

Improved Atoms-in-Molecule Charge Partitioning Functional for Simultaneously Reproducing the Electrostatic Potential and Chemical States in Periodic and Nonperiodic Materials

Thomas A. Manz* and David S. Sholl*

School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, 311 Ferst Drive N.W., Atlanta, Georgia 30332-0100, United States

S Supporting Information

ABSTRACT: We develop a nonempirical atoms-in-molecules (AIM) method for computing net atomic charges that simultaneously reproduce chemical states of atoms in a material and the electrostatic potential $V(r)$ outside its electron distribution. This method gives accurate results for a variety of periodic and nonperiodic materials including molecular systems, solid surfaces, porous solids, and nonporous solids. This method, called DDEC/c3, improves upon our previously published DDEC/c2 method (Manz, T. A.; Sholl, D. S. *J. Chem. Theory Comput.* **2010**, *6*, 2455–2468) by accurately treating nonporous solids with short bond lengths. Starting with the theory all AIM charge partitioning functionals with spherically symmetric atomic weights must satisfy, the form of the DDEC/c3 functional is derived from first principles. The method is designed to converge robustly by avoiding conditions that lead to nearly flat optimization landscapes. In addition to net atomic charges, the method can also compute atomic multipoles and atomic spin moments. Calculations performed on a variety of systems demonstrate the method's accuracy, computational efficiency, and good agreement with available experimental data. Comparisons to a variety of other charge assignment methods (Bader, natural population analysis, electrostatic potential fitting, Hirshfeld, iterative Hirshfeld, and iterative stockholder atoms) show that the DDEC/c3 net atomic charges are well-suited for constructing flexible force-fields for atomistic simulations.

1. INTRODUCTION

An atoms-in-molecule (AIM) theory divides the total electron density, $\rho(\vec{r})$, into non-negative atomic charge distributions. The quantum chemical topology (QCT) pioneered by Bader was one of the first AIM theories to gain widespread use.^{1,2} In QCT, space is divided into nonoverlapping compartments whose surfaces are perpendicular to the electron density gradient.^{1,2} The correspondence between Bader compartments and atoms is usually but not always one-to-one.³ Net atomic charges (NACs) computed using Bader's method are chemically meaningful and widely used. Bader NACs are not well-suited, however, in force fields for porous materials or molecules, because they do not accurately reproduce $V(\vec{r})$ in the material's pores or surrounding the molecule.⁴ This is due to the large deviations of the Bader atomic charge distributions from spherical symmetry, which give rise to large atomic multipole moments.^{4–6} Since $V(\vec{r})$ outside $\rho(\vec{r})$ can be formally written as an atomic multipole expansion, it will be most accurately described by NACs if atomic multipole moments are small.^{4–6} $V(\vec{r})$ outside a spherically symmetric charge distribution is equal to that of an equivalent net point charge. Thus, to minimize atomic multipole moments and accurately reproduce $V(\vec{r})$, an AIM theory should assign approximately spherically symmetric charge distributions to each atom.⁴ For systems containing several closely spaced atoms, the atomic charge distributions must overlap to achieve approximate spherical symmetry.

Several AIM methods partition $\rho(\vec{r})$ into overlapping atomic distributions.^{4,7–9} The Hirshfeld method, in which $\rho(\vec{r})$ is distributed among atoms in proportion to the densities of

(neutral) reference atoms, was one of the earliest methods for obtaining overlapping AIM distributions.⁸ Because the atomic distributions are optimized to resemble neutral atoms, the Hirshfeld method systematically underestimates atomic charge magnitudes leading to poor reproducibility of $V(\vec{r})$.^{4,7,10} To address this issue, Bultinck et al. introduced the iterative Hirshfeld (IH) method, in which $\rho(\vec{r})$ is distributed among atoms in proportion to the densities of isolated reference atoms having the same number of electrons.⁷ One limitation of the IH method is that some ions exist in the solid state that do not exist in isolation. For example, the O^{2-} ion exists in crystalline MgO, even though O^{2-} is not stable in a vacuum. This leads to the question of what is an appropriate reference ion to use for O^{2-} in the IH method. Lillestolen and Wheatley introduced the iterative stockholder atom (ISA) method which does not require reference ion densities.^{9,11} In the ISA method, each atomic weight is generated through an iterative process that spherically averages the density distribution assigned to that atom.^{9,11} This produces nearly spherically symmetric atomic distributions with small atomic dipole moments,⁹ so the ISA NACs accurately reproduce $V(\vec{r})$.⁴ We previously showed that the ISA and IH methods have limitations for modeling periodic solids.⁴ The IH method often overestimates the NAC magnitudes, due to the use of isolated reference anions that are more diffuse than corresponding anions in condensed materials.⁴ The primary limitation of the ISA method is that atomic electron distributions are not constrained to decay like

Received: March 16, 2012

Published: July 3, 2012

real atoms, leading to inaccurate and poorly transferable NACs for materials containing buried atoms.⁴

A general question is how to accurately partition $\rho(\vec{r})$ into overlapping atoms so that the NACs are both chemically meaningful and reproduce $V(\vec{r})$ in a variety of nonperiodic and periodic materials. Previously, we introduced the density derived electrostatic and chemical (DDEC) method for computing net atomic charges (NACs) for this task.⁴ The DDEC method is designed to combine and extend desirable properties of the ISA and IH methods. Like the ISA method, the DDEC method incorporates spherical averaging to minimize atomic multipole magnitudes so that $V(\vec{r})$ is accurately reproduced by the NACs. Like the IH method, the DDEC method uses reference ion densities to enhance the transferability and chemical meaning of computed NACs. To improve upon the IH method, the DDEC method includes the effects of charge compensation and dielectric screening in the reference ion densities. The first reference ion set, called c1, involves placing a periodic array of charged ions in the middle of a $10 \text{ \AA} \times 10 \text{ \AA} \times 10 \text{ \AA}$ unit cell with a uniformly distributed background charge to make the overall unit cell neutral.⁴ The second reference ion set, called c2, is a nonperiodic method that places each ion in the center of a spherical shell of compensating charge.⁴ The exponentially decaying tail of the reference ion density with increasing radius is captured by the c2 set but not by the c1 set, so the c2 set is preferable. In contrast to the isolated O^{2-} ion, the c2 O^{2-} ion is stable due to its compensating charge shell. The DDEC/c2 method gave reasonable results for a variety of periodic and nonperiodic materials including molecular systems, slabs, porous solids, and some nonporous solids.⁴ We found that two components are critical to achieve this good performance: (a) atomic weight factors that use an appropriate combination of spherical averaging and reference ion densities and (b) the use of charge compensated reference ions.⁴

Because they simultaneously reproduce chemical states of atoms in a material and $V(\vec{r})$ outside the material's electron distribution, DDEC NACs are well-suited for the development of force fields used in atomistic simulations. Recently, Watanabe et al.'s study of metal organic frameworks (MOFs) used atomistic simulations to compare the accuracy of different point charge models for reproducing CO_2 adsorption isotherms.¹⁰ DDEC charges gave isotherms in good agreement with isotherms produced from the ab initio electrostatic potential.¹⁰ In other recent studies, DDEC charges were used to study CO_2 adsorption, separation, and diffusion in MOFs and zeolites.^{12–15}

These successes are encouraging, but we have subsequently discovered a limitation of the DDEC/c1 and c2 methods. For some nonporous solids with short bond distances between diffuse atoms, the DDEC/c1 and c2 methods converge poorly if at all. This problem was observed, for example, in calculations for solid Mo_2C . Depending on the calculation details, either a large number (~ 1000) of iterations was required for convergence or else the net charge of the C atom exceeded C^{6-} , causing the calculation to terminate. The DDEC/c1 and c2 methods do not contain constraints to prevent assigned atomic charge distributions from becoming too diffuse, which allows the C atom to become arbitrarily diffuse, ultimately leading to unphysical results. A second example is the intermetallic Pd_3V . As explained in the Results and Discussion below, the DDEC/c2 method gives a net charge of -0.98 for V in Pd_3V , which is also a dense solid with short bond lengths.

Because the electronegativity of V (1.63) is less than that of Pd (2.20),¹⁶ we expect the V atomic charge to be positive. The Bader NAC for V in this material is 1.04, which also indicates a positive V atomic charge. We found the Hirshfeld (-0.25) and ISA (-1.48) NACs for V are negative in Pd_3V , so switching to these methods is not a solution.

In this article, we describe improvements in the DDEC method that solve the problems illustrated in the examples above. We first derive the optimization functional that all AIM methods with spherically symmetric atomic weights globally minimize. By reformulating this as a path action principle, we derive an analytic expression for the optimization landscape's curvature. This allows us to theoretically identify conditions under which problematic convergence may occur. Using these insights, we improve the DDEC method to make it converge accurately and robustly even for challenging materials. A constraint preventing assigned atomic charge distributions from becoming too diffuse is added to improve performance for nonporous solids with short bond distances. This constraint forces each atom to have an exponentially decaying electron density tail, as shown schematically in Figure 1. We

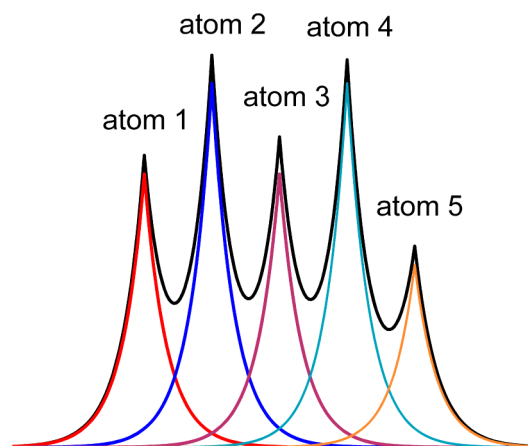


Figure 1. Schematic of DDEC/c3 electron density partitioning. Atoms 2, 3, and 4 are “buried” while atoms 1 and 5 are not. Each atom has an exponentially decaying tail to accurately describe atom–atom overlaps.

theoretically derive the parameters controlling this exponential decay. We also describe several improvements related to the reference densities, including a process that conditions the reference densities to the material of interest to improve accuracy. We derive the optimal proportion of spherical averaging to be used in conjunction with these conditioned reference densities. In addition, hybrid analytic-numeric integration routines are used that improve computational precision. We compare results using this new method, called DDEC/c3, to those computed using the previously published DDEC/c2 method. The new method performs much better where the previous method had difficulty, and changes are small where the previous method performed well. Moreover, the DDEC/c3 charges are in good agreement with available experimental data. Because the new method is more accurate, reliable, and computationally efficient, we recommend the DDEC/c2 method be entirely replaced by the DDEC/c3 method.

The remainder of this paper is organized as following. The theory underlying our new method is described in section 2. In section 2.1, we derive a universal path action for studying AIM

methods with spherically symmetric atomic weights. Section 2.2 describes improvements to the atomic reference densities that improve the curvature of this optimization landscape. Additional equations defining the DDEC/c3 method are contained in sections 2.3 and 2.4, and an iterative solution algorithm is given in section 2.5. Section 3 summarizes the ab initio calculations, the integration grids and routines, tests of integration accuracy, and the method for computing decay exponents of atom tails. Readers who are primarily interested in the application of our methods may find it useful to focus on section 4, which illustrates the use of our methods for a wide variety of molecular and bulk materials. In section 4.1, we compare the DDEC/c2 and DDEC/c3 methods and find that they give similar results for materials without buried atoms. In section 4.2, we find that the DDEC/c3 and experimentally extracted NACs for formamide are in excellent agreement. Also, the DDEC/c3 NACs for several iron-containing materials are found to be strongly correlated to core electron binding energy shifts and the Fe oxidation state. Section 4.3 demonstrates the applicability of the DDEC/c3 method for computing atomic multipoles and atomic spin moments. In section 4.4, we compare the DDEC/c3 method to the natural population analysis (NPA),¹⁷ Hirshfeld (HD),⁸ electrostatic potential (ESP) fitting,¹⁸ iterative Hirshfeld (IH),⁷ and iterative stockholder atom (ISA)⁹ methods. For 4-X-substituted bicyclo-[2,2,2]octane-1-carboxylic acids, the DDEC/c3 and IH methods gave a good combination of conformation transferability of the NACs and low values for the root mean squared error in $V(\vec{r})$. These properties make the DDEC/c3 method well-suited for constructing flexible force fields.

2. THEORY

2.1. Atoms-in-Molecules Formalism. Following Manz and Sholl,⁴ we begin by defining a material as a set of atoms $\{A\}$ located at positions $\{\vec{R}_A\}$, in a reference unit cell, U . For a nonperiodic system (e.g., a molecule), U is any parallelepiped enclosing the entire electron distribution. The reference unit cell has $k_1 = k_2 = k_3 = 0$, and summation over A means summation over all atoms in this unit cell. For a periodic direction, k_i ranges over all integers with the associated lattice vector \vec{v}_i . For a nonperiodic direction, $k_i = 0$ and \vec{v}_i is the corresponding edge of U . Using this notation, the vector and distance relative to atom A are given by

$$\vec{r}_A = \vec{r} - k_1\vec{v}_1 - k_2\vec{v}_2 - k_3\vec{v}_3 - \vec{R}_A \quad (1)$$

and $r_A = |\vec{r}_A|$.

We now describe a universal theory of AIM methods with overlapping atomic electron density distributions, $\{\rho_A(\vec{r}_A)\}$, that partition $\rho(\vec{r})$ proportional to spherically symmetric weight factors, $\{w_A(r_A)\}$:

$$\rho_A(\vec{r}_A) = w_A(r_A)\rho(\vec{r})/W(\vec{r}) \quad (2)$$

where

$$W(\vec{r}) = \sum_{k,A} w_A(r_A) \quad (3)$$

$$\rho(\vec{r}) = \sum_{k,A} \rho_A(\vec{r}_A) \quad (4)$$

$$\rho_A(\vec{r}_A) \geq 0 \quad (5)$$

and $\sum_{k,A} = \sum_{k_1} \sum_{k_2} \sum_{k_3} \sum_A$. Equations 2–5 describe a “stockholder partitioning” of electron density among atoms in a manner that is applicable to nonperiodic and periodic materials.^{4,8,9,11,19,20} The atomic weight factors are expressed as

$$w_A(r_A) = F_A(r_A, \{\rho_i(\vec{r}_i)\}) \quad (6)$$

where different choices for the charge partitioning functional, F_A , produce different AIM methods. In the Hirshfeld (HD) method,

$$F_A^{\text{HD}}(r_A) = \rho_A^{\text{ref}}(r_A, z_A) \quad (7)$$

is the spherically averaged density of the isolated neutral atom of atomic number z_A .⁸ In the iterative stockholder atom (ISA) method,⁹

$$F_A^{\text{ISA}}(r_A) = \rho_A^{\text{avg}}(r_A) \quad (8)$$

where the spherical average of a function $g_A(\vec{r}_A)$ is defined by

$$g_A^{\text{avg}}(r_A) = \frac{\oint g_A(\vec{r}'_A) \delta^{\text{dirac}}(r_A - r'_A) d^3\vec{r}'_A}{4\pi(r_A)^2} \quad (9)$$

Here, δ^{dirac} denotes the Dirac delta function. We use δ to denote the variational (i.e., full) derivative. In the iterative Hirshfeld (IH) method,⁷

$$F_A^{\text{IH}}(r_A) = \rho_A^{\text{ref}}(r_A, n_A) \quad (10)$$

with reference densities given by the linear combination

$$\rho_A^{\text{ref}}(r_A, n_A) = (1 - f)\rho_A^{\tau}(r_A) + f\rho_A^{\tau+1}(r_A) \quad (11)$$

between the spherically averaged ground state densities of isolated atoms of the same element having the closest lower (τ) and higher ($\tau + 1$) integer number of electrons, where $f = n_A - \text{floor}(n_A) = n_A - \tau$.⁷ In the DDEC/c2 method,

$$F_A^{\text{c2}}(r_A) = (\rho_A^{\text{ref}}(r_A, n_A))^{\chi} (\rho_A^{\text{avg}}(r_A))^{(1-\chi)} \quad (12)$$

with $\rho_A^{\tau}(r_A)$ computed via the c2 charge compensation scheme and $\chi \approx 0.1$.⁴

In practice, these AIM methods are solved iteratively. The converged solution corresponds to a global minimum of the optimization functional Ω defined as

$$\Omega = \sum_A \oint P(\zeta_A(\vec{r}_A)) d^3\vec{r}_A \quad (13)$$

$$\zeta_A(\vec{r}_A) = \frac{\rho_A(\vec{r}_A) F(\vec{r})}{F_A(r_A, \{\rho_i(\vec{r}_i)\}) \rho(\vec{r})} \quad (14)$$

$$F(\vec{r}) = \sum_{k,A} F_A(r_A, \{\rho_i(\vec{r}_i)\}) \quad (15)$$

As shown in Figure 2, $P(s)$ is the information distance function

$$P(s) = s \ln(s) - s + 1 \quad (16)$$

whose only stationary point is its global minimum $P(1) = 0$. Examining eq 13, a global minimum occurs when

$$\delta\Omega = \Omega = 0 \quad (17)$$

which occurs if and only if eqs 2 and 6 are both satisfied, i.e. $\zeta_A(\vec{r}_A) = 1$ for every atom.

A charge partitioning functional F_A leads to unique convergence if there exists only one set of converged atomic

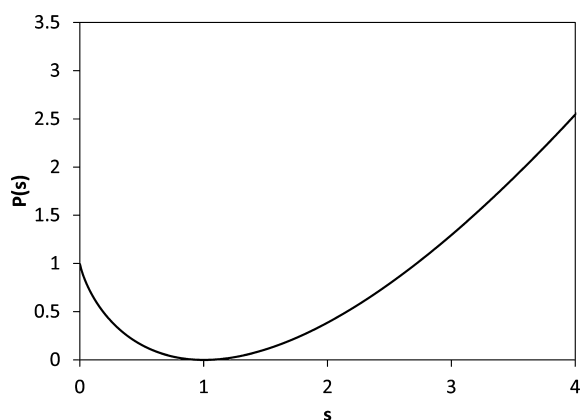


Figure 2. The information distance function $P(s) = s \ln(s) - s + 1$.

densities, $\{\rho_A(\vec{r}_A)\}$, that satisfies eq 17. In this case, the converged solution is the same regardless of the initial $\{w_A(r_A)\}$ estimates. Unique convergence is desirable. Convergence of the HD method is manifestly unique, since this is a noniterative method with $\{w_A(r_A)\}$ equal to the spherically averaged neutral atom densities. A completely variational method is advantageous for studying whether an iterative AIM method converges uniquely. A completely variational method minimizes the optimization functional's variational (i.e., full) derivative, while a partially variational method only minimizes the optimization functional's partial derivative. The ISA method (completely variational) has been shown to converge uniquely.^{11,19} A claimed proof that IH converges uniquely was inadequate because it was based on a partially variational method.^{21,22}

The information distance functional in eq 13 is different from Nalewajski and Parr's information distance functional.²⁰ Specifically, functional 13 measures the information distance between $\zeta_A(\vec{r}_A)$ and its optimal value 1, while Nalewajski and Parr's functional measures the information distance between $\rho_A(\vec{r}_A)$ and $w_A(r_A)$.²⁰ Nalewajski and Parr's functional was originally shown to be completely variational for the HD method.²⁰ Bultinck et al. showed the information distance approach of Nalewajski and Parr is also completely variational for the ISA method.¹⁹ While the HD and ISA methods were originally formulated in a completely variational manner,^{7,11,19,20} previous formulations of the IH and DDEC methods were only partially variational.^{4,7} Specifically, the information distance approach of Nalewajski and Parr²⁰ does not yield completely variational formulations of the IH and DDEC methods. Moreover, a completely variational alteration of the IH method which changed $w_A(r_A)$ was not entirely successful.²² Our method places the HD, ISA, IH, DDEC, and related AIM methods on a common and completely variational framework. Specifically, eq 13 describes the first completely variational method applicable to all AIM methods with spherical symmetric $\{w_A(r_A)\}$. Because the atomic weight factors of the HD, ISA, IH, and DDEC/c2 methods are unchanged, by definition the new optimization functional Ω gives HD, ISA, IH, and DDEC/c2 NACs identical to those arising from the previously reported optimization functionals for these methods. By using the common optimization functional Ω to describe all of these AIM methods, we can study the performance of any potential charge partitioning functional, F_A , by asking the questions: (a) What properties of F_A are necessary to give robust convergence? (b) What properties of F_A are necessary to give chemically meaningful

NACs? (c) What properties of F_A are necessary to give NACs that accurately reproduce $V(\vec{r})$ surrounding a material?

