JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

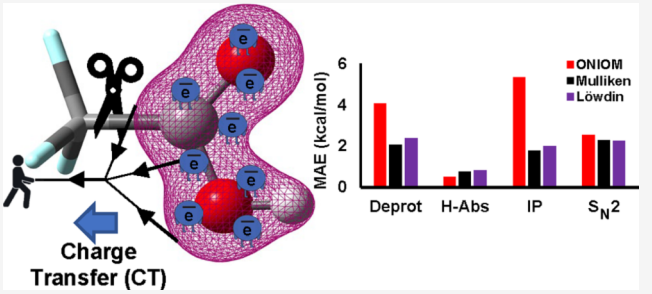
ONIOM Method with Charge Transfer Corrections (ONIOM-CT): Analytic Gradients and Benchmarking

Vikrant Tripathy, Nicholas J. Mayhall, and Krishnan Raghavachari*

Cite This: J. Chem. Theory Comput. 2022, 18, 6052−6064 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Hybrid methods such as ONIOM that treat different regions of a large molecule using different methods are widely used to investigate chemical reactions in a variety of materials and biological systems. However, there are inherent sources of significant errors due to the standard treatment of the boundary between the regions using hydrogen link atoms. In particular, an unbalanced charge distribution in the chemically important model region is a potential source of such problems. We have previously suggested ONIOM-CT (ONIOM with charge transfer corrections) which addresses this issue by applying a potential in the form of point charges to obtain a desired charge redistribution. The metric for charge redistribution relies on the type of population analysis used to obtain the charges. ONIOM-CT has been implemented using Mulliken and Löwdin population analyses and has been shown to improve computed reaction energies for illustrative chemical reactions. In this work, we derive and implement the analytic gradients for ONIOM-CT that requires solving two sets of coupled-perturbed self-consistent equations, one each for the model system and the full system. However, both are needed only at the low level of theory, allowing for an efficient formulation and implementation for both Mulliken and Löwdin population analyses. Benchmarking and illustrative geometry optimizations have been carried out for a previously studied set of reactions involving a single link atom between regions. Additionally, we have generalized our method for the treatment of model systems involving multiple link atoms to enable applications for a broader set of problems. The generalized methods are illustrated for both charge models. Furthermore, we have studied a set of three proton transfer reactions and demonstrate that significant improvement is achieved by ONIOM-CT over ONIOM using both Mulliken and Löwdin population analyses.



1. INTRODUCTION

Accurate quantum chemical calculations of the energies and properties of small molecules are currently achievable due to the systematic development of ab initio methods with ever-increasing levels of sophistication.[1,2] However, the steep scaling of these accurate methods with system size renders them impractical for many larger molecules of practical interest.[3] There is a clear need for the development of new methods which reduce the computational scaling without a significant compromise in accuracy. Hybrid methods are one such family of methods which use the concept of multiscale modeling to treat different parts of the molecule with different levels of theory depending on their chemical importance.[4,5] Molecular energies and properties are then derived as an appropriate sum of the corresponding components from individual calculations. ONIOM [our own N-Layer integrated molecular orbital (MO) molecular mechanics] is one such QM/QM method which treats the chemically important region with high accuracy and the rest of the molecule with lower accuracy.[6] The energy for the simplest two-layer ONIOM method is

$$E_{\mathrm{ONIOM}} = E_{\mathrm{RL}} + E_{\mathrm{MH}} - E_{\mathrm{ML}} \quad (1)$$

Here, we use the standard notation where "high" and "low" refer to the two levels of theory, and "model" and "real" refer to the active region and the full molecule, respectively. Thus RL, MH, and ML denote full molecule at low-level theory, model region at high-level theory, and model region at low-level theory, respectively. This extrapolation leads to three inexpensive calculations instead of one expensive calculation (high-level theory on the full molecule) and often results in a substantial computational speed-up, especially when the model region is much smaller than the full system. In addition, the ability to easily combine any two QM theories results in a broad scope for applications, making ONIOM a widely used method for the study of large molecules. In addition to energies, ONIOM has been implemented in conjunction with gradients,[7] force constants and vibrational intensities, ex-

Received: June 4, 2022
Published: September 26, 2022

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

citation energies,[8] implicit solvation models,[9] and so forth to enable a range of applications involving large molecules.

Division of a large molecule into a minimum of two regions is a necessary condition for the application of ONIOM.[10,11] If the regional boundary does not pass through any covalent bonds, the need for a special treatment of the boundary does not arise. However, if the boundary passes through any covalent bond (typically the case for most chemical problems), it requires special attention because the model region calculations may be unphysical with a dangling bond. Due to the simple formulation and often robust performance, hydrogen link atoms are the most widely used method to cap these dangling bonds in both ONIOM and QM/MM.[12,13] However, this simple solution might not be appropriate in cases where a charge imbalance is created by the replacement of the heavy-atom in the regional boundary with a hydrogen. There have been several different approaches in QM/MM to improve on this, such as the GHO (generalized hybrid orbital) approach by Gao et al.,[14] LSCF (local self-consistent field) method by Rivail and co-workers,[15] TBRC (tuned and balanced redistributed-charge) by Wang and Truhlar,[16] and covEPE (covalent elastic polarizable environment) by Nasluzov and co-workers[17] using a tuned fluorine pseudopotential. These methods use different properties of the model region to obtain the parameters for the respective methods. For example, the localized orbital methods use the local geometry and connectivity of the boundary atom, covEPE uses a combination of Mulliken charge and bond parameters of the boundary atom (Si in their case), and TBRC uses net Mulliken charge of the model region using a small basis set.

In an earlier paper, Mayhall and Raghavachari[18] demonstrated an improvement in performance of ONIOM for the prediction of reaction energies by the use of a potential in the form of a point charge placed on the link atom (i.e., effectively changing the link atom nuclear charge) to achieve charge balance in the model region. This method was termed ONIOM-CT (ONIOM with charge transfer) and was efficient in achieving the necessary charge redistribution. Because traditional ONIOM is already in widespread use,[4,19−23] the possibility of further improvement via an efficient formalism of ONIOM-CT advocates for further explorations and related developments. We will refer to ref 18 as "paper 1" throughout the remainder of the text.

The central idea in ONIOM-CT was to correct for the imbalance in the charge distribution in the chemically important model region introduced by the hydrogen link atom. The performances of three different charge models were analyzed in paper 1. All three charge models, viz. Mulliken,[24] Löwdin,[25] and Hirshfeld charges[26] performed better than ONIOM with Löwdin charges showing the best performance. In a later paper, Jose and Raghavachari[27] introduced a new method called ESP-ONIOM-CT which used the electrostatic potential of the model system to obtain the charge on the link atom needed to achieve charge balance. In a different approach, Beckett et al.[28] used a CDFT (constrained density functional theory) formalism to achieve charge balance and improve the performance of ONIOM.

In the present work, we formulate the analytic gradient expression of the ONIOM-CT energy and present an efficient algorithm for its evaluation. ONIOM-CT gradients have been implemented with two common charge analysis methods (Mulliken and Löwdin population analyses), and the method is also extended to multiple link atoms. One and two link atom methods are benchmarked using illustrative chemical reactions. The computational cost analysis shows that the additional computations required for the ONIOM-CT gradient are modest, particularly when compared to the full system calculation at the target method.

2. ONIOM-CT GRADIENT FORMALISM

This section describes, in detail, the theory and implementation of the ONIOM-CT gradient. We start by reiterating the basic principle of ONIOM-CT from paper 1, illustrated graphically in Figure 1.

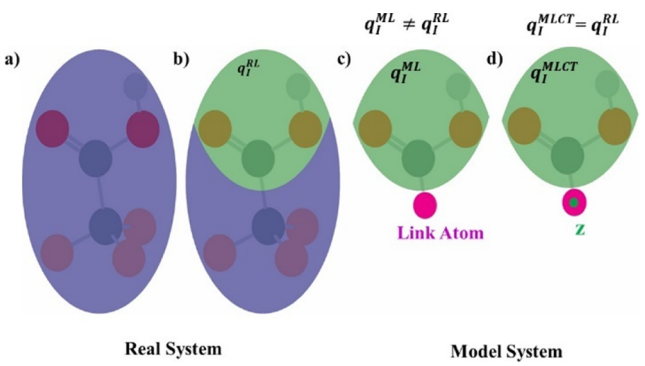


Figure 1. Graphical illustration of the workflow in model system preparation. (a) Full system, (b) appropriate division of molecule, (c) chemically important model region capped with a link atom (pink), and (d) appropriate charge placed on the nucleus of link atom (z). q_I^{RL} and q_I^{ML} correspond to net charge in the chemically important model region I in RL and ML calculations, respectively. q_I^{MLCT} is the net charge in the model region with the modified link atom nuclear charge.

The chemically important model region is defined by dividing the molecule into two parts using a hydrogen link atom to terminate the dangling bond. As noted earlier, this can cause a charge imbalance in the chemically important model region (denoted as region I) because, in general, $q_I^{ML} \neq q_I^{RL}$. Here, q_I^{ML} and q_I^{RL} represent the charges obtained by summing the charges on the atoms within the model region in the ML and RL calculations, respectively. The preparation of the model system is considered to be balanced if the charge in the model region is equal in both RL and ML calculations. In other words, the bond-cutting does not affect the number of electrons in the model region. This is achieved by modifying the nuclear charge on the link atom to achieve charge matching in the model region, i.e., $q_I^{MLCT} = q_I^{RL}$, where q_I^{MLCT} is the model region charge with the modified link atom nuclear charge. The nuclear charge on the link atom is iteratively modified until convergence is achieved within a predefined threshold, as illustrated in Figure 2.

Once the convergence is achieved, the same nuclear charge on the link atom is used in the MH calculation to obtain E_{MHCT}. Finally, the ONIOM-CT energy is given by

$$E_{ONIOM-CT} = E_{RL} + E_{MHCT} - E_{MLCT} \quad (2)$$

Here, we note that the ML calculations are much cheaper than both RL and MH calculations. Thus, a few iterations of this cheap step do not significantly increase the computational cost (vide infra).

