

# Classical Pauli repulsion: An anisotropic, atomic multipole model

Cite as: *J. Chem. Phys.* **150**, 084104 (2019); doi: [10.1063/1.5081060](https://doi.org/10.1063/1.5081060)

Submitted: 13 November 2018 • Accepted: 23 January 2019 •

Published Online: 22 February 2019

Joshua A. Rackers<sup>1</sup>  and Jay W. Ponder<sup>2,a)</sup> 

## AFFILIATIONS

<sup>1</sup>Program in Computational and Molecular Biophysics, Washington University, School of Medicine, Saint Louis, Missouri 63110, USA

<sup>2</sup>Department of Chemistry, Washington University in Saint Louis, Saint Louis, Missouri 63130, USA

<sup>a)</sup>Author to whom correspondence should be addressed: [ponder@dasher.wustl.edu](mailto:ponder@dasher.wustl.edu)

## ABSTRACT

Pauli repulsion is a key component of any theory of intermolecular interactions. Although Pauli or exchange repulsion has its origin in the quantum mechanical nature of electrons, it is possible to describe the resulting energetic effects via a classical model in terms of the overlap of electron densities. In fact, closed shell intermolecular repulsion can be explained as a diminution of electron density in the internuclear region resulting in decreased screening of nuclear charges and increased nuclear-nuclear repulsion. We provide a concise anisotropic repulsion formulation using the atomic multipoles from the Atomic Multipole Optimized Energetics for Biomolecular Applications force field to describe the electron density at each atom in a larger system. Mathematically, the proposed model consists of damped pairwise exponential multipolar repulsion interactions truncated at short range, which are suitable for use in compute-intensive biomolecular force fields and molecular dynamics simulations. Parameters for 26 atom classes encompassing most organic molecules are derived from a fit to Symmetry Adapted Perturbation Theory exchange repulsion energies for the S101 dimer database. Several applications of the multipolar Pauli repulsion model are discussed, including noble gas interactions, analysis of stationary points on the water dimer potential surface, and the directionality of several halogen bonding interactions.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5081060>

## I. INTRODUCTION

The beauty of classical physics models is not *that* they work for describing most of our world but rather *why* they work. While it is necessary for physics-based models to be accurate and predictive, these qualities alone are not sufficient. A true classical model must also be interpretable; that is to say, it must be a derivable approximation from first principles. Good classical models of everyday phenomenon are not just lucky; they are the true limiting behavior of fundamental physical laws. Nowhere is this principle more essential or more often forgotten than in molecular modeling. To solve difficult questions such as drug binding specificity or nanotube formation, fields from biology to materials science have come to rely on molecular mechanics models, or force fields, to generate hypotheses and make predictions. These predictions are only as good as the model used to make them, meaning that

every part of the force field must contain a sufficient level of accuracy. In particular, one of the most important parts of any force field is the term responsible for intermolecular Pauli repulsion. This term, which also goes by the names “steric” or “exchange” repulsion, is too often described as a mysterious “quantum mechanical (QM)” force. This could not be farther from the truth.<sup>1-3</sup> This paper intends to show that intermolecular Pauli repulsion is a simple consequence of Coulomb’s law and furthermore that this interpretation leads to an accurate classical model of Pauli repulsion.

The level of model accuracy needed is always a function of its intended use. For force fields, this standard is referred to as “chemical accuracy,” which we define here as a fidelity in computed energies to within 1 kcal/mol. While this requirement is not universal, and higher accuracy may well be required for many applications in molecular interactions, a particular example will serve to rationalize its importance. A primary,

current use for biomolecular force fields is prediction of drug binding affinities. Because of the relation  $\Delta G = RT \log(K_D)$ , at room temperature, every order of magnitude in the binding affinity translates into 1.36 kcal/mol in the free energy of binding. This is of great practical importance since a factor of 10 variance in a drug's binding affinity can be the difference between a medicine that hits a specific target and the one that binds non-specifically. One of the most important factors to achieving this level of accuracy for atom-based force fields is anisotropy. Work with the AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) force field has shown that adding atomic anisotropy via multipoles is an excellent way to make molecular mechanics models more accurate. Atomic multipoles are necessary to accurately predict the electrostatic potential around drug-like molecules<sup>4</sup> and they are needed to reproduce hydrogen bond geometries in water,<sup>5</sup> proteins,<sup>6</sup> and nucleic acids.<sup>7</sup> Recent work with AMOEBA, in the SAMPL6 challenge, showed that this more accurate model produces generally more accurate binding free predictions than its fixed charge counterparts.<sup>8</sup> All of this indicates that to reach the goal of "chemical accuracy," the next generation of force fields will need to account for atomic and molecular anisotropy.

While the importance of anisotropy is broadly recognized for the electrostatic portions of molecular mechanics models, it is widely overlooked in the other terms, particularly Pauli repulsion. The repulsion term is acutely important because in most force fields it is the only consistent source of positive energy in the system. This means that in the delicate balance between attraction and repulsion that exists in all condensed phase systems, the repulsion term shoulders the burden for most of the second half of that equation. In the canonical, minimum energy water dimer, for example, *ab initio* Symmetry Adapted Perturbation Theory (SAPT) energy decomposition calculations show that while the electrostatic, induction, and dispersion contributions to the interaction are all negative, the exchange-repulsion is the only source of positive energy in the system.<sup>9</sup> Despite this, nearly all common biomolecular force fields including AMOEBA use relatively simple, isotropic repulsion schemes. Most commonly these are the  $1/r^{12}$  repulsive Lennard-Jones potential, a Buckingham exponential form or, in the case of AMOEBA, Halgren's buffered 14-7 potential. Strong evidence is emerging that these isotropic Pauli repulsion functions may not be accurate enough to ensure "chemical accuracy" in biomolecular applications. Recent work by Anthony Stone has shown a strong angular dependence of the Pauli repulsion energy in halogen bonding interactions.<sup>4</sup> Furthermore, these sigma hole interactions are vitally important to the drug discovery process,<sup>10</sup> and evidence has emerged that this angular dependence is a large source of the selective binding geometries of sigma hole associated drug candidates.<sup>10-14</sup> In order to meet the standard needed for predictive biomolecular applications, we will need an accurate, physics-based, anisotropic Pauli repulsion model.

This work aims to present a classical Pauli repulsion model that is anisotropic and efficient to compute. In previous work, we have shown that a rough model of atomic charge

density can dramatically improve the accuracy of electrostatic and dispersion models for short-range intermolecular interactions.<sup>15</sup> Here we will show that this same simple density formulation, coupled to an atomic multipole model, yields a classical, physics-based model for Pauli repulsion. In addition to providing a qualitatively different level of accuracy compared to standard isotropic empirical models, this model also dispenses with the customary mysticism surrounding repulsion models. There is no attempt to write off another empirical model in terms of "quantum mechanical forces"; this model simply accounts for the loss in nuclear screening, relative to their isolated, unperturbed states, which molecules experience when their charge densities start to overlap. We will go about this in four stages. First, we will build out the theory underlying the model and its classical electrostatic interpretation. Second, we will describe the methods of the study. Third, we will demonstrate the accuracy of our Pauli repulsion model against benchmark SAPT data. And lastly, we will present pertinent discussion and conclusions.