Expanding the derivative,

$$\frac{\delta\Omega}{\delta\rho_A(\vec{r}_A)} = \sum_B \oint \frac{\delta\zeta_B(\vec{r}_B)}{\delta\rho_A(\vec{r}_A)} \ln(\zeta_B(\vec{r}_B)) d^3\vec{r}_B \quad (18)$$

Examining eq 11, $\rho_A^{\text{ref}}(r_A, n_A)$ is continuous with respect to n_A , but $d\rho_A^{\text{ref}}(r_A, n_A)/dn_A$ is not necessarily continuous with respect to n_A . Because

$$n_A = \oint \rho_A(\vec{r}_A) d^3\vec{r}_A \quad (19)$$

it follows $\delta\rho_A^{\text{ref}}(r_A, n_A)/\delta\rho_A(\vec{r}_A)$ is not necessarily continuous with respect to $\rho_A(\vec{r}_A)$. Consequently, $\delta\zeta_B(\vec{r}_B)/\delta\rho_A(\vec{r}_A)$ is not necessarily continuous with respect to $\rho_A(\vec{r}_A)$ when $\rho_B(\vec{r}_B)$ depends continuously on $\{\rho_A^{\text{ref}}(r_A, n_A)\}$. In this case, $\delta\Omega/\delta\rho_A(\vec{r}_A)$ is not necessarily continuous with respect to $\rho_A(\vec{r}_A)$. Nevertheless, $\delta\Omega/\delta\rho_A(\vec{r}_A)$ is continuous with respect to $\{\rho_A(\vec{r}_A)\}$ in the limit $\ln(\zeta_B(\vec{r}_B)) \rightarrow 0$ for every atom and all $\{\vec{r}_B\}$, because in this limit $\delta\Omega/\delta\rho_A(\vec{r}_A) \rightarrow 0$. Because $\delta\Omega/\delta\rho_A(\vec{r}_A)$ is continuous at all positions satisfying $\ln(\zeta_B(\vec{r}_B)) \rightarrow 0$, the slope discontinuities on the Ω optimization landscape are located away from and have no effect upon the solutions described by eq 17.

Accordingly, we can remove the slope discontinuities without affecting the optimization problem's solution. This is accomplished by defining the differential path action

$$dS = \sum_A \oint \delta\rho_A(\vec{r}_A) \ln(\zeta_A(\vec{r}_A)) d^3\vec{r}_A \quad (20)$$

The action, $S = \int dS$, takes a path in $\{\rho_A(\vec{r}_A)\}$ optimization space as its argument and has a real number as its result. Stationary points of the action occur where

$$\frac{\delta S}{\delta\rho_A(\vec{r}_A)} = \ln(\zeta_A(\vec{r}_A)) = 0 \quad (21)$$

for every atom. Since eqs 2 and 6 combine to give eq 21, stationary points of the action are equivalent to global minima of Ω and correspond to converged solutions of the AIM method. Since $\delta S/\delta\rho_A(\vec{r}_A)$ is continuous with respect to $\rho_A(\vec{r}_A)$ when $\zeta_A(\vec{r}_A)$ is continuous with respect to $\rho_A(\vec{r}_A)$, the slope discontinuity issue is resolved by the path action. Because the optimization landscape's slope is continuous, there is no pinning of NAC values by slope discontinuities.

Expanding the action's curvature gives

$$d^2S = \sum_A \oint \delta\rho_A(\vec{r}_A) \left(\frac{\delta\rho_A(\vec{r}_A)}{\rho_A(\vec{r}_A)} - \frac{\delta F_A(r_A, \{\rho_j(\vec{r}_j)\})}{F_A(r_A, \{\rho_j(\vec{r}_j)\})} + \frac{\delta F(\vec{r})}{F(\vec{r})} \right) d^3\vec{r}_A \quad (22)$$

If the curvature is positive everywhere, i.e., $d^2S > 0$, then the action has only one stationary point, and this stationary point is its global minimum. Since stationary points of the action correspond to converged solutions of the AIM method, the AIM method converges uniquely if $d^2S > 0$ everywhere. To the best of our knowledge, this is the first time a condition has been derived for unique convergence of a general charge partitioning functional $F_A(r_A, \{\rho_j(\vec{r}_j)\})$. Even when convergence is unique, problems can arise when substantial finite changes in $\{\rho_A(\vec{r}_A)\}$

produce almost negligible changes in the value of Ω near its minimum. In this case, the optimization landscape is nearly flat with $d^2S \approx 0$ even if $d^2S \geq 0$. As an example, the ISA method converges uniquely^{11,19} but gives NACs that are not chemically meaningful for many dense solids.⁴ The large number of iterations (~ 1000) required for convergence in these systems indicates the ISA optimization landscape is nearly flat. The key remaining question is how to choose the charge partitioning functional, F_A , to achieve the best performance for a wide variety of systems.

2.2. Improvements Related to the Reference Densities. Our approach for developing a better F_A is to start with the DDEC/c2 method and look for ways to increase the action's curvature to avoid nearly flat optimization landscapes. Inserting eq 12 into 22 gives

$$d^2S = (1 - \chi)\Lambda^I + \chi\Lambda^{II} + \Lambda^{III} \quad (23)$$

$$\Lambda^I = \sum_A \oint \delta\rho_A(\vec{r}_A) \left(\frac{\delta\rho_A(\vec{r}_A)}{\rho_A(\vec{r}_A)} - \frac{\delta\rho_A^{\text{avg}}(r_A)}{\rho_A^{\text{avg}}(r_A)} \right) d^3\vec{r}_A \quad (24)$$

$$\Lambda^{II} = \sum_A \oint \delta\rho_A(\vec{r}_A) \left(\frac{\delta\rho_A(\vec{r}_A)}{\rho_A(\vec{r}_A)} - \frac{\delta\rho_A^{\text{ref}}(r_A, n_A)}{\rho_A^{\text{ref}}(r_A, n_A)} \right) d^3\vec{r}_A \quad (25)$$

$$\Lambda^{III} = \sum_A \oint \delta\rho_A(\vec{r}_A) \left(\frac{\delta F^{c2}(\vec{r})}{F^{c2}(\vec{r})} \right) d^3\vec{r}_A \quad (26)$$

The first term simplifies to

$$\Lambda^I = \sum_A \oint \left(\frac{(\delta\rho_A(\vec{r}_A))^2}{\rho_A(\vec{r}_A)} - \frac{(\delta\rho_A^{\text{avg}}(r_A))^2}{\rho_A^{\text{avg}}(r_A)} \right) d^3\vec{r}_A \geq 0 \quad (27)$$

and Bultinck et al. proved this type of integral is non-negative.¹⁹

Examining eq 25, Λ^{II} will be non-negative if

$$\delta \ln(\rho_A^{\text{ref}}(r_A, n_A)) = \delta\rho_A^{\text{ref}}(r_A, n_A) / \rho_A^{\text{ref}}(r_A, n_A) \quad (28)$$

is not too large in magnitude. First, we note the electron density for every c2 reference ion decreases monotonically with increasing radius:

$$\frac{\partial \rho_A^{\tau}(r_A)}{\partial r_A} \leq 0 \quad (29)$$

Second, the reference densities integrate to the proper number of electrons

$$\int_0^\infty \rho_A^{\tau}(r_A) 4\pi(r_A)^2 dr_A = \tau \quad (30)$$

We expect the reference densities to increase monotonically at every value of r_A when an electron is added to the system

$$\rho_A^{\tau+1}(r_A) \geq \rho_A^{\tau}(r_A) \quad (31)$$

However, it is useful to adopt a stronger constraint than eq 31. Expanding

$$\delta \ln(\rho_A^{\text{ref}}(r_A, n_A))|_{n_A=\tau^+} = d n_A \left(\frac{\rho_A^{\tau+1}(r_A) - \rho_A^{\tau}(r_A)}{\rho_A^{\tau}(r_A)} \right) \quad (32)$$

we see that making $\rho_A^{\tau+1}(r_A) - \rho_A^{\tau}(r_A)$ monotonically decreasing with increasing r_A will help limit the magnitude of $\delta \ln(\rho_A^{\text{ref}}(r_A, n_A))$. This corresponds to the constraint

$$\frac{\partial^2 \rho_A^{\text{ref}}(r_A, n_A)}{\partial r_A \partial n_A} \leq 0 \quad (33)$$

The Supporting Information describes the algorithm we used to impose this constraint. Tests we performed showed constraint 33 significantly affected the values of the atomic charges only for systems in which the DDEC/c2 optimization landscape was nearly flat, i.e., only for dense solids with short bond distances.

Finally, Λ^{III} will be negligible if $F^{c2}(\vec{r})$ is consistently close to $\rho(\vec{r})$ since $\delta\rho(\vec{r}) = 0$ during optimization. An exact match between $F(\vec{r})$ and $\rho(\vec{r})$ is not generally possible, since $F(\vec{r})$ is a sum of spherically symmetric functions $F_A(r_A)$. We can, however, construct $F_A(r_A)$ from spherical averages of functions that do exactly sum to $\rho(\vec{r})$. Clearly, $\{\rho_A^{\text{avg}}(r_A)\}$ are spherical averages of the functions $\{\rho_A(\vec{r}_A)\}$ that exactly sum to $\rho(\vec{r})$. Suppose we create some new functions $\{Y_A^{\text{avg}}(r_A)\}$ that play a role similar to $\rho_A^{\text{ref}}(r_A, n_A)$ in the atomic weights

$$F_A(r_A, \{\rho_j(\vec{r}_j)\}) \approx (Y_A^{\text{avg}}(r_A))^\gamma (\rho_A^{\text{avg}}(r_A))^{(1-\gamma)} \quad (34)$$

except they are constructed as spherical averages of functions that exactly sum to $\rho(\vec{r})$,

$$\sum_{k,A} Y_A(\vec{r}_A) = \rho(\vec{r}) \quad (35)$$

The logical choice for $Y_A(\vec{r}_A)$ is the partitioning of $\rho(\vec{r})$ in which the weight assigned to each atom is $\rho_A^{\text{ref}}(r_A, n_A)$:

$$Y_A(\vec{r}_A) = \rho(\vec{r}) \rho_A^{\text{ref}}(r_A, n_A) / \rho^{\text{ref}}(\vec{r}) \quad (36)$$

$$\rho^{\text{ref}}(\vec{r}) = \sum_{k,A} \rho_A^{\text{ref}}(r_A, n_A) \quad (37)$$

$$Y_A^{\text{avg}}(r_A) = \rho_A^{\text{ref}}(r_A, n_A) \langle \rho(\vec{r}) / \rho^{\text{ref}}(\vec{r}) \rangle_{r_A} \quad (38)$$

$$Y(\vec{r}) = \sum_{k,A} Y_A^{\text{avg}}(r_A) \quad (39)$$

$\{Y_A^{\text{avg}}(r_A)\}$ are called the conditioned reference densities. In summary, we implemented several improvements related to the reference densities to improve the DDEC method's accuracy and convergence robustness.

2.3. Theoretical Derivation of the Optimal γ Value. We now derive an optimal value of γ that specifies the relative weight of conditioned reference densities to spherical averaging in the atomic weight factor of eq 34. Inserting eq 34 into 2 and taking the spherical average for radius r_A gives

$$\rho_A^{\text{avg}}(r_A) \approx Y_A^{\text{avg}}(r_A) \left(\left\langle \frac{\rho(\vec{r})}{W(\vec{r})} \right\rangle_{r_A} \right)^{1/\gamma} \quad (40)$$

Equation 40 can be rewritten as

$$\frac{\rho_A^{\text{avg}}(r_A)}{Y_A^{\text{avg}}(r_A)} \approx (\langle e^{\varepsilon_A(\vec{r}_A)} \rangle_{r_A})^{1/\gamma} \quad (41)$$

$$\varepsilon_A(\vec{r}_A) = \ln \left(\frac{\rho(\vec{r})}{W(\vec{r})} \right) \quad (42)$$

Expanding as a Taylor series about $\varepsilon_A(\vec{r}_A) = 0$,

$$\langle e^{\varepsilon_A(\vec{r}_A)} \rangle_{r_A} = \left\langle \sum_{j=0}^{\infty} \frac{(\varepsilon_A(\vec{r}_A))^j}{j!} \right\rangle_{r_A} \quad (43)$$

Because conditioned reference densities are spherical averages of functions that sum to $\rho(\vec{r})$ (eq 35), the average deviation is approximately zero

$$\varepsilon_A^{\text{avg}}(r_A) \approx 0 \quad (44)$$

which causes the even powers in eq 43 to dominate. This allows eq 40 to be rewritten as

$$\frac{\rho_A^{\text{avg}}(r_A)}{Y_A^{\text{avg}}(r_A)} \approx \exp \left(\frac{1}{\gamma} \ln \left(\sum_{j=0}^{\infty} \frac{\langle \varepsilon_A(\vec{r}_A)^{2j} \rangle_{r_A}}{(2j)!} \right) \right) \quad (45)$$

and expanded as a Taylor series about $\varepsilon_A(\vec{r}_A) = 0$ to give

$$\begin{aligned} \frac{\rho_A^{\text{avg}}(r_A)}{Y_A^{\text{avg}}(r_A)} \approx & 1 + \frac{\langle \varepsilon_A(\vec{r}_A)^2 \rangle_{r_A}}{2\gamma} + \left(\frac{\langle \varepsilon_A(\vec{r}_A)^4 \rangle_{r_A}}{24\gamma} \right. \\ & \left. + (\langle \varepsilon_A(\vec{r}_A)^2 \rangle_{r_A})^2 \left(\frac{1}{8\gamma^2} - \frac{1}{8\gamma} \right) \right) + \dots \end{aligned} \quad (46)$$

In the limit $\gamma \rightarrow 0$, higher order terms in the series dominate because they have infinitely large coefficients, so $\ln(\rho_A^{\text{avg}}(r_A)/Y_A^{\text{avg}}(r_A))$ can become arbitrarily large in magnitude. In this case, the converged solution has no dependence on the reference densities and produces approximately spherically symmetric $\rho_A(\vec{r}_A)$ whose NACs fit $V(\vec{r})$ accurately. In the limit $\gamma \rightarrow 1$, coefficients for the higher order terms are much smaller than those of lower order terms. In this case, the optimization retains as much information as possible about the reference densities, at the expense of potentially giving highly nonspherical $\rho_A(\vec{r}_A)$ whose NACs may not accurately reproduce $V(\vec{r})$. For example, the coefficient of the second order term is 10 and 0.1 times that of the fourth order term when γ is 0.9375 and 0.02459, respectively. A balance between these terms occurs when the second and fourth order terms have equal coefficients:

$$\frac{1}{2\gamma} = \frac{1}{24\gamma} + \left(\frac{1}{8\gamma^2} - \frac{1}{8\gamma} \right) \quad (47)$$

Thus, the ideal value is $\gamma = 3/14$. Tests we performed with varying γ for a range of materials confirmed this is an optimal value. Larger γ values gave larger than necessary atomic multipoles, causing a loss in accuracy for fitting $V(\vec{r})$, while smaller γ values caused a loss in chemical accuracy for some systems.

In the DDEC/c2 method using unconditioned reference densities, trial and error showed $\chi \approx 0.1$ is the approximate optimum.⁴ As shown in eq 12, χ controls the relative weighting of $\rho_A^{\text{ref}}(r_A, n_A)$ to $\rho_A^{\text{avg}}(r_A)$ in $F_A^{\text{c2}}(r_A)$. $\varepsilon_A^{\text{avg}}(r_A)$ is not necessarily approximately zero without conditioning, so in this case a Taylor series expansion for $\rho_A^{\text{avg}}(r_A)/\rho_A^{\text{ref}}(r_A, n_A)$ about $\varepsilon_A(\vec{r}_A) = \varepsilon_A^{\text{avg}}(r_A)$ is not necessarily approximately the same as one constructed about $\varepsilon_A(\vec{r}_A) = 0$. Consequently, we cannot be certain χ has the same optimal value for all systems. Because reference density conditioning involves one spherical averaging operation, an optimum value of χ should be less than or equal to γ , i.e., $0 < \chi \leq 3/14$, to produce a similar amount of spherical averaging. (The value $\chi \approx 0.1$ is in the middle of this range.) The key advantage of using conditioned reference densities is

that the above derivation shows the balanced value $\gamma = 3/14$ is system independent.

2.4. New Charge Partitioning Functional to Avoid Nearly Flat Optimization Landscapes. For a nearly flat optimization landscape, eq 2 is approximately satisfied for a range of $w_A(r_A)$ values, which implies

$$\begin{aligned} \delta\rho_A(\vec{r}_A) &\approx \delta \left(w_A(r_A) \frac{\rho(\vec{r})}{W(\vec{r})} \right) \\ &= \frac{\rho(\vec{r})}{W(\vec{r})} \left(\delta w_A(r_A) - \frac{w_A(r_A)}{W(\vec{r})} \delta W(\vec{r}) \right) \end{aligned} \quad (48)$$

Let us denote as $\{\phi\}$ those particular r_A values for which $w_A(r_A)$ has problematic convergence due to a nearly flat optimization landscape. Then, eq 48 must be satisfied for every \vec{r}_A having radius $r_A \in \{\phi\}$. Since $W(\vec{r}) \approx \rho(\vec{r})$ near the optimum and $\rho(\vec{r})$ does not change during optimization, $\delta W(\vec{r})$ will not have a large magnitude. Thus, if eq 48 is satisfied

$$\lim_{w_A(r_A) < W(\vec{r})} \delta\rho_A(\vec{r}_A) \approx \delta w_A(r_A) \quad (49)$$

For the ISA, IH, and DDEC methods, $\{w_A(r_A)\}$ are approximately normalized to contain n_A electrons,

$$n_A = \oint \rho_A(\vec{r}_A) d^3\vec{r}_A \approx \oint w_A(r_A) d^3\vec{r}_A \quad (50)$$

This implies that a nearly flat optimization landscape also requires

$$\delta n_A = \oint \delta\rho_A(\vec{r}_A) d^3\vec{r}_A \approx \oint \delta w_A(r_A) d^3\vec{r}_A \quad (51)$$

For $w_A(r_A) \ll W(\vec{r})$, eq 49 automatically satisfies eq 51. Points with $w_A(r_A) \approx W(\vec{r})$ must simultaneously satisfy the independent eqs 48 and 51. For the “exposed” part of a surface atom A, $w_A(r_A) \approx W(\vec{r})$, so surface atoms do not have any value p for which $w_A(r_A) \ll W(\vec{r})$ for every \vec{r}_A of length p . On the other hand, a “buried” atom A whose van der Waals (vdW) radius does not touch the vdW surface of the material at any point has some values p for which $w_A(r_A) \ll W(\vec{r})$ for every \vec{r}_A of length p . Buried atoms automatically satisfy eqs 49 and 51 for $r_A = p$.