To simplify the following derivations, we start with only one link atom and generalize to multiple link atoms later in this

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

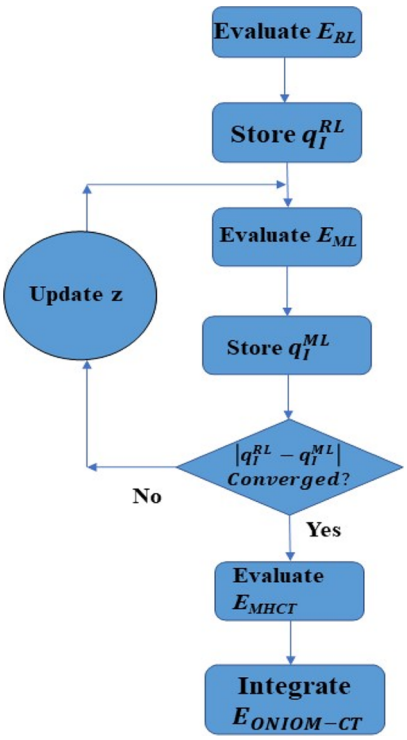


Figure 2. Flowchart for ONIOM-CT energy calculation.

section. The formulation of the ONIOM-CT gradient is illustrated more efficiently by recasting the energy expression in eq 2. This is achieved by separating the embedding contributions and writing it in the new form shown below

$$E_{\mathrm{ONIOM-CT}} = E_{\mathrm{RL}} + \tilde{E}_{\mathrm{MH}} - \tilde{E}_{\mathrm{ML}} + z\phi^{\Delta} \tag{3}$$

Here, $\tilde{E}_{\mathrm{MH}}$ and $\tilde{E}_{\mathrm{ML}}$ represent the MH and ML energies without their embedding contributions. z is the extra nuclear charge and ϕ^{Δ} is the difference in the electrostatic potential at the link atom between the MH and ML calculations, ϕ_{MH} and ϕ_{ML}, as shown in eqs 4–6.

$$\phi^{\Delta} = \phi_{\mathrm{MH}} - \phi_{\mathrm{ML}} \tag{4}$$

$$\phi_{\mathrm{MH}} = \sum_{\mathrm{A}} \frac{Z_{\mathrm{A}}}{R_{\mathrm{Al}}} - \sum_{\mu_{\mathrm{H}},\nu_{\mathrm{H}}} P_{\mu_{\mathrm{H}}\nu_{\mathrm{H}}} \langle \mu_{\mathrm{H}} | r_{\mathrm{l}}^{-1} | \nu_{\mathrm{H}} \rangle \tag{5}$$

$$\phi_{\mathrm{ML}} = \sum_{\mathrm{A}} \frac{Z_{\mathrm{A}}}{R_{\mathrm{Al}}} - \sum_{\mu_{\mathrm{L}},\nu_{\mathrm{L}}} P_{\mu_{\mathrm{L}}\nu_{\mathrm{L}}} \langle \mu_{\mathrm{L}} | r_{\mathrm{l}}^{-1} | \nu_{\mathrm{L}} \rangle \tag{6}$$

where Z_{A} is the nuclear charge of A^{th} nucleus of the model system, μ_{H} and ν_{H} are the basis functions of the high-level model system calculation, μ_{L} and ν_{L} are the basis functions of the low-level model system calculation, R_{Al} is the distance of the link atom from A^{th} nucleus, and r_{l} is the distance of the electron from the link atom which is integrated over all space. Because the nuclear contribution is the same in ML and MH, it cancels and does not appear in the gradient formalism. Adhering to the notation followed by the literature in the community, Greek letters α, β, γ, δ, σ, μ, ν represent atomic orbitals (AOs), i, j, k,... denote occupied MOs, a, b, c,... denote virtual MOs, p, q, r,... represent all MOs, and P, S, C matrices stand for the density, overlap, and MO coefficient matrices, respectively.

The new form makes it convenient to calculate the gradient as it separates the contribution of the extra nuclear charge (z). The gradient equation obtained from differentiating eq 3 is the following

$$E^{x}_{\mathrm{ONIOM-CT}} = E^{x}_{\mathrm{RL}} + \tilde{E}^{x}_{\mathrm{MH}} - \tilde{E}^{x}_{\mathrm{ML}} + z\phi^{\Delta x} + z^{x}\phi^{\Delta} \tag{7}$$

There are three different calculations from which we need to calculate the total gradient. Thus, it is necessary to group them into these three terms which correspond to RL, ML, and MH. E^{x}_{RL}, $\tilde{E}^{x}_{\mathrm{MH}}$, and $\tilde{E}^{x}_{\mathrm{ML}}$ can be evaluated in the same way as the regular ONIOM gradient. $\phi^{\Delta x}$ can also be grouped into $\tilde{E}^{x}_{\mathrm{MH}}$ and $\tilde{E}^{x}_{\mathrm{ML}}$ as it is a simple modification to the electron-nuclear integral derivative term. However, it is not obvious how to obtain z^{x} as z is obtained iteratively at each step and depends on both RL and ML calculations. Let us revisit the condition satisfied by z. The condition is the following

$$q_{\mathrm{I}}^{\mathrm{RL}} = q_{\mathrm{I}}^{\mathrm{MLCT}}(z) \tag{8}$$

$q_{\mathrm{I}}^{\mathrm{RL}}$ depends only on the geometry of the molecule. However, unlike regular ONIOM, $q_{\mathrm{I}}^{\mathrm{MLCT}}$ has dependence both on the geometry and the modified nuclear charge (z). Differentiating eq 8, we get

$$\frac{\mathrm{d}q_{\mathrm{I}}^{\mathrm{RL}}}{\mathrm{d}x} = \frac{\partial q_{\mathrm{I}}^{\mathrm{MLCT}}}{\partial x} + \frac{\partial q_{\mathrm{I}}^{\mathrm{MLCT}}}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}x} \tag{9}$$

Rearranging

$$\frac{\mathrm{d}z}{\mathrm{d}x} = B\left(\frac{\mathrm{d}q_{\mathrm{I}}^{\mathrm{RL}}}{\mathrm{d}x} - \frac{\partial q_{\mathrm{I}}^{\mathrm{MLCT}}}{\partial x}\right) \tag{10}$$

$$B = \left(\frac{\partial q_{\mathrm{I}}^{\mathrm{MLCT}}}{\partial z}\right)^{-1} \tag{11}$$

Here, B in eq 11 represents the inverse of the response of model region charge with respect to the link atom nuclear charge in the ONIOM-CT calculation. Thus, the derivative z^{x} in eq 10 can be expressed in terms of the gradients of the charges on the atoms in the model region and B. While analytic gradients of the charges with respect to geometry can be obtained for both Mulliken and Löwdin charges (vide infra), it is not possible to obtain an analytic expression for B though it can be easily evaluated numerically. Inserting eqs 4 and 10 into eq 7, we get

$$E^{x}_{\mathrm{ONIOM-CT}} = E^{x}_{\mathrm{RL}} + \tilde{E}^{x}_{\mathrm{MH}} - \tilde{E}^{x}_{\mathrm{ML}} + z\phi^{x}_{\mathrm{MH}} - z\phi^{x}_{\mathrm{ML}} + \left(\frac{\mathrm{d}q_{\mathrm{I}}^{\mathrm{RL}}}{\mathrm{d}x}\right)B\phi^{\Delta} - \left(\frac{\partial q_{\mathrm{I}}^{\mathrm{MLCT}}}{\partial x}\right)B\phi^{\Delta} \tag{12}$$

The various terms in eq 12 can be grouped into one of the three calculations (RL, ML, or MH) as shown in eqs 14–16, and the gradient of ONIOM-CT can be expressed in a simplified form as

$$E^{x}_{\mathrm{ONIOM-CT}} = \hat{E}^{x}_{\mathrm{RL}} + \hat{E}^{x}_{\mathrm{MH}} - \hat{E}^{x}_{\mathrm{ML}} \tag{13}$$

$$\hat{E}^{x}_{\mathrm{RL}} = E^{x}_{\mathrm{RL}} + \left(\frac{\mathrm{d}q_{\mathrm{I}}^{\mathrm{RL}}}{\mathrm{d}x}\right)B\phi^{\Delta} \tag{14}$$

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052–6064

$$\hat{E}^{x}_{\mathrm{ML}} = \tilde{E}^{x}_{\mathrm{ML}} + z\phi^{x}_{\mathrm{ML}} + \left(\frac{\partial q_{\mathrm{I}}^{\mathrm{MLCT}}}{\partial x}\right) B\phi^{\Delta} \tag{15}$$

$$\hat{E}^{x}_{\mathrm{MH}} = \tilde{E}^{x}_{\mathrm{MH}} + z\phi^{x}_{\mathrm{MH}} \tag{16}$$

Using these grouped expressions, all three calculations can be performed independently. However, unlike ONIOM, where the three calculations can be performed at any order, the order of calculations is critical. To calculate $\hat{E}^{x}_{\mathrm{RL}}$, we need both B andϕ^{Δ}. B is obtained numerically after the convergence of $q_{\mathrm{I}}^{\mathrm{MLCT}}$. ϕ^{Δ} needs both ML and MH calculation at the converged charge. $\hat{E}^{x}_{\mathrm{ML}}$ also requires both B and ϕ^{Δ}. $\hat{E}^{x}_{\mathrm{MH}}$ only requires the converged charge (z). Considering all these requirements, an optimum order of calculations has been designed and implemented, as shown in Figure 3.

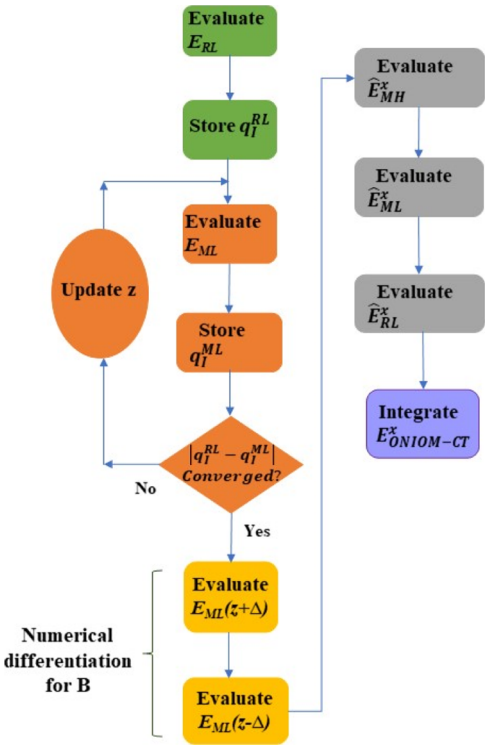


Figure 3. Flowchart for ONIOM-CT gradient evaluation. The different phases can be distinguished by the color of their boxes. Phases 1−5 (refer text) are performed in chronological order.