### A. A brief history of Pauli repulsion models

Well before the advent of modern quantum mechanics, scientists understood that molecules repel each other at short-range. Johannes van der Waals won a Nobel Prize in 1910 for "his work on the equation of state for gases and liquids," which postulated intermolecular interactions as the source of deviations from the ideal gas law.<sup>16</sup> It was not until the statement of the Pauli exclusion principle by Wolfgang Pauli in 1925<sup>17</sup> that physicists understood the explanation for the repulsive intermolecular interactions that keep molecules separated. The first model to approximate this Pauli repulsion phenomenon is due to Sir John Lennard-Jones, a man whose name is indelibly linked to the field of molecular modeling. Lennard-Jones first proposed a general polynomial form of the van der Waals potential in 1924<sup>18</sup> and suggested the now canonical 6-12 formulation in 1931.<sup>19</sup> While the  $1/r^6$  attractive term was taken from London's earlier work on dispersion,<sup>20</sup> the  $1/r^{12}$  repulsive term was chosen primarily out of convenience. Empirically, Lennard-Jones found that the  $1/r^{12}$  term provided an adequate estimate of the repulsive forces between closed-shell atoms near equilibrium. It is worth noting that Lennard-Jones himself had no illusions about the limitations of such a functional form. In his 1931 lecture, he acknowledges that for simple systems exchange energies fall off as  $e^{-\alpha r}/r$ . The Lennard-Jones potential, however, in its canonical form has been enormously influential. It was used in some of the very first molecular dynamics simulations of biological molecules<sup>21</sup> and continues to be the van der Waals function of choice for most popular biomolecular force fields including Amber (Assisted Model Building with Energy Refinement), CHARMM (Chemistry at HARvard Macromolecular Mechanics), GROMOS (GRONingen MOlecular Simulation), and OPLS (Optimized Potential for Liquid Simulations).

The next significant model function for intermolecular repulsion proposed was the simple exponential. Credited to Max Born and Joseph Mayer (1932),<sup>22</sup> and Richard Buckingham (1938),<sup>23</sup> this function has the form  $A e^{-\alpha r}$ . Both papers built on

the work of John Slater, who in 1928 worked out the repulsive force between two helium atoms.<sup>24</sup> Slater found the repulsive force to be exponential of the form  $P(r) e^{-\alpha r}$ , where  $P(r)$  is a polynomial. Slater proposed, however, that to a reasonable approximation  $P(r)$  could be replaced with a constant. Born and Mayer tested this hypothesis on ionic cubic lattices, and Buckingham extended their work to *ab initio* noble gas intermolecular forces. It is interesting to note that Lennard-Jones himself actually played a significant role in the development of these models. He, in fact, was responsible for communicating Buckingham's 1938 paper to the Royal Society. Despite its more substantial theoretical underpinnings, the Buckingham or Born-Mayer exponential functions have been less widely utilized in biomolecular force fields. Notably, the MM2, MM3, and MM4 force fields, designed for accurate conformational analysis and gas phase thermodynamics of hydrocarbons and simple organics, use an exponential repulsion function; however, wide-spread adoption for biomolecules was hampered in part due to the increased computational cost of exponential evaluation.<sup>25-27</sup>

There is one additional contribution to the class of simple, isotropic repulsion models that did not come until much later. In 1992, Thomas Halgren introduced the so-called "buffered 14-7" potential in an effort to raise the level of accuracy of van der Waals functions for organic and biomolecular force fields.<sup>28</sup> This function introduced the idea of buffering constants to fix the known systematic short-range over-repulsive nature of the canonical Lennard-Jones function. Halgren revisited the rare gas repulsion calculations that had guided Lennard-Jones's, Born and Mayer's, and Buckingham's model development with modern electronic structure methods and found that the buffered 14-7 form produced better fits than the 12-6 or exp-6 models he tested against. While the model yields good agreement with *ab initio* data, Halgren makes no attempt to justify it theoretically other than to present it as a perturbation on top of the accepted Lennard-Jones form. Henceforth it has commonly been accepted as a simple solution to the known problem of the excessive stiffness of the  $1/r^{12}$  repulsive wall. Despite its lack of rigorous theoretical justification, this model has been used with some success in the Merck molecular and the AMOEBA force fields.<sup>29-31</sup>

While these three models account for the vast majority of Pauli repulsion terms in biomolecular force fields, there are a large number of "boutique" molecular mechanics repulsive functions tailored to modeling specific types of compounds. Often intended for specific, high-accuracy applications, these models are frequently anisotropic with a larger number of parameters. While we shall refrain from dissecting every known such potential function, a handful of models are notable. Anthony Stone proposed a water model with an atom-atom exponential repulsion that varied according to the relative local geometries of the interacting water molecules.<sup>32</sup> Similarly, Misquitta and Stone developed a site-site anisotropic repulsive potential for pyridine that utilizes distributed densities to generate atomic repulsive "shape" parameters that enter the exponential.<sup>33</sup> In another example, the SAPT-5s water model uses an isotropic repulsive potential

with a polynomial prefactor but requires a number of off-atom sites for accuracy.<sup>34</sup> Lastly, an anisotropic short-range model that includes Pauli repulsion along with other short-range effects has been used to model polycyclic aromatic hydrocarbons.<sup>35</sup>

While the early models of Pauli repulsion acknowledged the general exponential nature of the term, it was not until the 1960s that more rigorous justification and specific functional dependence were provided. In 1961, Lionel Salem published the foundational paper for what would come to be known as the "orbital overlap" model of Pauli repulsion. Salem worked out the repulsive force experienced by two interacting helium atoms subject to the Hellmann-Feynman Theorem and showed that the repulsion energy can be accurately modeled by  $S^2/R$ , where  $S$  is the overlap integral between the interacting orbitals and  $R$  is the distance between the atoms.<sup>36</sup> Not only did he derive this dependence from first principles, but he also showed numerically that such an approximation is quite good for the He dimer example. This work was followed by further validation by Musher and Salem;<sup>37</sup> Murrell, Randic, and Williams;<sup>38</sup> and Murrell and Shaw.<sup>39</sup> The basis of this model is classical electrostatics. Salem showed unequivocally in the case of helium that the repulsive interaction experienced at close approach is caused by a depletion in electron density in the overlap region that de-screens the nuclei from each other, causing internuclear repulsion. Furthermore, Salem illustrated that the magnitude of this depletion is proportional to the square of the orbital overlap integral ( $S^2$ ). Because of the essential problem of defining an orbital in a classical force field, this framework for repulsive models has been less widely used. Two notable exceptions are the SIBFA (Sum of Interactions between Fragments) and EFP (Effective Fragment Potential) models. The SIBFA repulsive potential depends on the overlap between atom centers, bond centers, and lone pairs, as well as a prefactor that accounts for the relative orientation of the interacting pairs.<sup>40,41</sup> EFP uses monomer LMOs (Localized Molecular Orbitals), which makes the potential transferable but too expensive for large-scale biomolecular simulations.<sup>42,43</sup>