Since a nearly flat optimization landscape involves the swapping of electron density between two or more atoms, eqs 48 and 51 must hold for at least two atoms at every position \vec{r} where the density partitioning optimization is nearly flat. The easiest way to accomplish this is for a buried atom to be surrounded by other buried atoms on all sides with bond lengths short enough that the buried tails of next nearest neighbors overlap each other significantly. Near the nucleus of the atom between next nearest neighbors A and B, both $w_A(r_A) \ll W(\vec{r})$ and $w_B(r_B) \ll W(\vec{r})$, so eqs 49 and 51 are satisfied for both atoms. Since eq 48 must be satisfied for every \vec{r}_A having radius $r_A \in \{\phi\}$, a nearly flat landscape may occur when a buried atom significantly overlaps the buried tails of next nearest neighbor buried atoms in all directions. Thus, dense materials with short bond lengths may produce a nearly flat optimization landscape with the DDEC/c2 method. As mentioned in the Introduction, we noticed that the DDEC/c2 method converges poorly for systems of this type, including Mo₂C and Pd₃V. The DDEC/c2 method does not have any difficulties for solids like NaCl that have small overlaps between next nearest neighbors.

This problem can be fixed by adding a constraint on $F_A(r_A)$ that takes effect when $w_A(r_A) \ll \langle W(\vec{r}) \rangle_{r_A}$. For sufficiently large r_A , the electron density of an isolated atom or ion decays approximately exponentially with increasing r_A . Thus, buried tails will be most realistic if they decay approximately exponentially with increasing r_A :

$$\lim_{w_A(r_A) \ll \langle W(\vec{r}) \rangle} w_A(r_A) \approx w_A(r_A^0) \exp(-(r_A - r_A^0)t_A) \quad (52)$$

where $r_A \geq r_A^0$ is the buried tail region. The value of t_A is constrained to prevent a buried tail from becoming too diffuse. By directly relating $w_A(r_A)$ in the buried tail region to $w_A(r_A^0)$ in this manner, the indeterminacy in $w_A(r_A)$ that caused the flatness of the optimization landscape is removed. Equation 52 implies a constraint of the form

$$\frac{\partial F_A^{c3}(r_A)}{\partial r_A} + \eta_A(r_A) F_A^{c3}(r_A) \leq 0 \quad (53)$$

In the limit $\rho_A(\vec{r}_A) \rightarrow \rho(\vec{r})$, $F_A^{c3}(r_A)$ should only be constrained to monotonically decrease, so

$$\lim_{\rho_A(\vec{r}_A) \rightarrow \rho(\vec{r})} \eta_A(r_A) = 0 \quad (54)$$

As described in the Supporting Information, good transferability of the atomic charge distributions can be achieved using the tail constraint function

$$\eta_A(r_A) = b(1 - (\tau_A(r_A))^2) \quad (55)$$

$$\tau_A(r_A) = Y_A^{\text{avg}}(r_A) \langle Y(\vec{r})^{-1/2} \rangle_{r_A} / \langle Y(\vec{r})^{1/2} \rangle_{r_A} \quad (56)$$

The theoretically optimal value $b \approx 1.75/\text{bohr}$ is derived in the Supporting Information.

Taking into consideration eq 34, $F_A^{c3}(r_A)$ should be the spherically symmetric distribution that resembles

$$\sigma_A(r_A) = (Y_A^{\text{avg}}(r_A))^\gamma (\rho_A^{\text{avg}}(r_A))^{(1-\gamma)} \quad (57)$$

as closely as possible subject to constraint 53. To ensure chemical reasonableness, $F_A^{c3}(r_A)$ is also constrained by

$$n_A^{\text{val}} = n_A - n_A^{\text{core}} \geq 0 \quad (58)$$

These constraints can be applied more easily by rewriting $F_A^{c3}(r_A)$ as

$$F_A^{c3}(r_A) = \lambda_A G_A(r_A) \quad (59)$$

where

$$\lambda_A = \max \left(\left(n_A^{\text{core}} \left[\oint \frac{G_A(r_A)}{F_A^{c3}(\vec{r})} \rho(\vec{r}) d^3\vec{r}_A \right]^{-1} \right), 1 \right) \quad (60)$$

is the smallest $\lambda_A \geq 1$ satisfying constraint 58 and $G_A(r_A)$ is normalized to have the same integral as $\sigma_A(r_A)$:

$$\varphi_A = \int_0^\infty (G_A(r_A) - \sigma_A(r_A)) 4\pi(r_A)^2 dr_A = 0 \quad (61)$$

Constraint 53 now becomes

$$\phi_A(r_A) = \frac{dG_A(r_A)}{dr_A} + \eta_A(r_A) G_A(r_A) \leq 0 \quad (62)$$

To complete the definition of $F_A^{c3}(r_A)$, a functional $H(G_A(r_A))$ needs to be chosen such that minimizing H yields

$$G_A(r_A) \approx \sigma_A(r_A) \quad (63)$$

satisfying constraints 61 and 62. Because $G_A(r_A)$ is a reshaping of $\sigma_A(r_A)$ for some chosen $\{Y_A^{\text{avg}}(r_A), \rho_A^{\text{avg}}(r_A)\}$ inputs, the minimization of H must be performed holding $\{Y_A^{\text{avg}}(r_A), \rho_A^{\text{avg}}(r_A)\}$ constant. Specifically,

$$\frac{\partial H(G_A(r_A))}{\partial G_A(r_A)} = 0 \quad (64)$$

The functional form

$$H(G_A(r_A)) = \frac{1}{2} \int_0^\infty \frac{(G_A(r_A) - \sigma_A(r_A))^2}{h(\sigma_A(r_A))} 4\pi(r_A)^2 dr_A + \int_0^\infty \Gamma_A(r_A) \phi_A(r_A) 4\pi(r_A)^2 dr_A - \Phi_A \varphi_A \quad (65)$$

is convenient, because it transforms the choice of $H(G_A(r_A))$ into the choice of $h(\sigma_A(r_A))$. Φ_A and $\Gamma_A(r_A)$ are Lagrange multipliers enforcing constraints 61 and 62, respectively. Inserting eq 65 into 64 and evaluating with the help of integration by parts gives

$$G_A(r_A) = \sigma_A(r_A) + h(\sigma_A(r_A)) \left(\Phi_A - \Gamma_A(r_A) \eta_A(r_A) + \frac{d\Gamma_A(r_A)}{dr_A} + \frac{2\Gamma_A(r_A)}{r_A} \right) \quad (66)$$

For given $\{Y_A^{\text{avg}}(r_A), \rho_A^{\text{avg}}(r_A)\}$ inputs, $G_A(r_A)$ is uniquely determined if $H(G_A(r_A))$ is a convex functional. Because $\partial^2 H / \partial G_A(r_A) \partial G_A(r'_A) = 0$, $H(G_A(r_A))$ is convex if

$$\frac{\partial^2 H(G_A(r_A))}{\partial G_A(r_A)^2} = \frac{1}{h(\sigma_A(r_A))} \geq 0 \quad (67)$$

We are free to measure electron density using our choice of unit system. When changing from one unit system to another, the numerical value of the density is multiplied by a unit system conversion factor,

$$\{\sigma_A(r_A), G_A(r_A)\} \Rightarrow \{\lambda_{\text{units}} \sigma_A(r_A), \lambda_{\text{units}} G_A(r_A)\} \quad (68)$$

Comparing eqs 66 and 68, a change in unit system also involves

$$\{(h(\sigma_A(r_A))\Phi_A)\} \Rightarrow \{\lambda_{\text{units}} h(\sigma_A(r_A))\Phi_A\} \quad (69)$$

To be physically meaningful, $h(\sigma_A(r_A))$ must have the same form in all unit systems. Examining eq 69, this requires h to be a homogeneous function of $\sigma_A(r_A)$ such that

$$h(\lambda_{\text{units}} \sigma_A(r_A)) = (\lambda_{\text{units}})^\xi h(\sigma_A(r_A)) \quad (70)$$

Equation 70 requires

$$h(\sigma_A(r_A)) = (\sigma_A(r_A))^\xi \quad (71)$$

The Supporting Information theoretically derives the optimal value $\xi = 1/2$. Tests described in the Supporting Information confirmed $\xi = 1/2$ performs better than both $\xi = 0$ and 1.

This completes the specification of the charge partitioning functional $F_A^{c3}(r_A)$ that defines the DDEC/c3 method. The main features of this charge partitioning scheme are graphically summarized in Figure S1 of the Supporting Information. Our method is a first-principles approach to computing NACs and other AIM properties, because all of its parameters have been derived theoretically.

2.5. Iterative Solution Algorithm. We showed above that the task of assigning $\{\rho_A(\vec{r}_A)\}$ can be formulated as global minimization of Ω . We now describe an efficient computational algorithm for doing this. First,

$$\{\rho_A^{\text{core}}(\vec{r}_A) = w_A^{\text{core}}(r_A) \rho^{\text{core}}(\vec{r}) / \sum_{k,B} w_B^{\text{core}}(r_B)\} \quad (72)$$

were fit from $\rho^{\text{core}}(\vec{r})$ using a previously published method.²³ Then, iterations were performed to determine the atomic charges. Each iteration (aka “charge cycle”) involved several consecutive loops. The first loop runs over $\{A, \vec{r}_A\}$ for all grid points with $r_A \leq r_{\text{cutoff}}$. This loop generates and stores $W(\vec{r})$ (eq 3), $Y(\vec{r})$ (eq 39), and $\rho^{\text{ref}}(\vec{r})$ (eq 37). For the first charge cycle, $\{w_A(r_A)\}$, $\{Y_A^{\text{avg}}(r_A)\}$, and $\{\rho_A^{\text{ref}}(r_A, n_A)\}$ are set equal to the neutral atom densities, $\{\rho_A^{\text{ref}}(r_A, z_A)\}$. The second loop also runs over $\{A, \vec{r}_A\}$ where $r_A \leq r_{\text{cutoff}}$. This loop computes $\rho_A(\vec{r}_A)$ (eq 2) and then subtracts $\rho_A^{\text{core}}(\vec{r}_A)$ to obtain $\rho_A^{\text{val}}(\vec{r}_A)$. This loop also accumulates the sum $T(r_A)$ of valid grid points for each radial shell. For each radial shell, the sum of the following functions over valid grid points is also accumulated: (a) $\rho_A(\vec{r}_A)$, (b) $(Y(\vec{r}))^{1/2}$, (c) $Y_A^{\text{avg}}(r_A)/(Y(\vec{r}))^{1/2}$, and (d) $\rho(\vec{r})/\rho^{\text{ref}}(\vec{r})$. $\phi \rho_A^{\text{val}}(\vec{r}_A) d^3\vec{r}_A$ and

$$u_A = \frac{\partial n_A}{\partial \ln(\lambda_A)} = \oint \left(1 - \frac{w_A(r_A)}{W(\vec{r})}\right) \rho_A(\vec{r}_A) d^3\vec{r}_A \geq 0 \quad (73)$$

are also accumulated. After the second loop completes, n_A^{val} is computed by adding the spatial integral of $\rho_A^{\text{val}}(\vec{r}_A)$ to the valence occupancy corrections. (See the Supporting Information for a description of how these occupancy corrections are computed.) The third loop runs over $\{A, r_A\}$. In this loop, the accumulated sums (a) through (d) are divided by $T(r_A)$ to give the following spherical averages: (f) $\langle \rho(\vec{r})/\rho^{\text{ref}}(\vec{r}) \rangle_{r_A}$, (g) $\rho_A^{\text{avg}}(r_A)$, (h) $\langle (Y(\vec{r}))^{1/2} \rangle_{r_A}$, and (i) $\langle Y_A^{\text{avg}}(r_A)/(Y(\vec{r}))^{1/2} \rangle_{r_A}$. Quantity (f) is then used to update $Y_A^{\text{avg}}(r_A)$ (eq 38) and $\sigma_A(r_A)$ (eq 57). $\rho_A^{\text{ref}}(r_A, n_A)$ is subsequently updated using the current n_A^{val} value. To maximize computational efficiency by limiting the number of loops that must be performed over grid points, this computational algorithm updates $\rho_A^{\text{ref}}(r_A, n_A)$, $Y_A^{\text{avg}}(r_A)$, and $w_A(r_A)$ together (rather than in a fully sequential manner) during each charge cycle. Because it takes two subsequent charge cycles for changes in $\rho_A^{\text{ref}}(r_A, n_A)$ to fully propagate through the update sequence $\rho_A^{\text{ref}}(r_A, n_A)$ to $Y_A^{\text{avg}}(r_A)$ to $w_A(r_A)$, constraints 53 and 58 are only applied on the fourth and later charge cycles. On the first three charge cycles, $w_A(r_A)$ is updated using $\lambda_A = 1$ and $G_A(r_A) = \sigma_A(r_A)$. On the fourth and later charge cycles, $w_A(r_A)$ is updated using eqs 6 and 59 and the new estimates for λ_A and $G_A(r_A)$. The new estimate for λ_A is computed by

$$\lambda_{A,j} = \max(1, (\lambda_{A,j-1} + (n_A^{\text{core}} - n_A)/u_A)) \quad (74)$$

where the subscript j ($j-1$) denotes the estimate computed in this (the previous) charge cycle. Examining eq 73, $u_A = 0$ if and only if atom A does not overlap any other atoms. Because $\rho_A(\vec{r}_A) = \rho(\vec{r})$ in the space occupied by this atom, n_A is completely independent of the λ_A value for a nonoverlapping atom. Consequently, we set $\lambda_A = 1$ when u_A is less than some zero tolerance, to avoid division by zero errors in eq 74. The new estimate for $G_A(r_A)$ is computed by an iterative process in the fourth loop. This loop runs over three simultaneous indexes: (a) the index that cycles most rapidly is the radial shell number, r_A , (b) the number of the atom (i.e., 1 for the first

atom, 2 for the second atom, etc.) in the unit cell, and (c) the slowest cycling index is called the pass number, i . Starting with the initial estimate $G_A^0(r_A) = \sigma_A(r_A)$, constraint 62 is enforced by

$$G_A^i(r_A) = \min[G_A^i(r_A), (G_A^i(r_A) - \Delta r_A) e^{-\eta_A(r_A) \Delta r_A}] \quad (75)$$

Equation 75 is applied stepwise beginning with $r_A = \Delta r_A$ and increasing r_A by Δr_A increments until r_{cutoff} is reached. Here, $\Delta r_A > 0$ is the distance (~ 0.05 Å) between adjacent radial shells. After applying eq 75 for all r_A values for every atom, initial estimates for the next pass are computed by

$$\Delta_A^{i+1} = \frac{\int_0^{r_{\text{cutoff}}} 4\pi(r_A)^2 (\sigma_A(r_A) - G_A^i(r_A)) dr_A}{\int_0^{r_{\text{cutoff}}} 4\pi(r_A)^2 \sqrt{\sigma_A(r_A)} dr_A} \quad (76)$$

$$G_A^{i+1}(r_A) = G_A^i(r_A) + \Delta_A^{i+1} \sqrt{\sigma_A(r_A)} \quad (77)$$

Starting with these initial estimates, the next pass applies eq 75 to enforce constraint 62. Then eqs 76 and 77 are used to generate initial estimates for the subsequent pass. This process is repeated until $\Delta_A^i \rightarrow 0$ for every atom, indicating convergence of $\{G_A(r_A)\}$ and satisfaction of constraints 61 and 62. The converged $G_A(r_A)$ is solution 66 with $\xi = 1/2$ and

$$\Phi_A = \sum_i \Delta_A^i \quad (78)$$

Because constraint 75 never increases $G_A^i(r_A)$, it follows from eq 76 that $\Delta_A^i \geq 0$. Because each term in sum 78 is positive, subsequent passes always improve the estimate of Φ_A . Thus, Φ_A and $G_A(r_A)$ necessarily converge when enough passes have been performed. Thirty passes were enough to converge $\{\Phi_A\}$ to within $\sim 10^{-8}$ for all of the systems we studied. After $\{G_A(r_A)\}$ converges in the fourth loop, the program starts the next charge cycle by going back to the first loop. At least 10 charge cycles were performed in all cases. Additional charge cycles were performed until each $\{n_A\}$ and $\{\rho_A^{\text{avg}}(r_A)\}$ component for the current and previous charge cycle differed by less than a convergence tolerance (e.g., 10^{-5}) for four consecutive charge cycles.

We now show that this algorithm converges to a global minimum of Ω . Each charge cycle calculates the following key variables. $\{n_A\}$ determines $\{Y_A^{\text{avg}}(r_A)\}$ by eq 36, which with $\{\rho_A^{\text{avg}}(r_A)\}$ determines $G_A(r_A)$ by eqs 75 to 78. Multiplying $G_A(r_A)$ by λ_A estimated using eq 74 gives $\{w_A(r_A)\}$. The $\{w_A(r_A)\}$ determines $\{\rho_A(\vec{r}_A)\}$ by eqs 2 and 3. Finally, spherical averaging and integration of $\{\rho_A(\vec{r}_A)\}$ again yields $\{\rho_A^{\text{avg}}(r_A)\}$ and $\{n_A\}$. Because it takes two subsequent charge cycles for changes in $\rho_A^{\text{ref}}(r_A, n_A)$ to fully propagate through the update sequence $\rho_A^{\text{ref}}(r_A, n_A)$ to $Y_A^{\text{avg}}(r_A)$ to $w_A(r_A)$, a self-consistent solution has been found when $\{\rho_A^{\text{avg}}(r_A)\}$ and $\{n_A\}$ are unchanged for four consecutive charge cycles. Because $G_A(r_A)$ is an explicit functional of $\rho_A^{\text{avg}}(r_A)$ and $\{Y_A^{\text{avg}}(r_A, n_A)\}$, $\{G_A(r_A)\}$ is converged if $\{\rho_A^{\text{avg}}(r_A)\}$ and $\{n_A\}$ are converged. We now show that convergence of $\{n_A\}$ and $\{\rho_A^{\text{avg}}(r_A)\}$ can occur only if $\{\lambda_A\}$ are also converged. First, we define J_{AB} as

$$\begin{aligned} J_{AB} &= \frac{\partial n_A}{\partial \ln(\lambda_B)} \\ &= \oint \left(\frac{w_A(r_A)}{W(\vec{r})} \delta_{AB} - \frac{w_A(r_A)}{W(\vec{r})} \frac{w_B(r_B)}{W(\vec{r})} \right) \rho(r) d^3\vec{r} \end{aligned} \quad (79)$$

