm{vdW}}. \quad (51)$$

The first term is treated by the explicit electronic structural method discussed above, and the second term contains the vdW interactions. The latter account for the short-range exchange repulsion energy between atoms in different fragments and for long-range dispersion attractions that are omitted in the electronic energy term. The vdW energy is given by a sum of Lennard–Jones interactions as follows:

$$E_{\mathrm{vdW}} = \sum_{m}^{M}\sum_{n>m}^{M}\sum_{A\in m}\sum_{B\in n} 4\varepsilon_{AB}\left[\left(\frac{\sigma_{AB}}{r^{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{r^{AB}}\right)^{6}\right], \quad (52)$$

where standard combining rules are used to obtain the Lennard–Jones parameters from atomic parameters such that $\varepsilon_{AB}=\sqrt{\varepsilon_A\varepsilon_B}$ and $\sigma_{AB}=(\sigma_A+\sigma_B)/2$. Computation of the energy derivatives for these semiempirical terms is straightforward.

The electronic energy of a system with atoms treated by the generalized hybrid orbital[17] (GHO) approach can be written in the AO basis as

$$E_{\mathrm{elec}} = (1/2)\sum_{m=1}^{M}\sum_{pq\in m}^{N_{\mathrm{act}}+N_{\mathrm{aux}}} [P_{pq}(2H_{pq} + G_{pq}) + P'_{pq}I_{pq}] + E_{\mathrm{corr}}, \quad (53)$$

where density matrices $\mathbf{P}$ and $\mathbf{P}'$ in the AO basis are calcu-

lated by multiplying the auxiliary density matrix elements by 1/2 (to account for the 1/2 prefactor for the auxiliary orbitals) and setting the auxiliary density matrix elements to zero (auxiliary orbitals do not interact with MM atoms because the interactions are already determined when they are treated as active orbitals), then transforming to the AO basis. The energy correction term E_{corr} includes the interaction between auxiliary orbitals plus the nuclear attraction from the boundary atom where the auxiliary orbitals are located:

$$E_{\mathrm{corr}} = -\frac{1}{2}\sum_{m=1}^{M}\sum_{B\in m}\sum_{pq\in B} P_{pq}(2H'_{pq} + G'_{pq}), \tag{54}$$

where B is a boundary atom in fragment m. H'_{pq} corrects for the attraction from the corresponding boundary nuclei and orbital kinetic energies. In order to calculate G'_{pq}, we first set all active density to zero and scale the auxiliary density by 1/2, then transform the density matrix into the AO basis. Next, the two-electron matrix $\mathbf{G}$ is constructed using the resulting density matrix. Finally, G'_{pq} are obtained by transforming $\mathbf{G}$ into the hybrid basis.

The gradient of the total electronic energy is composed of the derivative of each term in Eq. (53). For convenience of discussion, we use E_1, E_2, and E_3 to denote QM, QM/MM, and correction energies, which are defined by

$$E_1 = \frac{1}{2}\sum_{m=1}^{M}\sum_{pq\in m} P_{pq}(2H_{pq} + G_{pq}),$$

$$E_2 = \frac{1}{2}\sum_{m=1}^{M}\sum_{pq\in m} P'_{pq}(2I_{pq}), \tag{55}$$

$$E_3 = E_{\mathrm{corr}}.$$

Consequently,

$$\frac{\partial E_{\mathrm{elec}}}{\partial X_A} = \frac{\partial E_1}{\partial X_A} + \frac{\partial E_2}{\partial X_A} + \frac{\partial E_3}{\partial X_A}, \tag{56}$$

where X_A is the Cartesian coordinate of atom A.

The derivative of E_1 can be explicitly written as follows:

$$\frac{\partial E_1}{\partial X_A} = \left(\frac{\partial E_1}{\partial X_A}\right)^{\mathrm{HF}} + \sum_{pq}\frac{\partial P^{\mathrm{AO}}_{pq}}{\partial X_A}(H^{\mathrm{AO}}_{pq} + G^{\mathrm{AO}}_{pq}). \tag{57}$$

The first term, which is the contribution from Hartree–Fock energy, is calculated in the usual fashion in the AO basis. The second term, which comes from the basis transformation of the density matrix,[16] can be expressed as

$$\frac{\partial \mathbf{P}^{\mathrm{AO}}}{\partial X_A} = \frac{\partial(\mathbf{T}^{-1})^{\dagger}}{\partial X_A}\mathbf{P}^{H}\mathbf{T}^{-1} + (\mathbf{T}^{-1})^{\dagger}\mathbf{P}^{H}\frac{\partial(\mathbf{T}^{-1})}{\partial X_A}. \tag{58}$$

The derivative of E_2 can be calculated in the same way by replacing $\mathbf{P}$ with $\mathbf{P}'$ and $2\mathbf{H}+\mathbf{G}$ with $2\mathbf{I}$.