The final class of Pauli repulsion models that merits consideration is "density overlap" models. In 1976, Kita, Noda, and Inouye performed molecular beam experiments with  $\text{Cl}^-X$  and  $\text{Br}^-X$ , where  $X = \text{He, Ne, and Ar}$ , and found that the repulsive energy was proportional to  $\Omega$ , the density overlap integral.<sup>44</sup> In 1981, Kim, Kim, and Lee arrived at the same conclusion upon examination of experimental noble gas repulsion data.<sup>45</sup> This observation was first turned into a molecular mechanics model by Wheatley and Price, whose use of anisotropic atomic densities yielded an anisotropic repulsion model.<sup>46</sup> This model influenced Stone's original work on the water model mentioned above. It is also the basis for the Pauli repulsion term in the GEM (Gaussian Electrostatic Model) force field<sup>47-50</sup> and recent force fields developed by Schmidt and co-workers.<sup>51,52</sup> Although this model seems to match experimental data well, it has little in the way of theoretical grounding. Despite its seeming similarity to the orbital overlap model, the density overlap model has a fundamental unit problem. This will be discussed in greater detail at the

end of Sec. II. Notwithstanding theoretical considerations, the density overlap model has been parameterized to meet a range of modeling needs.<sup>53-58</sup> Section II will shed more light on the nature of this model's empirical success.

This history of Pauli repulsion models shows no clear consensus as to which functional form is best. The choice of model has posed two typical trade-offs: first, between computational speed and model accuracy and, second, between model transferability and number of parameters. These trade-offs, however, are not endemic to Pauli repulsion models. Here we present an orbital overlap model that sidesteps both of them. The result is a fast-to-compute, anisotropic, transferable Pauli repulsion function.

## II. THEORY

An explanation of any Pauli repulsion model must start with a basic understanding of the Pauli exclusion principle. The Exclusion principle is a consequence of two simple facts: electrons are fermions and they are indistinguishable. Let us consider a system of two electrons, with wavefunctions,  $\phi_A(x_1)$  and  $\phi_B(x_2)$ . Since these electrons are indistinguishable, we must be able to swap labels and still end up with the same density. One can see that for the simple solution,

$$\phi(x_1, x_2) = \phi_A(x_1)\phi_B(x_2), \quad (1)$$

this condition is not met since it is not necessarily true that

$$\phi_A(x_1)\phi_B(x_2) \neq \phi_A(x_2)\phi_B(x_1). \quad (2)$$

However, since both sides of Eq. (2) are solutions to the Schrodinger equation, we can use a linear combination to produce the total wavefunction,

$$\Phi(x_1, x_2) = \phi_A(x_1)\phi_B(x_2) \pm \phi_A(x_2)\phi_B(x_1). \quad (3)$$

The positive and negative versions of Eq. (3) correspond to symmetric and antisymmetric wavefunctions, respectively, and define the difference between bosons and fermions. Because electrons in nature are always observed to have antisymmetric wavefunctions, they are classified as fermions. The requirement that fermionic wavefunctions be antisymmetric is the essence of the Pauli exclusion principle. For an antisymmetric wavefunction if  $\phi_A = \phi_B$ , the total wavefunction,  $\Phi$ , goes to zero, meaning that no two electrons may occupy the same state.

To illustrate how the Pauli exclusion principle leads to Pauli repulsion between molecules, let us consider the case of the helium dimer. As the simplest closed-shell dimer, He<sub>2</sub> provides a natural, general example for the repulsion between molecules. Our derivation will closely follow that of Salem's 1961 paper. All equations to follow are presented in atomic units, with  $4\pi\epsilon_0 = 1$ .

Consider two helium atoms separated by a large distance such that they do not interact. In this case, the two atoms have distinct wavefunctions,  $\phi_A$  and  $\phi_B$  where both are real, spherically symmetric, exponentially decaying functions. If these two atoms are brought close enough to each other to interact,

the total wavefunction, now a mix of  $\phi_A$  and  $\phi_B$ , must remain antisymmetric because of the Pauli exclusion rule. One way to do this is to construct orthonormal molecular orbitals from linear combinations of the atomic orbitals (LCAO),

$$\begin{aligned} \psi_g &= (2 + 2S)^{-1/2}(\phi_A + \phi_B), \\ \psi_u &= (2 - 2S)^{-1/2}(\phi_A - \phi_B), \end{aligned} \quad (4)$$

where  $S$  is the overlap integral,

$$S = \int \phi_A \phi_B dv, \quad (5)$$

needed for normalization. These molecular orbitals fulfill our requirement that the total wavefunction,

$$\Psi = \psi_g(x_1)\psi_u(x_2) - \psi_u(x_1)\psi_g(x_2) = -\Psi, \quad (6)$$

be antisymmetric.

After enforcing the Pauli exclusion rule, we can determine the total density that this antisymmetric wavefunction defines for the interacting helium dimer. It is worth noting that this is slightly different than the exact density because it lacks polarization effects. However, since the effect of polarization is small for the helium dimer. The density is simply the square of the wavefunction,

$$\rho = \Psi * \Psi = \psi_g^2 + \psi_u^2 \quad (7a)$$

$$= \frac{Z}{(1 - S^2)} (\phi_A^2 + \phi_B^2 + 2S\phi_A\phi_B), \quad (7b)$$

where  $Z$  is the nuclear charge. When  $S$ , the overlap integral, is small, we can approximate the prefactor in a binomial expansion,

$$\frac{1}{1 - S^2} = 1 + S^2 + S^4 + \dots \quad (8)$$

This gives us a good approximation to the total density,

$$\tilde{\rho} = Z(\phi_A^2 + \phi_B^2) - 2ZS\phi_A\phi_B + ZS^2(\phi_A^2 + \phi_B^2) + \dots, \quad (9)$$

where terms of order  $S^3$  or higher are dropped out. It is worth noting that this is slightly different than the exact density because it lacks polarization effects. However, since the effects of polarization are small for the helium dimer (as shown by Salem), this approximate density is accurate enough to ground a qualitative description of Pauli repulsion.