The gradient evaluation can be split into five phases.

(1) The model region charge was obtained from the RL calculation ($q_{\mathrm{I}}^{\mathrm{RL}}$)
(2) $q_{\mathrm{I}}^{\mathrm{MLCT}}$ was converged to obtain the charge to be added to the link atom nuclear charge (z)
(3) Numerical differentiation was performed to obtain B
(4) The gradients of all the three individual calculations were evaluated, i.e., $\hat{E}^{x}_{\mathrm{RL}}$, $\hat{E}^{x}_{\mathrm{ML}}$, and $\hat{E}^{x}_{\mathrm{MH}}$
(5) The individual gradients were summed appropriately to obtain $E^{x}_{\mathrm{ONIOM-CT}}$

The expression for $\hat{E}^{x}_{\mathrm{MH}}$ does not involve any charge gradient. Thus, it does not require any further simplification. However, evaluation of both $\hat{E}^{x}_{\mathrm{RL}}$ and $\hat{E}^{x}_{\mathrm{ML}}$ requires the analytic differentiation of the atomic charges. Thus, the choice of proper charge model is critical. Despite the widespread use in molecular modeling, atomic charges are not physical observables and thus do not have a unique definition. Charges are a coarse-grained form of molecular electronic density. Thus, the magnitude of the charge on atoms depends on how the coarse graining is performed: Mulliken and Löwdin charges depend strictly on the basis sets, ESP charges are obtained by fitting the electrostatic potential on a grid,[29] NPA charges employ NAOs (natural AOs),[30] Hirshfeld charges[26] partition molecular density into atomic densities using weights which depend on the isolated atomic densities, AIM or Bader charges partition space using the nature of molecular density gradients and assign the density in a particular partition to the atom present in it,[31] and so forth. Among all the above charges, Mulliken and Löwdin charges have analytic gradients due to their direct dependence on the underlying basis set (vide infra). In addition, analytic gradients of ESP charges have been obtained by Holden et al.[32] However, the calculation of analytic gradients of ESP charges is significantly more demanding compared to Mulliken and Löwdin charges. Thus, we have only considered the Mulliken and Löwdin charge models to quantify the charge transfer. The expressions of $\hat{E}^{x}_{\mathrm{RL}}$ and $\hat{E}^{x}_{\mathrm{ML}}$ clearly depend on the charge model as formulated below.

2.1. Mulliken Charges. The Mulliken charge on atom A (q_{A}) and its analytic gradient (q^{x}_{A}) are the following

$$q_{\mathrm{A}} = Z_{\mathrm{A}} - \sum_{\mu\epsilon\mathrm{A}}\sum_{\nu} P_{\mu\nu}S_{\mu\nu} \tag{17}$$

$$q^{x}_{\mathrm{A}} = -\sum_{\mu\epsilon\mathrm{A}}\sum_{\nu} (P^{x}_{\mu\nu}S_{\mu\nu} + P_{\mu\nu}S^{x}_{\mu\nu}) \tag{18}$$

Using eq 18 for the gradient of charge, the modified form of eqs 14 and 15 for RL and ML, respectively, are the following

$$\hat{E}^{x}_{\mathrm{RL}} = E^{x}_{\mathrm{RL}} - B\phi^{\Delta}\sum_{\mu\epsilon\mathrm{I}}\sum_{\nu} (P^{x}_{\mu\nu}S_{\mu\nu} + P_{\mu\nu}S^{x}_{\mu\nu}) \tag{19}$$

$$\hat{E}^{x}_{\mathrm{ML}} = \tilde{E}^{x}_{\mathrm{ML}} + z\phi^{x}_{\mathrm{ML}} - B\phi^{\Delta}\sum_{\mu\epsilon\mathrm{I}}\sum_{\nu} (P^{x}_{\mu\nu}S_{\mu\nu} + P_{\mu\nu}S^{x}_{\mu\nu}) \tag{20}$$

As noted earlier, I in eqs 19 and 20 represent all the atoms in the model region. Recasting the gradients in this form allows for the use of interchange theorem and, consequently, the z-vector method of Handy and Schaefer.[33] The general post-Hartree−Fock gradient formalism[34] can be used by transforming the density matrix gradient term from AO to MO basis, as shown in eqs 21 and 22.

$$\begin{aligned}\hat{E}^{x}_{\mathrm{RL}} = {} & E^{x}_{\mathrm{RL}} - B\phi^{\Delta}\sum_{\nu\epsilon\mathrm{I}}\sum_{\mu} P_{\mu\nu}S^{x}_{\mu\nu} \\ & + B\phi^{\Delta}\sum_{\nu\epsilon\mathrm{I}}\sum_{\mu} S_{\mu\nu}\sum_{\mathrm{i,j}} C_{\mu\mathrm{i}}C_{\nu\mathrm{j}}S^{x}_{\mathrm{ij}} \\ & - B\phi^{\Delta}\sum_{\nu\epsilon\mathrm{I}}\sum_{\mu} S_{\mu\nu}\sum_{\mathrm{i,a}} C_{\mu\mathrm{i}}C_{\nu\mathrm{a}}P^{x}_{\mathrm{ia}} \\ & - B\phi^{\Delta}\sum_{\nu\epsilon\mathrm{I}}\sum_{\mu} S_{\mu\nu}\sum_{\mathrm{i,a}} C_{\mu\mathrm{a}}C_{\nu\mathrm{i}}P^{x}_{\mathrm{ai}}\end{aligned} \tag{21}$$

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

$$\hat{E}^x_{ML} = E^x_{ML} + z\phi^x_{ML} - B\phi^{\Delta}\sum_{\nu\epsilon I}\sum_{\mu} P_{\mu\nu}S^x_{\mu\nu} + B\phi^{\Delta}\sum_{\nu\epsilon I}\sum_{\mu} S_{\mu\nu}\sum_{i,j} C_{\mu i}C_{\nu j}S^x_{ij} - B\phi^{\Delta}\sum_{\nu\epsilon I}\sum_{\mu} S_{\mu\nu}\sum_{i,a} C_{\mu i}C_{\nu a}P^x_{ia} - B\phi^{\Delta}\sum_{\nu\epsilon I}\sum_{\mu} S_{\mu\nu}\sum_{i,a} C_{\mu a}C_{\nu i}P^x_{ai} \quad (22)$$

The above equations are easily converted into a Lagrangian formalism suitable for a z-vector treatment to complete the efficient implementation of the analytic gradient. Further details can be found in the Appendix.

2.2. Löwdin Charges. The Löwdin charge on atom A (q_A) and its analytic gradient (q^x_A) are the following

$$q_A = Z_A - \sum_{\gamma\epsilon A}\sum_{\alpha,\beta} S^{1/2}_{\gamma\alpha}P_{\alpha\beta}S^{1/2}_{\beta\gamma} \quad (23)$$

$$q^x_A = -\sum_{\gamma\epsilon A}\sum_{\alpha,\beta} S^{1/2^x}_{\gamma\alpha}P_{\alpha\beta}S^{1/2}_{\beta\gamma} + S^{1/2}_{\gamma\alpha}P^x_{\alpha\beta}S^{1/2}_{\beta\gamma} + S^{1/2}_{\gamma\alpha}P_{\alpha\beta}S^{1/2^x}_{\beta\gamma} = -\sum_{\gamma\epsilon A}\sum_{\alpha}\left(\sum_{\beta} P_{\alpha\beta}S^{1/2}_{\beta\gamma} + S^{1/2}_{\gamma\beta}P_{\beta\alpha}\right)S^{1/2^x}_{\alpha\gamma} - \sum_{\alpha,\beta} P^x_{\alpha\beta}\sum_{\gamma\epsilon A} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma} \quad (24)$$

Transforming the gradients into a more useful form, we get

$$q^x_A = -\sum_{\alpha,\beta} N_{\alpha\beta}\sum_{\gamma,\delta} u_{\gamma\alpha}u_{\delta\beta}S^x_{\gamma\delta} - \sum_{\alpha,\beta} P^x_{\alpha\beta}\sum_{\gamma\epsilon A} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma} \quad (25)$$

$$N_{\alpha\beta} = \sum_{\sigma\epsilon A}\sum_{\mu,\nu}(P_{\mu\nu}S^{1/2}_{\nu\sigma} + S^{1/2}_{\sigma\nu}P_{\nu\mu})\frac{u_{\mu\alpha}u_{\sigma\beta}}{\sqrt{\lambda_\beta} + \sqrt{\lambda_\alpha}} \quad (26)$$

where λ_α and λ_β are eigenvalues of overlap matrix (S) and $u_{\mu\alpha}$ and $u_{\sigma\beta}$ are elements of corresponding eigenvectors. Detailed derivation of Löwdin charge gradient can be found in ref 35. The derivation of the form of $\hat{E}^x_{RL}$ and $\hat{E}^x_{ML}$ are similar to Mulliken charges. Using eq 25 for the charge gradient, the modified form of eqs 14 and 15 for RL and ML, respectively, are the following

$$\hat{E}^x_{RL} = E^x_{RL} - B\phi^{\Delta}\sum_{\alpha,\beta}\bar{N}_{\alpha\beta}\sum_{\gamma,\delta} u_{\gamma\alpha}u_{\delta\beta}S^x_{\gamma\delta} - B\phi^{\Delta}\sum_{\alpha,\beta} P^x_{\alpha\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma} \quad (27)$$

$$\hat{E}^x_{ML} = \tilde{E}^x_{ML} + z\phi^x_{ML} - B\phi^{\Delta}\sum_{\alpha,\beta}\bar{N}_{\alpha\beta}\sum_{\gamma,\delta} u_{\gamma\alpha}u_{\delta\beta}S^x_{\gamma\delta} - B\phi^{\Delta}\sum_{\alpha,\beta} P^x_{\alpha\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma} \quad (28)$$

Again, I in eqs 27 and 28 represent all the atoms in the model region. Thus, $N_{\alpha\beta}$, from eq 26, has been modified to account for all the atoms in the model region.