Table 1. Neutral and Charged Molecular Systems Studied

molecule	atoms	magnetism	geometry	XC theory	basis sets
single molecule magnet	Fe ₄ O ₁₂ N ₄ C ₄₀ H ₅₂	noncollinear	DFT	PW91	plane-wave
B ₄ N ₄	B ₄ N ₄	none	DFT	PW91	6-311+G*
[GdI] ²⁺	[GdI] ²⁺	S _Z = 7	DFT	PBE	SDD and plane-wave
Zr bisperoxy complex	ZrO ₄ N ₄ C ₅₂ H ₇₂	none	DFT	B3LYP	LANL2DZ
Zr complex	ZrN ₄ C ₅₂ H ₇₂	none	DFT	B3LYP	LANL2DZ
MgI	MgI	doublet	DFT	PBE	SDD and plane wave
[Cu ₂ pyridine complex] ²⁺	[Cu ₂ N ₁₀ C ₃₆ H ₅₂] ²⁺	triplet	DFT	B3LYP	LANL2DZ
MoI	MoI	S _Z = 5	DFT	PBE	SDD and plane-wave
SnI	SnI	doublet	DFT	PBE	SDD and plane-wave
TeI	TeI	doublet	DFT	PBE	SDD and plane-wave
TiI	TiI	S _Z = 3	DFT	PBE	SDD and plane-wave
[Cr(CN) ₆] ³⁻	[Cr(CN) ₆] ³⁻	S _Z = 3	DFT	B3LYP	6-311+G*
ozone singlet	O ₃	none	optimized	PW91, B3LYP, CCSD, SAC-Cl, CAS-SCF	6-311+G* (DFT), AUG-cc-pVTZ (CCSD, SAC-Cl, and CAS-SCF)
ozone triplet	O ₃	triplet	optimized	PW91, B3LYP, CCSD, SAC-Cl, CAS-SCF	6-311+G* (DFT), AUG-cc-pVTZ (CCSD, SAC-Cl, and CAS-SCF)
ozone cation	[O ₃] ¹⁺	doublet	optimized	PW91, B3LYP, CCSD	6-311+G* (DFT), AUG-cc-pVTZ (CCSD)
formamide	CHONH ₂	none	DFT	PW91, B3LYP, HF	6-311++G**, AUG-cc-pVDZ, AUG-cc-pVTZ, and plane-wave
Br substituted carboxylic acid	BrC ₉ H ₁₃ O ₂	none	4 configurations	B3LYP	6-311++G**
CN substituted carboxylic acid	C ₁₀ H ₁₃ NO ₂	none	4 configurations	B3LYP	6-311++G**
ester subst. carboxylic acid	C ₁₂ H ₁₈ O ₄	none	16 configurations	B3LYP	6-311++G**
OH substituted carboxylic acid	C ₉ H ₁₄ O ₃	none	8 configurations	B3LYP	6-311++G**
OH substituted carboxylic acid	C ₉ H ₁₄ O ₃	none	500 ab initio molecular dynamics steps	PW91 with D2 dispersion correction	plane-wave

Table 2. Porous Solids Studied with the PAW Method and Plane-Wave Basis Sets

solid	unit cell atoms	magnetism	geometry	XC theory
IRMOF-1	Zn ₃₂ C ₁₉₂ O ₁₀₄ H ₉₆	none	XRD and DFT	PW91
ZIF-90	Zn ₁₂ N ₄₈ C ₉₆ O ₂₄ H ₇₂	none	DFT	PW91
Zn(nicotinate) ₂	Zn ₆ N ₁₆ C ₉₆ O ₃₂ H ₆₄	none	DFT	PW91
ZIF-8	Zn ₁₂ N ₄₈ C ₉₆ H ₁₂₀	none	DFT	PW91
CuBTC	Cu ₁₂ C ₇₂ O ₄₈ H ₂₄	antiferromagnetic	DFT	PW91
natrolite	Na ₄ Al ₄ Si ₆ O ₂₄ H ₈	none	DFT	PBE
MIL-53(Al)	Al ₄ C ₃₂ O ₂₀ H ₂₀	none	XRD large pore	PW91
BN nanotubes	B ₂₀ N ₂₀	none	DFT	PW91

For a process in which $\{G_A(r_A)\}$ are converged, changes in $\{w_A(r_A)\}$ can arise only from changes in $\{\lambda_A\}$. For such a process,

$$\sum_A d \ln(\lambda_A) dn_A = \sum_A \sum_B d \ln(\lambda_A) J_{AB} d \ln(\lambda_B) \quad (80)$$

Inserting eq 79 into 80 and rearranging gives

$$\sum_A d \ln(\lambda_A) dn_A = \frac{1}{2} \sum_A \sum_B \left[(d \ln(\lambda_A) - d \ln(\lambda_B))^2 \right. \\ \left. \oint \left(\frac{w_A(r_A)}{W(\vec{r})} \frac{w_B(r_B)}{W(\vec{r})} \right) \rho(r) d^3\vec{r} \right] \\ \geq 0 \quad (81)$$

Each integral in eq 79 is positive if atoms A and B overlap and zero if atoms A and B do not overlap. A chemical system containing zero valence electrons has the trivial solution

$\{\rho_A(\vec{r}_A) = \rho_A^{\text{core}}(\vec{r}_A)\}$ and $\{n_A = n_A^{\text{core}}\}$. In a chemical system containing more than zero valence electrons, at least one atom must have $n_A^{\text{val}} > 0$ and hence $\lambda_A = 1$ according to eq 74. For this atom, λ_A does not change over successive charge cycles, i.e., $d \ln(\lambda_A) = 0$. $\{n_A\}$ is converged when $\{dn_A\} = \{0\}$ between successive charge cycles. Examining eq 81, $\{dn_A\} = \{0\}$ only if $d \ln(\lambda_A)$ is the same for every atom. Thus, for a chemical system containing more than zero valence electrons, $\{d \ln(\lambda_A)\} = \{0\}$ when $\{dn_A\} = \{0\}$. In other words, $\{\lambda_A\}$ are necessarily converged whenever $\{n_A\}$ and $\{\rho_A^{\text{avg}}(r_A)\}$ are converged. Examining eq 74, convergence of λ_A (i.e., $\lambda_{A,j} = \lambda_{A,j-1}$) is equivalent to constraint 58. Thus, all required equations and constraints are satisfied at convergence. As explained in the Theory section, this kind of self-consistent solution is a global minimum of Ω .

Our method is computationally efficient, numerically precise, and robust. Functions of \vec{r}_A only need to be stored for the current grid point. Since only complete functions of \vec{r} and r_A

Table 3. Nonporous Bulk Solids Studied

solid	unit cell atoms	magnetism	geometry	XC theory	basis sets
V ₃ Si	V ₆ Si ₂	none	XRD	PBE	plane-wave
magnetite	Fe ₈ O ₈	ferri-magnetic	DFT	PBE, PBE+U	plane-wave
Fe ₂ SiO ₄	Fe ₄ Si ₂ O ₈	antiferromagnetic	DFT (spinel phase)	PBE	plane-wave
Fe ₂ O ₃	Fe ₄ O ₆	antiferromagnetic	DFT	PBE	plane-wave
Fe ₃ Si	Fe ₃ Si	ferro-magnetic	XRD	PBE	plane-wave
Mo ₂ C	Mo ₈ C ₄	none	XRD (PBCN phase)	PBE	plane-wave
Mo ₂ C	Mo ₂ C	none	XRD (P6-M2 phase)	PBE	plane-wave
GaAs	Ga ₄ As ₄	none	DFT	PW91	LANL2DZ
h-BN	B ₂ N ₂	none	XRD	PW91	plane-wave
MgH ₂	Mg ₂ H ₄	none	DFT	PW91	plane-wave
NaCl	NaCl	none	DFT	PW91	plane-wave
NaF	NaF	none	DFT	PW91	plane-wave
MgO	MgO	none	DFT	PW91	plane-wave
H atom in Pd	Pd ₃₂ H	none	DFT	PW91	plane-wave
Pd ₃ V	Pd ₂₄ V ₈	none	DFT	PW91	plane-wave
H atom in Pd ₃ V ^{a,c}	Pd ₂₄ V ₈ H	none	DFT	PW91	plane-wave
H atom in Pd ₃ In ^{b,c}	Pd ₂₄ In ₈ H	none	DFT	PW91	plane-wave
H atom in Pd ₃ Hf ^{b,c}	Pd ₄₈ Hf ₁₆ H	none	DFT	PW91	plane-wave
SrTiO ₃	SrTiO ₃	none	DFT	PW91	plane-wave

^aNACs computed for H located in an interstitial octahedral site bounded by four Pd and two V atoms. ^bNACs computed for H located in an interstitial octahedral site bounded by six Pd atoms. ^cThese H interstitial sites had lower energy than a couple of alternatives we tested, but our objective was not to find the lowest energy H interstitial site in these materials.

Table 4. Slabs and Systems Studied with One or Two Periodic Dimensions

system	unit cell atoms	magnetism	geometry	periodic dimensions	XC theory	basis sets
Mo ₂ C(110) with K adatom	KMo ₇₂ C ₃₆	none	DFT	3	PBE	plane-wave
SrTiO ₃ (100)	Sr ₄ Ti ₃ O ₁₀	none	DFT	3	PW91	plane-wave
BN sheet	BN	none	DFT	2,3	PW91	LANL2DZ, plane-wave
neoprene	C ₄ H ₅ Cl	none	DFT	1	PW91	6-311++G**
NaF(100)	Na ₁₆ F ₁₆	none	DFT	3	PW91	plane-wave

must be stored, memory requirements scale linearly with increasing volume (number of grid points). Because a cutoff radius is used, the volume that must be evaluated for each atom is independent of the unit cell size. Consequently, after the valence and total electron density grids have been prepared, both the memory requirements and computational time required to optimize $\{\rho_A(\vec{r}_A)\}$ scale linearly with increasing system size. The method also has excellent precision if appropriate integration settings are used. For integrations involving r_A , we used 100 radial shells evenly spaced between 0 and $r_{\text{cutoff}} = 5$ Å. More information regarding convergence with respect to integration grids is given in section 3.2. Finally, the method is also robust. All systems in this paper converged in fewer than 200 charge cycles, and we have not observed any systems where convergence is problematic for the DDEC/c3 method.

3. METHODS

3.1. Geometry and Electron Density Generation. For the purposes of testing and illustrating our method, calculations were performed for a wide variety of materials, including neutral and charged molecules (Table 1), porous solids with periodic structures (Table 2), dense periodic solids (Table 3), and systems with one or two periodic dimensions and solid surfaces (Table 4). The one-dimensional neoprene, two-dimensional BN, and three-dimensional GaAs calculations performed with Gaussian basis sets were performed with the Gaussian software.²⁴ All other periodic DFT calculations were

performed in the Vienna Ab Initio Simulation Package (VASP)^{25,26} using the projector augmented wave (PAW) method²⁷ and a 400 eV cutoff. PAW is a frozen-core method that yields all-electron densities and includes relativistic corrections for the core electrons. For these periodic systems, the product of the number of k -points and the unit cell volume exceeded 4000 Å³. All nonperiodic calculations were performed with the Gaussian software.²⁴ The exchange-correlation functionals, basis sets, and additional details are summarized in Tables 1, 2, 3, and 4. LANL2DZ (SDD) basis sets included relativistic effective core potentials on atoms heavier than neon (argon). Experimental structures for the solids MIL-53 (Al),²⁸ h-BN,²⁹ IRMOF-1,³⁰ Fe₃Si,³¹ V₃Si,^{32,33} Mo₂C (PBCN phase),³⁴ and Mo₂C (P6-M2 phase)³⁵ were taken from the literature. DFT-optimized geometries for the metal–organic framework Cu₃(BTC)₂³⁶ and the Mo₂C(110) surface with a K adatom³⁷ were taken from the literature. Optimized geometries for many of the remaining structures were given in our previous reports.^{4,23} The three-dimensional slab geometries listed in Table 4 included ~10 Å of vacuum space. In VASP, geometries were optimized to give atomic forces <0.03 eV/Å. In Gaussian, geometries were optimized to give atomic forces and displacements smaller than 0.0025 au and 0.01 au, respectively. Bader atomic volumes were computed by the program of Henkelman and co-workers using the total electron density; then the valence and spin densities were integrated over these volumes to determine the Bader NACs and atomic spin moments (ASMs), respectively.^{38–40}

Table 5. Error in the Integrated Total Number of Electrons before Renormalization

#	system	grid spacing (bohr)	atoms	electrons	periodic dimensions	input type	occupancy correction	total integration error (e)
1	formamide	0.10	6	24	none	valence and total density grids	no	6×10^{-4}
2	formamide	0.14	6	24	none	natural orbitals and occupancies	yes	7×10^{-7}
3	formamide	0.14	6	24	three	valence, core, and pseudized density grids	yes	1×10^{-4}
4	formamide	0.14	6	24	three	valence and core density grids	no	6×10^{-4}
5	[GdI] ²⁺	0.06	2	115	none	valence and total density grids	no	0.03
6	[GdI] ²⁺	0.14	2	115	none	valence and total density grids	no	0.27
7	[GdI] ²⁺	0.14	2	115	none	natural orbitals and occupancies	yes	1×10^{-5}
8	Cu ₂ N ₁₀ C ₃₆ H ₅₂	0.15	100	394	none	valence and total density grids	no	2×10^{-3}
9	Cu ₂ N ₁₀ C ₃₆ H ₅₂	0.14	100	394	none	natural orbitals and occupancies	yes	2×10^{-8}
10	IRMOF-1	0.14	424	3040	three	valence and core density grids	no	1.1
11	IRMOF-1	0.14	424	3040	three	valence, core, and pseudized density grids	yes	8×10^{-5}
12	neoprene	0.14	10	46	one	natural orbitals and occupancies	yes	1×10^{-4}
13	BN sheet	0.14	2	12	two	natural orbitals and occupancies	yes	4×10^{-4}
14	GaAs	0.14	8	256	three	natural orbitals and occupancies	yes	7×10^{-9}

3.2. Hybrid Analytic-Numeric Integration. During charge partitioning, appropriate grids and integration routines must be used to achieve good computational efficiency and precision. DDEC charges are always computed using an all-electron density. When the electron density has been generated using a pseudopotential or effective core potential, the core electrons replaced by the pseudopotential or effective core potential are automatically reinserted at the beginning of charge analysis using stored core electron reference densities. To accurately integrate the electron density near the cusp in $\rho(\vec{r})$ near each nucleus, special techniques are used. Reasonably spaced uniform grids can be used with hybrid integration techniques that separate $\rho(\vec{r})$ into a “core” density $\rho^{\text{core}}(\vec{r})$ near each nucleus that is integrated analytically and a more delocalized “valence” density $\rho^{\text{val}}(\vec{r})$ that is integrated numerically. This separation is flexible and does not have to strictly correspond to the densities arising from distinct core and valence natural orbitals, although it often does. The specific allocation between core and valence densities depends on the type of input available (e.g., electron density grids or natural orbitals).

In addition to integrating core electrons analytically, we developed occupancy corrections that increase valence integration precision by orders of magnitude. These corrections are predicted differences between valence occupations computed using numerical integration for a chosen grid and the result which would be obtained if the integration were performed analytically using an infinitely fine grid. When the material's $\rho(\vec{r})$ was computed using Gaussian basis sets, the occupancy corrections normalize each basis set product to make sure its integration over the chosen grid matches the correct analytic value. For quantum chemistry calculations performed with Gaussian basis sets, the DDEC program (ddec.sourceforge.net) reads the natural orbitals and their occupancies (i.e., the eigenvectors and eigenvalues of the first order density matrix) and uses these to automatically compute suitable density grids and occupancy corrections.

When the material's $\rho(\vec{r})$ was computed using the projector augmented wave (PAW) method,^{27,41} occupancy corrections were computed as follows. The PAW method applies a projector function to the Kohn–Sham (i.e., nonpseudized) valence orbitals to generate pseudized valence orbitals.^{27,41} Because the pseudized density does not have cusps at the

nuclear positions and the pseudized and nonpseudized valence orbitals are identical outside a core radius enclosing each nucleus, an accurate valence electron count can be obtained by integrating the pseudized density.^{27,41} Accordingly, valence occupancy corrections for each atom were computed by integrating the difference between the pseudized and nonpseudized valence densities over nonoverlapping atomic volumes. This procedure gives accurate integration of the nonpseudized valence density for each atom despite the presence of an electron density cusp near each nucleus.

Occupancy corrections eliminate trial and error in the preparation of charge analysis input files. As shown in Table 5, when the grid spacing is 0.14 bohr (~ 0.07 Å) and occupancy corrections are included, the error in the integrated total number of electrons is typically $\leq \sim 10^{-4}e$ before renormalization. (After renormalization, the integrated total is exact.) A $10^{-4}e/\text{atom}$ error in the sum of integrated electrons before renormalization should have a maximum effect of $10^{-4}e \times 627$ (kcal/mol)/Hartree $\times 1/(1.5R^{\text{vdw}})$ on the root mean squared error (RMSE) shown in eq 82. For the typical value $R^{\text{vdw}} = \sim 3$ bohr, this corresponds to a maximum RMSE error of ~ 0.01 kcal/mol. A total integration error of $10^{-4}e$ before renormalization is expected to give smaller RMSE errors, because it is distributed over several atoms. Consequently, accurate charge partitioning can always be done using a 0.14 bohr (~ 0.07 Å) grid spacing with occupancy corrections. Except for some tests described in Table 5 and the CAS-SCF ozone calculations, all DDEC/c3 results in this paper included occupancy corrections. The Supporting Information contains further details on how these occupancy corrections were computed and lists the total integration errors for every system.

How do our integration grids compare to those typically used to integrate the exchange-correlation energy or electron density in DFT software packages? Uniformly spaced grids are the most convenient type for periodic systems. The grid spacing (0.14 bohr) we used is an excellent precision grid in comparison to those typically used in periodic DFT calculations. For nonperiodic calculations of molecular systems, many software packages (e.g., Gaussian 09) use atom-centered nonuniform grids with weights assigned by Becke's method.⁴² This type of grid is described by radial and angular grids, where the radial grid spacing is finer near atomic nuclei to provide accurate core electron integration. For molecular systems, these

Table 6. Integration and Basis Set Tests for the Formamide Molecule

test	XC	basis sets	AIM	grid type	net atomic charges					
					C	N	O	H1	H2	H3
1	HF ^a	AUG-cc-pVDZ	ISA	Becke ^b	0.878	−1.019	−0.690	0.413	0.434	−0.016
2	HF ^a	AUG-cc-pVDZ	ISA	uniform ^c	0.867	−1.011	−0.687	0.412	0.432	−0.013
3	HF ^a	AUG-cc-pVTZ	ISA	Becke ^b	0.869	−1.017	−0.689	0.414	0.435	−0.014
4	HF ^a	AUG-cc-pVTZ	ISA	uniform ^c	0.859	−1.010	−0.684	0.414	0.433	−0.012
5	B3LYP	AUG-cc-pVDZ	ISA	Becke ^d	0.74	−0.92	−0.59	0.39	0.41	−0.03
6	B3LYP	AUG-cc-pVDZ	ISA	uniform ^c	0.719	−0.903	−0.589	0.385	0.404	−0.017
7	B3LYP	AUG-cc-pVTZ	HD	Becke ^e	0.139	−0.132	−0.305	0.126	0.132	0.040
8	B3LYP	AUG-cc-pVTZ	HD	uniform ^c	0.139	−0.136	−0.304	0.128	0.133	0.040
9	PW91	6-311++G**	DDEC/c3	uniform ^f	0.550	−0.759	−0.513	0.351	0.364	0.008
10	PW91	6-311++G**	DDEC/c3	uniform ^g	0.550	−0.760	−0.513	0.352	0.364	0.008
11	PW91	plane-wave	DDEC/c3	uniform ^h	0.550	−0.755	−0.515	0.349	0.364	0.007
atomic dipole magnitude (au)										
7	B3LYP	AUG-cc-pVTZ	HD	Becke ^e	0.069	0.026	0.152	0.197	0.196	0.131
8	B3LYP	AUG-cc-pVTZ	HD	uniform ^c	0.071	0.027	0.147	0.200	0.199	0.135
9	PW91	6-311++G**	DDEC/c3	uniform ^f	0.101	0.053	0.018	0.033	0.032	0.071
10	PW91	6-311++G**	DDEC/c3	uniform ^g	0.103	0.053	0.017	0.033	0.032	0.071
11	PW91	plane-wave	DDEC/c3	uniform ^h	0.100	0.054	0.011	0.034	0.032	0.074

^aUsing the B3LYP optimized geometries for each basis set. ^bComputed by Wheatley (see Acknowledgments) using 300 radial and ~500 angular points per atom. ^c0.14 bohr uniform grid spacing with spherical averaging performed using 100 radial shells uniformly spaced between 0 and 5 Å. ^dFrom ref 11. ^eComputed using the HPAM software program described in ref 43, with 100 radial and 590 angular points per atom. ^fNonperiodic calculation using natural orbitals and occupancies as input (# 2 of Table 5). ^gNonperiodic calculation using density cube files as input (# 1 of Table 5). ^hPeriodic calculation using PAW method (15 Å × 15 Å × 15 Å unit cell) with valence, core, and pseudized density grid files used as program input (# 3 of Table 5).

