Ab initio and semiempirical methods for molecular dynamics simulations based on general Hartree–Fock theory

Sharon Hammes-Schiffer and Hans C. Andersen
Department of Chemistry, Stanford University, Stanford, California 94305

(Received 2 February 1993; accepted 22 March 1993)

We present two new methods for molecular dynamics simulations based on general Hartree–Fock (GHF) theory. The first method involves approximating ab initio STO-3G matrix elements with fitting functions to enable faster computation of the energy and forces for molecular dynamics simulations. The implementation of this method includes a frozen-core approximation. The second method involves developing semiempirical potentials by reparametrizing the fitting functions obtained in the first method to fit experimental data. This second method enables us to reproduce experimental quantities with only the computational effort of an STO-3G calculation. We successfully applied both of these methods in conjunction with the Car–Parrinello ab initio molecular dynamics method to the geometry optimization of lithium clusters, cationic and neutral, of up to five atoms.

I. INTRODUCTION

Ab initio molecular dynamics involves performing molecular dynamics on a system of nuclei without prior knowledge of the quantum mechanical potential energy surface. In other words, the potential energy surface is calculated “on the fly.” One method of doing this is to perform a fully converged variational quantum mechanical electronic calculation at each step of the classical trajectory of the nuclei. At the present time, this is computationally prohibitive for large systems. Therefore, Car and Parrinello introduced an alternative method in which the parameters describing the wave function are treated as dynamic variables so that classical dynamics can be simultaneously applied to the wave function parameters and the nuclear coordinates. They formulated their method in the framework of density functional theory with a plane wave basis set, and many examples of density functional theory applications of the Car–Parrinello (CP) method can be found in the literature. Several groups have also applied the CP method using Hartree–Fock (HF) theory. For example, Field applied the CP method using HF wave functions in conjunction with both the semiempirical neglect of diatomic differential overlap (NDDO) method and an ab initio method using a Gaussian basis set. Hartke and Carter also applied the CP method using ab initio HF theory with a Gaussian basis set and effective core potentials. In addition, they recently applied the CP method using Hartree–Fock (HF) theory. For example, Field applied the CP method using HF wave functions in conjunction with both the semiempirical neglect of diatomic differential overlap (NDDO) method and an ab initio method using a Gaussian basis set. Hartke and Carter also applied the CP method using ab initio HF theory with a Gaussian basis set and effective core potentials. In addition, they recently applied the CP method using Hartree–Fock (HF) theory.

The paper is organized as follows. In Sec. II we discuss the details of the methods, including a brief description of GHF theory, the frozen-core method, our method for fitting the STO-3G matrix elements, and the application of the CP method using GHF theory. In Secs. III and IV we present the results of fitting the STO-3G matrix elements and of reparametrizing to fit experimental data for lithium clusters. Finally, in Sec. V we summarize our results, compare with existing methods, and discuss the many possible applications of these new methods.

II. METHODS

A. General Hartree–Fock theory

In this subsection we discuss the GHF wave function and energy expression for a system of L nuclei with positions \( \{ \mathbf{R}_l \} \), charges \( \{ Z_l \} \), and masses \( \{ M_l \} \), and \( N \) electrons with positions \( \{ \mathbf{r}_j \} \) and spin coordinates \( \{ \omega_j \} \). In GHF theory, the wave function is represented by a single \( N \)-dimensional Slater determinant \( \Phi(x_1, x_2, ..., x_N) \) made up of \( N \) orthonormal spin orbitals \( \{ \chi(x) \} \), where \( x \) represents both the position \( r \) and the spin coordinate \( \omega \) of an electron. Each spin orbital can have both a spin up (\( \alpha \) and a spin down (\( \beta \)) spatial part, \( \psi^\alpha(r) \) and \( \psi^\beta(r) \), respectively:

\[
\chi_i(x) = \psi_i^\alpha(r) \alpha(\omega) + \psi_i^\beta(r) \beta(\omega).
\]

Each spatial orbital can be expanded in a basis of \( K \) atom-centered basis functions \( \{ \phi_k(r) \} \) so that
\[ \chi_i(x) = \sum_{\mu=1}^{K} [c_{\mu}^{\alpha} \phi_{\mu}(r) \alpha(\omega) + c_{\mu}^{\beta} \phi_{\mu}(r) \beta(\omega)]. \]  

Note that GHF is different from the traditional methods of unrestricted Hartree–Fock (UHF) and restricted Hartree–Fock (RHF) in that these two methods constrain each spin orbital to be either pure \( \alpha \) or pure \( \beta \). GHF has been shown to be advantageous over UHF and RHF for molecular dynamics since there are fewer energy local minima in the space of electronic states and fewer discontinuities in the ground state wave function. Moreover, the GHF method is size consistent, and the breaking of chemical bonds to form open shell fragments is properly described.