$$\bar{N}_{\alpha\beta} = \sum_{\sigma\epsilon I}\sum_{\mu,\nu}(P_{\mu\nu}S^{1/2}_{\nu\sigma} + S^{1/2}_{\sigma\nu}P_{\nu\mu})\frac{u_{\mu\alpha}u_{\sigma\beta}}{\sqrt{\lambda_\beta} + \sqrt{\lambda_\alpha}} \quad (29)$$

Similar to Mulliken charges, the density matrix gradient terms are transformed from AO to MO basis to take advantage of the general post−Hartree−Fock gradient formalism,[34] as shown in eqs 30 and 31.

$$\hat{E}^x_{RL} = E^x_{RL} - B\phi^{\Delta}\sum_{\alpha,\beta}\bar{N}_{\alpha\beta}\sum_{\gamma,\delta} u_{\gamma\alpha}u_{\delta\beta}S^x_{\gamma\delta} + B\phi^{\Delta}\sum_{\alpha,\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma}\sum_{i,j} C_{\alpha i}C_{\beta j}S^x_{ij} - B\phi^{\Delta}\sum_{\alpha,\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma}\sum_{i,a} C_{\alpha i}C_{\beta a}P^x_{ia} - B\phi^{\Delta}\sum_{\alpha,\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma}\sum_{i,a} C_{\alpha a}C_{\beta i}P^x_{ai} \quad (30)$$

$$\hat{E}^x_{ML} = E^x_{ML} + z\phi^x_{ML} - B\phi^{\Delta}\sum_{\alpha,\beta}\bar{N}_{\alpha\beta}\sum_{\gamma,\delta} u_{\gamma\alpha}u_{\delta\beta}S^x_{\gamma\delta} + B\phi^{\Delta}\sum_{\alpha,\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma}\sum_{i,j} C_{\alpha i}C_{\beta j}S^x_{ij} - B\phi^{\Delta}\sum_{\alpha,\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma}\sum_{i,a} C_{\alpha i}C_{\beta a}P^x_{ia} - B\phi^{\Delta}\sum_{\alpha,\beta}\sum_{\gamma\epsilon I} S^{1/2}_{\gamma\alpha}S^{1/2}_{\beta\gamma}\sum_{i,a} C_{\alpha a}C_{\beta i}P^x_{ai} \quad (31)$$

As in the case of Mulliken charges, these equations are converted into a Lagrangian formalism to complete the implementation via the z-vector method. Further details can be found in the Appendix.

2.3. Generalization to Multiple Link Atoms. Treatment of most realistic and chemically interesting problems through ONIOM will often require cutting more than just one bond during the formation of the model system. Thus, it is important not to restrict ourselves to just one link atom. Here, we will generalize the ONIOM-CT method to multiple link atoms and formulate its analytic gradient. The central problem in generalization to multiple link atoms stems from the non-unique way of rearranging the extra charge among link atoms. Additionally, the formulation of the analytic gradient may be challenging as we need one equation for each link atom to obtain the required gradient of the additional charge (z^x). Let us consider a generalization of eq 3 to multiple link atoms.

$$E_{ONIOM-CT} = E_{RL} + \tilde{E}_{MH} - \tilde{E}_{ML} + \sum_l z_l\phi_l^{\Delta} \quad (32)$$

Here, z_l and ϕ_l^{Δ} are the charge and net ESP, respectively, on the lth link atom (analogous to z and ϕ^{Δ}). The gradient takes the form

$$E^x_{ONIOM-CT} = E^x_{RL} + \tilde{E}^x_{MH} - \tilde{E}^x_{ML} + \sum_l z_l\phi_l^{\Delta^x} + \sum_l z^x_l\phi_l^{\Delta} \quad (33)$$

Analogous to the single link atom problem, we compare the model region charge gradients of RL and ML to obtain z^x_l, as shown below.

$$\frac{dq_I^{RL}}{dx} = \frac{\partial q_I^{MLCT}}{\partial x} + \sum_l \frac{\partial q_I^{MLCT}}{\partial z_l}\frac{dz_l}{dx} \quad (34)$$

However, the gradients of charges on individual link atoms cannot be easily separated in eq 34 to obtain the analytic

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

gradient. This issue can be handled in two unique ways. The first technique involves splitting the model region further into individual regions corresponding to each link atom, resulting in solving separate equations for each link atom charge gradient (z_l^x). Our preliminary tests of this method show issues with charge convergence in some cases, particularly for Mulliken charges which are usually of higher magnitude. This may be a result of interdependency of the charges on the atoms in individual regions on the atoms in other regions within the model system. Due to such difficulties, we decided not to follow this method further. The second technique is simpler, where the charges placed on the link atoms (z_l) are equal. This is a fair assumption in our applications because the bonds being cut are often similar in nature (typically C−C bonds). This will also hold in applications to most organic and biological systems as ONIOM model system boundaries usually pass through non-polar bonds. This results in simplification of eq 34 into a form similar to eq 9. This approach is followed in this paper. The resulting ONIOM-CT energy and gradient are

$$E_{\mathrm{ONIOM-CT}} = E_{\mathrm{RL}} + \tilde{E}_{\mathrm{MH}} - \tilde{E}_{\mathrm{ML}} + z\sum_l \phi_l^{\Delta} \tag{35}$$

and

$$E^x_{\mathrm{ONIOM-CT}} = E^x_{\mathrm{RL}} + \tilde{E}^x_{\mathrm{MH}} - \tilde{E}^x_{\mathrm{ML}} + z\sum_l \phi_l^{\Delta x} + z^x\sum_l \phi_l^{\Delta} \tag{36}$$

respectively. Because the link atom charges are equal, z_l is replaced by z and z^x will be obtained using eq 9. The final form of the analytic gradients remains unchanged with ϕ^{Δ} being replaced by $\sum_l \phi_l^{\Delta}$.

We should note in this context that our scheme of placing equal charges on all link atoms may not be appropriate in situations involving capping bonds to two (or more) atoms with significantly different electronegativities. A potential solution is to have a fixed ratio between the charges on link atoms based on the types of bonds being broken. In fact, a model very similar to this has been explored in our earlier work on "Generalized Asymmetric Mulliken Embedding".[36] The gradient implementation then will be very similar to our current implementation.

3. PARAMETERS USED IN THE IMPLEMENTATION AND VERIFICATION

ONIOM-CT calculation requires an iterative convergence of the charge on the link atom. The convergence criteria were chosen to be stringent, $10^{-7}e$. Numerical gradient tests showed that it was necessary to converge to higher accuracy than that for the ONIOM-CT energy evaluation in paper 1 ($10^{-5}e$). Each step in the cycle requires a two-point extrapolation from previous two steps to obtain the charge on the link atom (z) for the next step (refer to the Supporting Information for details). In other words, the q_1^{ML} and z corresponding to the previous two steps are used to obtain the z for the next step. The first iteration is regular ONIOM in the first step of optimization process or single point calculations. On the other hand, other steps start from converged z of previous step. However, a guess z is needed for the second step. We find that a small initial value is adequate, and the converged value of z is independent of the initial guess. Table S1 presents the value of converged Mulliken and Löwdin charges for TFAA and tBuAA (refer to Figure 4 for structures of TFAA and tBuAA) using various initial values of z. These two molecules are representative of X = F (electron withdrawing) and X = CH_3 (electron donating) systems employed to test our method. We have used an increment of $0.015e$ as our initial guess as it is not too small to be numerically insignificant or not too large to induce charge divergence.

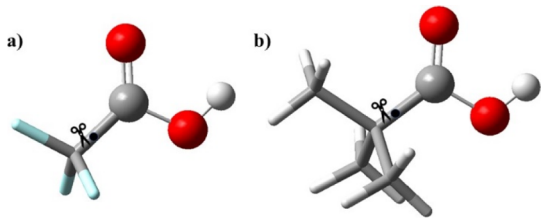


Figure 4. Structures of (a) 2,2,2-trifluoroacetic acid (TFAA) and (b) 2,2-dimethyl propanoic acid (tBuAA) are shown. The portion in ball and stick (to the right of the scissor) is the chemically important model region. The black circle represents the approximate location of the link atom (along the bond being cut at the same position as regular ONIOM).

The protocol developed to obtain the gradient of ONIOM-CT requires, in addition to the iteration for energy, a numerical differentiation of the model system charge with respect to the charge on the link atom nucleus (or nuclei) post-convergence. Choice of the appropriate step size is critical in numerical differentiation. Too small a step might not observe any change or observe only numerical noise. If the step size is too large, higher derivatives might have non-negligible contributions. Thus, numerical differentiation (symmetric finite difference) was carried out at varying step sizes for TFAA and tBuAA and the results are shown in Figure S2. The numerical derivative (B) is almost independent of the step size (Δz) for $\Delta z \leq 10^{-2}e$. Thus, $10^{-4}e$ is chosen as the step size, as it is neither too small (compared to $10^{-7}e$) nor too large. The step size can be easily modified or, if needed, more points could be used for higher precision.

In order to verify the correctness of the analytic gradients, numerical gradients with a step-size of 0.005 Å were evaluated for calibration. Four-point extrapolation was chosen to remove contributions from higher derivatives. Considering the numerical gradients as reference, the analytical gradient errors for two of the test cases are evaluated. MP2 and B3LYP are considered as high level of theory and the results are tabulated in Tables S2 and S3, respectively. We compare the numerical precision for the ONIOM-CT methods with that for simple ONIOM. Before performing such a comparison, we note that the ONIOM-CT analytic gradient requires a small numerical gradient component as well as two extra z-vector components. ONIOM force calculations, on the other hand, do not have any inherent need for the z-vector component. However, if post-SCF methods are used as the high level of theory, such as MP2, even ONIOM needs a post-SCF gradient (via z-vector treatment) for those individual components of ONIOM. Overall, ONIOM-CT gradients may be expected to have slightly less accurate gradients. However, the drop in accuracy is only minimal as seen in Tables S2 and S3. In particular, ONIOM-CT forces are comparatively more similar in accuracy to ONIOM while using MP2 rather than B3LYP.