Table 7. Effective Decay Exponents for Atom Tails in Selected Materials

material	periodic dimensions	buried atoms?	basis sets	decay exponents (bohr ^{−1})
formamide	none	no	6-311++G**	1.73(O), 1.68(N), 1.79(C), 1.60(H1), 1.67(H2), 1.56(H3)
MgI	none	no	SDD	1.40(Mg), 1.67(I)
[GdI] ²⁺	none	no	SDD	1.67(Gd), 2.20(I)
BN sheet	two	no	LANL2DZ	1.86(B), 1.78(N)
BN sheet	three	no	plane-wave	1.91(B), 1.82(N)
bulk NaF	three	yes	plane-wave	2.06(Na), 1.94(F)
bulk MgO	three	yes	plane-wave	2.02(Mg), 1.75(O)
bulk MgH ₂	three	yes	plane-wave	1.74(Mg), 1.75(H)
bulk SrTiO ₃	three	yes	plane-wave	1.75(Sr), 1.88(Ti), 1.75(O)
Mo ₂ C (PBCN)	three	yes	plane-wave	1.75(Mo), 1.76(C)
Mo ₂ C (P6-M2)	three	yes	plane-wave	1.75(Mo), 1.75(C)
bulk GaAs	three	yes	LANL2DZ	1.80(Ga), 1.75(As)
neoprene	one	no	6-311++G**	1.76–1.96(C), 1.59–1.73(H), 2.00(Cl)
MIL-53(Al)	three	almost	plane-wave	1.77(Al), 1.76–1.78(O), 1.74–1.80(H), 1.79–1.83(C)
[Cr(CN) ₆] ^{3−}	none	almost	6-311+G*	1.74(Cr), 1.70(C), 1.62(N)

nonuniform grids facilitate accurate integration using fewer grid points than uniformly spaced grids.⁴²

Although the integration grid affects computational efficiency and numerical precision, the mathematical value of the integrals being evaluated does not formally depend on the integration grid type. Thus, reasonably spaced uniform and nonuniform grids should give similar results. This can be tested by comparing results using our uniform grid hybrid integration method to results obtained using nonuniform Becke-style grids. Table 6 shows integration and basis set tests for the formamide molecule. Examining Table 6, the AUG-cc-pVDZ and AUG-cc-pVTZ basis sets give similar NACs. Comparing tests 1 to 2, 3 to 4, 5 to 6, and 7 to 8, NACs computed using Becke-style nonuniform grids are in good agreement with NACs computed using our uniformly spaced hybrid integration method. The last few rows of Table 6 show that both grid types give similar

atomic dipoles. Also, the total integration errors for tests 7 ($1.5 \times 10^{-6}e$) and 8 ($7 \times 10^{-7}e$) do not indicate a significant difference in accuracy between the two grid types. As a further example, the total integration errors for the ground state geometry of the Br substituted carboxylic acid (BrC₉H₁₃O₂) shown in Table 1 were $5 \times 10^{-5}e$ (Becke grids, HPAM program) and $8 \times 10^{-5}e$ (uniform grids, our hybrid integration), which also does not indicate a significant difference in accuracy between the two grid types.

For the PW91 functional, the DDEC/c3 NACs computed for formamide using three different input types agree with each other to within $\sim 10^{-3}e$. As shown in Table 6, calculations performed using a periodic unit cell and the PAW method (test 11) gave NACs and atomic dipole magnitudes virtually identical to those performed using a nonperiodic system and 6-311++G** basis sets (tests 9 and 10). The specific grid point

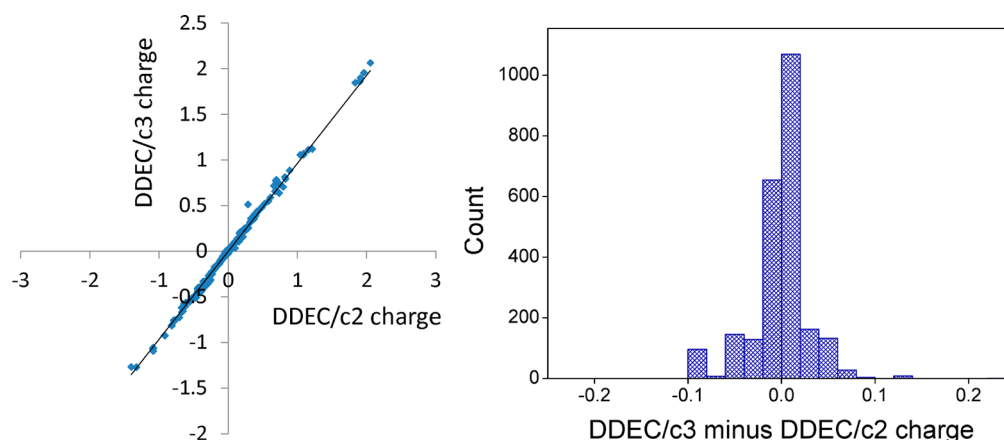


Figure 3. Comparison of DDEC/c3 charges to previously published DDEC/c2 charges, excluding nonporous solids and solid surfaces ($y = x$ line shown for comparison purposes).

locations were different for these three tests. The “core” electron density allocations for tests 9 (analysis from density matrix eigenstates) and 10 (analysis from valence and total density grids) were also different. The close agreement between all three sets of DDEC/c3 NACs shows that our hybrid integration approach is not unduly sensitive to (a) grid point locations, (b) the specific allocation of “core” electrons as long as the total density remains constant, and (c) the basis set type as long as sufficiently large basis sets are used. Additional tests were performed using natural orbitals and occupancies as input with (a) 0.10 bohr uniform grid spacing with 100 radial shells and a 5 Å cutoff, (b) 0.10 bohr uniform grid spacing with 75 radial shells and a 4 Å cutoff, and (c) 0.14 bohr uniform grid spacing with 75 radial shells and a 4 Å cutoff. As shown in the Supporting Information, these NACs were within a few thousandths of an electron of the values computed in test 9 using 0.14 bohr uniform grid spacing with 100 radial shells and a 5 Å cutoff.

3.3. Effective Decay Exponents. The DDEC program automatically computes the effective decay exponent for each atom after charge partitioning is completed. Specifically, $\ln(\rho_A^{\text{avg}}(r_A))$ is fit to $a_0 + a_1 r_A$ using least-squares fitting, where a_1 is the decay exponent. Typically, this regression is performed for $2 \text{ Å} \leq r_A \leq r_{\text{cutoff}}$ and has correlation coefficients $R^2 > 0.99$, indicating almost perfect fits. Table 7 shows results for selected materials. An atom is considered buried if the distance from the atom’s nucleus to the material’s vdW surface is larger than the atom’s vdW radius. The Al atom in MIL-53(Al) and the Cr atom in $[\text{Cr}(\text{CN})_6]^{3-}$ are almost buried, because only a small percentage of the vdW sphere of the Al or Cr atom lies on the material’s vdW surface. The value of the parameter b in eq 55 sets an approximate minimum on the decay exponent of buried and almost buried atoms. The theoretical value $b \approx 1.75/\text{bohr}$ is derived in the Supporting Information. Examining Table 7, all buried and almost buried atoms had decay exponents $\geq 1.74/\text{bohr}$. In contrast, exposed surface atoms may have decay exponents significantly less than b . For the formamide, MgI, and $[\text{GdI}]^{2+}$ molecules and the neoprene polymer, the decay exponents are >1.75 for some atoms and <1.75 for other atoms, indicating that the tail constraint has a negligible effect when a system contains only exposed surface atoms.

To further investigate the effects of the tail constraint, NACs and decay exponents were computed for formamide, MIL-

53(Al), bulk MgH_2 , bulk Mo_2C (PBCN and P6-M2 phases), and the $\text{Mo}_2\text{C}(110)$ surface with a K adatom using $b = 1.55, 1.65, 1.75, 1.85$, and $1.95/\text{bohr}$. Changing the b value (see Supporting Information) produced a corresponding change in the minimum decay exponent of buried and almost buried atoms. Buried atoms (e.g., Mg in bulk MgO) whose decay exponent is already significantly larger than b were unaffected by a change in b . The effects of changing b were small for many materials. Over the range $b = 1.55$ to 1.95 , all formamide NACs changed by less than $10^{-3}e$, demonstrating that the tail constraint has a negligible effect for this material. The $\text{Mo}_2\text{C}(110)$ surface with a K adatom was sensitive to the b value. For this system, the number of charge cycles required for convergence was 430 ($b = 1.55$), 218 ($b = 1.65$), 87 ($b = 1.75$), 76 ($b = 1.85$), and 65 ($b = 1.95$), which indicates the optimization landscape has substantially less curvature for $b \leq 1.65$ than for $b \geq 1.75$. For this material, the largest difference in NAC was $0.35e$ between $b = 1.55$ and $b = 1.75$ and $0.067e$ between $b = 1.75$ and $b = 1.95$. The NACs for $b = 1.55$ were problematic, because atoms that had essentially the same environments had NACs that differed by as much as $0.2e$. These results show that $b = 1.55$ is not adequate for correcting problematic convergence in the most difficult systems. For $b = 1.75$, atoms with essentially the same environments also had essentially the same NACs. In summary, the theoretically derived value $b = 1.75/\text{bohr}$ gives good performance.

Finally, we comment on the physical reasons why buried anions have higher effective decay exponents than exposed anions. A buried anion is surrounded by other atoms that screen its charge, unlike an exposed anion. Consequently, the density tail of an exposed anion is repelled more than that of an equivalent buried anion. This causes exposed (i.e., surface) anions to be more diffuse than their buried counterparts. Our charge method accounts for this by constraining the effective decay exponent to prevent buried atoms from becoming too diffuse.

4. RESULTS AND DISCUSSION

4.1. Comparison to DDEC/c2 Results. Figure 3 compares the DDEC/c2 and DDEC/c3 NACs for all systems in our three earlier reports, excluding the nonporous solids and solid surfaces.^{4,10,23} Specifically, this set was comprised of (a) all molecules and levels of theory shown in Table 1 except the substituted carboxylic acids and some of the basis set and XC

functional tests for formamide, (b) the BN sheet studied with a plane-wave basis set, and (c) all porous materials shown in Table 2 except natrolite. As shown in Figure 3, the DDEC/c2 and DDEC/c3 NACs were similar for all materials in this set and had an RMS difference of only 0.03e. The largest difference was 0.23e for the Cr atom in $[\text{Cr}(\text{CN})_6]^{3-}$, which had net charges of 0.242 (DDEC/c2) and 0.513 (DDEC/c3). As shown in Figure 4, this Cr atom is nearly buried. We consider

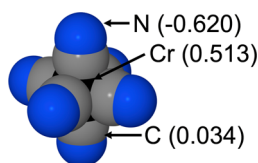


Figure 4. The vdW surface of $[\text{Cr}(\text{CN})_6]^{3-}$ showing the Cr atom is nearly buried. The DDEC/c3 charges are shown in parentheses.

the DDEC/c3 method to be more accurate than the DDEC/c2 method for this material, because the DDEC/c3 method is optimized to describe buried and nearly buried atoms. $[\text{Cr}(\text{CN})_6]^{3-}$ is a challenging molecule for which natural population analysis (NPA) and Mulliken populations projected onto a minimal basis set (MBS) give the unphysical Cr charges of -0.78 (NPA) and -11.97 (MBS).²³ These results show that changes introduced by the DDEC/c3 method do not adversely affect the computation of NACs in systems where the DDEC/c2 method already performed well.

Net atomic charges used in the development of force fields should approximately reproduce $V(\vec{r})$ outside a material's electron distribution. We now compare the accuracy of the DDEC/c2 and DDEC/c3 methods for fitting $V(\vec{r})$ as quantified by the root mean squared error (RMSE)

$$\text{RMSE} = \left[\frac{\sum_{\text{grid}} [(V(\vec{r}) - V(\vec{r}, \{q_A\}, \{\vec{\mu}_A\})) - \langle V(\vec{r}) - V(\vec{r}, \{q_A\}, \{\vec{\mu}_A\}) \rangle_{\text{grid}}]^2}{N_{\text{grid}}} \right]^{1/2} \quad (82)$$

where $V(\vec{r})$ is the ab initio electrostatic potential, $V(\vec{r}, \{q_A\}, \{\vec{\mu}_A\})$ is the electrostatic potential of the net atomic charges $\{q_A\}$ (and optionally, atomic dipoles $\{\vec{\mu}_A\}$), and N_{grid} is the number of grid points. A grid of uniformly spaced points was used between surfaces defined by γ_{inner} and γ_{outer} times the vdW radii, where $(\gamma_{\text{inner}}/\gamma_{\text{outer}}) = (1.4, 2.0)$ for nonperiodic materials and $(\gamma_{\text{inner}}/\gamma_{\text{outer}}) = (1.3, 20.0)$ for periodic materials.^{10,18,44} We used the same vdW radii as Watanabe et al.¹⁰ The RMSE value for an atomic charge model divided by the RMSE when all atomic charges are set to zero is called the relative root mean squared error, RRMSE.^{10,18,44,45} As shown in Table 8, the c2 and c3 methods gave almost the same RMSE, RRMSE, and atomic multipole magnitudes for each material. Previously, we compared the accuracy of various point charge models for reproducing $V(\vec{r})$ in these same 10 materials and found that the DDEC/c2 charges are suitable for constructing force fields for atomistic simulations of porous materials.^{4,10} Because the DDEC/c2 and DDEC/c3 methods have essentially the same accuracy for reproducing $V(\vec{r})$, these results show that DDEC/c3 charges are well suited for constructing force fields of porous materials.

The performance difference between the DDEC c2 and c3 methods is most apparent for nonporous solids with some covalent bonding and short bond lengths. As examples, we consider an interstitial H atom adsorbed in Pd and the ordered alloys Pd_3V , Pd_3Hf , and Pd_3In . The Pd_3V alloy without interstitial H was also studied. The chemical reasonableness of the computed NACs was assessed by comparison to the element electronegativities of 2.20 (H), 2.20 (Pd), 1.63 (V), 1.78 (In), and 1.3 (Hf).¹⁶ On the basis of these electronegativities, we expect V, In, and Hf to be positively charged in these alloys. Table 9 compares the average Bader, DDEC/c2, and DDEC/c3 NAC for each element in these materials. The DDEC/c2 calculation for interstitial H in Pd_3V failed because it tried to assign a V charge state more anionic than V^- , causing the calculation to terminate. The DDEC/c3 and Bader methods had no difficulties assigning NACs for any of these materials. The Bader and DDEC/c3 NACs followed the trend $\text{Hf} > \text{V} > \text{In} > (\text{Pd}, \text{H})$, which agrees with the order predicted by element electronegativities. In contrast, DDEC/c2 assigned negative charges to In, V, and Hf in these materials. DDEC/c2 failed because it does not contain appropriate constraints to prevent buried atoms like In, V, and Hf from becoming too diffuse in these materials. These examples show that the DDEC/c3 method is a clear improvement over the DDEC/c2 method for nonporous materials.

Table 8. Accuracy of Fitting the Electrostatic Potential (Values in Parentheses Include Atomic Dipoles)

material	RMSE (kcal/mol)		RRMSE		AM_{max}^a	
	c2	c3	c2	c3	c2	c3
B_4N_4	0.29 (0.33)	0.26 (0.33)	0.09 (0.10)	0.08 (0.10)	0.21d ^b	0.21d
BN tube	9.04 (2.19)	8.81 (2.40)	2.18 (0.53)	2.13 (0.58)	0.30q ^b	0.36q
BN sheet	0.09 (0.09)	0.09 (0.09)	0.74 (0.74)	0.73 (0.73)	0.36q ^b	0.42q
formamide	0.74 (0.40)	0.73 (0.43)	0.07 (0.04)	0.07 (0.04)	0.43q ^b	0.54q
lp-MIL-53	1.45 (0.49)	1.57 (0.59)	0.74 (0.25)	0.80 (0.30)	0.14q ^b	0.13q
IRMOF-1 (XRD)	0.81 (0.42)	0.83 (0.44)	0.38 (0.20)	0.39 (0.20)	0.30q ^b	0.31q
IRMOF-1 (PW91)	0.98 (0.55) ^c	0.65 (0.58)	0.37 (0.21) ^c	0.27 (0.24)	0.27q	0.32q
ZIF-8	0.76 (0.71) ^c	0.88 (0.72)	0.50 (0.46) ^c	0.57 (0.47)	0.12q	0.12q
ZIF-90	0.96 (0.95) ^c	0.81 (0.84)	0.13 (0.13) ^c	0.12 (0.12)	0.17q	0.19q
Zn-nicotinate	1.02 (0.39)	0.95 (0.45)	0.58 (0.22)	0.54 (0.26)	0.19q	0.19q

^aLargest atomic dipole magnitude (d) or quadrupole component magnitude (q) in atomic units. ^bData from ref 4. ^cData from ref 10.

Table 9. Average Atomic Charges for Interstitial H in Ordered Pd₃X Alloys

material	H charge			Pd charge			X charge		
	Bader	c2	c3	Bader	c2	c3	Bader	c2	c3
H atom in Pd	−0.04	−0.29	−0.25	0.00	0.01	0.01	n.a.	n.a.	n.a.
Pd ₃ V	n.a.	n.a.	n.a.	−0.35	0.33	−0.10	1.04	−0.98	0.31
H atom in Pd ₃ V	−0.22	failed	−0.32	−0.34	failed	−0.09	1.04	failed	0.32
H atom in Pd ₃ In	−0.05	−0.18	−0.18	−0.21	0.08	−0.08	0.64	−0.23	0.27
H atom in Pd ₃ Hf	−0.05	−0.1	0.02	−0.53	0.2	−0.31	1.58	−0.6	0.92

The DDEC/c2 method also failed to converge for a Mo₂C(110) surface with a K adatom. Charge partitioning is especially challenging for this system, because it contains both bulk and surface atoms, the K cation induces a surface image charge, and short Mo–C bond distances give large density overlaps between atoms. As shown in Figure 5, the DDEC/c3

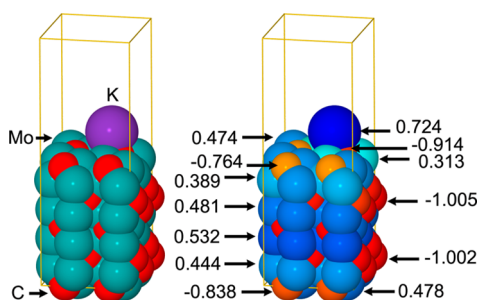


Figure 5. Image charge created by K cation adsorbed on a Mo₂C(110) surface. Left: Atoms colored by element: K (purple), Mo (light blue), C (red). Right: Atoms colored by net charge on a continuous color scale from dark red (−1) to green (0) to dark blue (+0.7). A negative image charge is visible below the K cation. Atomic charges are labeled for selected atoms.

method gave reasonable results for this system. Polarization of the electron cloud created a negative surface image charge under the K cation. Specifically, adjacent to the K cation, the surface C NAC is −0.914 instead of −0.838 and the surface Mo NACs are 0.248, 0.296, 0.313, and 0.319 instead of 0.478. These changes sum to −0.812, which is nearly the opposite of the K cation NAC of 0.724. A smaller effect of the K adatom on the Mo and C NACs penetrates into the first few subsurface layers. For comparison, bulk Mo₂C has NACs of 0.569 (Mo) and −1.139 (C) for the PBCN phase and 0.388 (Mo) and −0.776 (C) for the P6-M2 phase.