Equation (9) gives us an approximation of the true electron density of the helium dimer. Let us compare this result to the density that would have resulted if we had not imposed the Pauli exclusion principle,

$$\rho_0 = Z(\phi_A^2 + \phi_B^2). \quad (10)$$

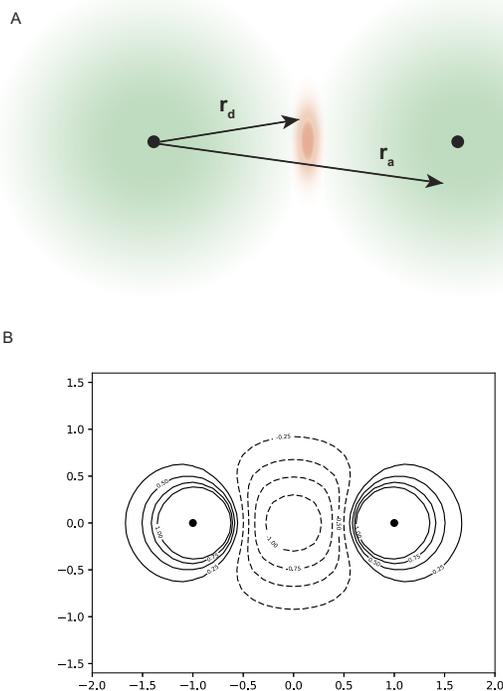
It is clear that this reference density differs from the true density of Eq. (9). It is also clear from the preceding derivation that the difference between the two is entirely due to

the imposition of the Pauli exclusion principle. We can calculate this difference by subtracting the true density from the reference,

$$\Delta\rho = \rho_0 - \tilde{\rho} = 2ZS\phi_A\phi_B - ZS^2(\phi_A^2 + \phi_B^2). \quad (11)$$

This gives us the change in density caused by the Pauli exclusion principle. Figure 1 shows how we can understand this change qualitatively.

There are two contributions to the change in density from the two terms in Eq. (11). The first term represents a depletion of electron density in the overlap region between nuclei, indicated by the red shaded region of Fig. 1(a). And the second term represents an accumulation of electron density around centers A and B, indicated by the green shaded regions of Fig. 1(a). The validity of this spatial decomposition is shown in Fig. 1(b). Coupled Cluster Singles-and-Doubles (CCSD) density difference calculations on the helium dimer show a clear pattern of depletion in the internuclear space accompanied by an accumulation near the nuclei. It is useful to note that since



**FIG. 1.** Change in electron density for interacting helium dimer. (a) Representation of the density difference between the interacting and superimposed non-interacting densities. The green and red regions denote areas of electron accumulation and depletion, respectively, upon enforcement of wavefunction antisymmetry. The distances  $r_a$  and  $r_d$  represent characteristic distances to each region from nucleus A. (b) Density difference from *ab initio* calculations. The difference between CCSD densities of the interacting and non-interacting dimers at an internuclear distance of 1.8 Å is shown as a contour plot. Values shown are in thousandths of electrons/bohr<sup>3</sup>.

$$\int \Delta\rho dv = 0, \quad (12)$$

these two changes are exactly equal in magnitude but opposite in sign. Of particular importance is the magnitude of the depleted charge in the overlap region. For the helium dimer shown in Fig. 1(b) with an internuclear separation of 1.8 Å (0.64 times the consensus He vdW diameter of 2.8 Å), the total depleted charge is only 0.01  $e^-$ . This depletion, however, is the dominant contributor to the total SAPT repulsion energy of ~5 kcal/mol. In fact, if one simply computes Coulomb's law between the nuclei and the depleted "positive" charge located at the midpoint between the two nuclei, the result is ~7 kcal/mol, a slight overestimate of the repulsion energy.

In addition to the qualitative description of how the density changes upon imposition of the Pauli repulsion principle, we can quantitatively assess how this change affects the energy of the system. The change in energy for nucleus A is

$$\Delta E_A = E_A(\rho_0) - E_A(\tilde{\rho}) = Z \int \frac{\Delta\rho}{r} dv = Z \left( \int \frac{\Delta\rho_d}{r} dv - \int \frac{\Delta\rho_a}{r} dv \right), \quad (13)$$

where  $\Delta\rho_d$  and  $\Delta\rho_a$  are the changes in density due to depletion and accumulation, respectively. Plugging in the two terms from Eq. (11), we find the change in energy,

$$\Delta E_A = Z \left( 2ZS \int \frac{\phi_A\phi_B}{r} dv - ZS^2 \int \frac{\phi_A}{r} dv - ZS^2 \int \frac{\phi_B}{r} dv \right) \quad (14a)$$

$$= Z \left( 2ZS \int \frac{\phi_A\phi_B}{r_d} dv - ZS^2 \int \frac{\phi_B}{r_a} dv \right), \quad (14b)$$

where the middle term of 14a is identically zero by symmetry. Equation (14b) introduces the notation for  $r$  illustrated in Fig. 1. The first integral in Eq. (14b), for small overlaps, is approximately zero everywhere except in that small overlap region. Similarly, the second integral of Eq. (14b) is approximately zero for all of space except the local density of atom B. One can see from Fig. 1 that generally  $r_d < r_a$ , meaning that the depletion term of Eq. (14b) outweighs the accumulation term, leaving

$$\Delta E_A \approx 2Z^2S \int \frac{\phi_A\phi_B}{r_d} dv \quad (15)$$

as a good approximation of the change in energy due to the Pauli exclusion principle. In his original paper, Salem confirmed the validity of this approximation for the helium dimer in the region of small overlap, showing that the energy of the depletion term is over 10 times larger than the accumulation term at the van der Waals minimum. Since we are concerned only with small overlaps, a further reasonable approximation is to assume  $r_d$  to be constant,

$$r_d = \frac{R}{2}, \quad (16)$$

where  $R$  is the internuclear distance. Plugging this into Eq. (15),

$$\Delta E_A \approx \frac{Z^2 S}{R} \int \phi_A \phi_B dv \quad (17)$$

simplifies the integral to give our final result,

$$U_{\text{Pauli}} = \frac{Z^2}{R} S^2. \quad (18)$$

This is the simple energy difference caused by the imposition of the Pauli repulsion principle on our unperturbed reference state.

Equation (18) constitutes the ‘‘Orbital Overlap’’ model of Pauli repulsion, and it is remarkable for three important reasons. First, it gives us a clear definition of Pauli repulsion. There is no mysterious ‘‘quantum force’’ that drives molecules apart—Eq. (18) reveals that the repulsion caused by enforcing the Pauli exclusion principle is electrostatic. The form of Eq. (18) bears a striking resemblance to Coulomb’s law,  $U = q_i q_j / R$ , only with a factor of  $S^2$  modulating the interaction. This similarity is not an accident. Figure 1 shows quite clearly that the main effect of requiring the wavefunction to be antisymmetric is a net loss of electron density, relative to the reference state, in the area between the two nuclei. This leads to an electrostatic repulsion between nuclei that is proportional to the overlap squared. Second, the form of Eq. (18) fits the asymptotic behavior we expect from a Pauli repulsion function. It is positive everywhere, making it indeed repulsive. Moreover, since  $S$  is proportional to an exponential, the repulsion energy goes to zero at a long range and becomes large when molecules strongly overlap. Third, the orbital overlap model lends itself to molecular mechanics models because of how it was derived. The above derivation relies on a choice of reference state, in this case the unperturbed electron densities of the separated helium atoms. This is exactly analogous to the strategy of most molecular mechanics models. Partial charges, multipoles, polarizabilities, etc., are all assigned to molecules in force fields as gas-phase monomer properties. In other words, the unperturbed molecular electron density is also the ‘‘reference state’’ of molecular mechanics, making the orbital overlap model inherently consistent with the rest of the force field model.