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

4. RESULTS AND DISCUSSION

4.1. Reaction Energies of One Link Atom Systems.

Paper 1 demonstrated the improved performance of ONIOM-CT using a set of 20 reactions consisting of deprotonation, hydrogen abstraction, ionization potential, and S_N2 reactions. While the systems are fairly small, they include both electron-withdrawing (F) and electron-donating substituents (CH_3) to check the performance of the new methods. In paper 1, the reactant and product energies were obtained at structures optimized with regular ONIOM due to the unavailability of analytic gradients for ONIOM-CT. Because we have now developed the analytic gradient of ONIOM-CT, we have used the same set of reactions using **optimized structures of both reactants and products** to test the performance of our methods (Table 1). However, we note that the variability of optimized structures between ONIOM and ONIOM-CT is modest for the molecules considered in our study. Nevertheless, we note that it is always preferable to compute energetics at equilibrium, especially if thermodynamic values are of interest. Moreover, the implementation of analytic derivatives will have critical importance if one wishes to do dynamics.

Table 1. Test Set of Reactions with One Link Atom[a]

	Deprotonation
1)	$CX_3-\{-CH_2OH_2^+ \rightarrow CX_3-\{-CH_2OH + H^+$
2)	$CX_3-\{-CH_2OH \rightarrow CX_3-\{-CH_2O^- + H^+$
3)	$CX_3-\{-CH_2NH_3^+ \rightarrow CX_3-\{-CH_2NH_2 + H^+$
4)	$CX_3-\{-CH_2NH_2 \rightarrow CX_3-\{-CH_2NH^- + H^+$
5)	$CX_3-\{-COOH \rightarrow CX_3-\{-COO^- + H^+$
	H-Abstraction
6)	$CX_3-\{-CH_2OH \rightarrow CX_3-\{-CH_2O^\cdot + H^\cdot$
7)	$CX_3-\{-CH_2NH_2 \rightarrow CX_3-\{-CH_2NH^\cdot + H^\cdot$
	Ionization Potential
8)	$CX_3-\{-CH_2O^- \rightarrow CX_3-\{-CH_2O^\cdot + e$
9)	$CX_3-\{-CH_2NH^- \rightarrow CX_3-\{-CH_2NH^\cdot + e$
	S_N2 Reactions
10)	$CX_3-\{-CH_2F + Cl^- \rightarrow CX_3-\{-CH_2Cl + F^-$

[a]All reactions have been carried out with X = F and CH_3. The location of the chemically important model system is to the right of the wavy line.

Table 2. List of Combinations of High and Low Theoretical Models Used in This Study

serial	high	low
1A	MP2/6-311+G(d,p)	HF/3-21G
1B	MP2/6-31+G(d)	HF/3-21G
2A	B3LYP/6-311+G(d,p)	HF/3-21G
2B	B3LYP/6-31+G(d)	HF/3-21G
3A	ωB97XD/6-311+G(d,p)	HF/3-21G
3B	ωB97XD/6-31+G(d)	HF/3-21G

Table 2 presents six different combinations of theoretical methods used to test the performance of our method on these systems. A smaller basis set (3-21G) is used at the low-level to avoid the well-known issue of unphysical Mulliken charges from larger basis sets.[37] Additionally, the use of smaller basis sets keeps the computational cost in check as the low-level calculation needs to be performed several times at each step. At the high-level, three different correlated methods, including two DFT functionals and a post-SCF method (MP2), are employed. Moreover, two different larger basis sets are used to represent more realistic applications.

Results of ONIOM-CT using Mulliken and Löwdin charges and its comparison with ONIOM are summarized in Tables 3 and 4, respectively. Because ONIOM considers the high-level method as the reference, errors are with respect to a full calculation at the high-level theory. MAE (mean absolute error), standard deviation, max errors are over all the 20-reaction sets. MAE (X = F) and MAE (X = CH_3) are over the 10-reaction set, each consisting of only X = F and X = CH_3, respectively.

The performance of ONIOM-CT using either of the charge models is better than ONIOM for each combination of theories employed. Overall, the average performance over all the combination of theories is marginally better for the Mulliken charge model than Löwdin. The improvements in ONIOM-CT reaction energy with Mulliken and Löwdin charge models are 48.8 and 42.8%, respectively. The standard deviation of the errors with ONIOM-CT is also smaller than that of ONIOM. Because the MAE and standard deviation are both smaller for ONIOM-CT, it provides more accurate and precise reaction energies than ONIOM. Additionally, the maximum error is also reduced, indicating fewer outliers.

Table 3. MAE, in kcal/mol, of ONIOM and ONIOM-CT Using Mulliken Charges with Respect to High-Level Full System Calculation

	MAE[a]		standard deviation		max error		MAE (X = F)[a]		MAE (X = CH_3)[a]	
theory combination	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]
1A	3.38	1.81	3.93	2.30	6.65	5.86	3.81	1.33	2.94	2.29
1B	3.10	2.26	3.60	2.69	5.64	4.83	3.07	1.43	3.14	3.09
2A	3.57	1.92	4.09	2.40	7.16	6.32	3.45	1.49	3.69	2.34
2B	3.48	1.49	4.00	2.01	6.19	6.22	3.52	1.46	3.44	1.52
3A	3.62	1.75	4.19	2.14	6.99	3.98	3.71	1.38	3.52	2.12
3B	3.61	1.36	4.19	1.77	7.08	3.44	3.72	1.36	3.49	1.37
average	3.46	1.77	4.00	2.22	6.62	5.11	3.55	1.41	3.37	2.12

[a]MAE is mean absolute error over all reactions. MAE(X = F) and MAE(X = CH_3) correspond to only reactions with X = F and X = CH_3, respectively. [b]ONIOM-CT corresponding to Mulliken charges.

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

Table 4. MAE, in kcal/mol, of ONIOM and ONIOM-CT Using Löwdin Charges with Respect to High-Level Full System Calculation

	MAE[a]		standard deviation		max error		MAE (X = F)[a]		MAE (X = CH_3)[a]	
theory combination	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]
1A	3.38	1.74	3.93	2.26	6.65	5.43	3.81	1.75	2.94	1.73
1B	3.10	1.45	3.60	1.96	5.64	4.86	3.07	1.25	3.14	1.64
2A	3.57	2.42	4.09	2.83	7.16	6.98	3.45	2.26	3.69	2.58
2B	3.48	2.00	4.00	2.45	6.19	7.08	3.52	2.06	3.44	1.95
3A	3.62	2.29	4.19	2.60	6.99	4.58	3.71	2.33	3.52	2.25
3B	3.61	2.00	4.19	2.33	7.08	4.28	3.72	2.12	3.49	1.88
average	3.46	1.98	4.00	2.41	6.62	5.54	3.55	1.96	3.37	2.01

[a]MAE is mean absolute error over all reactions. MAE(X = F) and MAE(X = CH_3) correspond to only reactions with X = F and X = CH_3, respectively. [b]ONIOM-CT corresponding to Löwdin charges.

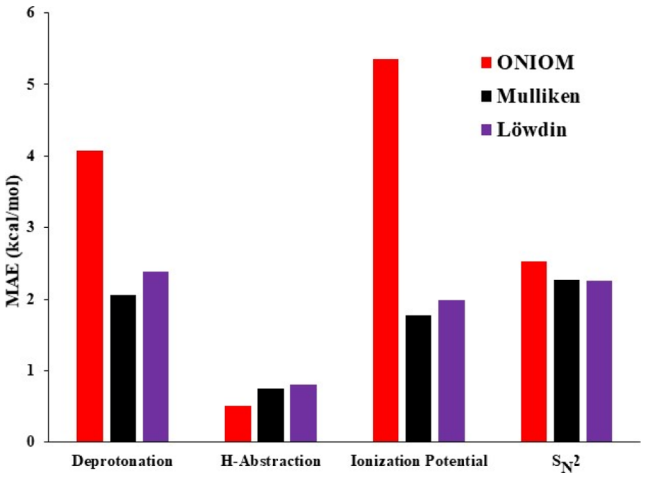


Figure 5. Comparison of reaction energies obtained with ONIOM, ONIOM-CT(Mulliken), and ONIOM-CT(Löwdin) charge models for different types of one link atom reactions shown in Table 1. The results shown represent an average over all six combinations of theories in Table 2 for all the reactions in that category including both X = F and X = CH_3. The errors are with respect to full system calculation at high-level theory.

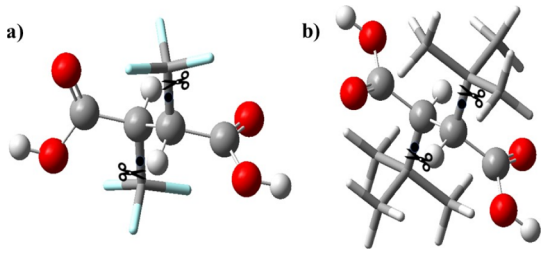
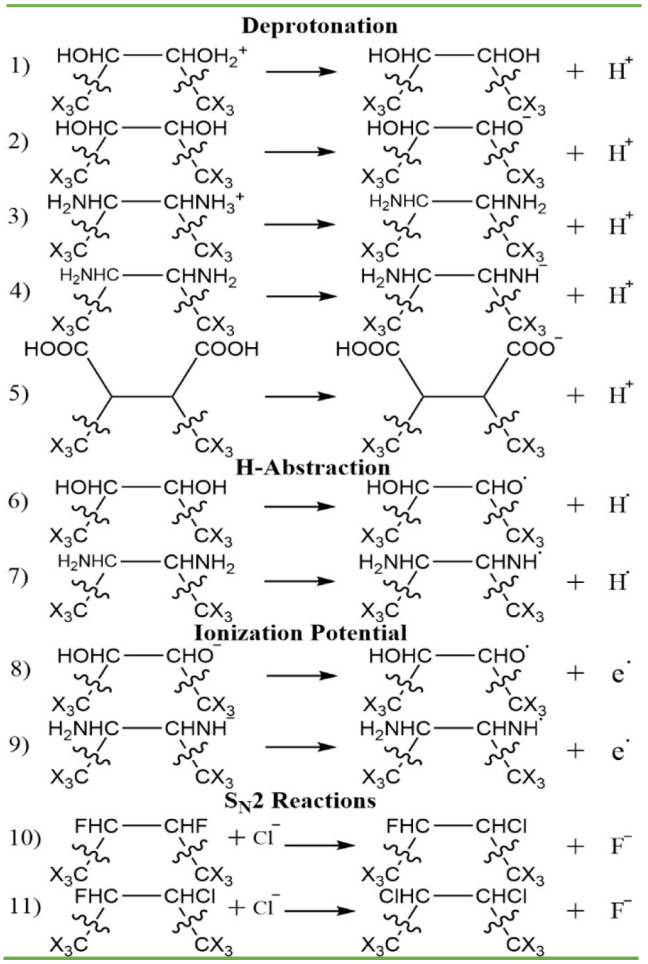


Figure 6. Structures of (a) 2,3-bis(trifluoromethyl) butanedioic acid (2TFAA) and (b) 2,3-bis(1,1-dimethylethyl) butanedioic acid (2tBuAA) are shown. The portion in ball and stick is the chemically important model region. The black circle represents the approximate location of the link atom (along the bond being cut at the same position as regular ONIOM).

