A generalization of the charge equilibration method for non-metallic materials

Razvan A. Nistor

Department of Physics & Astronomy, University of Western Ontario, London, Ontario N6A 3K7, Canada

Jeliazko G. Polihronov and Martin H. Müser[*]

Department of Applied Mathematics, University of Western Ontario, London, Ontario N6A 5B7, Canada

Nicholas J. Mosey

Department of Chemistry, University of Western Ontario, London, Ontario N6A 5B7, Canada

(Dated: May 25, 2006)

Assigning effective atomic charges that properly reproduce the electrostatic fields of molecules is a crucial step in the construction of accurate interatomic potentials. We propose a new approach to calculate these charges, which as previous approaches, is based on the idea of charge equilibration. However, we only allow charge to flow between covalently bonded neighbors by using the concept of so-called split charges. The semi-empirical fit parameters in our approach do not only reflect atomic properties (electronegativity and atomic hardness) but also bond-dependent properties. The new method contains two popular but hitherto disjunct approaches as limiting cases. We apply our methodology to a set of molecules containing the elements silicon, carbon, oxygen, and hydrogen. Effective charges derived from electrostatic potential surfaces can be predicted more than twice as accurately as with previous works, at the expense of one additional fit parameter per bond type controlling the polarizability between two bonded atoms. Additional bond-type parameters can be introduced, but barely improve the results. An increase in accuracy of only 30% over existing techniques is achieved when predicting Mulliken charges. However, this could be improved with additional bond-type parameters.

I. INTRODUCTION

Atomistic computer simulations are often hindered by the absence of suitable force fields describing the interactions between atoms in chemically complex environments[1,2]. One important reason for this shortcoming is the difficulty in assigning effective atomic charges that properly generate the electrostatic fields produced by atoms, molecules, clusters, or solids. Electrostatic fields dominate the long-range interactions between atoms, and hence play a crucial role in many processes, such as protein folding and drug delivery. Force-field simulations of complex ferroelectric and piezoelectric solids, as well as other dielectric materials, are also impeded due to the difficulties in assigning partial charges. Our interest in the topic is motivated by the desire to better understand how the presence of calcium reduces the functionality of anti-wear films that form on rubbing cast iron surfaces when it replaces zinc as a charge-balancing counter ion in lubricant additives[3]. Such simulations are hindered by the insufficient accuracy of current model potentials for commercial lubricants on one hand, and by the numerical complexity for many-body first-principle calculations on the other hand.

Charge equilibration (QE) methods were designed to predict effective atomic charges in chemically inhomogeneous systems so that long-range interactions between atoms can be properly captured with empirical force fields[4–8]. The main idea is that atomic charges can be calculated self-consistently for a given atomic configuration using a set of empirical parameters and the electrostatic interactions between all atoms, as will be outlined below. QE methods have been applied to a variety of systems. However, their successful applications appear to be limited to relatively homogeneous chemical compositions. This may be due to the fact that the optimum QE parameterization of elements depends on the local chemical bonds[9], implying limited transferability.

Moreover, the QE approach predicts that two unlike molecules in vacuum, such as CH_4 and H_2O, each carry opposite non-zero charges even if there is a large separation between the molecules. The violation of molecular charge neutrality contradicts chemical intuition and our quantum chemical calculations, which find two electrically neutral molecules at large separation. Similar comments apply when pulling an HF molecule apart in vacuum into neutral hydrogen and fluorine atoms. The only possibility to ensure that molecules with closed electron shells remain neutral within the QE description is to (artificially) impose constraints for these molecules[6]. Adjustments of this kind, however, become questionable as soon as bonds are allowed to break.

In this study, we intend to alleviate the aforementioned shortcomings of the current QE formulation, while maintaining the idea of charge equilibration. For this purpose, we will rewrite the QE method in terms of so-called split charges by expressing the effective charge Q_i of atom i as

$$Q_i = \sum_j \bar{q}_{ij}, \tag{1}$$

where the "split charge" $\bar{q}_{ij}$ represents the charge flown from a covalently bonded neighbor atom j to atom i. Generalizing the original QE method to a

new split charge equilibration (SQE) approach ensures electric neutrality of isolated atoms or molecules and promises to enhance transferability. Other advantages are the to-be-expected rapid convergence of fast multipole methods[10–12] (charges can be binned into neutral boxes) and the possibility of unambiguously defining electric polarization when using periodic boundaries conditions. This latter aspect is particularly crucial when fluctuating box geometries are used[13].

We note that split charges have been introduced earlier in the atom-atom charge transfer (AACT) model[14] and in previous versions of the molecular mechanics (MM) method[15,16], see equation (9) in Ref. [15]. However, unlike our approach, neither MM nor AACT contains the regular QE method as a limiting case.

In the remainder of this manuscript, we compare how well different methods reproduce partial charges obtained from quantum chemical calculations. In section II, we give a brief overview of the QE method and rewrite the model in terms of split charges. In section III, we discuss the minimization techniques used to solve for the split charges, a methodology analogous to the original QE method. In the following section IV, we introduce extensions which are made possible through the SQE approach. In section V, we discuss the methodology in applying different parameterization schemes to a set of molecules containing the elements C, Si, O, and H before highlighting the results in section VI. In this first study, we will confine our attention to molecules or clusters, in which each atom is in a well-defined bonding state, i.e., all C and Si atoms are tetra-coordinate and all O are bi-coordinate.
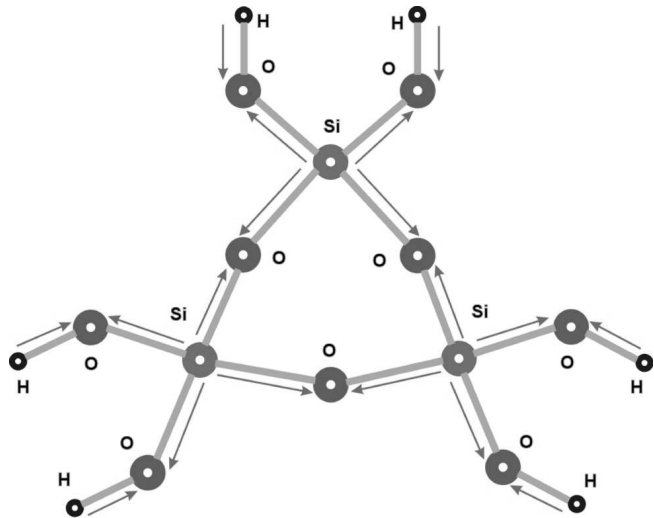

FIG. 1: Schematic representation of the charge flow due to split charges in a three-membered silica ring. Labels H, O, and Si represent hydrogen, oxygen, and silicon atoms, respectively. The total charge on an oxygen, for example, results from the amount of charge "flown" to the atom from its covalently bonded neighbours.

arguments[17,18]. However, the QE approach neglects the self-interaction induced restoring forces (hardnesses) on the charges that are non-diagonal in the index of the atoms. In fact, it had been argued that the hardness matrix should be modified to reflect better reflect molecular structures[19].