This orbital overlap model of Pauli repulsion is by no means new, but it has not been widely taken up by molecular mechanics models. The reason for this is the challenge of determining the  $S^2$  term for a classical model. Models like SIBFA and EFP have taken the strategy of explicitly calculating molecular orbital overlaps to directly obtain  $S^2$ . These models are certainly accurate and have the feature of giving realistic anisotropic repulsion, but they can be too slow for large-scale molecular dynamics simulation and pose problems for parameterization of biological macromolecules. An alternative is to use an empirical model of orbital overlap, but this leaves open the question of how to determine parameters defining the anisotropy. We present here a novel approach to this issue—an anisotropic model of Pauli repulsion can be faithfully constructed from the anisotropy encoded in an atom’s multipole moments.

## A. Multipole overlap Pauli repulsion

A model for orbital overlap requires some method for describing the electron distribution around a molecule. Previous work has shown that a simple, hydrogen-like model of charge density can be used to accurately predict the electrostatic interactions of dimers at a short-range.<sup>15</sup> In this model, the point Coulomb potential,  $V = q/r$  in atomic units, was replaced by

$$V(r) = \frac{q}{r} (1 - e^{-\alpha r}), \quad (19)$$

where  $\alpha$  is a parameter introduced to describe the width of the electron density. For the present model, we modify this slightly, as suggested by Slipchenko and Gordon.<sup>59</sup> The potential generated by the electron of a hydrogen-like atom is

$$V(r) = \frac{q}{r} [1 - (1 + Zr)e^{-2Zr}], \quad (20)$$

where  $Z$  is the nuclear charge. The form of Eq. (20) differs slightly from Eq. (19), suggesting a better approximation for the model,

$$V(r) = \frac{q}{r} \left[ 1 - \left( 1 + \frac{1}{2} \alpha r \right) e^{-\alpha r} \right], \quad (21)$$

which more accurately captures the asymptotic behavior of the potential. From this potential, we wish to build a model electron distribution. To do this, we can apply Poisson’s equation,

$$\nabla^2 V = -4\pi\rho, \quad (22)$$

to obtain the charge density that generates the potential of Eq. (19),

$$\rho(r) = \frac{q\alpha^3}{8\pi} e^{-\alpha r}. \quad (23)$$

The density, however, does not directly give the information we need. In order to use the orbital overlap model of Pauli repulsion, we must have a model for orbitals. To get from a density model to an orbital model, we apply

$$\rho = \phi * \phi. \quad (24)$$

If we impose the restriction that the orbitals be real, then we are left with model pseudo-orbitals,

$$\phi = \sqrt{\rho} = \sqrt{\frac{q\alpha^3}{8\pi}} e^{-\frac{\alpha r}{2}}. \quad (25)$$

It is important to be clear about the purpose of these model orbitals. What we are interested in for Pauli repulsion is the regime of small overlap. Correspondingly, these orbitals are simply meant to approximate the form of the outermost extent of an atom’s electron distribution.

Within this pseudo-orbital model, we can evaluate the overlap integral between interacting orbitals, A and B,

$$S = \int \phi_A \phi_B dv = \frac{\sqrt{q_A q_B \alpha_A^3 \alpha_B^3}}{8\pi} \int e^{-\frac{\alpha_A r_A - \alpha_B r_B}{2}} dv. \quad (26)$$

Performing the integration according to the method of Coulson<sup>60</sup> and refactoring yield the result

$$S = \sqrt{\frac{q_A q_B \alpha_A^3 \alpha_B^3}{R}} f_{damp}(R), \quad (27)$$

where,

$$f_{damp}(R) = \begin{cases} \frac{\sqrt{R}}{\alpha^3} \left( 1 + \frac{\alpha R}{2} + \frac{1}{3} \left( \frac{\alpha R}{2} \right)^2 \right) e^{-\frac{\alpha R}{2}}, & \alpha_A = \alpha_B \\ \frac{1}{2X^3 \sqrt{R}} \left[ \alpha_A (RX - 2\alpha_B) e^{-\frac{\alpha_B R}{2}} + \alpha_B (RX + 2\alpha_A) e^{-\frac{\alpha_A R}{2}} \right], & \alpha_A \neq \alpha_B. \\ X = \left( \frac{\alpha_A}{2} \right)^2 - \left( \frac{\alpha_B}{2} \right)^2 \end{cases} \quad (28)$$

There are two important features to note about  $S$ . First, it is asymptotically exponential, matching our general intuition about orbital overlap. Second, it depends on the respective atomic multipoles of A and B. We can elucidate this fact by writing  $S^2$  in “Coulombic” form,

$$S^2 = q_A T_{pauli} q_B, \quad (29)$$

with

$$T_{pauli} = \frac{\alpha_A^3 \alpha_B^3}{R} f_{damp}^2. \quad (30)$$

Equation (29) represents the charge-charge overlap term of our Pauli repulsion model. If we wish our model to be isotropic, we simply stop here and compute the repulsion energy according to Eq. (18). However, for multipolar force fields, we are not bound to simply using the charge-charge component of the overlap. Following the example of Slipchenko and Gordon,<sup>59</sup> we can show how to compute the charge-dipole and higher-order multipole terms for orbital overlaps as well.

Consider the overlap at distance,  $R$ , of a charge density,  $Q$ , with a finite dipole,  $\mu$ , where the dipole is represented by two equal and opposite charges,  $q^-$  and  $q^+$ , separated by a distance,  $d$ . For this interaction,

$$\begin{aligned} S_{charge-dipole}^2 &= S_{Q-q^-}^2 + S_{Q-q^+}^2 \\ &= QT_{pauli} \left( R - \frac{d}{2} \right) q^- + QT_{pauli} \left( R + \frac{d}{2} \right) q^+. \end{aligned} \quad (31)$$

If we define the dipole moment,

$$\mu = qd, \quad (32)$$

then Eq. (31) becomes

$$S_{charge-dipole}^2 = Q\mu \frac{T_{pauli} \left( R + \frac{d}{2} \right) - T_{pauli} \left( R - \frac{d}{2} \right)}{d}. \quad (33)$$

If we take the limit as  $d \rightarrow 0$ , then

$$S_{charge-dipole}^2 = Q \frac{\partial T_{pauli}}{\partial R} \mu. \quad (34)$$

Note that this is exactly analogous to the derivation of the electrostatic multipole interaction,

$$\begin{aligned} U_{electrostatic} &= q_A T q_B + q_A \nabla T \mu_B - \mu_A \nabla T q_B + \mu_A \nabla \nabla T \mu_B + \dots \\ &\text{with } T = \frac{1}{R}, \end{aligned} \quad (35)$$

where the only difference is the kernel,  $T$ .