Considering X = F and X = CH_3 separately, ONIOM-CT performs better for both cases. The improvement is more pronounced for X = F than X = CH_3. Mulliken charges impart 60.3 and 37.1% improvement for X = F and X = CH_3, respectively. 44.8 and 40.4% improvements are observed for X = F and X = CH_3, respectively, using Löwdin charges. For Mulliken charges, the performance difference between X = F and X = CH_3 is significantly higher than for Löwdin charges.

Table 5. Test Set of Reactions with Two Link Atoms[a]



[a]The partition is marked by wavy lines.

Our test set consists of four different types of reactions. Here, we consider them separately. Figure 5 summarizes these results for ONIOM-CT using both charge models and compares them with ONIOM. All the results are averaged over all the reactions of the same type, including the six different combinations of theories and X = F and X = CH_3. Both charge models provide improvement over ONIOM for all the reactions except H-abstraction for which the ONIOM error is already very low (0.51 kcal/mol), and ONIOM-CT error still stays below 1 kcal/mol using either charge model. Deprotonation and ionization potential reactions experienced

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

Table 6. MAE for Two Link Atom Reactions in Table 5, in kcal/mol, of ONIOM and ONIOM-CT Using Mulliken Charges with Respect to High-Level Full System Calculation

	MAE[a]		standard deviation		max error		MAE (X = F)[a]		MAE (X = CH_3)[a]	
theory combination	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]	ONIOM	Mulliken[b]
1A	6.20	3.40	7.79	4.00	14.11	8.05	5.58	2.65	6.81	4.15
1B	6.04	1.69	7.78	2.05	15.40	4.36	4.91	1.42	7.16	1.96
2A	6.44	5.33	8.19	6.46	16.83	14.05	5.18	3.33	7.70	7.32
2B	6.61	4.20	8.44	4.85	17.52	10.34	5.46	3.25	7.76	5.15
3A	6.40	5.05	8.07	6.01	15.42	12.46	5.30	3.28	7.50	6.81
3B	6.48	4.00	8.30	4.58	16.30	9.08	5.49	3.17	7.48	4.82
average	6.36	3.95	8.10	4.66	15.93	9.72	5.32	2.85	7.40	5.04

[a]MAE is mean absolute error over all reactions. MAE(X = F) and MAE(X = CH_3) correspond to only reactions with X = F and X = CH_3, respectively. [b]ONIOM-CT corresponding to Mulliken charges.

Table 7. MAE for Two Link Atom Reactions in Table 5, in kcal/mol, of ONIOM and ONIOM-CT Using Löwdin Charges with Respect to High-Level Full System Calculation

	MAE[a]		standard deviation		max error		MAE (X = F)[a]		MAE (X = CH_3)[a]	
basis sets	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]	ONIOM	Löwdin[b]
1A	6.20	3.80	7.79	4.48	14.11	8.68	5.58	3.17	6.81	4.42
1B	6.04	2.70	7.78	3.36	15.40	7.62	4.91	1.88	7.16	3.51
2A	6.44	5.83	8.19	6.82	16.83	14.38	5.18	4.19	7.70	7.46
2B	6.61	5.31	8.44	6.21	17.52	13.40	5.46	3.99	7.76	6.62
3A	6.40	5.39	8.07	6.27	15.42	12.54	5.30	4.09	7.50	6.68
3B	6.48	5.09	8.30	5.92	16.30	12.19	5.49	3.91	7.48	6.27
average	6.36	4.69	8.10	5.51	15.93	11.47	5.32	3.54	7.40	5.83

[a]MAE is mean absolute error over all reactions. MAE(X = F) and MAE(X = CH_3) correspond to only reactions with X = F and X = CH_3, respectively. [b]ONIOM-CT corresponding to Löwdin charges.

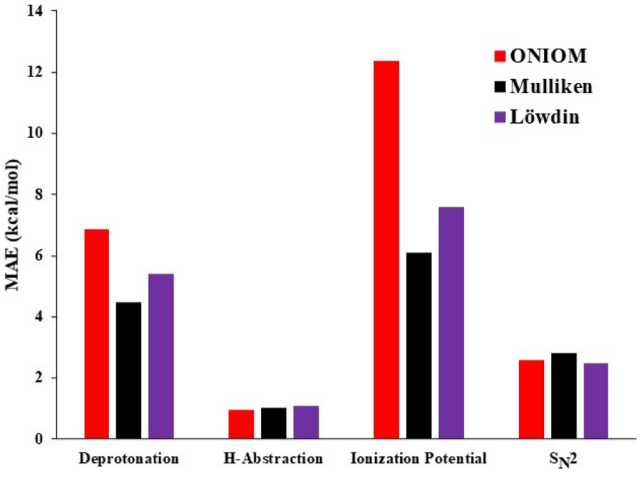


Figure 7. Comparison of reaction energies obtained with ONIOM, ONIOM-CT(Mulliken), and ONIOM-CT(Löwdin) charge models for different types of two link atom reactions shown in Table 5. The results shown represent an average over all six combinations of theories in Table 2 for all the reactions in that category including both X = F and X = CH_3. The errors are with respect to full system calculation at high-level theory.

largest improvements in performance. On the other hand, S_N2 reactions saw little improvement. Overall, both the charge models performed equally well for all the four different types of reactions.

4.2. Reaction Energies of Two Link Atom Systems. As discussed in Section 2, application to real-world problems would involve breaking multiple bonds while forming model systems. To study this, we have considered larger systems (Figure 6) which are two link atom analogues of those considered earlier for the one link atom. Table 5 presents the set of reactions used for benchmarking our extension of ONIOM-CT to multiple link atom systems. Again, the reactions include deprotonation, ionization potential, hydrogen abstraction, and S_N2 reactions. The choice of these systems is natural due to the good performance of ONIOM-CT for the corresponding one link atom systems. The reference molecules (neutral molecules before substitutions) are symmetric, thus requiring us to study the reaction on only one half because the reaction on the other half would be equivalent. The same six combinations of methods are used, as shown in Table 2. With the increase in system size, the use of a smaller basis set at the low-level becomes even more important. As discussed in Section 2, we will be treating both link atoms uniformly.

Results of ONIOM-CT using Mulliken and Löwdin charges and their comparison with ONIOM are summarized in Tables 6 and 7, respectively. Similar to the one link atom case, all the results are averaged over all the six different combinations of theories used. The improvements in ONIOM-CT reaction energy with Mulliken and Löwdin charge models are 37.9 and 26.3%, respectively. The standard deviation of the errors with ONIOM-CT is smaller than that of ONIOM. For two link atom systems, the performance of Mulliken charges is better. Both standard deviation and max error are lower for ONIOM-CT than ONIOM using either of the charge models. Considering X = F and X = CH_3 separately, ONIOM-CT performs better for both cases. Larger improvement is observed for X = F than X = CH_3. Mulliken charge model improved the performance by 46.43 and 31.89% for X = F and X = CH_3, respectively. 33.46 and 21.22% improvements are observed for X = F and X = CH_3, respectively, using Löwdin charges.

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

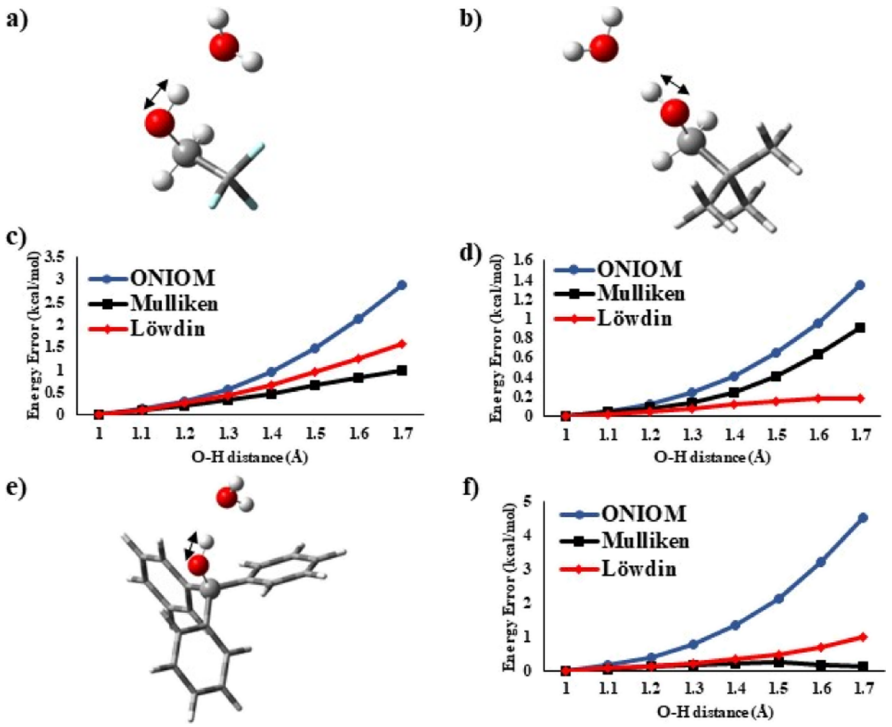


Figure 8. Models used for the PES study of proton transfer. (a,b) are two representative molecules from our one link atom data set in the presence of a water molecule with (c,d) presenting the absolute error in the corresponding PES scans, respectively. (e,f) represent a three link atom system and the corresponding error in PES scan. ONIOM, ONIOM-CT(Mulliken), and ONIOM-CT(Löwdin) PES absolute errors are presented here. The errors are relative to full system calculation at high-level theory. Calculations are performed at the B3LYP/6-311+G(d,p):HF/3-21G level of theory. The portion in ball and stick is the chemically important model region.