4.2. Comparison to Experimental Results. Although it is not straightforward to experimentally measure NACs, a variety of spectroscopic techniques can yield information about them. High resolution X-ray diffraction (XRD) and electron beam diffraction are two approaches for experimentally extracting

NACs.^{46–48} Typically, the diffraction data are refined using a multipolar model (e.g., Hansen-Coppens⁴⁹) to determine atomic coordinates, thermal parameters, and an electron density map.^{47–52} Because a multipolar basis set is over-complete, refinement must be done with a more restrictive model to determine NACs.^{46,51} NACs are commonly extracted from XRD data by following the multipolar refinement with a second refinement that includes only monopole terms with spherical pseudoatoms.^{46,48} Using the above terminology, this procedure, called “kappa refinement”,⁴⁶ models the electron density for each spherical pseudoatom as

$$\rho_{\text{at}}(r_A) = \rho_A^{\text{core}}(r_A) + n_A^{\text{val}} \kappa^3 p_A^{\text{val}}(\kappa r_A) \quad (83)$$

where $p_A^{\text{val}}(r_A)$ is the normalized valence density of the neutral reference atom

$$p_A^{\text{val}}(r_A) = (\rho_A^{\text{ref}}(r_A, z_A) - \rho_A^{\text{core}}(r_A)) / (z_A - n_A^{\text{core}}) \quad (84)$$

This model includes only two adjustable parameters for each atom: (a) the number of valence electrons, n_A^{val} , and (b) a radial factor κ that determines whether the atom in the material is contracted ($\kappa > 1$) or expanded ($\kappa < 1$) relative to the neutral atom. Kappa refinement has two key limitations. First, the pseudoatom densities do not necessarily sum to $\rho(\vec{r})$. Second, charged atoms are represented as expanded or contracted versions of the neutral atom, which may or may not be a useful approximation. For purely ionic materials like MgO, refinement can be performed using reference ions instead of neutral atoms.⁴⁷ These limitations suggest that further improvements are needed to make extracting accurate NACs from experimental data a routine procedure.

Table 10 compares experimentally derived and theoretically computed NACs for the formamide molecule shown in Figure 6. Stevens performed high resolution XRD of two independently grown crystals.⁴⁸ For each crystal, two or three symmetry related forms of each independent reflection were measured. The two data sets were then compared and combined to reduce noise and yielded 1125 independent reflections for subsequent refinement. This is one of the most extensive and careful studies of this kind ever done.^{48,50} The atomic coordinates,

Table 10. Experimental and Theoretical Formamide Atomic Charges

atom	high res. XRD		computed with B3LYP/AUG-cc-pVTZ level of theory						
	fixed H radial factor ^a	optimized H radial factors ^b	Bader	DDEC/c3	ESP	HD	IH	ISA	NPA
O	−0.50 ± 0.04	−0.55 ± 0.04	−1.149	−0.557	−0.562	−0.304	−0.537	−0.593	−0.605
N	−0.71 ± 0.11	−0.78 ± 0.07	−1.183	−0.788	−0.923	−0.136	−0.862	−0.911	−0.808
C	0.37 ± 0.11	0.51 ± 0.08	1.469	0.624	0.680	0.139	0.644	0.726	0.534
H1	0.35 ± 0.05	0.39 ± 0.03	0.411	0.352	0.389	0.128	0.360	0.389	0.394
H2	0.37 ± 0.04	0.40 ± 0.03	0.426	0.369	0.429	0.133	0.377	0.407	0.388
H3	0.12 ± 0.06	0.03 ± 0.03	0.026	0.000	−0.012	0.040	0.018	−0.019	0.096

^aKappa refinement with κ_{H} fixed at 1.40, reference 46. ^bReference 48.

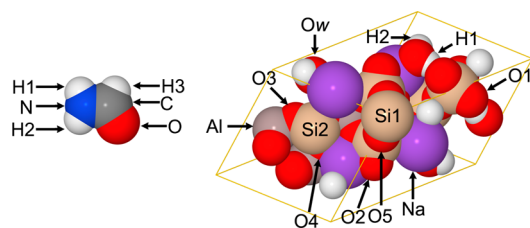


Figure 6. The structures of formamide and natrolite. For natrolite, lines indicate the unit cell boundaries.

thermal parameters, and electron density map were determined using multipolar refinement.^{48,50} The XRD data were collected at a low temperature (90 K) to improve the accuracy of this multipolar refinement by limiting thermal vibrations. Because it does not have any core electrons, H scatters X-rays weakly. Consequently, H parameters are often estimated rather than refined. NACs were extracted by spherical pseudoatom refinement with radial factors for H fixed at 1.40 (column 2 in Table 10) or completely optimized (column 3).^{46,48} The experimentally derived NACs using optimized H radial factors are nearly indistinguishable from the theoretically computed DDEC/c3 NACs using either the B3LYP (Table 10) or PW91 (Table 6) functionals. Examining Table 10, the NPA atomic charges were also nearly indistinguishable from the experimentally extracted ones. Except for a slight overestimation of the N and C charge magnitudes, the IH, ISA, and ESP atomic charges were also close to the experimentally extracted ones. Finally, the Bader and HD atomic charges were similar in sign but not in magnitude to the experimentally extracted ones.

Next, we consider the ionic solids MgO, NaF, and NaCl. The spherical atom approximation is reasonable for these materials because the atomic dipoles and quadrupoles must be zero by crystal symmetry. A high resolution electron beam diffraction study showed that MgO is completely ionic,⁴⁷ which agrees with the DDEC/c3 charge magnitude of 2.000 (2.005) using 10 (two) frozen core electrons for Mg. Kappa refinement of XRD data gave a charge magnitude of 0.95 ± 0.01 for crystalline NaF,⁵³ while the DDEC/c3 charge magnitude is 1.000 (using 10 frozen core electrons for Na). For crystalline NaCl, the DDEC/c3 charge magnitude is 0.960, which suggests that NaCl is slightly less ionic than NaF.

A direct comparison to experimental data can also be made for natrolite ($\text{Na}_2\text{Al}_2\text{Si}_3\text{O}_{10} \cdot 2\text{H}_2\text{O}$), a naturally occurring zeolite with a relatively well-ordered Al–Si distribution.^{54,55} Ghermani et al. used multipolar refinement with spherical H pseudoatoms and Ow–H bond lengths fixed at 0.96 Å to determine structural and thermal parameters and the electron density map.⁵⁶ NACs were extracted by subsequent kappa refinement with the Na charge set to +1 and the H_2O net charge assumed to be zero.⁵⁶ Table 11 compares these experimentally derived NACs and geometric parameters to those computed with the DDEC/c3 method and PW91 functional. We compute a net charge of $-0.045e$ for the H_2O group. The atom labels shown in Figure 6 above follow Ghermani et al.⁵⁶ While the geometries and Na charges are in excellent agreement, the theoretically computed NACs for Si, Al, O, and H have significantly larger magnitudes than the experimentally extracted ones. We could not compute RMSE values for these two sets of NACs, because Na^+ and water fill the material's pores, causing no grid points to be located outside $1.3\times$ the vdW radii. Although the Al/Si distribution in natrolite is reasonably well ordered, some structural disorder is inevitable

Table 11. Experimental and Theoretical Natrolite Atomic Charges

	high res. XRD ^a	DDEC/c3 PW91 opt. geometry
net atomic charges		
Si1	1.84 ± 0.12	2.172
Si2	1.65 ± 0.10	2.207
Al	1.51 ± 0.11	2.067
O1	-0.90 ± 0.05	-1.227
O2	-1.21 ± 0.05	-1.318
O3	-1.03 ± 0.05	-1.337
O4	-1.07 ± 0.05	-1.320
O5	-0.87 ± 0.05	-1.113
Na	1.00	1.000
Ow	-0.59 ± 0.03	-0.926
H1	0.24 ± 0.03	0.446
H2	0.36 ± 0.03	0.435
selected bond lengths (Å)		
Ow–H	0.96	0.98, 0.99
Na–Ow	2.37, 2.39	2.34, 2.35
Na–O3	2.37	2.34
Na–O4	2.39	2.37
H1–O1	1.89	1.80
H2–O5	2.06	2.01
Si–O	1.61–1.64	1.62–1.65
Al–O	1.73–1.75	1.74–1.76
Ow–O1	2.84	2.77

^aFrom ref 56.

in any zeolite. Our calculations show a small deviation from the constraint on the net H_2O charges used by Ghermani et al. It is unclear what impact these two factors have on the charges determined in the experimental analysis of Ghermani et al. Because the XRD data of Ghermani et al. were collected at room temperature, vibrational effects may be harder to correct for than in the low temperature formamide example described above.

NACs are most difficult to determine for dense solids with short bond lengths where next nearest neighbors have significant electron density overlaps. We considered V_3Si as an example, because its XRD-derived electron density map and NACs have been previously reported.^{32,33,46} This material is a type II superconductor below ~ 17 K even in the presence of high magnetic fields.^{57–62} V_3Si is representative of A15 compounds having the Weaire–Phelan structure. Kelvin's conjecture that a body centered cubic form has the highest ratio of volume to cell–cell interfacial area was accepted for over 100 years until this ratio was shown to be higher for the Weaire–Phelan structure.^{63,64} The short bond lengths (e.g., 2.64 (V–Si) and 2.36 (V–V) Å) and low atom–atom interfacial area of V_3Si produce strongly overlapping atoms that make charge partitioning difficult. Even the direction of electron transfer, from V to Si or vice versa, has been disputed. The electronegativities of Si (1.90) and V (1.63) predict electron transfer from V to Si to give a positive V charge and a negative Si charge.¹⁶ However, kappa refinement of room temperature XRD data gave net charges of -0.4 (V) and 1.2 (Si), suggesting electron transfer from Si to V.⁴⁶ Since kappa refinement was not performed on the low-temperature superconducting state, the discussion below focuses on the high temperature normal state. Polyhedral volume partitioning using XRD data indicated a total transfer of 1.5 to 2.5 electrons from each Si to three V's.³³ Spectroscopic measurements have

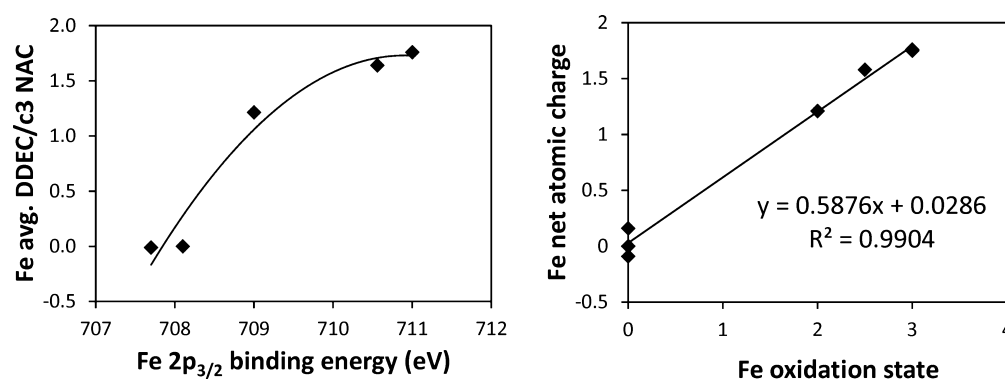


Figure 7. Left: Correlation between 2p_{3/2} binding energy and average Fe net atomic charge for Fe, Fe₂SiO₄, Fe₂O₃, Fe₃O₄, and Fe₃Si. Right: Correlation between Fe oxidation state and Fe net atomic charge for these materials.

Table 12. DDEC/c3 Atomic Population Analysis of Several Ozone Spin States^a

input type ^b	XC theory	molecular properties		DDEC/c3 population analysis							
		S _z	μ	NAC inner	ASM inner	ASM outer	Δμ _{NAC}	Δμ _{dip}	ΔQ _{NAC}	ΔQ _{dip}	ΔQ _{quad}
singlet (<i>q</i> _{net} = 0)											
orbitals	PW91	0	0.253	0.358	0	0	0.189	0.000	1.008	0.776	0.000
orbitals	B3LYP	0	0.273	0.390	0	0	0.201	0.000	0.981	0.749	0.000
orbitals	CCSD	0	0.251	0.364	0	0	0.192	0.000	0.862	0.671	0.000
orbitals	SAC–CI	0	0.270	0.390	0	0	0.201	0.000	0.840	0.647	0.000
density	CAS(12,9)	0	0.215	0.309	0	0	0.177	−0.001	0.836	0.633	0.008
doublet (<i>q</i> _{net} = 1) ^d											
orbitals	PW91	1/2	0.014	0.484	0.202	0.399	0.123	0.000	0.322	0.318	0.000
orbitals	B3LYP	1/2	0.008	0.519	0.217	0.391	0.156	0.000	0.308	0.308	0.000
orbitals	CCSD	1/2	0.014	0.482	0.239	0.381	0.118	0.000	0.325	0.240	0.000
triplet (<i>q</i> _{net} = 0)											
orbitals	PW91	1	0.294	0.206	0.375	0.812	0.046	0.000	0.556	0.681	0.000
orbitals	B3LYP	1	0.290	0.205	0.377	^c	0.041	0.000	0.551	0.690	0.000
orbitals	CCSD	1	0.323	0.230	0.425	0.787	0.054	0.000	0.651	0.601	0.000
orbitals	SAC–CI ^d	0	0.296	0.204	0	0	0.040	0.000	0.671	0.615	0.000
density	CAS(12,9) ^d	0	0.217	0.159	0	0	0.048	0.001	0.511	0.622	0.007

^aAll calculations used a 0.14 bohr grid spacing. ^bInput type was either (i) natural orbitals and occupancies (“orbitals”) or (ii) valence and total density grids (“density”). ^cB3LYP broke symmetry for the ozone triplet, leading to NACs of −0.092 and −0.113 and ASMs of 0.860 and 0.763 for the outer O atoms. ^dCAS and SAC–CI are normally used to find S_z = 0 states. Because the doublet state does not contain a spin eigenstate with S_z = 0, we do not report CAS or SAC–CI results for the +1 cation doublet.

given conflicting results. Kβ-to-Kα X-ray intensity ratios indicated that V₃Si has 0.54 ± 0.25 fewer 3d electrons than pure V metal, suggesting electron transfer from V to Si.⁶⁵ On the other hand, γ-ray Compton scattering profiles indicated a total transfer of 0.4 electrons from each Si to three V's.⁶⁶ The core electron binding energy shift between vanadium in V₃Si and the pure metal was too small to readily measure by X-ray photoelectron spectroscopy (XPS), which indicates that the V net charge is small in magnitude.^{67,68} For this material, DDEC/c3 gives NACs of 0.203 (V) and −0.609 (Si), which is consistent with the electronegativities. Using Bader's method of volume partitioning, we calculated NACs of 0.38 (V) and −1.14 (Si), which are also consistent with the electronegativities. This example suggests that theoretically computed NACs are sometimes more consistent and easier to obtain than experimentally extracted ones.

Core electron binding energy shifts are a way to assess the chemical relevance of computed NACs. Since removal of electrons from an atom raises the electrostatic potential near its nucleus, the core electron binding energy increases as the atom's net charge becomes more positive. This binding energy

is also affected by the potential exerted by other atoms in the system. Specifically,

$$\Delta E_A^{\text{core}} = \left(\frac{\partial E_A^{\text{core}}}{\partial q_A} \right) q_A + V^{\text{ext}}(\vec{R}_A) \quad (85)$$

where E_A^{core} is the binding energy for a core electron (e.g., 1s) for atom A, q_A is the net charge of atom A, and $V^{\text{ext}}(\vec{R}_A)$ is the electrostatic potential exerted at \vec{R}_A due to other atoms in the system.⁶⁹ Figure 7 shows the correlation between DDEC/c3 atomic charge and XPS measured^{70,71} core electron binding energies for the solids Fe, Fe₃O₄, Fe₂SiO₄, FeO, and Fe₃Si. (XPS results for Fe₂SiO₄ were for the fayalite phase.⁷¹) As expected, the Fe NAC increased with increasing 2p_{3/2} binding energy. This correlation shows that the computed NACs are chemically meaningful and exhibit good transferability. As discussed by Bader, AIM transferability embodies the concept that the charge distributions assigned to an atom should be similar in two situations for which the atom's chemical state and environment are similar.^{72,73}

Figure 7 also shows a linear correlation between Fe atomic charge and oxidation state for these same materials. This type of correlation is useful for determining the oxidation states of transition metal atoms at different sites in mixed oxidation state materials. For example, magnetite (Fe_3O_4) is an inverse spinel structure with Fe atoms in tetrahedral and octahedral sites. Tetrahedral Fe has +3 oxidation, and for these we calculated a DDEC/c3 charge of 1.75 to 2. Above $T_V = \sim 120$ K, the octahedral Fe atoms are equivalent with an average oxidation of +2.5. Below $T_V = \sim 120$ K, the symmetry of the unit cell is reduced, and not all octahedral Fe atoms are equivalent.^{74,75} As noted in earlier publications, the DFT+U method can be used to simulate charge ordering transitions in magnetite.^{23,76,77} $U_{\text{eff}} = 0$ simulates the high temperature behavior by allowing Fe d electrons to delocalize,⁷⁶ and for these conditions we find all octahedral Fe atoms are equivalent with a NAC of 1.58, which indicates +2.5 oxidation. Larger U_{eff} values (e.g., 3.2 eV) localize Fe d electrons and separate the octahedral Fe atoms into two groups: (a) those with NAC of ~ 1.5 indicating a +2 oxidation state and (b) those with NAC of ~ 2.0 indicating a +3 oxidation state. Recently, we used the DDEC/c3 method to assign oxidation states of individual Fe atoms in an antiphase domain boundary arising from subsurface charge ordering in the reconstructed magnetite (001) surface.⁷⁸

4.3. Atomic Multipoles and Atomic Spin Moments. In addition to NACs, AIM methods can also compute atomic spin moments (ASMs) for magnetic systems. Previously, we published a method for computing ASMs in periodic and nonperiodic systems with collinear and noncollinear magnetism.²³ This method takes the spin density distribution and $\{w_A(r_A)\}$ as inputs and solves for the atomic spin densities and ASMs.²³ For all systems in that publication, we recomputed the ASMs using $w_A(r_A) = F_A^3(r_A)$ using $\chi_{\text{spin}} = 1/2$ and $r_{\text{cutoff}} = 5$ Å with 100 radial shells. For natural orbital and PAW input types, spin occupancy corrections (see Supporting Information) ensured the ASMs summed to within 10^{-3} bohr magneton of the unit cell's ab initio total spin moment. For all systems, differences between the DDEC/c3 ASMs and the previously published DDEC/c2 results were small, and all of the main conclusions of our earlier report were confirmed by the new results. For the ozone states in Table 12, all DDEC/c3 ASMs differed by less than 0.01 from the previously published DDEC/c2 results. This is not surprising, because spin partitioning is known to be less sensitive than charge partitioning to the choice of atomic population analysis method.^{23,79} For the $\text{Fe}_4\text{O}_{12}\text{N}_4\text{C}_{40}\text{H}_{52}$ single molecule magnet with noncollinear magnetism studied in our earlier report,²³ DDEC/c3 (DDEC/c2) gave ASM magnitudes of 2.325 (2.10) for each Fe atom and ≤ 0.053 (≤ 0.12) for all other atoms. The Fe ASM vector angles differed by $< 0.1^\circ$ between the DDEC/c3 and DDEC/c2 methods. In summary, our previously published method for computing ASMs²³ works reliably with the DDEC/c3 optimized charge distributions.