II. QE AND SQE METHODS

A. Original QE method

The starting point of the QE method is an expression for the potential energy V_Q of a molecule or cluster dependent on the effective atomic charges assigned to the nucleic centers[4,5], given by

$$V_Q = \sum_i \left\{ \frac{1}{2} \kappa_i Q_i^2 + \left(\chi_i + \Phi_i^{\text{ext}} \right) Q_i \right\} + \sum_{i,j>i} \frac{Q_i Q_j}{4\pi\epsilon_0 R_{ij}}, \quad (2)$$

where κ_i and χ_i are the electronic hardness and atomic electronegativity of atom i, respectively, Φ_i^{ext} is an external electrostatic potential at the location of atom i, and R_{ij} denotes the distance between atoms i and j. The electronic hardness and atomic electronegativity may be parameterized from the ionization potential and electron affinity of a bare atom, however, are treated as fit parameters in this work. The charges Q_i are obtained for a given atomic configuration by minimizing V_Q under the constraint of conserving the total net charge on the system. The existence of a charge restoring term in Eq. (2) related to κ_i can be motivated from first-principle

B. Extension to split-charge formalism

Using the split-charge representation from Eq. (1), Eq. (2) can be rewritten as

$$V_Q = \sum_{i,j,k} \frac{1}{2} \kappa_i \bar{q}_{ij} \bar{q}_{ik} + \sum_{i,j} \frac{1}{2} \left(\chi_i - \chi_j \right) \bar{q}_{ij} + V_{\text{C}}, \quad (3)$$

where V_{C} denotes the last two terms on the right-hand side of Eq. (2), which are the Coulomb interaction between the atomic charges plus the coupling to an external field.

The $1/R_{ij}$ dependence of the Coulomb potential is often replaced with Slater integral interactions, which are represented by functions $J_{ij}(R_{ij})$[5]. These functions take into account the effect of the overlap of the electronic wave functions of atoms i and j, thereby eliminating the Coulomb singularity when R_{ij} approaches zero. In the following, we will confine our attention to Slater orbital interactions.

Eq. (3) invites for a variety of generalizations. For example, one can constrain a split charge $\bar{q}_{ij}$ to be zero if R_{ij} exceeds a threshold value, which would best be chosen to slightly exceed the chemical bond length. This

alteration would automatically remedy the problem related to the charge neutrality of molecules that are separated by a large distance. (Possible schemes to avoid such "hard cut-offs," which would automatically lead to undesirable, discontinuous charge changes during bond breaking, will be discussed further below.) A graphical representation of the effect of locally restricting the charge flow is given in Fig. 1. As long as a cluster is fully connected via chemical bonds, the cut-off radius does not have any effect on the final charge. Another generalization is to replace expressions of the type $(\chi_i - \chi_j)$ with an arbitrary value $\bar{\chi}_{ij} = -\bar{\chi}_{ji}$. This makes it possible to include the effect of chemical induction.

Furthermore, for reasons of symmetry, the atomic hardness of atom i might be affected by the chemical nature of the atoms to which the central atom is bonded, suggesting the generalization of κ_i to κ_i^{jk} in Eq. (3). For instance, it was proposed that both electronegativity and electronic hardness of a hydrogen atom depend on whether it is bonded to a carbon or an oxygen atom[9]. Of course, it is desirable to keep the number of adjustable parameters to a minimum and to avoid parameters whose value can depend on three elements, such as the triply indexed coefficient κ_i^{jk}, unless there exists a justified motivation, as in addressing the delocalized nature of aromatic carbon-carbon bonds (e.g., sp^2 versus sp^3 hybridization).

C. Final split-charge potential

For the reasons mentioned above, we need a well-motivated choice of how to systematically generalize Eq. (3). One possibility of avoiding triply indexed coefficients is to rewrite Eq. (3) as

$$V_Q = \frac{1}{2} \sum_{i,j} \left(\kappa_{ij}^{(s)} \bar{q}_{ij} + \kappa_{ij} \bar{Q}_{ij} + \bar{\kappa}_{ij} Q_{ij} + \bar{\chi}_{ij} \right) \bar{q}_{ij} + V_C, \tag{4}$$

where $Q_{ij} = (Q_i + Q_j)/2$ and $\bar{Q}_{ij} = (Q_i - Q_j)/2$. As long as a cluster is fully connected, this expression is isomorphic to Eq. (2) provided that we maintain what we call the "charge-equilibration rules"

$$\begin{aligned} \kappa_{ij} &= (\kappa_i + \kappa_j)/2, \\ \bar{\kappa}_{ij} &= (\kappa_i - \kappa_j)/2, \\ \bar{\chi}_{ij} &= (\chi_i - \chi_j), \end{aligned} \tag{5}$$

and if the new parameter κ_{ij}^s is set to zero. This new parameter can be interpreted as a bond hardness.

By expressing the potential in terms of a bond parametrization scheme, not an atomic parametrization, we may, as a generalization, break away from the QE rules and incorporate additional flexibility into the model. An interesting feature of Eq. (4) is that charge-restoring forces that are non-diagonal in the index of atoms can be implicitly parameterized as well, e.g., by choosing the terms $\kappa_{ij}^{(s)}$ different than zero. This allows

for new avenues along which to parameterize the charge distribution in SQE. For example, if we set the atomic κ_i's, and hence implicitly the κ_{ij}'s and the $\bar{\kappa}_{ij}$'s, to zero, and allow the $\bar{\chi}_{ij}$ to differ from $\chi_i - \chi_j$, then we reproduce the AACT schemes. Moreover, the appealing aspect of the SQE method is that Eq. (4) not only makes it possible to reproduce the two limits of a pure atomic or a pure bond-based treatment. Eq. (4) also allows one to combine both approaches within one single framework. Furthermore, any generalization suggested to improve on either method that is known to us can be incorporated into the current scheme as well, e.g., a split charge $\bar{q}_{ij}$ can be rewritten as a dipole, which may be allowed to have components orthogonal to the vector connecting atoms i and j, as in the molecular mechanics methods[15,16] (In this work, we do not pursue polarization effects in the plane normal to the chemical bond, because it is probably more efficient to cast these effects in atomic polarizabilities as done, for example, in Refs. [20–24]).

Writing the model in terms of split charges, as in Eq. (4), allows three previous methods to be studied in one framework: (a) the original QE method, (b) the AACT method, and (c) a variant of the way in which electrostatics are incorporated within the molecular mechanics method. See Fig. 2 as a guide to which terms in Eq. (4) can be used to obtain these three previous methods. In addition, generalizations and combinations of pure atomic and pure bond based treatments can also be studied using the same potential. Several of these new parameterization schemes will be discussed further below. In the next section, we first establish the method of split charge equilibration.