There is an improvement in performance over all four types of reactions as shown in Figure 7 with Ionization Potential reactions showing most improvement in performance followed by deprotonation reactions. H-abstraction and S_N2 reactions did not show any improvement in performance due to the ONIOM errors already being low. All the results are averaged over all the reactions of the same type, including the six different combinations of theories and X = F and X = CH_3.

4.3. PES Study. Apart from the study of reaction energies, we have also tested the performance of our method for the investigation of potential energy surfaces (PESs). To illustrate this, proton transfer reactions between a molecule and a neighboring water molecule are considered with the link atom being in close proximity to the bond being broken. Figure 8 presents three test molecules and corresponding absolute error in PES calculations using ONIOM and ONIOM-CT. CF_3CH_2OH and $C(CH_3)_3CH_2OH$ are part of our one link atom data set and trityl hydroxide [$C(Ph)_3OH$] is a representative example of a molecule needing multiple link atoms (three link atoms). In all the three cases, the link atom is only two bonds away from the O−H bond undergoing dissociation, warranting the use of ONIOM-CT. In both ONIOM and ONIOM-CT calculations, the water molecule, the OH group, and the connected C or CH_2 group are considered in the model region. The molecules are first optimized at the high-level theory [B3LYP/6-311+G(d,p)]. The PES is obtained by performing a rigid scan about the O−H bond, i.e., the other atoms except the H atom undergoing transfer, are kept fixed. In all three cases, ONIOM-CT using either of the charge models performed better than ONIOM. As expected, the ONIOM error is largest for trityl hydroxide due to the presence of three link atoms. The MAEs over all the points on the PES (except the first point) are 1.79, 0.16, and 0.41 for ONIOM and ONIOM-CT using Mulliken and Löwdin charges, respectively. This suggests that the presence of a few appropriate charges to correct the charge distribution in the model region results in improved performance in proton transfer reactions.

5. COMPUTATIONAL COST ANALYSIS

The aim of hybrid methods is to accelerate calculations with minimal loss in accuracy. Therefore, it is pivotal to check the efficiency of our method with respect to high-level full system calculation. The comparison is performed on both small (one link atom) and large (two link atom) systems using either charge models. The small systems are TFAA and tBuAA. The large systems are the corresponding two link atom analogues which will be called 2TFAA and 2tBuAA, as shown earlier in Figure 5. While the test systems are not very large molecules, the use of two different sizes of systems provides us some estimate of the scaling of the method. Force calculations are carried out on each of the systems considered at the given level of theory and the results are summarized in Table S4. Timing comparisons are shown for both MP2 and B3LYP as high-level methods while small basis set HF was used as the low-level method.

As expected, the use of hybrid methods results in higher gains over a computationally expensive method such as MP2 than a hybrid DFT functional such as B3LYP. This can be seen from nearly 95% savings provided by MP2:HF compared to 86−90% by B3LYP:HF for 2tBuAA (the largest system considered for timings). The computational savings are much more prominent for larger systems due to the target calculations being significantly more expensive, demonstrating

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

the importance of developing hybrid methods for large systems. Computational cost of ONIOM-CT calculations is the same for both Mulliken and Löwdin charge models. ONIOM-CT calculations are marginally more expensive than ONIOM as compared to full system calculation at high-level theory.

6. CONCLUSIONS

In the present work, we formulated the analytic gradient of ONIOM-CT (ONIOM with Charge Transfer). The analytic gradients are formulated efficiently for two different simple, but widely used, charge models, viz. Mulliken and Löwdin. The extra work needed to obtain Löwdin charge model gradient did not result in increased computational cost over Mulliken charge model. Performance of ONIOM-CT method using either charge models was benchmarked against that of ONIOM using high-level as the reference. Reaction energies of a set of 10 reactions are obtained using optimized structures of both the reactant and product using our new implementation. Improvement in performance was observed for ONIOM-CT using either charge models compared to ONIOM, with a slightly larger improvement for Mulliken charges.

Furthermore, we have extended the ONIOM-CT method to multiple link atoms which expands its domain of application significantly. Analytic gradient is also formulated for multiple link atom systems. The performance of ONIOM-CT using either Mulliken or Löwdin charge models was benchmarked against that of ONIOM. A set of 11 reaction energies obtained similar to the one link atom case. Improvement in performance was analogous to the one link atom case, indicating generalizability of the method to multiple link atoms. The insights obtained from these benchmark studies were tested on a set of three proton transfer reactions consisting of both single and multiple link atom systems. ONIOM-CT performance was better than that of ONIOM at each point in the PES, with the multiple link atom system showing the most improvement.

The improvement over a popular and widely used method such as ONIOM for a variety of reactions using an inexpensive method such as ONIOM-CT is promising. In the future, we intend to combine the ONIOM-CT method with electrostatic embedding in ONIOM to form a more complete, yet computationally efficient, formalism accounting for two of the major deficiencies of ONIOM. We hope it will lead us further toward maintaining high accuracy while keeping the model system size small.

APPENDIX

The generalized gradient for both RL and ML needs to be written in the following form

$$E^x = \sum_{\alpha\beta\gamma\delta} \Gamma^{\mathrm{eff}}_{\alpha\beta\gamma\delta}(\alpha\beta|\gamma\delta)^x + \sum_{\alpha\beta} P^{\mathrm{eff}}_{\alpha\beta} H^x_{\alpha\beta} + \sum_{\alpha\beta} W^{\mathrm{eff}}_{\alpha\beta} S^x_{\alpha\beta} + V^x_{\mathrm{nuc}} \tag{A1}$$

$$\Gamma^{\mathrm{eff}}_{\alpha\beta\gamma\delta} = P^{\Delta}_{\alpha\beta} P^{\mathrm{HF}}_{\gamma\delta} - P^{\Delta}_{\alpha\delta} P^{\mathrm{HF}}_{\gamma\beta} \tag{A2}$$

$$P^{\mathrm{eff}}_{\alpha\beta} = P^{\mathrm{HF}}_{\alpha\beta} + P^{\Delta}_{\alpha\beta} \tag{A3}$$

$$W^{\mathrm{eff}}_{\alpha\beta} = W^{\mathrm{HF}}_{\alpha\beta} + W^{\Delta}_{\alpha\beta} \tag{A4}$$

where $P^{\mathrm{HF}}_{\gamma\delta}$ and $W^{\mathrm{HF}}_{\alpha\beta}$ are HF density and energy weighted density matrix, respectively. $P^{\Delta}_{\alpha\beta}$ and $W^{\Delta}_{\alpha\beta}$ are corrections to the density and energy weighted density matrix, respectively.

It is convenient to define the corrections in the orthonormal MO basis.

$$P^{\Delta}_{\alpha\beta} = \sum_{\mathrm{pq}} C_{\alpha\mathrm{p}} C_{\beta\mathrm{q}} P^{\Delta}_{\mathrm{pq}} \tag{A5}$$

$$P^{\Delta}_{\mathrm{ij}} = 0 \tag{A6}$$

$$P^{\Delta}_{\mathrm{ab}} = 0 \tag{A7}$$

Both the occupied−occupied (oo) and virtual−virtual (vv) blocks of the density matrix correction are zero. The occupied−virtual (ov) block needs solving a single set of SCF response equation

$$\sum_{\mathrm{bj}} [(\mathrm{ij}\|\mathrm{ab}) - (\mathrm{ib}\|\mathrm{ja})] P^{\Delta}_{\mathrm{bj}} + (\varepsilon_{\mathrm{a}} - \varepsilon_{\mathrm{i}}) P^{\Delta}_{\mathrm{ai}} = L_{\mathrm{ai}} \tag{A8}$$

The exact form of L_{ai} depends on the charge model used for ONIOM-CT.

$$\text{Mulliken: } L_{\mathrm{ai}} = -B\phi^{\Delta} \sum_{\nu\in \mathrm{I}} \sum_{\mu} S_{\mu\nu} (C_{\mu\mathrm{i}} C_{\nu\mathrm{a}} + C_{\mu\mathrm{a}} C_{\nu\mathrm{i}}) \tag{A9}$$

$$\text{Löwdin: } L_{\mathrm{ai}} = -B\phi^{\Delta} \sum_{\alpha,\beta} \sum_{\gamma\in \mathrm{I}} S^{1/2}_{\gamma\alpha} S^{1/2}_{\beta\gamma} (C_{\alpha\mathrm{i}} C_{\beta\mathrm{a}} + C_{\alpha\mathrm{a}} C_{\beta\mathrm{i}}) \tag{A10}$$

The energy weighted density matrix corrections will also depend on the charge model under consideration. Thus, we will formulate them separately.

$$\text{Mulliken: } W^{\Delta}_{\alpha\beta} = \begin{cases} -B\phi^{\Delta} P^{\mathrm{HF}}_{\alpha\beta} + \sum_{\mathrm{pq}} C_{\alpha\mathrm{p}} C_{\beta\mathrm{q}} \bar{W}^{\Delta}_{\mathrm{pq}} & \alpha \in \mathrm{I} \text{ and } \beta \in \mathrm{I} \\ -1/2 B\phi^{\Delta} P^{\mathrm{HF}}_{\alpha\beta} + \sum_{\mathrm{pq}} C_{\alpha\mathrm{p}} C_{\beta\mathrm{q}} \bar{W}^{\Delta}_{\mathrm{pq}} & \alpha \in \mathrm{I} \text{ or } \beta \in \mathrm{I} \\ \sum_{\mathrm{pq}} C_{\alpha\mathrm{p}} C_{\beta\mathrm{q}} \bar{W}^{\Delta}_{\mathrm{pq}} & \alpha \notin \mathrm{I} \text{ and } \beta \notin \mathrm{I} \end{cases} \tag{A11}$$