From standard HF theory, as discussed in Ref. 8, the electronic energy \( E_{\text{elec}} \) for a given nuclear configuration and a given set of orthonormal spin orbitals is (using atomic units)

\[ E_{\text{elec}} = \sum_{i=1}^{N} \left\langle \chi_i | h | \chi_i \right\rangle + \sum_{i=1}^{N} \sum_{j<i}^{N} \left\{ \left\langle \chi_i \chi_j | \chi_i \chi_j \right\rangle - \left\langle \chi_i \chi_j | \chi_j \chi_i \right\rangle \right\}. \]

where

\[ h(r_i, \{R_j\}) = -\frac{1}{2} \nabla_i^2 + \sum_{j=1}^{L} V_f(r_i, R_j), \]

\[ V_f(r_i, R_j) = \left( \frac{Z_f}{|r_i - R_j|} \right), \]

and the two-electron integrals are specified using standard notation. According to the Born–Oppenheimer (BO) approximation, the nuclei move on a potential energy surface given by

\[ E(\{c_{\mu}^{\alpha}, c_{\mu}^{\beta}\}, \{R_j\}) = E_{\text{elec}}(\{c_{\mu}^{\alpha}, c_{\mu}^{\beta}\}, \{R_j\}) + \sum_{i,j}^{L} \left( \frac{Z_i Z_j}{|R_i - R_j|} \right). \]

For a given basis set and a given nuclear configuration, the BO ground state energy is found by minimizing \( E_{\text{elec}} \) with respect to the coefficients \( \{c_{\mu}^{\alpha}, c_{\mu}^{\beta}\} \) subject to the orthonormality constraints.

In the next two subsections we discuss two ways to speed up the calculation of the BO ground state energy. The first is a frozen-core approximation, which decreases the number of coefficients in the minimization procedure, and the second is a faster method of evaluating the matrix elements in the electronic energy. For simplicity we will discuss these approximations in terms of lithium atoms and an STO-3G basis set, but the methods are easily generalizable to other atoms and other basis sets.

### B. Frozen-core approximation

In the frozen-core approximation, the core electrons are considered as part of the nucleus so that the electronic energy expression explicitly involves only the valence electrons. Various formulations of frozen-core approximations have been discussed in the literature. We now present the frozen-core method we used in our calculations.

First an STO-3G RHF calculation was performed on a lithium atom, and the doubly occupied atomic orbital was assigned to be the 1s basis function, which we will also call the spatial core orbital, while the singly occupied atomic orbital was assigned to be the 2s valence basis function. For a system of \( L \) lithium atoms, \( 2L \) of the spin orbitals, which we will call the core spin orbitals, were chosen to be the doubly occupied \( L \) spatial core orbitals. The other \( L \) spin orbitals, which we will call the valence spin orbitals, are made up of linear combinations of the 2s, 2p\(_x\), 2p\(_y\), and 2p\(_z\) valence basis functions centered on each atom. These valence basis functions \( \{\phi_{\mu}^{\text{val}}\} \) must be orthogonalized to the \( L \) spatial core orbitals \( \{\psi_{\sigma}^{\text{core}}\} \) in order to ensure that the valence spin orbitals are orthogonal to the core spin orbitals. The orthogonalized valence basis functions \( \{\phi_{\mu}^{\text{orth}}\} \) are simply

\[ |\phi_{\mu}^{\text{orth}}\rangle = |\phi_{\mu}^{\text{val}}\rangle - \sum_{\sigma} \langle \phi_{\mu}^{\text{core}} | \phi_{\mu}^{\text{val}} \rangle |\psi_{\sigma}^{\text{core}}\rangle. \]

We neglect the lack of orthogonality among the core spin orbitals, which is small for the interatomic distances of interest.

The terms of the Hartree–Fock energy containing core orbitals can be manipulated so that the energy can be written in terms of only the \( L \) valence spin orbitals \( \{\chi_i^{\text{val}}\} \), which are linear combinations of the orthogonalized valence basis functions:

\[ E = \text{E}_{\text{core}} + \sum_{I,I'<=I} \frac{Z_I^2 Z_{I'}^2}{|R_I - R_{I'}|} + \sum_{I=1}^{L} \left( \chi_i^{\text{val}} | h_{\text{eff}} | \chi_i^{\text{val}} \right) \]

subject to the orthonormality constraints on the valence spin orbitals. \( \text{E}_{\text{core}} \) is the energy of a lithium cation with a charge of plus one, \( Z_I^2 \) is the charge of this core (i.e., plus one), and \( h_{\text{eff}} \) is the same as \( h \) defined in Eqs. (4) and (5) except \( V_f \) is replaced by \( V_f^\text{eff} \) (where \( Z_I \) is replaced with \( Z_I^2 \)). Note that for a general pair of basis functions \( \phi_{\mu} \) and \( \phi_{\nu} \),

\[ \langle \phi_{\mu} | V_f^\text{eff} | \phi_{\nu} \rangle = \langle \phi_{\mu} | V_f | \phi_{\nu} \rangle + 2 \langle \phi_{\mu} | \psi_{\nu}^{\text{core}} | \phi_{\nu} \rangle \]

where \( \psi_{\nu}^{\text{core}} \) is the core spatial orbital centered on atom \( I \).

Since the valence spin orbitals are linear combinations of the orthogonalized valence basis functions, the matrix elements in the electronic energy can be calculated from the matrix elements between orthogonalized valence basis functions. From Eq. (7) we can see that the matrix elements between the orthogonalized valence basis functions consist of the matrix elements between the original valence basis functions plus a sum of core terms. In our calculations we neglected terms higher than second order in the overlap between core spatial orbitals and valence basis functions \( \langle \psi_{\nu}^{\text{core}} | \phi_{\mu}^{\text{val}} \rangle \), and we neglected all terms containing the overlap between different core spatial orbitals.
Thus all of the STO-3G matrix elements still must be calculated, but the terms are combined in a way that allows the energy minimization to be performed on only the valence spin orbitals. We point out that this approximation will break down at small internuclear distances, but we have found that it is valid for the distances of interest for lithium clusters.