III. MINIMIZATION TECHNIQUES

The principle of charge equilibration, or equivalently that of electronegativity equalization[19,25,26], states that charge will flow between the N atoms in a molecule until the chemical potentials of the atoms are at the same level,

$$\mu_1 = \mu_2 = \ldots \mu_k \cdots = \mu_N. \tag{6}$$

Using the fact that $\mu_i = \partial V/\partial N_i = e \partial V/\partial Q_i$, we can rewrite the above expression as

$$\frac{\partial V}{\partial Q_1} = \frac{\partial V}{\partial Q_2} = \cdots \frac{\partial V}{\partial Q_i} \cdots = \frac{\partial V}{\partial Q_N}. \tag{7}$$

A set of linear equations to be solved for the partial charges can then be constructed by equating the derivatives of Eq. (2), along with enforcing the added constraint of charge conservation. The system can be written as $\mathbf{C} \cdot \mathbf{Q} = -\tilde{\chi}$, where $\mathbf{C}$ is a matrix of coefficients given in[5], $\mathbf{Q}$ is a vector of the charges, and $\tilde{\chi}$ is a vector array of the difference in atomic electronegativity values. The system can be solved using self consistent methods, linear solvers, or extended Lagrangian schemes[6].

In the original QE method, equating the chemical potential of an atom i and an atom j is equivalent to writing

$$\frac{\partial V_Q}{\partial Q_i} - \frac{\partial V_Q}{\partial Q_j} = 0\,. \tag{8}$$

The split-charge analogue of Eq. (8) is

$$\frac{\partial V_Q}{\partial \bar{q}_{ij}} = 0, \tag{9}$$

and the identity $\partial V_Q/\partial \bar{q}_{ij} = \partial V_Q/\partial Q_i - \partial V_Q/\partial Q_j$ holds exactly for all terms in the potentials of Eqs. (2) and (4) as long as the V_Q on the left-hand side is completely isomorphic to the V_Q on the right-hand side of this identity. If we break away from the original QE rules such that the potentials in Eqs. (2) and (4) are no longer isomorphic, then we still use Eq. (9) as the split-charge equivalent of chemical potential equalization.

Taking the derivative of Eq. (4) with respect to a generic split charge labeled $\bar{q}_{kk'}$ and invoking the minimization principle of Eq. (9) results in

$$
\begin{aligned}
2\kappa_{kk'}^{(\mathrm{s})} \bar{q}_{kk'} \;+\;& \frac{1}{2}\sum_{i=1}^{N}(\kappa_{kk'} + \bar{\kappa}_{kk'} + \kappa_{ki} + \bar{\kappa}_{ki})\bar{q}_{ki} \\
-\;& \frac{1}{2}\sum_{i=1}^{N}(\kappa_{k'k} + \bar{\kappa}_{k'k} + \kappa_{k'i} + \bar{\kappa}_{k'i})\bar{q}_{k'i} \\
+\;& \frac{\partial V_{\mathrm{C}}}{\partial \bar{q}_{kk'}} = -\bar{\chi}_{kk'}\,,
\end{aligned}
\tag{10}
$$

where by choice, the symmetry relations of Eq. (5) may or may not be enforced. [In the parameterization schemes introduced below, the symmetry relations are always applied, though in one scheme the values for κ_i, and χ_i are altered depending on the chemical environment of atom i.] If the external field $\Phi_{\mathrm{ext}}^i = 0$, then it is possible to re-express the electrostatic term V_{C}, defined below Eq. (3), in the split-charge formalism as $V_{\mathrm{C}} = \sum \bar{q}_{ij} J_{ij}^{i'j'} \bar{q}_{i'j'}$, where the sum runs over all $\{i,j,i',j'\}$ and $J_{ij}^{i'j'} = (J_{ii'} - J_{ij'} - J_{ji'} + J_{jj'})$. If any of the indices in the regular electrostatic term J_{ij} are equal, then $J_{kk} = 0$, irrespective of whether or not the divergent Coulombic or Slater overlap integrals are used. Differentiating V_{C} with respect to $\bar{q}_{kk'}$ gives

$$\frac{\partial V_{\mathrm{C}}}{\partial \bar{q}_{kk'}} = \sum_{i<j}[J_{ik} - J_{ik'} - J_{jk} + J_{jk'}]\bar{q}_{ij}. \tag{11}$$

Eq. (10) and Eq. (11) can be used to construct a system of linear equations in the unknown $\bar{q}_{ij}$'s of the form $\mathbf{A}\cdot\mathbf{q} = -\tilde{\chi}$. The solution can be found as in the QE method, using linear solvers or extended Lagrangians. It is important to note, however, the system may be underdetermined. In this case, different realizations of split charges can add to give the same optimimum atomic charges as calculated in Eq. (1). In an extended Lagrangian formalism, an underdetermined system does not

cause any difficulty, as the motion of eigenvectors of the $\mathbf{A}$ matrix with eigenvalue zero remains stable. For linear solvers, singular value decompositions avoid any numerical ambiguity associated with these eigenvectors. Note that only solutions for those $\bar{q}_{ij}$ need to be found for which $i < j$ due to the relation $\bar{q}_{ij} = -\bar{q}_{ji}$.

Since Eq. (11) is independent of the atomic or bond parameterizations of the electronic hardness and electronegativity, the identity

$$\frac{\partial V_{\mathrm{C}}}{\partial \bar{q}_{ij}} = \frac{\partial V_{\mathrm{C}}}{\partial Q_i} - \frac{\partial V_{\mathrm{C}}}{\partial Q_j} \tag{12}$$

always holds and can be used to efficiently calculate the forces on the split charges if the split-charge potential of Eq. (4) is incorporated into a molecular dynamics simulation. Eq. (12) makes it possible for the current split-charge scheme to be easily incorporated into standard molecular dynamics packages.

IV. GENERALIZATIONS OF THE SQE METHOD

In the following, we specify a variety of possible parameterization schemes of V_Q, whose ability to accurately reproduce quantum chemical charges will later be tested. We will review previous methods and how they are formulated within the framework of the split-charge potential of Eq. (4). Following this brief discussion, new generalizations of the QE method will be presented. Their ability to reproduce quantum chemical data will be investigated in the next section.

A. Review of previous methods

1. The original QE method

We have discussed the isomorphism between the original QE potential[4,5] of Eq. (2) and the split-charge potential of Eq. (4) in Section II. In the split-charge formalism, the QE potential is obtained by setting the pure bond hardness term to zero, $\kappa_{ij}^{(\mathrm{s})} = 0$, and enforcing the "QE rules" of Eq. (5) such that

$$V_Q = \frac{1}{2}\sum_{i,j}\left(\kappa_{ij}\bar{Q}_{ij} + \bar{\kappa}_{ij}Q_{ij} + \bar{\chi}_{ij}\right)\bar{q}_{ij} + V_{\mathrm{C}}, \tag{13}$$

where $\kappa_{ij} = (\kappa_i+\kappa_j)/2$, $\bar{\kappa}_{ij} = (\kappa_i-\kappa_j)/2$, and $\bar{\chi}_{ij} = (\chi_i-\chi_j)$. Under the "QE rules", this potential is isomorphic to the orignal QE model.