In the same way, we can compute  $S^2$  for arbitrary order multipole orbital overlaps. In the current model, we shall take this through quadrupole-quadrupole repulsion, yielding

$$\begin{aligned} S_{total}^2 &= q_A T_{pauli} q_B + q_A \nabla T_{pauli} \mu_B - \mu_A \nabla T_{pauli} q_B + \mu_A \nabla^2 T_{pauli} \mu_B \\ &\quad + q_A \nabla^3 T_{pauli} \Theta_B - \Theta_A \nabla^3 T_{pauli} q_B + \mu_A \nabla^4 T_{pauli} \Theta_B \\ &\quad - \Theta_A \nabla^4 T_{pauli} \mu_B + \Theta_A \nabla^5 T_{pauli} \Theta_B. \end{aligned} \quad (36)$$

This is the source of anisotropy in our model. Rather than introducing any new parameters, we simply use the shape of the atom encoded in the multipole moments to tell us about the anisotropy of repulsion. This result provides a total multipole overlap model of Pauli repulsion,

$$U_{pauli} = \frac{K_A K_B}{R} S_{total}^2. \quad (37)$$

The parameter,  $K$ , is introduced to set the relative sizes of different atom classes. This model will be referred to throughout the remainder of the paper as the “Multipolar Pauli Repulsion” model. A variant that uses only the first, charge-charge term of Eq. (36) will also be discussed. We refer to this as the “Isotropic Pauli Repulsion” model.

We intend to use this model on a broad array of complicated biomolecular intermolecular interactions. To determine the parameters, particularly  $K$ , that accurately describe these interactions, we have chosen to use *ab initio* SAPT exchange repulsion calculations to generate reference data. It should be noted that this is technically an approximation. While our derivation relies on a Hellmann-Feynman theorem analysis of density differences, the SAPT exchange repulsion term has no associated density. This approximation, however, is necessary, accurate, and consistent with our model. The SAPT exchange repulsion energy is a direct approximation of the energy increase required to antisymmetrize the wavefunctions of two monomer reference states—exactly the quantity that our model is built to reproduce.

## B. A note on “density overlap” models

The other major class of overlap models for Pauli repulsion utilizes density overlap. In these models, the Pauli repulsion energy is modeled as

$$U_{pauli} = K_{ij} \Omega, \quad (38)$$

where

$$\Omega = \int \rho_i \rho_j dv. \quad (39)$$

This model is supported with some experimental evidence,<sup>45,61</sup> but it has serious flaws as an interpretable model

for Pauli repulsion energy. This can be simply illustrated with straightforward dimensional analysis.  $\Omega$  is the one-center integral of two densities which means that it has dimension (in atomic units),

$$\Omega = \frac{e}{a_0^3} \frac{e}{a_0^3} a_0^3 = \frac{e^2}{a_0^3}. \quad (40)$$

These are not the units of a Coulombic energy ( $e^2/a_0$  in atomic units). Although  $K_{ij}$  can be given units that yield an energy, this does not make the model Coulombic since the term is a constant that does not depend on distance. This is a problem for two reasons. First, it makes the model inconsistent with the Hellmann-Feynman (Electrostatic) theorem. According to the Hellmann-Feynman theorem, every intermolecular force is the result of applying Coulomb's law to a change in electron density. This dimensional analysis demonstrates that there exists no Hellmann-Feynman-based justification for the density overlap model of Pauli repulsion since any such rationale must necessarily be Coulombic. It should be noted that there are other Pauli repulsion energy models that do not have Coulombic interpretations as well, with the simple Lennard-Jones model among them. However, these models are, by and large, unitless which means that they do not suffer from the second and more important reason that the density overlap model is problematic—the density overlap model has the wrong distance dependence. Since the density overlap is proportional to charge squared over distance cubed, applying the electric constant no longer gives units of energy but energy over distance squared. This is problematic because this formulation now explicitly depends on the unit chosen for distance. A consequence of this is that the radial dependence of the Pauli repulsion energy will be qualitatively incorrect. This is illustrated in Fig. 2 for the case of helium dimer repulsion. Both the  $S^2/R$  and density overlap models are governed largely by exponentials, as shown by the nearly straight lines on the semi-log plot. However, the slope of the density overlap exponential function clearly differs from the SAPT exchange repulsion.

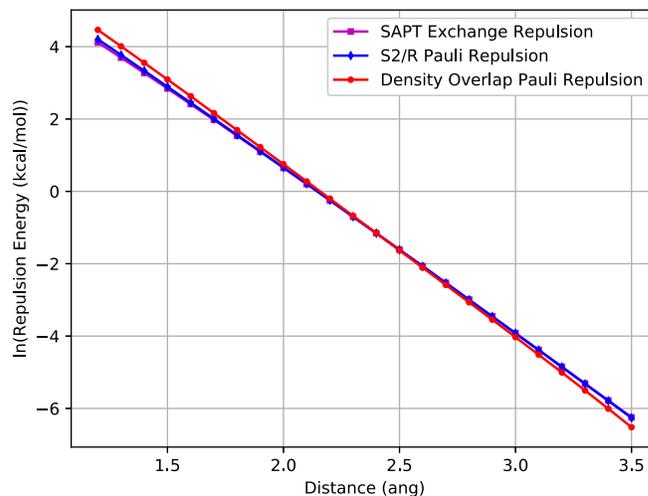
A simple model system explains this difference in radial dependence of the  $S^2/R$  and density overlap curves. If we assume that both atoms have an isolated electron density described by Eq. (23) ( $\sim e^{-\alpha r}$ ), the resulting Pauli repulsion energies of the density overlap and  $S^2/R$  models, respectively, will be

$$U_{\text{density overlap}} = K_{ij} \int \rho_i \rho_j d\mathbf{v} = K_{ij} \frac{q_i q_j}{(8\pi)^2} \frac{\pi}{a^3} \left(1 + \alpha R + \frac{1}{3}(\alpha R)^2\right) e^{-\alpha R} \quad (41)$$

and

$$\begin{aligned} U_{S^2/R} &= K_{ij} \frac{1}{R} \left( \int \phi_i \phi_j d\mathbf{v} \right)^2 \\ &= K_{ij} \frac{1}{R} \frac{q_i q_j}{(8\pi)^2} \frac{(8\pi)^2}{a^6} \left(1 + \frac{1}{2}\alpha R + \frac{1}{12}(\alpha R)^2\right)^2 e^{-\alpha R}. \end{aligned} \quad (42)$$

For this simple model, the arguments of the exponentials of both models are identical due to the orbital overlap,  $S$ , being



**FIG. 2.** Radial dependence of the  $S^2/R$  (blue) and density overlap (red) models for the helium dimer. The repulsion energy of both models computed from monomer wavefunctions determined with the aug-cc-pVQZ basis set is compared to the SAPT2+ repulsion energy (magenta) and plotted on a semi-log scale. The proportionality constant of each model was fixed to reproduce the SAPT repulsion energy at 2.4 Å. The SAPT exchange repulsion curve is almost entirely obscured by the  $S^2/R$  curve. While the slope of the  $S^2/R$  curve matches that of the SAPT repulsion energy, the slope of the density overlap curve does not.

squared. This explains the similarities in the curves in Fig. 2. However, the  $R$ -dependent, polynomial prefactor of Eq. (41) clearly differs from Eq. (42). It is this difference that causes the radial divergence illustrated in Fig. 2. We will further illustrate this phenomenon in Sec. IV with additional noble gas and water dimer calculations.