C. Fitting the matrix elements

The traditional basis functions used in HF theory are linear combinations of Gaussians. Evaluation of the matrix elements and their derivatives involves calculating many Gaussian integrals, which becomes computationally unfeasible for large systems. In order to avoid this problem, we have devised a scheme to fit the matrix elements with a small set of simple analytical functions. These analytical functions explicitly depend on the positions of the basis functions involved. Moreover, they contain parameters that are determined by the types of basis functions involved. These parameters are calculated for each particular type of atom by fitting to some of the matrix elements of small molecules involving that atom. This fitting procedure has to be performed only once for each type of atom, and then any system involving this atom can be studied.

Before we discuss the analytical functions that we used in our calculations, we will introduce some notation. Note that for a given basis function \( \phi_{\mu} \), the label \( \mu \) contains information about both the type of basis function (i.e., \( 1s, 2s, 2p_x, 2p_y, \) or \( 2p_z \)) and the atom on which it is centered. If we let \( \mu_a \) be the type of basis function and \( \mu_b \) be the atom on which it is centered, then we can use the following notation: \( \mu \equiv (\mu_a, \mu_b) \). For clarity we will often use the notation \( \phi(\mu, \nu) \) for the basis functions. The subscripts \( i, j, \) and \( k \) will always indicate \( x, y, \) and \( z \) components. Moreover, \( R_{AB} \) and \( R_{\mu_B} \) are defined as follows.

For each pair of basis functions \( \phi_{\mu_a} \) and \( \phi_{\nu_B} \), we used the STO-3G basis set to calculate the charge transfer \( C^{\mu_\nu A}_{\nu B}(R_{\mu_B} \rho) \) and the dipole vector \( D^{\mu_\nu A}_{\nu B}(R_{\mu_B} \rho) \) with components \( D^{\mu_\nu A}_{ij}(R_{\mu_B} \rho) \) and the quadrupole tensor \( Q^{\mu_\nu A}_{ij}(R_{\mu_B} \rho) \) with components \( Q^{\mu_\nu A}_{ijkl}(R_{\mu_B} \rho) \). (Note that this definition of quadrupole tensor is not the standard, traceless definition of a quadrupole.) The dipole and quadrupole moments were calculated relative to the nominal center of the overlap between the two basis functions concerned. When the two basis functions were on the same atom, the nominal center was chosen to be the position of the nucleus of this atom. In the case of two valence basis functions on different atoms, the nominal center was chosen as the midpoint of the line between the two nuclei on which the basis functions were centered. In the case of one valence basis function and one spatial core orbital, the nominal center was chosen to be the position of the nucleus corresponding to the spatial core orbital.

The one-center integrals can be divided into three categories: one-center integrals, resonance integrals, and all other integrals. We now discuss the analytical functions used to fit each of these types of matrix elements.

The one-center integrals are simply constants and were set to their \textit{ab initio} STO-3G values. Some of the one-center integrals are related to the ionization energies of a lithium atom. For example, if \( IE_1 \) is the frozen-core STO-3G first ionization energy of a lithium atom, then

\[
\langle \phi_(2s_A) | -\nabla^2/2 + V_A^\text{eff} | \phi_(2s_A) \rangle = -IE_1.
\]

We used the frozen-core STO-3G ionization energies for this type of matrix element.

A resonance integral is of the type

\[
\langle \phi_(\mu_A, A) | -\nabla^2/2 + V_A^\text{eff} + V_B^\text{eff} | \phi_(\nu_B, B) \rangle,
\]

where \( A \) and \( B \) represent different atoms. The resonance integrals involve the following types of parameters: \( b_{pm} \) (units of energy), \( a_{pm} \) (units of length), and \( k_{pm} \) (unitless). The subscripts for these parameters and all other parameters in this paper are defined as follows. The first subscript \( p \) on each parameter indicates a pair of orbitals \( \mu/\nu \), and can have integer values of 1–6, where the values are defined as follows: 1, \( 1s/1s \); 2, \( 1s/2s \); 3, \( 1s/2p \); 4, \( 2s/2s \); 5, \( 2s/2p \); and 6, \( 2p/2p \). (Note that this notation does not distinguish between the three components of the 2p orbital.) The second subscript \( m \) on each parameter indicates the type of multipole involved and is defined as follows: \( c \), charge; \( d \), dipole; and \( g \), quadrupole. All resonance integrals except those involving two 2p basis functions were fit to the product of the charge associated with the overlap of the two basis functions and a quadratic function of the internuclear separation:

\[
\langle \psi_(\mu_A, A) | -\nabla^2/2 + V_A^\text{eff} + V_B^\text{eff} | \phi_(\nu_B, B) \rangle = \frac{R_{AB}}{a_{pc}} + k_{pc} \left( \frac{R_{AB}}{a_{pc}} \right)^2.
\]

Note that for these types of resonance integrals, the first subscript \( p \) can have values 1–5, and the second subscript \( m \) is always \( c \). Resonance integrals involving two 2p basis functions were fit to the sum of two quadratic functions, the first multiplied by the charge and the second multiplied by the trace of the quadrupole tensor associated with the two basis functions:

\[
\langle \phi_(2p, A) | -\nabla^2/2 + V_A^\text{eff} + V_B^\text{eff} | \phi_(2p, B) \rangle
\]

\[
= C^{2p,2p}(R_{AB}) \frac{R_{AB}}{a_{pc}} + k_{pc} \left( \frac{R_{AB}}{a_{pc}} \right)^2
\]

\[
+ \sum_k Q^{2p,2p}(R_{AB}) \left( \frac{R_{AB}}{a_{eq}} \right)^2.
\]
We fit the resonance integral parameters using the values of the STO-3G integrals at three [Eq. (11)] or six [Eq. (12)] different vectors $\mathbf{R}_{AB}$.