2. The Molecular Mechanics variant

The previously mentioned way in which electrostatics are incorporated into the molecular mechanics (MM)

$$V = \frac{1}{2}\sum_{ij}\left(\kappa_{ij}^{(s)}\,\bar{q}_{ij} + \kappa_{ij}\,\bar{Q}_{ij} + \bar{\kappa}_{ij}\,Q_{ij} + \bar{\chi}_{ij}\right)\bar{q}_{ij} + V_C$$

Original QE Method ······· **AACT**

······ **MM Method** − − **New Combinations**

FIG. 2: Graphic representation of the terms in Eq. (4) that are used to obtain previous charge-equilibration methods as limiting cases. Solid Line - The original QE method[5] iff κ_{ij}, $\bar{\kappa}_{ij}$, and $\bar{\chi}_{ij}$ follow the QE rules in Eq. (5). Dotted Line - The molecular mechanics variant[15,16] in which a fixed split charge is assigned per bond type. The split charge is in fact determined by the ratio of $\bar{\chi}_{ij}/\kappa_{ij}^{(s)}$ which is a fixed value per bond type. Dashed Line - The AACT method[14] which is similar to Eq. (4), apart from absence of terms κ_{ij} and $\bar{\kappa}_{ij}$. Long-Dash Line - The split charge equilibration method of the present study that can generalize and combine each of the previous methods as described in Section IV.

variant[15,16] can also be formulated in the split-charge potential of Eq. (4). To achieve this, a fixed dipole is placed along the line of two covalently bonded atoms. In the split-charge formalism, this amounts to placing a fixed split charge on each type of bond, since the dipole is proportional to the split charge. Eq. (4) then becomes

$$V_Q = \frac{1}{2}\sum_{i,j}\left(\kappa_{ij}^{(s)}\bar{q}_{ij} + \bar{\chi}_{ij}\right)\bar{q}_{ij}\,, \qquad (14)$$

where $\bar{\chi}_{ij}$ need not obey the QE rules. Since there is no structural information of the molecule contained in this potential due to the absence of the V_C term, invoking the minimization condition of Eq. (9) results in $\bar{q}_{ij} = -\bar{\chi}_{ij}/\kappa_{ij}^{(s)}$, which remains a fixed value for each bond type.

3. The atom-atom charge transfer model

Chelli *et al.*[14] have used the concept of split charges to capture the correct polarization response of conjugated hydrocarbon molecules. Their atom-atom charge transfer (AACT) model contained the essential ingredients of the split-charge potential in Eq. (4), except it did not contain the original QE method as a limiting case. The AACT model can be thought of as the MM model plus the V_C structural term containing the Coulomb or Slater overlap integrals,

$$V_Q = \frac{1}{2}\sum_{i,j}\left(\kappa_{ij}^{(s)}\bar{q}_{ij} + \bar{\chi}_{ij}\right)\bar{q}_{ij} + V_C\,, \qquad (15)$$

where $\bar{\chi}_{ij}$ does not need to obey the "QE rules". More recent versions of molecular mechanics[27] also allow for charge transfer and polarizability of bonds.

B. Introduction of new generalizations

The split-charge potential of Eq. (4) is a compact representation in which the three previously discussed methods are contained. Moreover, because the formalism is encapsulated in the one equation, we may combine various aspects of the previous methods to investigate how accurately different parameterization schemes of V_Q reproduce given quantum chemical data. In this subsection, four parameterization schemes will be considered. The first two methods have been previously studied, but are refitted in this work for comparison. The last two methods are new parameterization schemes made possible by the split-charge formalism.

(i) Refitted QE method

The first parameterization scheme considered is the original QE method[4,5], which is used as a performance benchmark for the other schemes below. In this parameterization method, see Eq. (13), each atom type is characterized by two parameters, the electronegativity χ_i and electronic hardness κ_i, which are treated as fit parameters in this work.

(ii) Refitted MM variant

The second parameterization method considered is the refitted MM variant. This scheme is similar to that used in molecular mechanics methods in which a fixed split charge $\bar{q}_{ij}$ is placed on a bond, i.e., one fit parameter per type of bond (see Eq. (14)). Surprisingly, this simple scheme gave distinctly better results than the full AACT model[14] as long as the values of electronegativity differences $\bar{\chi}_{ij}$ and (inverse) bond polarizations $\kappa_{ij}^{(s)}$ were constrained to physically meaningful numbers. Once the bond charges are not determined self-consistently, the AACT model becomes close in spirit to models built on fixed split charges such as the AM1-BCC model[28,29].

(iii) QE + AACT

The third parameterization method considered makes full use of the split-charge formalism and its ability to combine various aspects of different methods. In this so-called QE+AACT parameterization scheme, the pure bond hardness term from the AACT method is added to the QE model. In other words, the potential has the full form of Eq. (4)

$$V_Q = \frac{1}{2}\sum_{i,j}\left(\kappa_{ij}^{(s)}\bar{q}_{ij} + \kappa_{ij}\bar{Q}_{ij} + \bar{\kappa}_{ij}Q_{ij} + \bar{\chi}_{ij}\right)\bar{q}_{ij} + V_C\,,$$

$$(16)$$

where the "QE rules" of Eq. (5) are enforced, $\kappa_{ij} = (\kappa_i + \kappa_j)/2$, $\bar{\kappa}_{ij} = (\kappa_i - \kappa_j)/2$, and $\bar{\chi}_{ij} = (\chi_i - \chi_j)$. The pure bond type term obeys the symmetry relations $\kappa_{ij}^{(s)} = \kappa_{ji}^{(s)}$. In this method, there are two parameters per atom type, the atomic electronegativities χ_i and the atomic hardnesses κ_i, plus one parameter, the pure bond hardness $\kappa_{ij}^{(s)}$, per bond type.

(iv) Perturbations added to QE + AACT

The final parameterization method is a generalization of (iii) in which chemically induced perturbations of atomic hardness and electronegativity are added, i.e.,

$$\chi_i = \chi_i^0 + \sum_{\langle j \rangle} \Delta\chi_{ij} \tag{17a}$$

$$\kappa_i = \kappa_i^0 + \sum_{\langle j \rangle} \Delta\kappa_{ij}, \tag{17b}$$

where j runs over all atoms bonded to atom i. The parameters with the superscript 0 are atomic values, which only depend on the nature of atom i. Conversely, $\Delta\chi_{ij}$ and $\Delta\kappa_{ij}$ depend on the type of bonds formed by atom i and are not taken to obey any symmetry relations, i.e., $\Delta\kappa_{ij} \not\equiv \Delta\kappa_{ji}$ and $\Delta\chi_{ij} \not\equiv -\Delta\chi_{ji}$. In this parameterization scheme, which does obey the "QE-rules", there are two fit parameters per atom type, the χ_i and κ_i, one fit parameter for the pure bond hardness $\kappa_{ij}^{(s)}$, and four additional delta corrections per type of bond (since the corrections are not symmetric).