Similarly, the derivative of correction energy E_3 can be divided into a Hartree–Fock contribution and the contribution from the derivative of the transformation matrix. The Hartree–Fock contribution can be evaluated by calculating the force between two boundary atoms using standard methods except that the density matrix in the AO basis is obtained by the transformation of density matrix in the HO basis with active density matrix elements set to zero and auxiliary density matrix elements multiplied by 1/2. The contribution from the derivative of transformation matrix can be evaluated in the same way as for E_1.

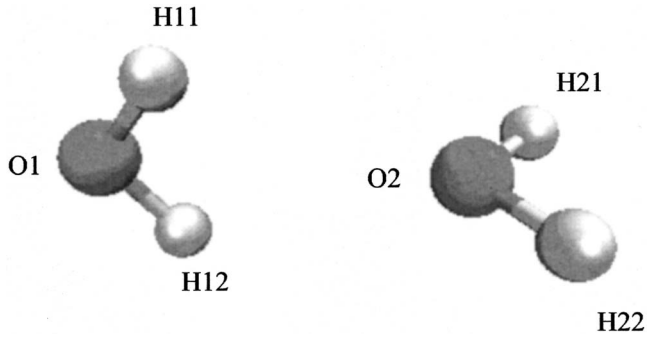

FIG. 1. Atom numbers assigned to the water dimer with each of the monomers treated as a QM fragment.

IV. RESULTS AND DISCUSSION

We have implemented the variational X-POL potential in the CHARMM (Ref. 18) package (version c33a1). In the current implementation of variational X-POL, Austin Model 1 (AM1) (Ref. 16) is chosen as the QM method and Mulliken charges are used to approximate the electrostatic potential for QM/MM interactions. Standard Lennard–Jones parameters in the CHARMM force field[18] are used for amino acids and the values reported in Ref. 8 are adopted for water.[8] The validity and implementation of the variational X-POL potential were tested for a water dimer (see Fig. 1). In a subsequent paper, we present results from molecular dynamics (MD) simulation of a solvated protein employing the X-POL potential. Table I lists the Mulliken charges calculated by AMI by using the self-consistent but not variational X-POL potential,[1,7] and by using the variational X-POL potential using the same, fixed input Cartesian coordinates. In X-POL calculations, each

TABLE I. Mulliken population charges for the water dimer from X-POL. AM1 denotes a full AM1 calculation for the water dimer, SX-POL denotes the self-consistent but not variational X-POL potential. X-POL denotes the variational result obtained using the equations of the present article. Cartesian coordinates (in Å) used for the water dimer are O1(−6.302 51, 2.127 23, 0.823 74), H11(−6.891 82, 2.556 39, 1.444 04), H12(−5.736 54, 1.587 07, 1.375 23), O2(−5.610 42, 3.122 60, −1.661 72), H21(−6.136 34, 2.539 42, −2.209 02), and H22(−5.804 20, 2.830 43, −0.771 04).

Atom	AM1	SX-POL	X-POL
O1	−0.4008	−0.4056	−0.4095
H11	0.2056	0.2028	0.2047
H12	0.2056	0.2028	0.2047
O2	−0.4104	−0.4001	−0.4109
H21	0.1840	0.1873	0.1852
H22	0.2160	0.2129	0.2257

TABLE II. Analytic and numerical forces (kcal mol^{-1} Å^{-1}) for the water dimer calculated by X-POL. Anal. denotes analytic gradient, Numer. denotes forces calculated by finite differences from displacements of atom positions, and Diff. denotes the deviation of analytic gradient from the numerical one. SX-POL and X-POL are defined in Table I. Cartesian coordinates for the water dimer are also given in Table I.

	SX-POL		X-POL		
Dim.	Atom	Numer.	Anal.	Numer.	Diff.
X	O1	1.939 851	1.909 172	1.909 175	−0.000 003
Y	O1	2.290 872	2.253 172	2.253 176	−0.000 004
Z	O1	0.029 756	0.056 810	0.056 811	−0.000 001
X	H11	2.678 915	2.671 901	2.671 899	0.000 002
Y	H11	−1.948 362	−1.939 838	−1.939 839	0.000 001
Z	H11	−3.559 570	−3.536 848	−3.536 849	0.000 001
X	H12	−2.623 117	−2.611 412	−2.611 413	0.000 001
Y	H12	−2.501 712	−2.494 665	−2.494 663	0.000 002
Z	H12	−3.246 610	−3.224 742	−3.224 742	0.000 000
X	O2	−5.501 806	−5.510 548	−5.510 546	−0.000 002
Y	O2	−7.480 413	−7.492 851	−7.492 850	−0.000 001
Z	O2	−13.727 177	−13.757 475	−13.757 482	−0.000 007
X	H21	1.798 564	1.798 540	1.798 539	0.000 001
Y	H21	2.002 657	2.002 396	2.002 397	0.000 000
Z	H21	1.760 802	1.764 064	1.764 064	0.000 000
X	H22	1.707 371	1.742 347	1.742 346	0.000 001
Y	H22	2.633 766	2.682 457	2.682 455	0.000 001
Z	H22	−8.711 621	−8.816 759	−8.816 765	0.000 006