All other integrals represent the Coulombic interaction of two charge distributions, each of which is either a point charge for a nucleus or a smooth distribution corresponding to the product or overlap of two basis functions. (Integrals with one point charge and one smooth distribution correspond to one-electron integrals, and integrals with two smooth distributions are two-electron integrals. In the integrals under consideration here, at least one of the two is a smooth distribution.) As discussed above, for each overlap distribution given by the product of the two basis functions $\phi_\mu$ and $\phi_\nu$, we choose a nominal center and calculate the net amount of charge $C^{\mu \nu}(\mathbf{R}_{\mu \nu})$, the net dipole moment $D^{\mu \nu}(\mathbf{R}_{\mu \nu})$, and the net quadrupole moment tensor $Q^{\mu \nu}(\mathbf{R}_{\mu \nu})$. (Higher multipole moments could be calculated if necessary.) If the two charge distributions in an integral are nonoverlapping and far apart, the usual multipole expansion can be used to estimate the value of the integral using these nominal centers, charges, dipole moments, and traceless parts of the quadrupole tensor in the usual way, without any further knowledge of the details of the distributions. If the two charge distributions are closer together and overlapping, the multipole expansion breaks down and becomes singular, but the actual Coulombic interaction is finite and nonsingular. We want to develop approximate formulas for these integrals that have the correct multipolar behavior at large distances, that are nonsingular at short distances, and that are flexible enough to fit the actual integrals for the distances that are important for low energy nuclear configurations.

In order to do this, we imagine expressing an overlap charge distribution as the sum of three parts, each of which is a smooth distribution: a spherically symmetric part with the same net charge as the true overlap (but no net dipole moment and no isotropic quadrupole tensor), a distribution with the same dipole moment as the true overlap (but no net charge and no nonzero quadrupole tensor elements), and a distribution with the same quadrupole moment tensor as the true distribution (but no net charge and no net dipole moment). The sum of all three parts has the same net charge, the same net dipole moment, and the same traceless part of the quadrupole tensor as the true overlap. As a result of this separation, the interaction between two charge distributions can be expressed as the sum of the interactions between the various parts in one distribution with the various parts in the other. Similarly, the interaction between a point charge nucleus and a charge distribution is the sum of the interactions of the point charge with the various parts of the charge distribution.

After experimenting with several different functions, we decided to use the function $h(R) = q_1 q_2 \tanh(R/a)/R$ for the interaction between two spherical charge distributions with total charges of $q_1$ and $q_2$ a distance $R$ apart. This function has the correct limiting behavior since it approaches the finite value $q_1 q_2/a$ as $R \to 0$ and it behaves like $q_1 q_2/R$ as $R \to \infty$. The same function was used for the interaction of a point charge with a spherical charge distribution. In order to calculate the interaction of charges with dipoles, dipoles with dipoles, and charges with quadrupoles, we performed a multipole expansion using this function $h(R)$ rather than the usual $q_1 q_2/R$. We point out that the multipole expansion has no true physical meaning but is rather simply a way to obtain consistent interaction functions with the correct limiting behavior.

The resulting interaction energy for a distribution with charge $C$ centered at $\mathbf{R}_1$ interacting with a distribution with dipole $\mathbf{D}$ centered at $\mathbf{R}_2$ is

$$C(\mathbf{D} \cdot \tilde{\mathbf{R}}_{12}) \left[ \frac{\text{sech}^2(R_{12}/a) \tanh(R_{12}/a)}{a R_{12}} \right],$$

where $\tilde{\mathbf{R}}_{12} = (\mathbf{R}_2 - \mathbf{R}_1)/R_{12}$.

The interaction energy for a distribution with dipole $\mathbf{D}_1$ centered at $\mathbf{R}_1$ interacting with a distribution with dipole $\mathbf{D}_2$ centered at $\mathbf{R}_2$ is

$$-f(R_{12})(\mathbf{D}_1 \cdot \mathbf{D}_2) - g(R_{12})(\mathbf{D}_1 \cdot \tilde{\mathbf{R}}_{12})(\mathbf{D}_2 \cdot \tilde{\mathbf{R}}_{12}),$$

where

$$f(R_{12}) = \frac{\frac{1}{2} \text{sech}^2(R_{12}/a) \tanh(R_{12}/a)}{a R_{12}}$$

and

$$g(R_{12}) = -3f(R_{12}) - \frac{2 \text{sech}^2(R_{12}/a) \tanh(R_{12}/a)}{a^2 R_{12}}.$$

The interaction energy for a distribution with charge $C$ centered at $\mathbf{R}_1$ interacting with a distribution with quadrupole tensor $\hat{Q}$ centered at $\mathbf{R}_2$ is

$$\frac{C}{2} \left[ f(R_{12}) \sum_i Q_{ii} + g(R_{12}) (\tilde{\mathbf{R}}_{12} \cdot \hat{\mathbf{Q}} \cdot \tilde{\mathbf{R}}_{12}) \right].$$

We found that we only had to include terms up to charge-quadrupole interactions to obtain an adequate fit to the matrix elements for studying lithium clusters. If necessary in other situations, higher-order terms can also be calculated.

Each of these functions has the desired properties of giving the correct interaction for large distances between the nominal centers of the distributions (provided the $C$, $\mathbf{D}$, and $\hat{\mathbf{Q}}$ parameters are those of the actual distributions) and a finite interaction at small distances. Note that the behavior at large distances is not affected by the length parameter $a$, so we can use these functions to fit the matrix elements by regarding the length parameters as adjustable parameters to be chosen to obtain a good fit. We point out that in this procedure, we are free to use different length parameters for the different interactions.