Additional generalizations of the SQE formulation were tested, but did not result in any significant improvement over the aforementioned schemes. For example, in a pure bond parametrization where we break away from the QE rules that maintain atomic parameters on each atom type, i.e., in what we call a pure split-charge parametrization, the system is described by three parameters per bond type (κ_{ij}, $\bar{\kappa}_{ij}$, and $\bar{\chi}_{ij}$ which do not obey the "QE rules"). This pure split-charge scheme produced fits of comparable quality to method (iii) at the expense of additional fit parameters. In all, there are thirty-two possible parametrization schemes contained in Eq. (4), with the inclusion of the bond perturbation terms of Eq. (17): κ_{ij} and $\bar{\kappa}_{ij}$ can obey or break away from the QE rules; χ_{ij} can obey or break away from the QE rules; $\kappa_{ij}^{(s)}$ can be on or off; $\Delta\kappa_{ij}$ and $\Delta\chi_{ij}$ can each be on or off. However, the four methods presented above contain adequate generalizations of the QE method to demonstrate the versatility of Eq. (4).

In this study, the four aforementioned parameterization schemes only apply to situations where the bonding is well defined, i.e., all parameters with double indices apply to covalently bonded or nearest neighbor atoms. Generalizations to bond-breaking situations appear to be possible, e.g., by increasing the values of the $\kappa_{ij}^{(s)}$'s with decreasing orbital overlap between atoms i and j.

Training Set		Test Set	
1.	Si_3H_8	2.	$H_3SiOSiH_3$
4.	$(H_3Si)_3SiH$	3.	$(HO)_3SiOSiH_3$
8.	Si_6H_{12}	5.	$((HO)_2SiO)_3$
12.	$Si_5O_5H_{10}$	6.	SiH_4
13.	$(HO)_3SiSiH_3$	7.	Si_2H_6
14.	$((HO)_3SiO)_2Si(OH)_2$	9.	$(H_3SiO)_2SiH_2$
21.	$Si_4O(OH)_8$	10.	$Si_4O_4H_8$
23.	$((CH_3)_3Si)_2O$	11.	$(OSiH_2)_3$
25.	$(CH3)_2SiHC_2H_5$	15.	$Si4O_3(OH)_8$
27.	$(CH3)_2Si(CH_2)_2Si(CH_3)_2$	16.	$Si_2(OH)_6$
30.	$((CH_3)_2Si)_5$	17.	$Si_3(OH)_8$
31.	C_5H_12	18.	$[(HO)_2SiO]_4$
32.	C_3H_8	19.	$[(HO)_3SiO](HO)_2SiSi(OH)_3$
34.	CH_3OH	20.	$Si_4O_2(OH)_8$
35.	CH_4	22.	$(CH_3)_3\text{-}SiH$
37.	$CH(OH)_3$	24.	$((CH3)_3Si)_2$
38.	$C(OH)_2(CH_3)_2$	26.	$(CH_3)_3SiC_2H_5$
41.	$C(CH_3)_3(OH)$	28.	$(CH_3)SiH(CH_2)_3$
		29.	$((CH_3)_2Si)_6$
		33.	$HO\text{-}CH_2\text{-}OH$
		36.	$CH_3CH(CH_3)_2$
		39.	$CH_3\text{-}C(OH)_3$
		40.	$C(OH)_4$
(i)	28%		28%
(ii)	34%		24%
(iii)	13%		13%
(iv)	8.3%		11%

TABLE I: Training and Test sets of molecules used in this study. Molecular structures are also listed in the auxilary material[31]. The error functions over each individual set is listed for the four parameterization methods used in this study.

Finding the proper functional dependence for this spatially varying function is a difficult task, which is beyond the scope of the present work. However, we do expect the values of $1/\kappa_{ij}^{(s)}$, and thus the $\bar{q}_{ij}$'s, to be negligibly small if i and j are separated by a "next-nearest neighbor distance," while we expect the values to be similar to the stated values if R_{ij} is close to a chemical bond length. When applying SQE to bond breaking and hybridization changes, it will probably also be necessary to employ ideas used in so-called reactive empirical bond-order potentials[30]. Given the advances of SQE demonstrated further below in this work, one may anticipate similar progress for bond-breaking situations.

V. METHODOLOGY: FITTING TO *AB INITIO* CHARGES

A crucial step in this work is to first assign partial atomic charges to molecular structures from *ab initio* calculations and then to determine the adjustable parameters in each approach that best reproduce these quantum-chemical target charges Q_i^{qc}. The Q_i^{qc}'s themselves are obtained via a least-square fit of the electrostatic potential (ESP) on a surface constructed around the molecule

of interest[32,33]. The ESP is calculated using density functional theory (DFT)[34,35] at the B3LYP/6-31G(d,p) level of theory[36,37]. Final results are taken from the calculations of the Jaguar package[38]. The Q_i^{qc}'s are constrained to give the proper dipole for the molecule and to have net zero charge. We follow the argument that ESP partial charges are more meaningful than Mulliken charges,[39] which are known to be dependent on the basis set. Once the Q_i^{qc}'s are known, we select approximately 50% of our molecules to be in a training set, and the others for a test set. The molecules are listed in Table I. More details, such as exact structures and partial charges are available on the auxiliary electronic material[31]. Using the various parameterization methods (i)-(iv) of Eq. (4), we fit the semi-empirical parameters in each method to the training set.

To quantify the quality of fits, we define the normalized standard deviation σ_n for one molecule according to

$$\sigma_n^2 = \frac{\sum_i (Q_i - Q_i^{\mathrm{qc}})^2}{\sum_i (Q_i^{\mathrm{qc}})^2} . \qquad (18)$$

The overall quality of a fit is a measure of the average values of σ^2 over the total N_{C} number of molecules in the learning and test set defined by

$$\langle \sigma \rangle = \sqrt{\frac{1}{N_{\mathrm{C}}} \sum_{n=1}^{N_{\mathrm{C}}} \sigma_n^2} . \qquad (19)$$

Minimizing $\langle \sigma \rangle$ over the training set gives charges that are in the best agreement with the ESP charges. Once the optimum parameters for a particular method are found using the training set, we calculate the charges for the remaining molecules in the test set. Generally, we observe very similar values for $\langle \sigma \rangle$ evaluated over the training set and $\langle \sigma \rangle$ for the test set.

VI. RESULTS

All (DFT-relaxed) molecular structures, ESP charges, fit parameters, Slater integral parameters, and fitted charges [methods (i) through (iv)] are reported within the auxiliary electronic material[31]. This information may prove useful to users of (S)QE methods. Slater integral parameters for the elements C and O are taken from the original QE method by Rappé and Goddard[5], and updated Slater integral parameters for Si and H are taken from Ref.[40]. The set of forty-one molecules that we consider consists of three subsets. The first subset (a), labeled 1-21 on the auxiliary electronic material, only contains the elements Si, O, and H. The second set (b), structures 22-30, contains all elements of this study, while the third set (c), structures 31-41, contains (sp³-hybridized) C, O, and H.