Despite this incorrect radial functional dependence, the density overlap model can still provide a serviceable empirical model. There are two ways that this has been performed in prior efforts. The first is to choose the  $K_{ij}$  proportionality for a representative interaction distance. Since, like the orbital overlap model, the density overlap model is dominated by an exponential term, if the  $K$  constant is chosen for a suitable interaction distance, the radial error may not become large within the range sampled in application. The other method is to include a distance dependent prefactor ( $1/R$ ,  $1/R^2$ , etc.) in the function. This has been proposed by Nyeland and co-workers.<sup>62–64</sup> These models explicitly address the units of the function, albeit empirically, and have been shown to be valid over a wider range of interaction distances than the pure density overlap model.

As opposed to unitless models like Lennard-Jones or Buckingham functions which sidestep the question, the orbital overlap model explicitly satisfies the dimensional analysis test. If we take Eq. (36) and for simplicity, only consider the charge-charge term of  $S^2$  (the same holds for the higher-order terms), this gives,

$$U_{\text{pauli}} = \frac{e^2}{a_0}, \quad (43)$$

the atomic units of a Coulomb energy. This is fully consistent with the Hellmann-Feynman theorem interpretation of the Pauli repulsion energy. It is also what we should expect, given that the derivation of the orbital overlap model is built on a set of electrostatic interaction arguments. This makes the multipole overlap Pauli repulsion model interpretable and, as we shall show in Sec. IV, this interpretability bestows on the model some measure of transferability.

### III. METHODS

#### A. Parameterization

To fit the multipole Pauli repulsion model, we utilized the previously published S101x7 database.<sup>65</sup> This database is meant to capture intermolecular interactions that are important for biomolecular applications and has been used to parameterize electrostatics and dispersion models.<sup>15,66</sup> The database contains 101 unique sets of molecular dimers with seven different points along the dissociation curve at 0.7, 0.8, 0.9, 0.95, 1.0, 1.05, and 1.1 times the equilibrium distance. See Ref. 65 for a complete description of the generation of database geometries. We augmented the database with a set of methane and formaldehyde homodimers with geometries generated in the same manner. We used the publicly available SAPT2+ energy decomposition analysis calculations we previously published to extract the exchange repulsion (Pauli repulsion) energy of each dimer pair in the database. The SAPT2+ exchange repulsion energies were computed using the so-called  $S^2$  approximation which has been previously shown to be accurate for biomolecular fragment interactions.<sup>67</sup>

To fit the parameters of our model, we defined 26 unique atom classes. These classes are assigned according to the qualitative chemical environment of each atom and are listed in Table I. They are adopted from the atom classes used in a previously published study of van der Waals energies of the S101x7 database.<sup>68</sup>

For each atom class, the model as defined by Eq. (36) requires three parameters: the size of the atom,  $K$ , the shape of the atom,  $\alpha$  [as defined in Eq. (25)], and the number of valence electrons,  $q$  [as defined in Eq. (36)]. The purpose of  $K$  and  $\alpha$  is straightforward; together they set the strength and shape of the exponential repulsion between atoms. The purpose of  $q$  as a parameter is slightly more nuanced. In point force fields (charge-only or multipolar), the atomic charges are a combination of the nuclear charge with the net electronic charge on that atom. This definition will not work for our model because only the electrons are involved in overlap. Furthermore, since we are only interested in the region of small overlaps, it does not make sense to use all of the electrons on each atom, as only the outermost part of the electron density is involved in overlap. Thus, the parameter,  $q$ , is best thought of as the maximum (not necessarily an integer) number of electrons that are involved in overlap for a particular atom. This turns out to be an important parameter because it sets how anisotropic the Pauli repulsion of a specific atom will be. A large  $q$  will make the first, isotropic term of Eq. (36) large, while a small

TABLE I. Atom classes and parameter values for the anisotropic multipole Pauli repulsion model. Classes are taken from those in Ref. 68.

N	Pauli repulsion class	K	$\alpha$	$q$
1	H (nonpolar)	2.25	4.63	1.00
2	H (nonpolar, alkane)	1.82	4.23	1.00
3	H (polar, N-H/N aromatic)	1.11	4.21	1.00
4	H (polar, O-H)	1.18	4.20	1.00
5	H (aromatic, C-H)	1.24	4.43	1.00
6	H (polar, S-H)	1.20	4.01	1.00
7	C ( $sp^3$ )	2.62	4.64	3.41
8	C ( $sp^2$ , alkene)	1.42	3.56	3.92
9	C ( $sp^2$ , C=O)	1.31	3.50	2.02
10	C (aromatic, C-C)	1.37	3.77	4.00
11	C (aromatic, C-X)	1.37	3.70	3.70
12	N ( $sp^3$ )	3.61	4.16	2.00
13	N ( $sp^2$ )	4.62	4.27	2.15
14	N (aromatic)	3.93	4.40	2.48
15	O ( $sp^3$ , hydroxyl, water)	3.57	4.74	3.00
16	O ( $sp^2$ , carbonyl)	1.43	4.14	6.00
17	O ( $O^-$ in $AcO^-$ )	1.19	3.77	5.87
18	O ( $O^-$ in $HPO_4^{2-}$ )	1.25	3.73	5.82
19	O ( $O^-$ in $H_2PO_4^-$ )	1.47	4.02	5.75
20	O (O in $H_3PO_4$ )	1.63	4.15	5.78
21	P (phosphate)	1.74	4.40	4.98
22	S (sulfide, R-SH)	3.40	3.62	3.39
23	S (sulfur IV, DMSO)	1.52	3.33	6.00
24	F (organofluorine)	1.38	4.72	5.05
25	Cl (organochloride)	1.91	3.76	5.91
26	Br (organobromine)	2.02	3.52	6.63

$q$  will make the first term small relative to the higher-order anisotropic terms. To prevent overfitting,  $q$  for all hydrogen atoms is set to be the negative of the total number of electrons (-1.0 plus the partial charge). Additionally,  $q$  for heavy atoms is constrained to lie between 2 and the number of valence electrons of the element. For the isotropic Pauli repulsion model [only using the first term of Eq. (36)],  $q$  is set to be the number of valence electrons since in the isotropic model,  $q$  and  $K$  are redundant.