Thus, each pair of basis functions making up a charge distribution was assigned different length parameters $a$ associated with its charge, dipole, and quadrupole for one-electron and two-electron interactions. We point out that in this case $a$ has the physical meaning of an approximate width of the charge distribution. When the pair of basis functions involved at least one core orbital, then $a = v_p n$ for
one-electron matrix elements and $a = u_{pm}$ for two-electron matrix elements, where the subscripts are defined as above. When the pair of basis functions did not involve a core orbital, $a$ was calculated in a slightly more sophisticated way and increased with the distance $R$ between the centers of the atoms on which the two basis functions were centered. In this case there were actually two length parameters associated with each type of interaction, so

$$a = v_{pm}(1 + R/w_{pm})$$

for one-electron matrix elements and

$$a = u_{pm}(1 + R/s_{pm})$$

for two-electron matrix elements. Thus, a particular pair of basis functions could have as many as 12 length parameters associated with it, not including the resonance integral parameters. Most pairs of basis functions had fewer length parameters associated with them, however, since the dipole moment was not included unless the charge distribution involved at least one $2p$ function, and the quadrupole moment was only included for charge distributions involving two $2p$ functions.

The matrix elements were evaluated using the length parameters of the pair or pairs of basis functions involved. For example, a general one-electron matrix element $\langle \phi_{p} | V | \phi_{q} \rangle$ was represented as the interaction between a point charge at atom $I$ and a charge distribution given by the overlap of the basis functions $\phi_{p}$ and $\phi_{q}$. In this case, the length parameter used for each interaction function was simply the one associated with the pair of basis functions involved. A general two-electron matrix element $\langle \phi_{p} | \phi_{q} | \phi_{s} | \phi_{t} \rangle$ was represented by the interaction between two charge distributions, one given by the overlap of the two basis functions $\phi_{p}$ and $\phi_{s}$ and the other given by the overlap of the two basis functions $\phi_{q}$ and $\phi_{t}$. In this case, the length parameter that we used for each interaction function was the average of the two parameters associated with the two pairs of basis functions.

We determined all of the parameters by fitting to some of the $ab$ initio STO-3G matrix elements for diatomic and triatomic lithium at the internuclear distances of interest. This procedure had to be done only once in order to perform any calculations involving lithium. The parameters we used are given in Table I. In addition, Fig. 1 compares some of the fit matrix elements with the exact ones for Li$_2$. We point out that the relevant internuclear distances for lithium clusters are larger than 4.9 a.u. Figures 1(a) and 1(b) show that the one-electron valence matrix elements are fit well for the distances of interest. Due to space limitations, we are unable to show all of the two-electron matrix elements, so Fig. 1(c) shows a representative sample of the many types of two-electron matrix elements, including the worst fit that we have seen to date (curve 8). Note that curves 2, 6, 7, and 8 in Fig. 1(c) were not fit directly but were calculated using parameters that were fit to other matrix elements. Moreover, even the curves that were fit directly were only fit at one distance (or three distances for resonance integrals). Thus, the functional forms that we are using model the behavior of the matrix elements well.

This method of fitting the matrix elements contains a great deal of freedom as far as the number of interaction terms and the number of parameters included. We point
D. Geometry optimization

In order to test our fitting method, we performed geometry optimizations on lithium clusters. In this section we will discuss the procedure we used for these geometry optimizations. One possibility would have been to perform a full minimization at each different nuclear configuration and apply a conjugate gradient method. Since we are interested in applications to \textit{ab initio} molecular dynamics, however, we chose to use the CP method for geometry optimization. We now briefly describe the CP method based on GHF theory with an atom-centered basis. The reader is referred to the review articles of Madden and co-worker and Payne and co-workers for a more complete discussion of the CP method.¹¹

The CP method involves viewing the coefficients in the wave function as the coordinates of particles with a fictitious mass $m$ and performing classical dynamics on a system consisting of the coefficients $\{c_\alpha, c_\beta\}$ and the nuclear coordinates $\{R\}$, with a potential energy

$$E(\{c_\alpha, c_\beta\}, \{R\}) = E_{\text{elec}}(\{c_\alpha, c_\beta\}, \{R\}) + \sum_{i,j} \frac{Z_i Z_j}{|R_i - R_j|}.$$  (20)

The Lagrangian for this system is

$$L = \sum_{\mu} \frac{1}{2} m (|\dot{c}_\mu|^2 + |\dot{\phi}_\mu|^2) + \frac{1}{2} \sum_j M_j \dot{R}_j^2 - E(\{c_\alpha, c_\beta\}, \{R\})$$  (21)

subject to the orthonormality constraints

$$\sigma_{ij}(\{c_\alpha, c_\beta\}, \{R\}) = \sum_{\mu \nu} (c_{\alpha \mu} c_{\beta \nu} + c_{\alpha \nu} c_{\beta \mu}) S_{\mu \nu}(\{R\}) - \delta_{ij} = 0,$$  (22)

where $S_{\mu \nu}(\{R\})$ is the overlap integral between basis functions $\phi_\mu$ and $\phi_\nu$.

This Lagrangian gives rise to the following equations of motion:

$$\dot{c}_\mu = -\frac{\partial E}{\partial \phi_\mu} + \sum_j \lambda_{ij} \frac{\partial \sigma_{ij}}{\partial \phi_\mu},$$  (23)

$$M_j \ddot{R}_j = -\frac{\partial E}{\partial R_j} + \sum_{i,j<i} \lambda_{ij} \frac{\partial \sigma_{ij}}{\partial R_j},$$  (24)

where the $\lambda_{ij}$ are Lagrange multipliers and $k$ indicates the spin ($\alpha$ or $\beta$).