One goal in developing model potentials is to ensure transferability. In our case, we would like to predict the

Method	a. Si-O-H	b. C-O-H	c. Si-C-O-H	a+b	a+b+c
(i)	12%	33%	13%	27%	28%
(ii)	15%	45%	27%	29%	29%
(iii)	9.5%	12%	12%	11%	13%
(iv)	9.4%	6.8%	12%	8.6%	10%
N_C	21	11	9	32	41

TABLE II: Overall fit percentages $\langle \sigma \rangle$ for each parameterization method (i)-(iv) of Section IV . The first three columns of data give the results for each family of molecules fitted separately: (a) silicates, (b) hydrocarbons, and (c) carbonsilicates. The fourth column shows the results for simultaneous fits on both the silicate and hydrocarbon families. The final column gives $\langle \sigma \rangle$ when all three families are fitted.

electrostatic interactions between, say oxygen and hydrogen atoms, with the same adjustable parameters, irrespective of whether we are concerned with subset (a), which contains silicon but not carbon, or whether we are concerned with subset (c), which contains carbon but no silicon. To test transferability of the various approaches, we determined the optimum parameters in each approach by fitting the parameters to either (a), (b), (c), (a+b), or (a+b+c). The quality of these fits is shown in Table II. The fit parameters obtained on the complete test family (a+b+c) are shown in Table III. The optimum fit parameters in the chemically homogeneous cases (a) and (b) are displayed on the auxiliary web pages[31].

In all cases, the errors evaluated on the complete set of molecules are equal to or higher than the means of the isolated fits. The results for methods (iii) and (iv) appear to change the least when new bond types are added. It is interesting to note that the original QE method does very well if applied only to subset (a). However, when other bonds are introduced, the results become significantly less satisfactory. This behavior may be interpreted as poor transferability.

It is instructive to compare the numerical values of the fit parameters for methods (i) and (iii). One can note that the values of the χ_i's do not change significantly when introducing the concept of bond hardness. This trend is consistent with the interpretation that the original QE method is the leading-order term in an expansion of effective atomic charges into atomic (one-body), bond (two-body), and other many-body terms. The trend could furthermore suggest that method (iii) is the next step in that expansion. The values for the atomic hardnesses κ_i systematically decrease from method (i) to method (iii). This is due to the fact that there is no analogue of the bond hardness in the original QE approach. It is also interesting to observe that the electronic hardness of oxygen decreases the most significantly from method (i) to (iii) and that this decrease is "compensated" by large values of κ^{s} for Si-O and C-O bonds.

Since the motivation in generalizing the QE method is to more accurately reproduce the electrostatic potential (ESP) of a molecule, we can use the charges from each parameterization scheme to generate a potential surface

and compare how accurate this surface is to the quantum chemical ESP. To quantify the difference between a calculated and a quantum ESP surface, we define a relative standard deviation on a molecule between the predicted ESP (V^{SQE}) and the quantum chemical ESP (V^{QC}) by

$$\sigma_n^2 = \frac{\sum_i [V^{\mathrm{SQE}}(\mathbf{r}_i) - V^{\mathrm{QC}}(\mathbf{r}_i)]^2}{\sum_i V^{\mathrm{QC}}(\mathbf{r}_i)^2} , \qquad (20)$$

where the sums run over all the grid points. The overall error function $\langle \sigma \rangle$ for the molecules in a set is defined as in Eq. (19). Using the parameters obtained by fitting to the ESP charges as mentioned above, we can generate V^{SQE}'s for a set of molecules and compare to the quantum chemical V^{QC}'s. For example, for the set of all molecules (i.e. the (a+b+c) set mentioned in Table II), the ESP surface errors for the different parameterization methods are: (i) 84%, (ii) 82%, (iii) 54%, (iv) 48%. The parameters in methods (i)-(iv) were fit to ESP charges, which themselves have an ESP surface error of 30%. Alternatively, we can directly fit to the quantum chemical ESP surface, instead of the ESP charges. In this case, again for the entire set of molecules (a+b+c), the final errors for the different parameterization schemes are: (i) 51%, (ii) 50%, (iii) 42%, (iv) 39%. Although the difference in errors between the refitted QE method (i), and the generalized SQE methods (iii) and (iv), is less severe when fitting to the quantum chemical ESP surface, the results show that the additional fit parameters inherent in methods (iii) and (iv) allow these schemes to reproduce quantum chemical data more accurately. Similar trends can be established for fitting to Mulliken charges. For example, the results for the (a+b+c) set of molecules are: (i) 12%, (ii) 10%, (iii) 7.2%, (iv) 4.8%.

The incorporation of a cut-off in the SQE formalism has a significant impact on the predicted ESP when a system of more than one molecule is considered. It remedies the artifact of having a resultant net charge on the molecules when they are separated by a large distance in vacuum, as predicted by the original QE method. For example, the error in the total ESP surface around a water-methane dimer, when the molecules are separated by 5 Å, improves by nearly a factor of 2 with the inclusion of a cut-off for the same set of parameters: 29% with cut-off as opposed to 59% without cut-off. These results were obtained using the QE model as formulated in the split-charge representation of method (i) using the original QE parameters.[5] Similar orders of improvement are obtained using the optimized sets of parameters. Parameters optimized on the ESP surfaces and the Mulliken charges are also listed in the auxiliary web material.

In determining the optimum SQE method between (iii) and (iv), one must balance between accuracy, transferability, and the number of fit parameters. Methods (iii) improves the results with respect to schemes (i) and (ii) by more than a factor of two when fitting to the ESP charges. However, the lowest error functions across each family of molecules is obtained using scheme (iv), which adds perturbation terms to the electronic hardness and

Method	Atom	χ_i	κ_i	Bond	
		Atomic-Type		Bond-Type	
(i) 8	H	5.0780	17.3351	NA	NA
	C	5.3039	8.6530		
	O	8.3440	14.1631		
	Si	4.4264	7.8226		
Method	**Atom**	χ_i	κ_i	**Bond**	$\bar{q}_{ij}$
(ii) 8	NA	NA	NA	H-C	0.0908
				H-O	0.3775
				H-Si	-0.0770
				C-C	0.0000
				C-O	0.2918
				C-Si	-0.1897
				O-Si	-0.2986
				Si-Si	0.0000
Method	**Atom**	χ_i	κ_i	**Bond**	$\kappa_{ij}^{(\mathrm{s})}$
(iii) 16	H	5.0780	16.1954	H-C	1.2698
	C	5.2086	8.1313	H-O	0.0627
	O	8.5220	12.4062	H-Si	2.1629
	Si	4.3850	6.7348	C-C	1.4719
				C-O	4.9727
				C-Si	2.7455
				O-Si	4.0194
				Si-Si	4.9988