In regions of space where $\rho(\vec{r}) = 0$, $V(\vec{r})$ can be formally expressed as a distributed multipole expansion.^{4–6} Because the m th order molecular multipole moment can be formed directly from atomic multipoles (AMs) up to order m including AMs up to order m in a force field exactly reproduces the molecular multipole moments up to order m .^{5,6} AMs of any order can be computed as linear combinations of integrals of the form

$$M_A^{l_x l_y l_z} = -\oint (X - X_A)^{l_x} (Y - Y_A)^{l_y} (Z - Z_A)^{l_z} \rho_A(\vec{r}_A) d^3\vec{r}_A \quad (86)$$

(Capital Z_A is the z coordinate of atom A, and lowercase z_A is its atomic number.) Because the time-consuming step of optimizing $\{\rho_A(\vec{r}_A)\}$ is already performed during charge partitioning, the additional computational time required to compute AMs is small. Although these general features of AMs are common to all AIM methods, the occupancy corrections described in this article give significantly improved precision for computing both NACs and AMs. We consider different spin states of ozone as an example. In Table 12, $\Delta\mu_{\text{NAC}}$ is the absolute difference between the ab initio dipole moment (μ) and the dipole moment of the NACs, and $\Delta\mu_{\text{dip}}$ is the absolute difference between μ and the dipole moment of the NACs plus atomic dipoles. The quadrupole moment error was quantified by

$$\Delta Q = \sqrt{(\Delta Q_{x^2-r^2/3})^2 + (\Delta Q_{y^2-r^2/3})^2 + (\Delta Q_{z^2-r^2/3})^2} \quad (87)$$

where $\Delta Q_{x^2-r^2/3}$ is the difference between the ab initio $Q_{x^2-r^2/3}$ output by Gaussian and $Q_{x^2-r^2/3}$ for the DDEC/c3 truncated multipole expansion. Truncating the multipole expansion at $m = 0$ (NACs), $m = 1$ (atomic dipoles), and $m = 2$ (atomic quadrupoles) gives the errors ΔQ_{NAC} , ΔQ_{dip} , and ΔQ_{quad} . Calculations using the natural orbitals and their occupancies as input included occupancy corrections and had negligible integration errors as evidenced by $\Delta\mu_{\text{dip}} = \Delta Q_{\text{quad}} = 0.000$ values. Calculations using the valence and total density grids as input did not include occupancy corrections and had small but non-negligible integration errors as evidenced by the $\Delta\mu_{\text{dip}}$ and ΔQ_{quad} values. This example illustrates the excellent integration precision achieved by using occupancy corrections. This example also shows the effects of XC theory on the NACs, ASMs, and AMs.

An interesting area in which to consider multipole effects is the surfaces of ionic solids. The $\text{SrTiO}_3(100)$ and $\text{NaF}(100)$ surfaces shown in Figure 8 are examples. SrTiO_3 is composed

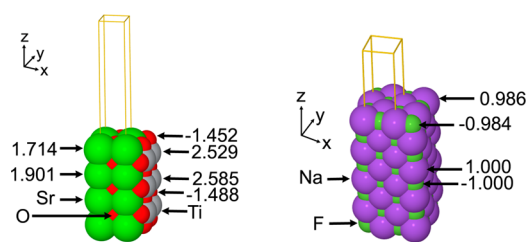


Figure 8. $\text{SrTiO}_3(100)$ and $\text{NaF}(100)$ slabs with coordinate axes and unit cell dimensions displayed. Atomic charges are given for selected atoms.

of alternating SrO and TiO_2 planes. The two oxygen atoms in the Ti plane have opposite $Q_{x^2-y^2}$ values and the same $Q_{2z^2-x^2-y^2}$ values. Table 13 lists results for $\text{SrTiO}_3(100)$ and bulk SrTiO_3 . For the bulk material, atomic dipoles are zero for all atoms, and only oxygen atoms have nonzero quadrupoles. The electron density map of bulk SrTiO_3 measured by high resolution γ -ray diffraction also shows this.⁸⁰ Unfortunately, reliable NACs could not be extracted from the experiments, because reflections for neutral Sr and Sr^{2+} were not distinguishable.⁸⁰ Our calculations for $\text{SrTiO}_3(100)$ show that atomic dipoles are induced near the surface. In addition, the NAC for surface Sr

Table 13. NACs and AMs (in Atomic Units) for SrTiO₃(100) and Bulk SrTiO₃^a

layer	atom	NAC	μ_z	$Q_{x^2-y^2}$	$Q_{2z^2-x^2-y^2}$
1	O	-1.452	0.172	0.000	-0.057
1	Sr	1.714	-0.151	0.000	0.010
2	O	-1.496	-0.033	± 0.221	0.156
2	Ti	2.529	-0.002	0.000	-0.091
3	O	-1.488	0.016	0.000	-0.387
3	Sr	1.901	-0.018	0.000	0.005
4	O	-1.505	0.001	± 0.218	0.191
4	Ti	2.585	-0.003	0.000	-0.012
bulk	O (Sr)	-1.494	0.000	0.000	-0.421
bulk	O (Ti)	-1.494	0.000	± 0.211	0.211
bulk	Sr	1.901	0.000	0.000	0.000
bulk	Ti	2.582	0.000	0.000	0.000

^aFor all atoms, $\mu_x = \mu_y = Q_{xy} = Q_{xz} = Q_{yz} = 0.000$.

was 1.714 compared to 1.901 for the bulk-like layers. This reduced the surface layer charge ($1.714 - 1.452 = 0.262$) compared to the bulk material's SrO layer charge ($1.901 - 1.494 = 0.407$). Thus, NAC changes reduced the surface's polarity. A Sr atom in bulk-like layers is surrounded by negatively charged oxygen atoms, while a Sr atom on the surface does not have any oxygen atoms above it. Because the Sr net charge is stabilized less at the material's surface layer than in the bulk-like layers, the Sr charge magnitude is lower for the surface layer.

As shown in Figure 8, similar calculations were performed for the NaF(100) surface. NACs for surface (bulk) atoms were 0.986 (1.000) for Na and -0.984 (-1.000) for F. While all atomic dipoles and quadrupoles were zero for bulk NaF, the surface Na (F) atoms had $\mu_z = -0.001$ (0.040) and $Q_{2z^2-x^2-y^2} = 0.009$ (0.085). Once again, we see lower charge magnitudes for the surface layer than for the bulk-like layers. However, the surface induced atomic population changes were much smaller for NaF(100) than for SrTiO₃(100). Because F has the highest electronegativity of any element on the Pauling scale,¹⁶ it retains ionic character more readily than many other elements.

4.4. Comparison of Different Charge Assignment Methods. In this section, we compare the DDEC/c3 method to several methods for assigning charges from quantum chemistry calculations. We do not consider the Mulliken method here, because of its explicit dependence on the choice of basis set used to generate $\rho(\vec{r})$.¹⁷ Because the Bader, HD, IH, ISA, ESP, and DDEC methods are functionals of $\rho(\vec{r})$, they are

formally independent of the choice of basis set used to generate $\rho(\vec{r})$. The NPA method also has low sensitivity to the basis set choice.¹⁷ While different charge assignment methods may be suitable for different purposes, NACs are commonly used to describe (i) the chemical states of atoms in materials and (ii) the electrostatic potential $V(\vec{r})$ surrounding a material. Both of these properties are important for constructing transferable force fields in which a single set of parameters can describe a variety of geometric conformations.

Force fields used in molecular dynamics, Monte Carlo, and other types of atomistic simulations typically contain several types of interactions: electrostatic interactions due to the atomic charges and (optionally) atomic multipoles; van der Waals interactions; and bond, angle, and torsion force parameters to describe the material's flexibility.^{81–84} In some cases, the individual chemical species have fixed geometries so the electrostatic potential grid from an ab initio calculation can be directly used as input to compute the electrostatic interactions during subsequent atomistic simulations.¹⁰ In flexible force fields, the conformation of an individual chemical species changes during the atomistic simulation. NACs are commonly used in flexible force fields to describe electrostatic properties. Constructing a flexible force field is easier when NACs are not very sensitive to the molecular conformation.

In this section, we quantify the suitability of different sets of NACs for constructing flexible force fields by computing RMSE and RRMSE values for several geometric conformations of 4-X-substituted bicyclo[2,2,2]octane-1-carboxylic acids, with X = -H, -Br, -CN, -OH, and -CO₂C₂H₅. We chose these examples because a prior experimental investigation allows us to independently assess the chemical relevance of the assigned NACs. Specifically, Roberts and Moreland found that the apparent ionization constants of these acids is experimentally correlated to their chemical reactivity.⁸⁵ Because these ionization constants followed the same trend as the analogously substituted acetic acids X-CH₂CO₂H, Roberts and Moreland inferred that they correlate to the electron withdrawing ability of the substituent.⁸⁵ Hammett parameters derived from the rate or equilibrium constants of substituted benzene derivatives were the original substituent constants in organic chemistry.⁸⁶ Roberts and Moreland used the saturated bicyclo[2,2,2]octane ring instead of benzene to isolate the intrinsic electron withdrawing ability of the substituent from any resonance effects that might occur in the aromatic benzene derivatives. Their modified substituent constants, $\sigma' = 0.000$ (-H), 0.283 (-OH), 0.297 (-CO₂C₂H₅), 0.454 (-Br), and 0.579 (-CN),

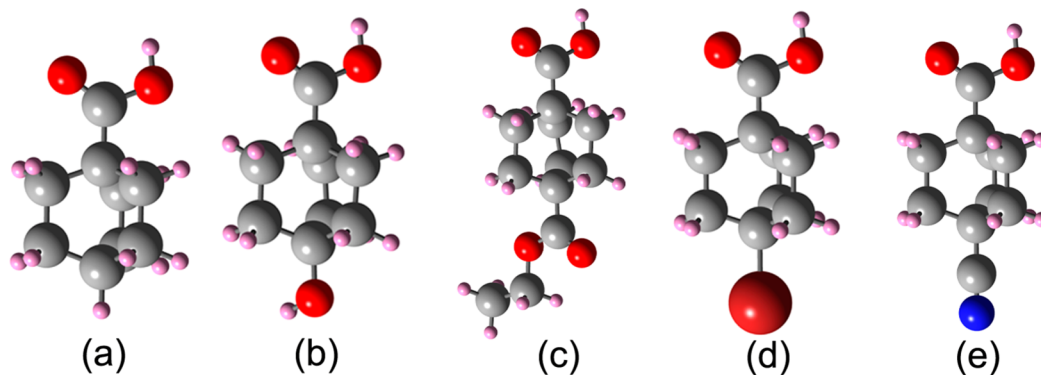


Figure 9. Lowest energy conformations (B3LYP/6-311++G** level of theory) of 4-X-substituted bicyclo[2,2,2]octane-1-carboxylic acids: X = (a) -H, (b) -OH, (c) -CO₂C₂H₅, (d) -Br, and (e) -CN.

Table 14. Fragment Charges for the Lowest Energy Conformation

X	σ'^a	substituent net charge						weighted sum of eq 88					
		DDEC	ESP	HD	IH	ISA	NPA	DDEC	ESP	HD	IH	ISA	NPA
H	0.000	0.04	−0.02	0.03	0.07	−0.01	0.21	−0.03	−0.01	−0.01	−0.06	0.02	−0.20
OH	0.283	−0.25	−0.33	−0.08	−0.25	−0.31	−0.29	−0.12	−0.10	−0.05	−0.15	−0.09	−0.34
CO ₂ C ₂ H ₅	0.297	−0.02	−0.04	−0.03	0.06	−0.07	0.02	−0.08	−0.02	−0.04	−0.09	−0.04	−0.29
Br	0.454	−0.25	−0.19	−0.10	−0.11	−0.29	−0.02	−0.16	−0.18	−0.07	−0.13	−0.15	−0.30
CN	0.579	−0.16	−0.31	−0.16	−0.12	−0.25	−0.02	−0.14	−0.10	−0.11	−0.16	−0.10	−0.32
corr. coeff. (R ²)		0.44	0.44	0.92	0.26	0.51	0.17	0.81	0.47	0.93	0.69	0.71	0.59

^aFrom ref 85.

are considered one of the most accurate estimates of the relative electron withdrawing ability of these substituents in organic chemistry.⁸⁵

First, we computed the lowest energy conformation for each molecule at the B3LYP/6-311++G** level of theory, as shown in Figure 9. To do this, we computed energies for several different conformations of each molecule and selected the one with lowest energy. We have not ruled out the possibility that a lower energy conformation might exist. We then computed the net charges of the substituent group for each molecule using the DDEC, ESP, HD, IH, ISA, and NPA methods, as shown in Table 14. For each charge method, the correlation coefficients (R²) for the line $q_{\text{net}} = a_0 + a_1\sigma'$ were determined from linear regression. As shown in Table 14, the σ' values were strongly correlated to the substituent net charges computed by the HD method, but not those of the other charge methods. The primary reason for this is that a negatively charged atom will repel electrons from adjacent atoms, so that the effect of the substituent's charge propagates through the molecule's bonds. Therefore, it is best to perform a weighted sum of the net charges of all atoms in the molecule, where atoms in the substituent group receive a weight of one and the remaining atoms receive a weight less than one that decreases as the atom is further removed from the substituent group. For example, we can use the weighted sum

$$\sum_A (0.75)^{N_{\text{bonds}}} q_A \quad (88)$$

where q_A is the NAC for atom A and N_{bonds} is the number of bonds in the shortest chain connecting the atom to the substituent group. As shown in Table 14, the σ' values are more closely correlated to this weighted sum than to the substituent net charges. The best correlations were obtained for the HD and DDEC/c3 methods. Additional tests (not shown) using values between 0.4 and 0.95 instead of 0.75 in eq 88 did not change the ordering for the DDEC/c3 method (H < CO₂C₂H₅ < OH < CN < Br), which differed from the experimental ordering (H < OH < CO₂C₂H₅ < Br < CN). These results show that the DDEC/c3 charges accurately described large changes (but not small changes) in σ' .

To assess the suitability of these charge assignment methods for constructing flexible force fields, we first optimized the geometries of several stationary points on the B3LYP/6-311++G** potential energy surface. This was accomplished by performing 180° rotations of selected bonds followed by geometry optimization until the forces on every atom were zero. This corresponds to two angular positions for each intramolecular rotational degree of freedom. The two rotational degrees of freedom of the −COOH group gave at least four conformations for every molecule. We considered one (two) additional rotational degrees of freedom within the X = −OH

(−CO₂C₂H₅) substituent, for a total of eight (16) representative conformations. Other stationary points may exist in addition to those we considered.

Table 15 lists the mean absolute deviation (MAD) of the NACs over the set of conformations considered. A small MAD

Table 15. Assessment of the Conformational Transferability of Different Charge Assignment Methods

substituent:	mean absolute deviation of NACs				
	H	Br	CN	OH	ester
conformations considered:	4	4	4	8	16
DDEC/c3	0.007	0.007	0.007	0.010	0.008
ESP	0.038	0.057	0.051	0.045	0.041
HD	0.002	0.002	0.002	0.003	0.002
IH	0.006	0.006	0.006	0.007	0.007
ISA	0.015	0.015	0.016	0.025	0.016
NPA	0.004	0.004	0.004	0.006	0.006

occurs when NACs are similar for different geometric conformations. Examining Table 15, the HD and NPA methods gave the smallest MADs, while the ESP and ISA methods gave the largest MADs. The IH and DDEC/c3 methods gave intermediate MADs. After computing the NACs, the RMSE and RRMSE values were calculated for each conformation using the method described in section 4.1. Table 16 lists the average of these values for each substituent group. Three different sets of NACs were considered: NACs optimized separately for each representative conformation, the average NACs of the representative conformations, and using just the NACs of the low energy conformation. The results showed clear and expected trends. Because the ESP NACs mathematically minimize the RMSE over a chosen set of grid points,¹⁸ ESP gave the lowest RMSE values of all charge assignment methods when NACs were separately optimized for each conformation. Achieving the lowest possible RMSE for each conformation and the lowest possible MAD of NACs between conformations are inherently antagonistic constraints. Consequently, it is not surprising that the ESP method also gave the highest MAD of NACs between conformations. At the other extreme, the Hirshfeld method gave the smallest MAD of NACs and the largest RMSE values. Both of these extremes are undesirable for the construction of flexible force fields.

The DDEC/c3 method tries to achieve a compromise between low RMSE values for each conformation and low MAD of NACs between conformations. Examining Table 16, the DDEC/c3 method gave the second or third lowest RMSE values when NACs were optimized separately for each conformation. When conformation averaged NACs were used, only the ESP method gave lower RMSE values than the

Table 16. Average RMSE and RRMSE Values for Various Charge Assignment Methods (the Best Values Shown in Boldface Type)

substituent:	avg. RMSE (kcal/mol)					avg. RRMSE				
	H	Br	CN	OH	ester	H	Br	CN	OH	ester
NACs optimized separately for each conformation										
DDEC/c3	0.81	1.15	0.87	0.89	0.77	0.13	0.18	0.10	0.13	0.10
ESP	0.49	0.93	0.38	0.48	0.42	0.07	0.14	0.04	0.07	0.04
HD	2.85	3.26	3.67	3.73	3.27	0.41	0.48	0.40	0.51	0.41
IH	1.12	2.49	1.60	1.35	1.05	0.18	0.38	0.18	0.20	0.14
ISA	0.73	1.45	0.74	0.71	0.70	0.11	0.22	0.08	0.10	0.09
NPA	1.71	3.23	2.56	1.94	2.98	0.25	0.49	0.29	0.27	0.31
conformation averaged NACs										
DDEC/c3	1.27	1.48	1.29	1.40	1.25	0.18	0.22	0.14	0.19	0.16
ESP	1.10	1.36	1.00	1.37	1.38	0.15	0.20	0.11	0.19	0.14
HD	2.88	3.31	3.70	3.71	3.31	0.41	0.50	0.41	0.50	0.42
IH	1.48	2.65	1.84	1.75	1.39	0.23	0.41	0.21	0.25	0.18
ISA	1.33	1.81	1.31	1.57	1.43	0.18	0.26	0.14	0.21	0.18
NPA	2.12	3.47	2.86	2.42	3.71	0.30	0.52	0.32	0.33	0.38
all conformations use NACs from low energy conformation										
DDEC/c3	1.39	1.61	1.38	1.73	1.44	0.19	0.23	0.15	0.24	0.18
ESP	1.49	1.73	1.26	2.12	1.91	0.19	0.24	0.13	0.28	0.20
HD	2.97	3.24	3.68	3.74	3.31	0.42	0.48	0.41	0.51	0.42
IH	1.55	2.61	1.85	1.98	1.48	0.23	0.40	0.21	0.28	0.19
ISA	1.46	1.97	1.41	1.98	1.74	0.19	0.29	0.15	0.27	0.21
NPA	2.23	3.50	2.93	2.57	4.05	0.31	0.52	0.32	0.35	0.42

DDEC/c3 method. Finally, when NACs from the low energy conformation were used, the DDEC/c3 method usually gave lower RMSE values than all of the other charge assignment methods. These characteristics are favorable for the construction of flexible force fields.

As a further test, ab initio molecular dynamics was performed in VASP using the PBE functional with Grimme's D2 dispersion correction to account for long-range interactions.^{87,88} The OH substituted molecule was placed in the center of a $15 \times 15 \times 15$ Å cubic unit cell and simulated for 500 steps at 300 K (Nosé thermostat) with a step size of 1 fs. This short trajectory generated a set of geometries for which the atomic forces were not necessarily zero, as typically occurs in atomistic simulations. After these geometries were generated, single point calculations were performed on every 20th step using the B3LYP/6-311++G** method in Gaussian 09. This allowed the RMSE values to be computed using the same NACs as Tables 15 and 16. As shown in Table 17, the DDEC/c3 and IH methods gave essentially identical results and performed better than the other methods when the low energy conformation NACs were used. When the conformation averaged NACs were used, the DDEC/c3, IH, and ESP methods gave essentially identical results and performed better

Table 17. Average RMSE and RRMSE Values for Geometries Generated Using ab Initio Molecular Dynamics (the Best Values Are Shown in Boldface Type)

	DDEC/c3	ESP	HD	IH	ISA	NPA
using the low energy conformation NACs of Table 16						
RMSE (kcal/mol)	2.51	3.23	4.38	2.55	3.04	3.65
RRMSE	0.27	0.35	0.47	0.27	0.33	0.39
using the conformation averaged NACs of Table 16						
RMSE (kcal/mol)	2.08	2.11	4.31	2.11	2.44	3.39
RRMSE	0.23	0.23	0.47	0.23	0.26	0.37

than the other methods. These results confirm the suitability of DDEC/c3 NACs for constructing flexible force fields.