In order to do the constrained dynamics, we used the RATTLE algorithm,¹² which is based on the velocity Verlet

out that we included as many terms and parameters as were necessary to reproduce lithium cluster bond lengths, which are very long, and bond energies, which are very small. Since the potential energy surfaces are extremely flat, many terms must be included in order to obtain reasonable results. Thus, lithium is an especially difficult atom to parametrize. Further evidence of this is that a literature survey revealed no studies of lithium clusters using the modified neglect of diatomic overlap (MNDO) method.¹⁰

---

FIG. 1. Examples of fit matrix elements for Li₃, where atom A is at the origin and atom B is on the y axis with y=R. Long dashes represent exact STO-3G matrix elements and short dashes represent fits. (a) Resonance integrals. 1: \(2s_A | h_{\alpha f} | 2p_B \); 2: \(2p_A | h_{\alpha f} | 2p_B \); 3: \(2p_A | h_{\alpha f} | 2p_B \); 4: \(2s_A | h_{\alpha f} | 2p_B \). (b) Other one-electron integrals. 1: \(2s_A | h_{\alpha f} | 2p_B \); 2: \(2p_A | h_{\alpha f} | 2p_B \); 3: \(2p_A | h_{\alpha f} | 2p_B \); 4: \(2s_A | h_{\alpha f} | 2p_B \). (c) Some two-electron integrals. 1: \(2p_A 2s_A | h_{\alpha f} 2p_B \); 2: \(2s_A 2s_A | h_{\alpha f} 2p_B \); 4: \(2s_A 2s_A | h_{\alpha f} 2p_B \); 5: \(2s_A 2p_A | h_{\alpha f} 2p_B \); 6: \(2s_A 2s_A | h_{\alpha f} 2p_B \); 7: \(2s_A 2p_A | h_{\alpha f} 2p_B \); 8: \(2s_A 2p_A | h_{\alpha f} 2p_B \).
Since we were only doing geometry optimizations, whenever the energy increased during a time step we returned to the previous positions and coefficients, zeroed all of the velocities, and restarted the dynamics. This procedure was repeated until the energy no longer changed by more than a given tolerance. Since we were not interested in true dynamics, we were able to use a larger time step and fictitious mass than could be used for dynamics simulations. We found that values of 5000 a.u. for the fictitious mass and 125 a.u. for the time step worked well for our applications. We also developed some simple methods to speed up the minimization procedure. For example, after each time step that had started with all zero velocities, we scaled up the velocities by a common factor.

Before concluding this section, we point out that the main advantage of using the fit matrix elements is that the derivatives of the matrix elements are straightforward and fast to calculate. This allowed us to use the full forces as opposed to the Hellman–Feynman forces in our implementation of the CP method. As discussed in Ref. 6, the Hellman–Feynman forces without Pulay corrections do not conserve total energy or total linear and angular momentum. Thus, full forces are necessary to obtain accurate results.

III. AB INITIO FITTING RESULTS

In this section we present the results of applying the methods discussed in the preceding section to lithium clusters. For more information on lithium clusters the reader is referred to a recent review article on the quantum chemistry of small clusters.13

As discussed in Sec. II C, we first fit the lithium STO-3G matrix elements. Then, using these fit matrix elements, we applied the geometry optimization scheme described in Sec. II D to lithium clusters, both neutrals and cations, of up to five atoms. We used previous minimal basis set ab initio calculations on lithium clusters4 to help us choose starting geometries, but to avoid being trapped in saddle points of high symmetry we always started each optimization with asymmetric, nonplanar geometries and random coefficients. Moreover, for each cluster we performed the geometry optimization with at least two different starting geometries to be sure that the minimum nuclear configuration was independent of the starting conditions. We point out that our procedure has the usual drawback that it does not necessarily lead to the global minimum nuclear configuration.

We did RHF, UHF, and GHF calculations using the fit matrix elements and the methods described above. In addition, for comparison we used the program pel3 to perform RHF STO-3G geometry optimizations of the lithium clusters using the self-consistent-field method with the exact matrix elements. All of these results are presented in Fig. 3 and Table II, which refers to the geometries depicted in Fig. 2. For cases where more than one minimum geometry was found, we present the geometry with the lowest GHF energy.

We point out that the RHF results using the exact integrals are very similar to the RHF results using the fit integrals, which validates both the frozen-core approximation and our fitting procedure. We emphasize that only a small number of matrix elements were used to fit the parameters. Moreover, although the functional forms were chosen so as to be flexible enough to qualitatively reproduce the bond lengths and energies of small lithium clusters, these quantities were not used in the actual fitting of the parameters. Note that the RHF method using the fit integrals predicts a scalene rather than an isosceles triangle as the minimum for Li₃. Since the difference in energy between the scalene triangle minimum and the closely related isosceles triangle is only on the order of 10⁻⁵ a.u., however, the dynamics will not be significantly affected by this discrepancy.

Moreover, note that the GHF results are similar to the UHF results and do not give any unphysical results. The
TABLE II. Bond lengths and binding energies (BE) for lithium clusters using exact integrals with RHF method and fit integrals with RHF, UHF, and GHF methods. All quantities given in atomic units. Geometry and bond labels refer to Fig. 2.