TABLE III: Parameters used in the first three parameterization methods (NA - not applicable for current scheme). Listed below the method number is the total number of fit parameters used in each method for molecules containing H, C, O, and Si: (i) Refitted QE method - 2 parameters per atom type, (ii) Molecular Mechanics variant - 1 parameter per bond type (the fixed split charge placed on the bond), (iii) A combination between the QE and AACT methods - 2 parameters per atom type plus 1 parameter per bond type. All methods are contained within the framework of Eq. (4).

electronegativity as defined in Eq. (17). Thus, along with the bond hardness, an additional two corrections to the electronic hardness and two corrections to the atomic electronegativity are added per bond type (recall the perturbations are not symmetric). This gives a total of five parameters per type of bond and two parameters per type of element. In defense of such a parameter laden model, we would argue that this is the first attempt to incorporate in a QE based method a strategy that actively adapts to incorporate 3-body effects in a given chemical configuration, adjusting the hardness and electronegativity of atoms based on their covalently bonded neighbors. In comparison, method (iii) adds only one additional fit parameter to the model, namely the pure bond hardness $\kappa_{ij}^{(\mathrm{s})}$, per new bond type. Thus, considering the relative performance of method (iii), we suggest that this scheme offers the greatest improvement over the original QE method while keeping the number of additional parameters at a minimum.

There are several possible explanations for the discrepancy between the fits including all possible bond parameters and the quantum chemical data. First, the ESP fitting routine for the *ab initio* charges may not be accurate within 10%, in particular for buried atoms. A second

	ab initio	(i)	(ii)	(iii)	(iv)
1 Si	0.9382	0.9141	0.8677	0.9402	0.9952
2 O	-0.6471	-0.5384	-0.5972	-0.5377	-0.5990
3 Si	0.9367	0.9143	0.8677	0.9405	0.9955
4 C	-0.5901	-0.4587	-0.4621	-0.5086	-0.5578
5 C	-0.5623	-0.5065	-0.4621	-0.5370	-0.5928
6 C	-0.5784	-0.4742	-0.4621	-0.5177	-0.5691
7 C	-0.5936	-0.4558	-0.4621	-0.5069	-0.5558
8 C	-0.5616	-0.5031	-0.4621	-0.5352	-0.5906
9 C	-0.5713	-0.4808	-0.4621	-0.5217	-0.5740
10 H	0.1226	0.0798	0.0908	0.0924	0.1050
11 H	0.1300	0.0870	0.0908	0.0980	0.1125
12 H	0.1328	0.0879	0.0908	0.0996	0.1143
13 H	0.1150	0.0913	0.0908	0.1018	0.1163
14 H	0.1208	0.0934	0.0908	0.1024	0.1184
15 H	0.1218	0.0930	0.0908	0.1025	0.1187
16 H	0.1297	0.0903	0.0908	0.1014	0.1167
17 H	0.1248	0.0882	0.0908	0.0984	0.1134
18 H	0.1191	0.0836	0.0908	0.0955	0.1087
19 H	0.1335	0.0874	0.0908	0.0991	0.1136
20 H	0.1237	0.0791	0.0908	0.0919	0.1044
21 H	0.1317	0.0868	0.0908	0.0981	0.1125
22 H	0.1149	0.0905	0.0908	0.1012	0.1156
23 H	0.1200	0.0929	0.0908	0.1019	0.1178
24 H	0.1221	0.0927	0.0908	0.1025	0.1187
25 H	0.1224	0.0890	0.0908	0.0988	0.1140
26 H	0.1172	0.0852	0.0908	0.0968	0.1103
27 H	0.1274	0.0911	0.0908	0.1020	0.1175
$\langle\sigma\rangle$		14.89%	15.84%	10.11%	6.02 %

TABLE IV: Atomic charges on the molecule $((CH_3)_3Si)_2O$ shown in Fig. 3 as obtained by the four parameterization methods used in this study. The results at the bottom of the table indicate the overall fit of methods (i)-(iv) to the *ab initio* ESP charges listed in the third column.

and probably more likely explanation, is that there are non-negligible contributions from the polarizability tensor that our potential in Eq. (4) does not address. For example, our potential is unable to capture any polarizability components perpendicular to a bond. When we examine how our calculated charges differ from the quantum chemical charges, we begin to see certain trends arise. If the calculated charges on hydrogens connected to carbons in a molecule are more negative than the ESP charges, then the charges on the carbons will conversely tend to be more positive than the ESP charges. As an example, consider the partial charges for a molecule from the Si-C-O-H family are listed in Table IV for each method of calculation. The molecule itself is shown in Fig. 3. Method (iii) which underestimates the charges on all the hydrogens in the molecule (i.e., Q_H is more negative than the ESP hydrogens), overestimates the charges on all the carbons (i.e., Q_C is more positive than the ESP carbons). Method (iv) tends to correct the trend and provide a better fit in the end. The example may illustrate possible improvements to the split-charge potential in future work. It may be beneficial to consider how a dipole reflecting the core polarizability attached to each atom in a molecule corrects the polarization response as pursued
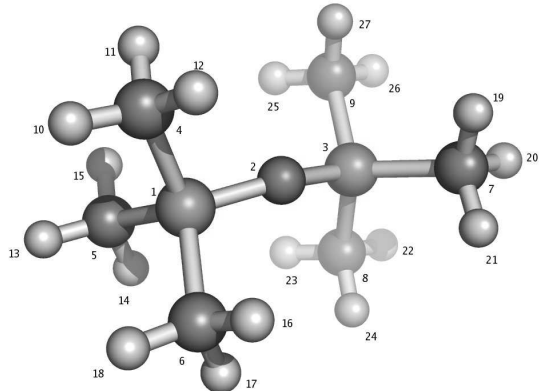

FIG. 3: Enumeration of atoms in the $((CH_3)_3Si)_2O$ molecule with their specific charges listed in Table IV as calculated by the *ab initio* Jaguar code and the four parameterization methods of this study.

in Refs. [20–24].

VII. CONCLUSIONS

In this work, we proposed a phenomenological generalization of the charge-equilibration (QE) method, which was made popular by Mortier, van Genechten, and Gasteiger[4] and also by Rappé and Goddard[5]. In the QE model, effective charges are assigned to atoms by minimizing a potential energy function in which the adjustable parameters are atomic parameters. In our approach, we represented atomic charges as split charges, which represent the amount of charge flown from one atom to another. This representation allows one to rewrite the potential in terms of bond parameters rather than atomic parameters, which in turn makes possible further generalizations that are not contained in the original QE model. One example is that the electrostatic hardness of a bond between two elements can be altered. Concepts similar to that of split charges have been used before such as in previous versions of the molecular mechanics force field[15,16] and the AACT model[14]. However, those approaches were pure bond-type descriptions and did not contain the original QE potential as a limiting case. Our approach incorporates simultaneously atom-based and bond-based descriptions within a unified framework.