Previous studies have examined the suitability of the IH method for reproducing the electrostatic potential. Van Damme et al. computed RMSE values for a large set of rigid molecules and found that the IH method gives lower values than the HD, NPA, and Mulliken methods.⁸⁹ We previously computed RMSE values for a variety of rigid periodic and nonperiodic systems and found that the IH method gives lower values than the HD, NPA, Mulliken, and Bader methods but higher values than the DDEC and ISA methods.⁴ We found that the IH method overestimates charge magnitudes for mostly ionic systems, because it uses gas phase reference densities computed without charge compensation.⁴ Verstraelen et al. computed IH charges for periodic silicates and reported this same problem.⁹⁰ In another study, Verstraelen et al. compared the performance of the Mulliken, NPA, IH, restrained electrostatic potential (RESP), electronegativity equalization, and split charge equalization NACs for modeling the flexibility of the protein fragment penta-alanine.⁹¹ For these methods, the IH NACs gave "the best compromise between transferability (robustness with respect to conformation changes) and the ability to reproduce electrostatic properties of penta-alanine."⁹¹ These findings are consistent with our results above. In addition, our results indicate that the DDEC/c3 and IH methods have similar performances for modeling the flexibility of molecular systems during atomistic simulations.

The DDEC/c3 method gives excellent performance for a wider range of periodic and nonperiodic materials than some other charge assignment methods. The NPA method is not yet available for periodic materials. Bader's method is applicable to both periodic and nonperiodic materials but gives NACs that do not accurately reproduce $V(\vec{r})$ surrounding a material.⁴ Thus, while Bader NACs are chemically meaningful, they are not optimal for constructing force fields of porous materials. The Repeating Electrostatic Potential Extracted ATomic

(REPEAT) method⁴⁵ finds NACs that minimize the RMSE in porous periodic materials. However, the ESP,¹⁸ RESP,⁹² and REPEAT⁴⁵ methods cannot be applied to nonporous materials. The ISA method also gives poor results for nonporous materials, because it does not include any constraints or reference densities to force atom tails to decay sufficiently rapidly.⁴ The HD method systematically underestimates charge magnitudes in both molecular and periodic systems.^{4,7} Application of the IH method to periodic materials would require the use of charge compensated reference densities, because ions like O^{2-} are stable in solids but not in isolation. In contrast to the DDEC/c3 method, the IH method does not condition the reference densities to account for expansion or contraction of the same type of ion in different materials. This raises the question as to whether it would be possible to obtain accurate results with the IH method using the same set of reference densities for a wide variety of molecular and solid systems. Applications of the IH method to periodic materials have been only partly successful, because of its tendency to overestimate atomic charge magnitudes for ionic materials.^{4,90} The DDEC/c3 method is a practical approach that gives excellent results through the process of conditioning charge compensated reference densities to fit the particular material of interest.

5. CONCLUSIONS

We have shown that all atoms-in-molecules methods that partition $\rho(\vec{r})$ using spherically symmetric atomic weights, $\{w_A(r_A)\}$, converge to a global minimum of an optimization functional, Ω . We also showed that global minimization of Ω is equivalent to finding stationary points of a path action, S . By expanding the curvature d^2S , we explained why convergence is most challenging for dense solids with short bond lengths and significant density overlaps between nearby atoms. Several improvements to the DDEC method were then introduced to facilitate convergence robustness by increasing d^2S . Specifically, we showed that convergence is improved by (a) using conditioned reference densities, $\{Y_A^{avg}(r_A, n_A)\}$, that are spherical averages of functions that exactly sum to $\rho(\vec{r})$ and (b) constraining $w_A(r_A)$ to decay approximately exponentially in the buried tail region. This charge partitioning functional's parameters were derived from first principles. To satisfy chemical reasonableness, the number of valence electrons assigned to each atom is constrained to be non-negative. These improvements gave a charge partitioning method with excellent chemical accuracy, computational efficiency, and convergence robustness.

We used this method, called DDEC/c3, to study a variety of periodic and nonperiodic materials: (a) 20 charged and neutral molecules, (b) eight porous materials, (c) 19 nonporous bulk solids, and (d) five slabs and systems with one or two periodic dimensions. Fifteen of these materials had collinear magnetism, one had noncollinear magnetism, and the remaining were nonmagnetic. These materials contained 31 different chemical elements. The DDEC/c3 method gave reliable net atomic charges, atomic spin moments, and atomic multipoles for all of the systems studied. For a number of materials, comparisons were made to available experimental data. For several iron-containing materials, DDEC/c3 NACs were correlated to core electron binding energy shifts measured by XPS. For formamide and natrolite, DDEC/c3 NACs were in reasonable agreement with NACs extracted from XRD using spherical pseudoatom refinement. Our method was also used to study

the image charge produced by adsorption of a K^+ cation on a Mo_2C (110) surface.

For molecules and porous periodic materials, the DDEC/c3 NACs had reasonable accuracy for reproducing $V(\vec{r})$ outside the electron distribution. The conformational transferability of DDEC/c3 NACs was studied for a series of substituted bicyclo[2,2,2]octane-1-carboxylic acids. The DDEC/c3 NACs exhibited a compromise between the accuracy of fitting $V(\vec{r})$ for an individual conformation and transferability of the NACs between conformations. When NACs from the low energy conformation were used to construct a force field, the average RMSE value for a set of geometric conformations generated by either force minimization or ab initio molecular dynamics was lower for the DDEC/c3 method than for the ESP, HD, IH, ISA, and NPA methods. However, RMSE differences between the DDEC/c3, ESP, ISA, and IH methods were small. These characteristics are especially attractive for the development of force fields for modeling flexible materials.

For selected systems, tests showed that the DDEC/c3 method works with a variety of exchange-correlation theories: DFT, DFT+U, coupled cluster methods like CCSD and SAC-CI, and multireference methods like CAS-SCF. Irrespective of whether the ab initio calculation used explicitly relaxed core electrons (all-electron), frozen core electrons (e.g., PAW), or an effective core potential (pseudopotential), all core and valence electrons are automatically included in DDEC charge partitioning. The DDEC/c3 method has excellent numerical precision and is not unduly sensitive to the choice of basis set or grid points. It can be readily applied to systems containing a large number of atoms in the unit cell. All of these attributes make the DDEC/c3 method well suited for atomic population analysis of a wide variety of periodic and nonperiodic materials, especially for obtaining net atomic charges that simultaneously reproduce chemical states of atoms in a material and $V(\vec{r})$ outside its electron distribution.

■ ASSOCIATED CONTENT

📄 Supporting Information

Additional computational details and .xyz files containing the atomic coordinates, net atomic charges, atomic dipoles and quadrupoles, atomic spin moments, and fitted decay exponents. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: thomasamanz@gmail.com, david.sholl@chbe.gatech.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This material is based on work supported as part of the Center for Atomic Level Catalyst Design, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001058 (Center for Atomic Level Catalyst Design). Supercomputing resources were provided by the Georgia Institute of Technology and the Extreme Science and Engineering Discovery Environment (XSEDE project grant TG-CTS100027). XSEDE is funded by NSF grant OCI-1053575. We thank Nita Chandrasekhar for providing the DFT

optimized geometries for H adsorbed in Pd, Pd₃In, Pd₃Hf, and Pd₃V. We thank Dr. Richard Wheatley for computing the ISA NACs of formamide for Becke-style integration grids, as shown in Table 6.

REFERENCES

- (1) Bader, R. F. W. *J. Phys. Chem. A* **2007**, *111*, 7966–7972.
- (2) Bader, R. F. W.; Macdougall, P. J.; Lau, C. D. H. *J. Am. Chem. Soc.* **1984**, *106*, 1594–1605.
- (3) Cao, W. L.; Gatti, C.; Macdougall, P. J.; Bader, R. F. W. *Chem. Phys. Lett.* **1987**, *141*, 380–385.
- (4) Manz, T. A.; Sholl, D. S. *J. Chem. Theory Comput.* **2010**, *6*, 2455–2468.
- (5) Laidig, K. E. *J. Phys. Chem.* **1993**, *97*, 12760–12767.
- (6) Kosov, D. S.; Popelier, P. L. A. *J. Phys. Chem. A* **2000**, *104*, 7339–7345.
- (7) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.
- (8) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129–138.
- (9) Lillestolen, T. C.; Wheatley, R. J. *J. Chem. Commun.* **2008**, 5909–5911.
- (10) Watanabe, T.; Manz, T. A.; Sholl, D. S. *J. Phys. Chem. C* **2011**, *115*, 4824–4836.
- (11) Lillestolen, T. C.; Wheatley, R. J. *J. Chem. Phys.* **2009**, *131*, 144101.
- (12) Erucar, I.; Keskin, S. *J. Phys. Chem. C* **2011**, *115*, 13637–13644.
- (13) Haldoupis, E.; Nair, S.; Sholl, D. S. *J. Am. Chem. Soc.* **2012**, *134*, 4313–4323.
- (14) Zheng, B.; Sant, M.; Demontis, P.; Suffritti, G. B. *J. Phys. Chem. C* **2012**, *116*, 933–938.
- (15) Fang, H.; Kamakoti, P.; Zang, J.; Cundy, S.; Paur, C.; Ravikovitch, P. I.; Sholl, D. S. *J. Phys. Chem. C* **2012**, *116*, 10692–10701.
- (16) Allred, A. L. *J. Inorg. Nucl. Chem.* **1961**, *17*, 215–221.
- (17) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.
- (18) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.
- (19) Bultinck, P.; Cooper, D. L.; Van Neck, D. *J. Phys. Chem. Chem. Phys.* **2009**, *11*, 3424–3429.
- (20) Nalewajski, R. F.; Parr, R. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8879–8882.
- (21) Bultinck, P.; Ayers, P. W.; Fias, S.; Tiels, K.; Van Alsenoy, C. *Chem. Phys. Lett.* **2007**, *444*, 205–208.
- (22) Ghillemijn, D.; Bultinck, P.; Van Neck, D.; Ayers, P. W. *J. Comput. Chem.* **2011**, *32*, 1561–1567.
- (23) Manz, T. A.; Sholl, D. S. *J. Chem. Theory Comput.* **2011**, *7*, 4146–4164.
- (24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakasugi, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakamijima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Ciolowski, J.; Fox, D. J. *Gaussian 09*, Gaussian, Inc.: Wallingford, CT, 2009.
- (25) Hafner, J. *J. Comput. Chem.* **2008**, *29*, 2044–2078.
- (26) Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169–11186.
- (27) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758–1775.
- (28) Loiseau, T.; Serre, C.; Huguenard, C.; Fink, G.; Taulelle, F.; Henry, M.; Bataille, T.; Ferey, G. *Chem.—Eur. J.* **2004**, *10*, 1373–1382.
- (29) Wyckoff, R. W. G. *Crystal Structures*; John Wiley & Sons: New York, 1963; Vol. 1, pp 85–237.
- (30) Li, H.; Eddaoudi, M.; O’Keeffe, M.; Yaghi, O. M. *Nature* **1999**, *402*, 276–279.
- (31) Zuxiang, Y. *Rock Miner. Anal.* **1984**, *3*, 231–238.
- (32) Staudenmann, J. L. *Solid State Commun.* **1978**, *26*, 461–468.
- (33) Staudenmann, J. L.; Coppens, P.; Muller, J. *Solid State Commun.* **1976**, *19*, 29–33.
- (34) Parthe, E.; Sadagopan, V. *Acta Crystallogr.* **1963**, *16*, 202–205.
- (35) Fries, R. J.; Kempter, C. P. *Anal. Chem.* **1960**, *32*, 1898–1898.
- (36) Watanabe, T.; Sholl, D. S. *J. Chem. Phys.* **2010**, *133*, 094509.
- (37) Han, J. W.; Li, L. W.; Sholl, D. S. *J. Phys. Chem. C* **2011**, *115*, 6870–6876.
- (38) Henkelman, G.; Arnaldsson, A.; Jonsson, H. *Comput. Mater. Sci.* **2006**, *36*, 354–360.
- (39) Sanville, E.; Kenny, S. D.; Smith, R.; Henkelman, G. *J. Comput. Chem.* **2007**, *28*, 899–908.
- (40) Tang, W.; Sanville, E.; Henkelman, G. *J. Phys.: Condens. Matter* **2009**, *21*, 084204.
- (41) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953.
- (42) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547–2553.
- (43) Elking, D. M.; Perera, L.; Pedersen, L. G. *Comput. Phys. Commun.* **2012**, *183*, 390–397.
- (44) Chen, D. L.; Stern, A. C.; Space, B.; Johnson, J. K. *J. Phys. Chem. A* **2010**, *114*, 10225–10233.
- (45) Campana, C.; Mussard, B.; Woo, T. K. *J. Chem. Theory Comput.* **2009**, *5*, 2866–2878.
- (46) Coppens, P.; Gururow, T. N.; Leung, P.; Stevens, E. D.; Becker, P. J.; Yang, Y. W. *Acta Crystallogr., Sect. A* **1979**, *35*, 63–72.
- (47) Zuo, J. M.; O’Keeffe, M.; Rez, P.; Spence, J. C. H. *Phys. Rev. Lett.* **1997**, *78*, 4777–4780.
- (48) Stevens, E. D. *Acta Crystallogr., Sect. B* **1978**, *34*, 544–551.
- (49) Hansen, N. K.; Coppens, P. *Acta Crystallogr., Sect. A* **1978**, *34*, 909–921.
- (50) Stevens, E. D.; Rys, J.; Coppens, P. *J. Am. Chem. Soc.* **1978**, *100*, 2324–2328.
- (51) Abramov, Y. A.; Volkov, A. V.; Coppens, P. *Chem. Phys. Lett.* **1999**, *311*, 81–86.
- (52) Kuntzinger, S.; Ghermani, N. E. *Acta Crystallogr., Sect. B* **1999**, *55*, 273–284.
- (53) Su, Z. W.; Coppens, P. *Acta Crystallogr., Sect. A* **1995**, *51*, 27–32.
- (54) Alberti, A.; Vezzalini, G. *Acta Crystallogr., Sect. B* **1981**, *37*, 781–788.
- (55) Kirfel, A.; Orthen, M.; Will, G. *Zeolites* **1984**, *4*, 140–146.
- (56) Ghermani, N. E.; Lecomte, C.; Dusauroy, Y. *Phys. Rev. B* **1996**, *53*, 5231–5239.
- (57) Dew-Hughes, D. *Cryogenics* **1975**, *15*, 435–454.
- (58) Muto, Y.; Toyota, N.; Noto, K.; Akutsu, K.; Isino, M.; Fukase, T. *J. Low Temp. Phys.* **1979**, *34*, 617–640.
- (59) Testardi, L. R. *Rev. Mod. Phys.* **1975**, *47*, 637–648.
- (60) Perucchi, A.; Nicoletti, D.; Ortolani, M.; Marini, C.; Sopracase, R.; Lupi, S.; Schade, U.; Putti, M.; Pallecchi, I.; Tarantini, C.; Ferretti, M.; Ferdeghini, C.; Monni, M.; Bernardini, F.; Massidda, S.; Dore, P. *Phys. Rev. B* **2010**, *81*, 092509.
- (61) Mukherjee, S.; Agterberg, D. F. *Phys. Rev. B* **2011**, *84*, 134520.
- (62) Reinert, F.; Nicolay, G.; Eltner, B.; Ehm, D.; Schmidt, S.; Hufner, S.; Probst, U.; Bucher, E. *Phys. Rev. Lett.* **2000**, *85*, 3930–3933.
- (63) Thomson, W. *Philos. Mag.* **1887**, *24*, 503.
- (64) Weaire, D.; Phelan, R. *Philos. Mag. Lett.* **1994**, *69*, 107–110.
- (65) Raj, S.; Padhi, H. C.; Polasik, M.; Basa, D. K. *Solid State Commun.* **1999**, *110*, 275–279.
- (66) Sharma, B. K.; Manninen, S.; Paakkari, T.; Richardson, M. W.; Rundqvist, S. *Philos. Mag. B* **1984**, *49*, 363–370.
- (67) Bisi, O.; Chiao, L. W. *Phys. Rev. B* **1982**, *25*, 4943–4948.

- (68) Werfel, F.; Drager, G.; Brummer, O.; Jurisch, M. *Phys. Status Solidi B* **1977**, *80*, K95–K99.
- (69) Jolly, W. L.; Perry, W. B. *J. Am. Chem. Soc.* **1973**, *95*, 5442–5450.
- (70) Shabanova, I. N.; Trapeznikov, V. A. *J. Electron Spectrosc. Relat. Phenom.* **1975**, *6*, 297–307.
- (71) Yamashita, T.; Hayes, P. *Appl. Surf. Sci.* **2008**, *254*, 2441–2449.
- (72) Bader, R. F. W. *Acc. Chem. Res.* **1975**, *8*, 34–40.
- (73) Bader, R. F. W. *Chem. Rev.* **1991**, *91*, 893–928.
- (74) Blasco, J.; Garcia, J.; Subias, G. *Phys. Rev. B* **2011**, *83*, 104105.
- (75) Senn, M. S.; Wright, J. P.; Attfield, J. P. *Nature* **2012**, *481*, 173–176.
- (76) Wenzel, M. J.; Steinle-Neumann, G. *Phys. Rev. B* **2007**, *75*, 214430.
- (77) Lodziana, Z. *Phys. Rev. Lett.* **2007**, *99*, 206402.
- (78) Parkinson, G. S.; Manz, T. A.; Novotny, Z.; Sprunger, P. T.; Kurtz, R. L.; Schmid, M.; Sholl, D. S.; Diebold, U. *Phys. Rev. B* **2012**, *85*, 195450.
- (79) Ruiz, E.; Cirera, J.; Alvarez, S. *Coord. Chem. Rev.* **2005**, *249*, 2649–2660.
- (80) Jauch, W.; Reehuis, M. *Acta Cryst. A* **2005**, *61*, 411–417.
- (81) Hill, J. R.; Freeman, C. M.; Subramanian, L. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2000; Vol. 16, pp 141–216.
- (82) Kaminski, G.; Jorgensen, W. L. *J. Phys. Chem.* **1996**, *100*, 18010–18013.
- (83) Shan, T. R.; Devine, B. D.; Hawkins, J. M.; Asthagiri, A.; Phillpot, S. R.; Sinnott, S. B. *Phys. Rev. B* **2010**, *82*, 235302.
- (84) Ponder, J. W.; Case, D. A. In *Protein Simulations*; Daggett, V., Ed.; Elsevier: Amsterdam, 2003; Vol. 66, pp 27–85.
- (85) Roberts, J. D.; Moreland, W. T. *J. Am. Chem. Soc.* **1953**, *75*, 2167–2173.
- (86) Hammett, L. P. *J. Am. Chem. Soc.* **1937**, *59*, 96–103.
- (87) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (88) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (89) Van Damme, S.; Bultinck, P.; Fias, S. *J. Chem. Theory Comput.* **2009**, *5*, 334–340.
- (90) Verstraelen, T.; Sukhomlinov, S. V.; Van Speybroeck, V.; Waroquier, M.; Smirnov, K. S. *J. Phys. Chem. C* **2012**, *116*, 490–504.
- (91) Verstraelen, T.; Pauwels, E.; De Proft, F.; Van Speybroeck, V.; Geerlings, P.; Waroquier, M. *J. Chem. Theory Comput.* **2012**, *8*, 661–676.
- (92) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.