water molecule is defined as a QM fragment. All three calculations give similar Mulliken charges with the largest deviation being 0.01.

Table II shows excellent agreement of analytic and numerical forces for the water dimer. Since the variational and nonvariational X-POL potentials give very similar Mulliken charges, it is not surprising that forces calculated by finite differences are very similar for both potentials.

It is instructive to compare the present method to some previous approaches. For example, one can compare the present method to the NDDO fragment SCF method of Ferenczy *et al.*[19] in which a system is divided into a subsystem (*S*) and an environment (*E*) with a block diagonal Fock matrix. The environmental orbitals are obtained from a zeroth-order calculation similar to our first SCF iteration. In the fragment SCF method, the density matrix of the environment and the electronic integrals of both the subsystems and the environment are used to construct the final Fock matrix of the subsystem. This is analogous to the second iteration of the subsystem in X-POL, but in the fragment SCF method, the self-consistent subsystem is not rotated through the entire system to make the whole system self-consistent.[19] In the fragment SCF method, one would not get an analytic gradient if one switches *S* and *E* because the method is not variational. A second fundamental difference between these methods is that X-POL uses combined QM/MM methodology where the MM representation is obtained from a self-consistent QM calculation. The MM representation of the environment, which has also been employed in the local SCF approximation[20] and the generalized hybrid orbital method[17,21,22] (which have some similarities to the present method in the treatment of a QM system with its environment), makes the method efficient for large molecules. On the other hand, all electronic integrals are explicit computed in the fragment SCF calculation.[19] In a different fragment orbital approach in Ref. 11, the fragmental electronic structures were iteratively converged, but the system energy was not obtained variationally. The use of QM/MM methodology also distinguishes the present method from the divide-and-conquer procedure of Lee *et al.*, which constructs a fully antisymmetric wave function for the entire system along with explicit evaluation of all one- and two-electron integrals.[23]

Another possible comparison is to polarized MM, which is reviewed[24,25] elsewhere. Empirical treatments of molecular polarization are not unique, giving rise to various representations and functional forms of polarization energy. Notable examples include the atomic interacting point dipole method,[26–28] the shell model,[29] and the chemical potential equalization approach employing fluctuating charges.[30,31] The present explicit polarization method treats molecular polarization, which is a property of the electronic structure, by a wave-function approach.[1] In MM, it is very difficult, if not impossible, to include charge transfer energy terms, whereas some charge transfer components are already incorporated in the present X-POL potential through interactions via the boundary orbitals.[1] Charge transfer, when its explicit consideration becomes necessary or desirable,[32–35] can be conveniently modeled by the X-POL potential by including the charge transfer partners in the same fragment. The X-POL method also has the advantage over polarized MM that the QM treatment employed in X-POL can be applied to bond breaking and chemical reactions for which polarized MM is inapplicable.[34,35]

V. CONCLUSION

We have presented a variational methodology to be employed in a new generation force field, called X-POL, to be ultimately used for complex problems such as protein simulations. The methodology involves using electronic structure theory and the combined QM/MM method in conjunction with a variational principle. The system is partitioned into fragments, and interactions within each fragment are treated using electronic structure theory, while interactions between fragments are treated using a combined QM/MM method. The semiemprical AM1 model is chosen as the QM method in our current implementation, although the theory and future work are not restricted to semiempirical models and can be applied to *ab initio* molecular orbital and density function theory. Mulliken charges are used in the present implementation to generate the electrostatic potential for calculating QM/MM interactions, although dipoles and multipolar terms could be included straightforwardly for systems in which anisotropic polarization beyond that which can be adequately modeled by distributed monopoles is a concern. Because the wave function of the system is variationally optimized, the gradient of the electronic energy can be calculated analytically, which provides a stable and efficient method for calculating the forces needed for MD simulations. Tests for a water dimer show that the methodology presented here gives correct analytic forces in the current implementation.