<table>
<thead>
<tr>
<th>Cluster (geometry)</th>
<th>Quantity</th>
<th>Exact integrals</th>
<th>Fit integrals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RHF</td>
<td>RHF</td>
</tr>
<tr>
<td>Li$_3^-$</td>
<td>a</td>
<td>5.73</td>
<td>5.78</td>
</tr>
<tr>
<td>(1)</td>
<td>BE</td>
<td>0.055</td>
<td>0.054</td>
</tr>
<tr>
<td>Li$_3$</td>
<td>a</td>
<td>5.10</td>
<td>5.31</td>
</tr>
<tr>
<td>(1)</td>
<td>BE</td>
<td>0.0077</td>
<td>0.0022</td>
</tr>
<tr>
<td>Li$_3^-$</td>
<td>a, b</td>
<td>5.67</td>
<td>5.85</td>
</tr>
<tr>
<td>(2)</td>
<td>c</td>
<td>5.67</td>
<td>5.85</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>0.092</td>
<td>0.089</td>
</tr>
<tr>
<td>Li$_4$</td>
<td>a</td>
<td>5.21</td>
<td>5.38</td>
</tr>
<tr>
<td>(2)</td>
<td>b</td>
<td>5.21</td>
<td>5.49</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>6.56</td>
<td>7.13</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>0.016</td>
<td>0.012</td>
</tr>
<tr>
<td>Li$_4^-$</td>
<td>a</td>
<td>5.78</td>
<td>5.92</td>
</tr>
<tr>
<td>(3)</td>
<td>b</td>
<td>5.78</td>
<td>5.92</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>5.47</td>
<td>5.67</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>5.77</td>
<td>5.79</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>0.120</td>
<td>0.118</td>
</tr>
<tr>
<td>Li$_5$</td>
<td>a</td>
<td>5.73</td>
<td>5.91</td>
</tr>
<tr>
<td>(4)</td>
<td>b</td>
<td>4.80</td>
<td>5.06</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>0.048</td>
<td>0.044</td>
</tr>
<tr>
<td>Li$_5^-$</td>
<td>a</td>
<td>5.14</td>
<td>5.49</td>
</tr>
<tr>
<td>(5)</td>
<td>b</td>
<td>3.14</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>6.06</td>
<td>6.21</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>6.06</td>
<td>6.21</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>0.146</td>
<td>0.141</td>
</tr>
<tr>
<td>Li$_6$</td>
<td>a</td>
<td>5.66</td>
<td>5.83</td>
</tr>
<tr>
<td>(6)</td>
<td>b</td>
<td>5.73</td>
<td>5.84</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>5.51</td>
<td>5.71</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>5.48</td>
<td>5.43</td>
</tr>
<tr>
<td></td>
<td>BE</td>
<td>0.075</td>
<td>0.077</td>
</tr>
</tbody>
</table>

GHF prediction of an equilateral rather than an isosceles triangle for Li$_3$ is simply a mathematical artifact and will not affect the dynamics significantly since the difference between the minimum UHF energy and the minimum GHF energy is only on the order of $10^{-3}$ a.u., and the UHF equilibrium geometry is close to an equilateral triangle. Since GHF has many important advantages over RHF and UHF for molecular dynamics and does not lead to unphysical results, we used GHF for the semiempirical extensions of this method, which will be discussed in the next section.

Thus, our results show that our method of fitting the \textit{ab initio} matrix elements works well in the case of lithium. Since lithium is a particularly difficult case, as discussed in Sec. II C, this suggests that our method will work for most elements. Moreover, the same procedure can be used to fit larger basis set calculations. This method should be much faster than traditional \textit{ab initio} methods when doing molecular dynamics calculations on large systems since the energy and forces are simple analytical functions rather than integrals over Gaussian basis functions.

IV. SEMIEMPIRICAL RESULTS

The STO-3G calculations discussed in the preceding section could be improved by increasing the basis set and by adding configuration interaction, both of which are computationally costly. As an alternative to this, we adjusted the parameters in the fitting functions to fit some experimental data. In doing this, we are assuming that the fit STO-3G method contains the basic physics and that by scaling the parameters we will not destroy this physics but will increase the agreement with experiment. In the following fitting procedures, we only used the GHF method rather than the RHF or UHF methods.

Recall that the one-center matrix elements are simply constants and that we set these equal to the STO 3G values. When fitting to experiment, we replaced the constant parameters that are related to the ionization energies with the experimental values rather than the STO-3G values. Then we scaled the other constant parameters so that they increased by similar factors. After scaling the one-center matrix elements in this way, we no longer changed them.

We divided the other parameters into four groups: the length parameters for the resonance integrals, the energy parameters for the resonance integrals, the unitless parameters for the resonance integrals, and all of the length parameters for the other integrals. In our fitting procedure, we changed one parameter in each group and scaled the other parameters to maintain a constant ratio within each group. Since all of the parameters within the same group were multiplied by the same factor, only four parameters were actually adjusted. In order to fit these four parameters to experimental data, we used the simplex method to minimize a function that was a weighted combination of the percentage error in the equilibrium bond lengths and bond energies of Li$_3^+$, Li$_4^-$, and Li$_5^-$ as compared to experimental values.

Since there is not complete agreement between different experiments and different high-level theoretical calculations, there is quite a bit of freedom in which values to use. Moreover, the weighting of the different pieces of experimental data influences the results. We emphasize that our goal was not to obtain the best possible fit but rather to present the method. One set of results is shown in Table III and Fig. 4. Note that the agreement for Li$_3$, Li$_4^+$, and Li$_5$ is very good considering we did not fit any of these quantities. The bond lengths of all clusters studied are within 4% of some of the high-level theoretical results. The binding energies of the cations were already fairly good and did not improve much, but the binding energies of the neutral clusters dramatically improved, as shown in Fig. 4. Moreover, these results could be improved further if more than four parameters were varied and if more experimental data were used in the fitting procedure. Thus, the semiempirical method based on scaling the parameters used to fit the STO-3G matrix elements is very promising.