$$\bar{W}^{\Delta}_{\mathrm{ij}} = \begin{cases} B\phi^{\Delta} \sum_{\alpha\beta} S_{\alpha\beta} C_{\alpha\mathrm{i}} C_{\beta\mathrm{j}} - \sum_{\mathrm{ak}} P^{\Delta}_{\mathrm{ak}} (\mathrm{aj}\|\mathrm{ki}) & \alpha \in \mathrm{I} \text{ and } \beta \in \mathrm{I} \\ 1/2 B\phi^{\Delta} \sum_{\alpha\beta} S_{\alpha\beta} C_{\alpha\mathrm{i}} C_{\beta\mathrm{j}} - \sum_{\mathrm{ak}} P^{\Delta}_{\mathrm{ak}} (\mathrm{aj}\|\mathrm{ki}) & \alpha \in \mathrm{I} \text{ or } \beta \in \mathrm{I} \\ - \sum_{\mathrm{ak}} P^{\Delta}_{\mathrm{ak}} (\mathrm{aj}\|\mathrm{ki}) & \alpha \notin \mathrm{I} \text{ and } \beta \notin \mathrm{I} \end{cases} \tag{A12}$$

$$\bar{W}^{\Delta}_{\mathrm{ai}} = -P^{\Delta}_{\mathrm{ai}} \varepsilon_{\mathrm{i}} \tag{A13}$$

$$\bar{W}^{\Delta}_{\mathrm{ab}} = 0 \tag{A14}$$

$$\text{Löwdin: } W^{\Delta}_{\alpha\beta} = -B\phi^{\Delta} \sum_{\gamma,\delta} \bar{N}_{\gamma\delta} u_{\alpha\gamma} u_{\beta\delta} + \sum_{\mathrm{pq}} C_{\alpha\mathrm{p}} C_{\beta\mathrm{q}} \bar{W}^{\Delta}_{\mathrm{pq}} \tag{A15}$$

$$\bar{W}^{\Delta}_{\mathrm{ij}} = B\phi^{\Delta} \sum_{\alpha,\beta} \sum_{\gamma\in \mathrm{I}} S^{1/2}_{\gamma\alpha} S^{1/2}_{\beta\gamma} C_{\alpha\mathrm{i}} C_{\beta\mathrm{j}} - \sum_{\mathrm{ak}} P^{\Delta}_{\mathrm{ak}} (\mathrm{aj}\|\mathrm{ki}) \tag{A16}$$

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064

$$\bar{W}_{ai}^{\Delta} = -P_{ai}^{\Delta}\varepsilon_i \quad \text{(A17)}$$

$$\bar{W}_{ab}^{\Delta} = 0 \quad \text{(A18)}$$

For more information on how these equations are implemented, see refs 38 and 39.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00584.

Two-point extrapolation scheme during ONIOM-CT iterations, accuracy of our analytic gradient, computaional cost analysis and plots supporting our choice of parameters (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Krishnan Raghavachari − *Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States;* orcid.org/0000-0003-3275-1426; Email: kraghava@indiana.edu

Authors

Vikrant Tripathy − *Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States;* orcid.org/0000-0002-3246-0680

Nicholas J. Mayhall − *Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States;* orcid.org/0000-0002-1312-9781

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00584

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge support from the National Science Foundation grants CHE-1665427 and CHE-2102583 at Indiana University.

■ REFERENCES

(1) Dunning, T. H. A Road Map for the Calculation of Molecular Binding Energies. *J. Phys. Chem. A* **2000**, *104*, 9062−9080.

(2) Fang, Z.; Vasiliu, M.; Peterson, K. A.; Dixon, D. A. Prediction of Bond Dissociation Energies/Heats of Formation for Diatomic Transition Metal Compounds: CCSD(T) Works. *J. Chem. Theory Comput.* **2017**, *13*, 1057−1066.

(3) Schlegel, H. B.; Frisch, M. J. Computational Bottlenecks in Molecular Orbital Calculations. In *Theoretical and Computational Models for Organic Chemistry*; Formosinho, S. J., Csizmadia, I. G., Arnaut, L. G., Eds.; Springer Netherlands: Dordrecht, 1991; pp 5−33.

(4) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z. F.; Liu, F. Y.; Li, H. B.; Ding, L. N.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678−5796.

(5) Warshel, A. Computer simulations of enzyme catalysis: methods, progress, and insights. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.

(6) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO +MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and $Pt(P(t\text{-}Bu)_3)_2+H_2$ Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357−19363.

(7) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct.: THEOCHEM* **1999**, *461−462*, 1−21.

(8) Caricato, M.; Vreven, T.; Trucks, G. W.; Frisch, M. J.; Wiberg, K. B. Using the ONIOM hybrid method to apply equation of motion CCSD to larger systems: Benchmarking and comparison with time-dependent density functional theory, configuration interaction singles, and time-dependent Hartree−Fock. *J. Chem. Phys.* **2009**, *131*, 134105.

(9) Vreven, T.; Mennucci, B.; da Silva, C. O. d.; Morokuma, K.; Tomasi, J. The ONIOM-PCM method: Combining the hybrid molecular orbital method and the polarizable continuum model for solvation. Application to the geometry and properties of a merocyanine in solution. *J. Chem. Phys.* **2001**, *115*, 62−72.

(10) Kuno, M.; Hannongbua, S.; Morokuma, K. Theoretical investigation on nevirapine and HIV-1 reverse transcriptase binding site interaction, based on ONIOM method. *Chem. Phys. Lett.* **2003**, *380*, 456−463.

(11) Froese, R. D. J.; Morokuma, K. Accurate calculations of bond-breaking energies in C_{60} using the three-layered ONIOM method. *Chem. Phys. Lett.* **1999**, *30S*, 419−424.

(12) Harrison, M. J.; Burton, N. A.; Hillier, I. H. Catalytic Mechanism of the Enzyme Papain: Predictions with a Hybrid Quantum Mechanical/Molecular Mechanical Potential. *J. Am. Chem. Soc.* **1997**, *119*, 12285−12291.

(13) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(14) Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J. A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.

(15) Théry, V.; Rinaldi, D.; Rivail, J. L.; Maigret, B.; Ferenczy, G. G. Quantum mechanical computations on very large molecular systems: The local self-consistent field method. *J. Comput. Chem.* **1994**, *15*, 269−282.

(16) Wang, B.; Truhlar, D. G. Combined Quantum Mechanical and Molecular Mechanical Methods for Calculating Potential Energy Surfaces: Tuned and Balanced Redistributed-Charge Algorithm. *J. Chem. Theory Comput.* **2010**, *6*, 359−369.

(17) Nasluzov, V. A.; Ivanova, E. A.; Shor, A. M.; Vayssilov, G. N.; Birkenheuer, U.; Rösch, N. Elastic Polarizable Environment Cluster Embedding Approach for Covalent Oxides and Zeolites Based on a Density Functional Method. *J. Phys. Chem. B* **2003**, *107*, 2228−2241.

(18) Mayhall, N. J.; Raghavachari, K. Charge Transfer Across ONIOM QM:QM Boundaries: The Impact of Model System Preparation. *J. Chem. Theory Comput.* **2010**, *6*, 3131−3136.

(19) Humbel, S. Integrated molecular orbital+molecular orbital (IMOMO) treatment of single electron transfer (SET) via a two-center three-electron (2c-3e) bonded complex: substituted carbonyl-ammonia example. *J. Mol. Struct.: THEOCHEM* **1999**, *461−462*, 153−166.

(20) Li, Q. S.; Zhao, Q. H.; Zhang, S. W. An Integrated Molecular Orbital + Molecular Orbital (IMOMO) Study of the O + H−R → OH + R Reaction Class. *J. Phys. Chem. A* **2004**, *108*, 6430−6436.

(21) Truong, T. N.; Truong, T. T. T. A reaction class approach with the integrated molecular orbital+molecular orbital methodology. *Chem. Phys. Lett.* **1999**, *314*, 529−533.

(22) Karadakov, P. B.; Morokuma, K. ONIOM as an efficient tool for calculating NMR chemical shielding constants in large molecules. *Chem. Phys. Lett.* **2000**, *317*, 589−596.

(23) Bobuatong, K.; Limtrakul, J. Effects of the zeolite framework on the adsorption of ethylene and benzene on alkali-exchanged zeolites: an ONIOM study. *Appl. Catal., A* **2003**, *253*, 49−64.

(24) Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I. *J Chem Phys* **1955**, *23*, 1833−1840.

(25) Löwdin, P. O. On the Nonorthogonality Problem. *Advances in Quantum Chemistry*; Elsevier, 1970; Vol. *5*, pp 185−199.

(26) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta* **1977**, *44*, 129−138.
(27) Jose, K. V. J.; Raghavachari, K. Electrostatic Potential-Based Method of Balancing Charge Transfer Across ONIOM QM:QM Boundaries. *J. Chem. Theory Comput.* **2014**, *10*, 4351−4359.
(28) Beckett, D.; Krukau, A.; Raghavachari, K. Charge redistribution in QM:QM ONIOM model systems: a constrained density functional theory approach. *Mol. Phys.* **2017**, *115*, 2813−2822.
(29) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431−439.
(30) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural population Analysis. *J. Chem. Phys.* **1985**, *83*, 735−746.
(31) Bader, R. F. W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18*, 9−15.
(32) Holden, Z. C.; Rana, B.; Herbert, J. M. Analytic gradient for the QM/MM-Ewald method using charges derived from the electrostatic potential: Theory, implementation, and application to ab initio molecular dynamics simulation of the aqueous electron. *J. Chem. Phys.* **2019**, *150*, 144115.
(33) Handy, N. C.; Schaefer, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031−5033.
(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian Development Version*: Wallingford, CT, 2016.
(35) Mayhall, N. J.; Raghavachari, K.; Hratchian, H. P. ONIOM-based QM:QM electronic embedding method using Löwdin atomic charges: Energies and analytic gradients. *J. Chem. Phys.* **2010**, *132*, 114107.
(36) Parandekar, P. V.; Hratchian, H. P.; Raghavachari, K. Applications and assessment of QM:QM electronic embedding using generalized asymmetric Mulliken atomic charges. *J. Chem. Phys.* **2008**, *129*, 145101.
(37) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons Ltd.: Chichestor, England, 2007; pp 293−296.
(38) Hratchian, H. P.; Parandekar, P. V.; Raghavachari, K.; Frisch, M. J.; Vreven, T. QM:QM electronic embedding using Mulliken atomic charges: Energies and analytic gradients in an ONIOM framework. *J. Chem. Phys.* **2008**, *128*, 034107.
(39) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135−149.

https://doi.org/10.1021/acs.jctc.2c00584
J. Chem. Theory Comput. 2022, 18, 6052−6064