Although the X-POL method is presented and named here in light of the previously stated goal[1] of developing a next-generation force field, using direct electronic structure methods to go beyond MM, it will sometimes be convenient to distinguish the underlying QM method from the X-POL force field. In that case, the underlying method may be classified as a combined QM/QM method to distinguish it from combined QM/MM methods, and we have used a double self-consistent-field (DSCF) method to optimize the wave function.

Key elements in the present implementation of the X-POL and DSCF algorithms are the GHO boundary and the interaction of a given fragment with its environment by a QM/MM formalism. Although the GHO method was originally developed at the NDDO level,[17,21] it was later extended to *ab initio* Hartree–Fock theory,[36] to a tight-binding density functional theory including overlap,[37] and to density functional theory.[22] The DSCF optimization technique presented here at the NDDO level, can also be extended to higher levels of theory and used in various contexts.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health (GM46736), the Office of Naval Research under Grant No. N00012-05-01-0538, and the National Science Foundation under Grant No. CHE07-04974.

[1] W. Xie and J. Gao, J. Chem. Theory Comput. **3**, 1890 (2007).
[2] J. Gao, Rev. Comput. Chem. **7**, 119 (1995).
[3] A. Warshel and M. Karplus, J. Am. Chem. Soc. **94**, 5612 (1972).
[4] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[5] M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).
[6] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[7] J. Gao, J. Phys. Chem. B **101**, 657 (1997).
[8] J. Gao, J. Chem. Phys. **109**, 2346 (1998).
[9] S. J. Wierzchowski, D. A. Kofke, and J. Gao, J. Chem. Phys. **119**, 7365 (2003).
[10] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem., Quantum Chem. Symp. **13**, 225 (1979).
[11] K. Kitaura, S. Sugiki, T. Nakano, Y. Komeiji, and M. Uebayasi, Chem. Phys. Lett. **336**, 163 (2001).
[12] P. Pulay, Mol. Phys. **17**, 197 (1969).
[13] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
[14] J. Gao, J. Comput. Chem. **18**, 1062 (1997).
[15] J. A. Pople, D. P. Santry, and G. A. Segal, J. Chem. Phys. **43**, S129 (1965).
[16] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).
[17] P. Amara, M. J. Field, C. Alhambra, and J. Gao, Theor. Chem. Acc. **104**, 336 (2000).
[18] B. R. Brooks, R. E. Bruccoleni, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[19] G. G. Ferenczy, J.-L. Rivail, P. R. Surján, and G. Náray-Szabo, J. Comput. Chem. **13**, 830 (1992).
[20] V. Thery, D. Rinaldi, J.-L. Rivail, B. Maigret, and G. G. Ferenczy, J. Comput. Chem. **15**, 269 (1994).
[21] J. Gao, P. Amara, C. Alhambra, and M. J. Field, J. Phys. Chem. A **102**, 4714 (1998).
[22] J. Pu, J. Gao, and D. G. Truhlar, ChemPhysChem **6**, 1853 (2005).
[23] T.-S. Lee, J. P. Lewis, and W. Yang, Comput. Mater. Sci. **12**, 259 (1998).
[24] S. W. Rick and S. J. Stuart, Rev. Comput. Chem. **18**, 89 (2002).
[25] J. W. Ponder and D. A. Case, Adv. Protein Chem. **66**, 27 (2003).
[26] F. J. Vesely, J. Comput. Phys. **24**, 361 (1977).
[27] B. T. Thole, Chem. Phys. **59**, 341 (1981).
[28] J. Gao, D. Habibollahzadeh, and D. Shao, J. Phys. Chem. **99**, 16460 (1995).
[29] M. Sprik and M. L. Klein, J. Chem. Phys. **89**, 7556 (1988).
[30] A. K. Rappé and W. A. Goddard III, J. Phys. Chem. **95**, 3358 (1991).
[31] D. M. York and W. Yang, J. Chem. Phys. **104**, 159 (1996).
[32] A. Van der Vaart and K. M. Merz, Jr., J. Am. Chem. Soc. **121**, 9180 (1999).
[33] Y. Mo and J. Gao, J. Phys. Chem. B **110**, 2916 (2006).
[34] J. Gao, S. Ma, D. T. Major, K. Nam, J. Pu, and D. G. Truhlar, Chem. Rev. (Washington, D.C.) **106**, 3188 (2006).
[35] A. V. Marenich, R. M. Olson, A. Chamberlin, C. J. Cramer, and D. G. Truhlar, J. Chem. Theory Comput. **3**, 2055 (2007).
[36] J. Pu, J. Gao, and D. G. Truhlar, J. Phys. Chem. A **108**, 632 (2004).
[37] J. Pu, J. Gao, and D. G. Truhlar, J. Phys. Chem. A **108**, 5454 (2004).