Before concluding this section, we point out that several of the concepts inherent in our semiempirical method are related to ideas contained in the MNDO method. The MNDO method is simpler, but as a result it has less flexibility than our method. For example, the MNDO method...
TABLE III. Bond lengths and binding energies (BE) for lithium clusters using our semiempirical GHF method, high-level theoretical calculations including correlation, and experimental results. All quantities given in atomic units. Geometry and bond labels refer to Fig. 2. (Note that the experimental dissociation energies quoted from Refs. 30 and 34 are not corrected for the zero vibrational energy, but this correction will not be significant for purposes of comparison with our method.)

<table>
<thead>
<tr>
<th>Cluster (geometry)</th>
<th>Quantity</th>
<th>Semi-empirical</th>
<th>High-level theory with correlation</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li$_2^+$</td>
<td>a</td>
<td>5.86</td>
<td>6.09,$^a$ 5.85,$^b$ 5.91,$^c$ 5.72$^d$</td>
<td>0.0478,$^a$ 0.0477,$^b$ 0.0601 ± 0.0032$^a$</td>
</tr>
<tr>
<td>(1)</td>
<td>BE</td>
<td>0.0451</td>
<td>0.0441,$^a$ 0.0476,$^b$ 0.047$^c$ 0.0562$^d$</td>
<td>0.55$^b$</td>
</tr>
<tr>
<td>Li$_2$</td>
<td>a</td>
<td>4.98</td>
<td>5.05,$^a$ 5.04,$^b$ 5.00,$^c$ 5.09$^d$</td>
<td>0.0375,$^a$ 0.0385,$^b$ 0.0366,$^c$ 0.0378$^e$</td>
</tr>
<tr>
<td>(1)</td>
<td>BE</td>
<td>0.0370</td>
<td>0.118,$^a$ 0.0926$^b$</td>
<td>0.104 ± 0.007$^a$</td>
</tr>
<tr>
<td>Li$_7^+$</td>
<td>a,b,c</td>
<td>5.58</td>
<td>5.66,$^a$ 5.83,$^b$</td>
<td>0.0562$^a$ 0.0478,$^b$ 0.0477,$^c$ 0.0406 ± 0.0024$^a$</td>
</tr>
<tr>
<td>(2)</td>
<td>BE</td>
<td>0.116</td>
<td>0.118,$^a$ 0.0926$^b$</td>
<td>0.0661 ± 0.0064$^a$</td>
</tr>
<tr>
<td>Li$_3$</td>
<td>a,b</td>
<td>5.05</td>
<td>5.07,$^a$ 5.6,$^b$ 5.22,$^b$ 5.3,$^c$ 5.35$^d$</td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>c</td>
<td>6.88</td>
<td>6.17,$^a$ 6.78,$^b$ 6.11,$^b$ 6.0,$^c$ 6.22$^d$</td>
<td></td>
</tr>
<tr>
<td>Li$_4^+$</td>
<td>a</td>
<td>5.51</td>
<td>5.84,$^a$</td>
<td>0.150 ± 0.008$^a$</td>
</tr>
<tr>
<td>(4)</td>
<td>b</td>
<td>5.28</td>
<td>5.24,$^d$</td>
<td>0.124 ± 0.003$^a$</td>
</tr>
<tr>
<td>Li$_4$</td>
<td>a</td>
<td>5.39</td>
<td>5.55,$^a$ 6.00$^b$</td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>b</td>
<td>4.76</td>
<td>4.94,$^a$ 5.05$^b$</td>
<td></td>
</tr>
<tr>
<td>BE</td>
<td></td>
<td>0.132</td>
<td>0.1076,$^d$ 0.1031$^d$</td>
<td>0.1</td>
</tr>
</tbody>
</table>
S. Hammes-Schiffer and H. C. Andersen: Molecular dynamics simulations

FIG. 4. Binding energy per atom as a function of cluster size (number of atoms). Circles represent fit ab initio STO-3G GHF; squares represent experimental results from Refs. 29, 30, 33, and 34. Open symbols represent neutral lithium clusters, and solid symbols represent cationic lithium clusters.

mental data or accurately known quantum mechanical results, so in practice it may be at least as accurate as density functional methods.

The methods presented in this paper have a wide variety of applications. Both the ab initio and the semiempirical methods can easily be applied to different elements and different basis sets. In addition to studying clusters, we are currently applying both methods to energy calculations of the lithium crystal. Finally, these methods could be used to generate model potentials for materials. For this purpose, the intermediate step of fitting to STO-3G results could be excluded, and the parameters could be fit directly to experimental data. In practice, however, this intermediate step is useful in that it can be used to determine the relative values of many parameters, thus reducing the number of parameters that must be adjusted to fit to experiment. In either case, the potential surface generated could be viewed simply as an empirical potential that is loosely based on GHF theory. This development of model potentials would have many interesting applications in simulations of materials.

ACKNOWLEDGMENTS

We thank Allan L. L. East for help with the ab initio calculations using PSI. We also thank Dr. W. Kob for helpful programming advice. This work was supported by the National Science Foundation Grant No. CHE-8918841 and benefited from computer resources provided by the National Science Foundation through Grant No. CHE 8821737. S. H.-S. was supported by a National Science Foundation graduate fellowship and a grant from the AT&T Bell Laboratories Graduate Research Program for Women.

7 S. Hammes-Schiffer and H. C. Andersen (in press).
15 psi 1.0 (PSITECH Inc., Watkinsville, GA, 1989).