Four different parameterization schemes were investigated in this study: (i) the original QE method, which was refitted to our training set of molecules, (ii) a variant of the way in which electrostatics are incorporated in the molecular mechanics methods, where fixed split charges are assigned to each type of bond, (iii) a combination of the original QE and the AACT methods, comprising of two parameters per atom type (electronic hardness and electronegativity) as well as one parameter per bond type

(bond hardness), and (iv) a generalization of the previous scheme, in which perturbative corrections were added to the electronic hardness and atomic electronegativity. These corrections depended on the nature of the elements to which the "central" atom was covalently bonded.

The parameterization schemes (i) and (ii) were able to reproduce effective atomic charges obtained from quantum chemical calculations within an accuracy of approximately 30%. Using the most simple variant of a split-charge approach, i.e., method (iii), reduced this error by a factor of 2.3 to 13%. Allowing for four more bond-dependent parameters in scheme (iv) decreases the error further to 10%. In all, method (iii) improved upon the results of the original QE method with the fewest number of additional fit parameters to the model.

Besides its higher accuracy, the split charge approach has several other advantages over the original QE method. First, the polarization of the system is defined unambiguously even if periodic boundary conditions are employed. Second, it is possible to use fixed split charges, such as the ones used in method (ii), as approximants within multiple time step schemes[41]. Third, being able to group charges into neutral entities in the summation of electrostatic interactions will be computationally advantageous in large-scale simulations. The potentially most-important benefit of our approach is that distance-dependent terms can be introduced, i.e, we expect the hardness of a bond to diverge continuously as the distance between two bonded atoms increases and the overlap of their electron shells decreases. Thus, our approach allows one to effectively introduce a bandgap within the charge equilibration method without having to artificially impose charge-neutrality constraints on individual molecules. Incorporating the suggested distance-dependent hardnesses will reveal whether the split-charge approach is also a beneficial extension of the original QE approach in bond-breaking situations.

[*] mmuser@uwo.ca; http://publish.uwo.ca/~mmuser/; MHM thanks Prof. Woo for computational resources and Prof. Baines for providing useful names for the molecular structures. Financial support from the Natural Sciences and Engineering Research Council of Canada and from SHARCnet is gratefully acknowledged.

[1] J. W. Ponder and D. A. Case. *Adv. Protein Chem.*, 66:27, 2003.

[2] S. Yip, editor. *Handbook of Materials Modeling*. Springer, Dordrecht, 2005.

[3] N. J. Mosey, M. H. Müser, and T. K. Woo. *Science*, 307:1612, 2005.

[4] W. J. Mortier, K. van Genechten, and J. Gasteiger. *J. Am. Chem. Soc.*, 107:829, 1985.

[5] A. K. Rappé and W. A. Goddard III. *J. Phys. Chem.*, 95:3358, 1991.

[6] S. W. Rick, S. J. Stuart, and B. J. Berne. *J. Chem. Phys.*, 101:6141, 1994.

[7] S. W. Rick and S. J. Stuart. *Rev. Comp. Chem.*, 18:89, 2002.

[8] A. D. MacKerell Jr. In S. Yip, editor, *Handbook of Materials Modeling*, page 509. Springer, Dordrecht, 2005.

[9] T. Ogawa, N. Kurita, H. Sekino, O. Kitao, and S. Tanaka. *Chem. Phys. Lett.*, 397:382, 2004.

[10] L. Greengard and V. Rokhlin. *J. Comput. Phys.*, 73:325, 1987.

[11] J. Carrier, L. Greengard, and V. Rokhlin. *SIAM J. Sci. Stat. Comput.*, 9:669, 1988.

[12] C. A. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon. *Chem. Phys. Lett.*, 230:8, 1994.

[13] D. Herzbach and M. H. Müser. *Comput. Phys. Comm.*, 174:17, 2006.

[14] R. Chelli, P. Procacci, R. Righini, and S. Califano. *J. Chem. Phys.*, 111:8569, 1999.

[15] L. Dosen-Micovic, D. Jeremic, and N. L. Allinger. *J. Am. Chem. Soc.*, 105:1716, 1983.

[16] L. Dosen-Micovic, D. Jeremic, and N. L. Allinger. *J. Am. Chem. Soc.*, 105:1723, 1983.

[17] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, Th. Frauenheim, S. Suhai, and G. Seifert. *Phys. Rev. B*, 58:7260, 1998.

[18] G. Tabacchi, C. J. Mundy, J. Hutter, and M. Parrinello. *J. Chem. Phys.*, 117:1416, 2002.

[19] D. M. York and W. Yang. *J. Chem. Phys.*, 104:159, 1996.

[20] H. A. Stern, G. A. Kaminski, J. L. Banks, R. Zhou, B. J. Berne, and R. A. Friesner. *J. Phys. Chem. B*, 103:4730, 1999.

[21] B. Mannfors, K. Palmo, and S. Krimm. *J. Mol. Struct.*, 556:1, 2000.

[22] H. A. Stern, F. Rittner, B. J. Berne, and R. A. Friesner. *J. Chem. Phys.*, 115:2237, 2001.

[23] C. Sagui, L. G. Pedersen, and T. A. Darden. *J. Chem. Phys.*, 120:73, 2004.

[24] C. Sagui, P. Pomorski, T. A. Darden, and C. Roland. *J. Chem. Phys.*, 120:4530, 2004.

[25] R. T. Sanderson. *Science*, 114:670, 1951.

[26] R. T. Sanderson. *Chemical Bonds and Bond Energy*. Academic Press, New York, 1976.

[27] B. Ma, J.-H. Lii, and N. L. Allinger. *J. Comput. Chem.*, 21:813, 2000.

[28] A. Jakalian, B. L. Bush, D. B. Jack, and C. I. Bayly. *J. Comput. Chem.*, 21:132, 2000.

[29] A. Jakalian, D. B. Jack, and C. I. Bayly. *J. Comput. Chem.*, 23:1623, 2002.

[30] D. W. Brenner, O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni, and S. B. Sinnott. *J. Phys.: Condens. Matt*, 14:783, 2002.

[31] See EPAPS Document No. **to be filled in by editor**. The electronic material is currently also located on: http://www.apmaths.uwo.ca/~mmuser/Molecules/ It contains all molecular structures, ESP charges, model parameters, and calculated charges, along with standard deviations.

[32] L. E. Chirlian and M. M. Francl. *J. Comp. Chem.*, 8:894, 1987.

[33] C. M. Breneman and K. B. Wiberg. *J. Comp. Chem.*, 11:361, 1990.

[34] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B864, 1964.

[35] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.

[36] A. D. Becke. *J. Chem. Phys.*, 98:5648, 1993.

[37] C. Lee, W. Yang, and R. G. Parr. *Phys. Rev. B*, 37:785, 1988.

[38] *Jaguar* 5.0, Schrödinger, LLC: Portland, Oregon, 2002.

[39] R. S. Mulliken. *J. Chem. Phys.*, 23:1833, 1955.

[40] J. Sefcik, E. Demiralp, T. Cagin, and W. A. Goddard III. *J. Comput. Chem.*, 23:1507, 2002.

[41] M. Tuckerman, B. J. Berne, and G. J. Martyna. *J. Chem. Phys.*, 97:1990, 1992.