It is important to emphasize what is not being fit in this model. The dipole and quadrupole moments of each atom are taken directly from the Distributed Multipole Analysis (DMA) based procedure detailed in the appendix of Ref. 69. This does two important things. First, it makes the multipolar Pauli repulsion model consistent with the AMOEBA model and future AMOEBA-like models. In particular, this means that the model can be used along with previously published AMOEBA-like electrostatics<sup>15</sup> and dispersion<sup>66</sup> models. Second, this insulates the Pauli repulsion model from the most common problem of anisotropic repulsion models: overfitting. The dipole moment of each atom has three independent components, and the traceless quadrupole has five. Empirically fitting these parameters would result in a massive overfitting problem. Past anisotropic models have either fit very specific models (e.g., water) to large datasets<sup>32,33</sup> or used explicit atomic orbitals.<sup>40,43</sup> This model sidesteps the troubles associated with both of those approaches by using the

multipoles that come from directly fitting the electrostatic potential around a molecule.

Because the Pauli repulsion energy exhibits strong exponential character, we choose to perform a natural log fit to obtain parameters for the multipolar and isotropic Pauli repulsion models. To do this, we minimized the residual,  $\log(\text{SAPT Exchange}) - \log(\text{model})$ , for each dimer data point in the S101x7 dataset using a Levenberg-Marquardt least squares routine in the Tinker molecular mechanics package.<sup>70</sup> This prevents the closest, but rarely accessible, dimer points from biasing the fit. A third model, termed vdW2017, was also fit. This model uses the AMOEBA standard buffered 14-7 functional form,

$$U_{vdW} = U_{\text{repulsion}} + U_{\text{dispersion}} = \epsilon_{ij} \left( \frac{1 + \delta}{\rho_{ij} + \delta} \right)^7 \left( \frac{1 + \gamma}{\rho_{ij}^7 + \gamma} - 2 \right), \quad (44)$$

where  $\delta$  and  $\gamma$  are global shape parameters, and  $\epsilon$  and  $\rho$  are set by the Waldman-Hagler and arithmetic combining rules, respectively, as suggested in Ref. 68. All four parameters were allowed to vary in the fit with  $\epsilon$  and  $\rho$  for each atom being set by the same atom classes presented in Table I. Additionally, a fixed “hydrogen-reduction factor” of 0.9 was applied to all hydrogens as described in Ref. 31. Equation (44) can be split into a positive and a negative part which represent the contributions to repulsion and dispersion, respectively. However, since the buffered 14-7 parameters for the two terms are not independent, a (non-natural log weighted) least squares fit was carried out that simultaneously minimized (SAPT exchange–vdW2017 repulsion) and (SAPT dispersion–vdW2017 dispersion) for each dimer data point. In both fits, S101 dimers including the triple-bonded ethyne molecule were excluded as was performed previously, eliminating the S101 class for *sp*-hybridized carbon from the fitted parameters.

## B. Computational details

The multipolar and isotropic Pauli repulsion models have been implemented in publicly available versions of the Tinker molecular mechanics package.<sup>70–72</sup> It is worth noting for future force field development that the additional overhead to compute the multipolar Pauli repulsion model on top of an existing multipole electrostatic calculation is small. As the similarity between Eqs. (35) and (36) suggests, the intermediate quantities necessary for energy and forces are largely identical between the two models.

Lastly, we explored calculating the orbital overlap and density overlap directly from quantum mechanical calculations. Since the components of interacting dimers each have multiple occupied orbitals rather than single model pseudo-orbitals, we must define  $S^2$  for this situation. Because we are working with orthogonal molecular orbitals in the LCAO (Linear Combination of Atomic Orbitals) convention, we use the sum-of-squares definition,

$$S^2 = \sum_i^A \sum_j^B \langle \psi_i | \psi_j \rangle^2, \quad (45)$$

where the A and B represent the two monomers with sums over  $i$  and  $j$ , the occupied orbitals on A and B, respectively.<sup>62</sup> We express the occupied orbitals in terms of atomic basis functions so that Eq. (45) is invariant under orthogonal transformations of the molecular orbitals. The density overlap is calculated on a grid according to Eq. (38). The QM orbital and density overlap calculations using SCF monomer orbitals with an aug-cc-pVDZ basis were performed using the Psi4NumPy program.<sup>73</sup>

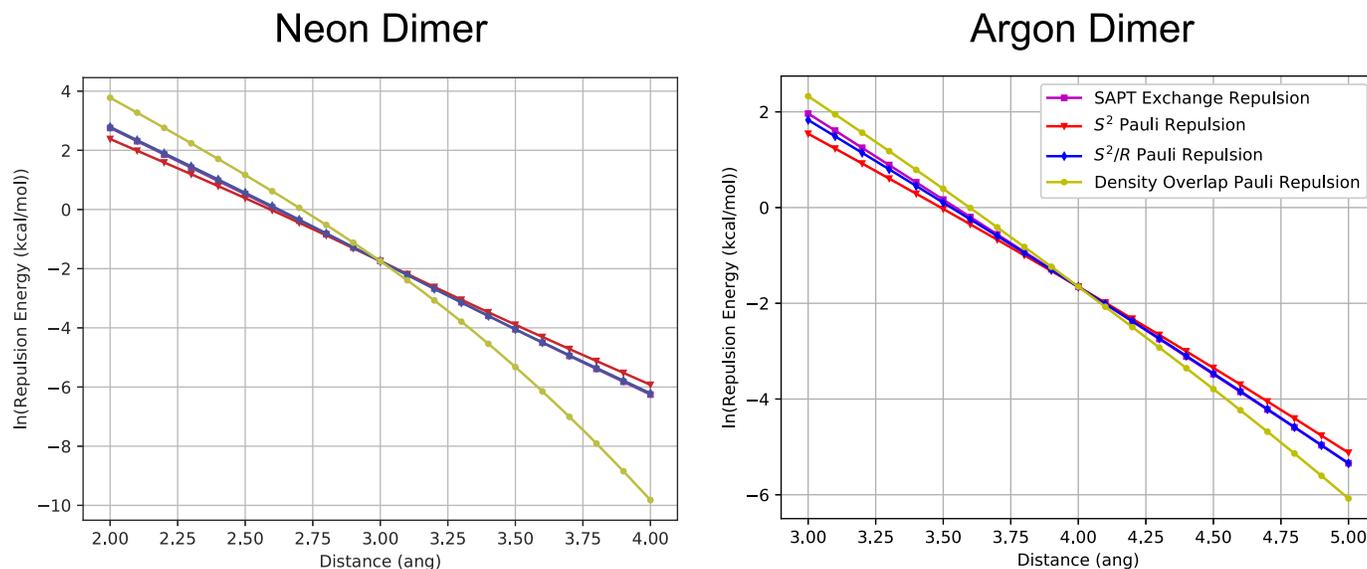
The computational cost of the multipolar Pauli repulsion model was evaluated on a typical computer workstation. The time to complete 100 energy and force calculations was measured using a four-core 3.4 GHz Intel Core i7 processor. The tests were run on a  $25 \times 25 \times 25$  Å water box with 500 water molecules. The cutoff distance for Pauli repulsion is set to 5 Å, and dispersion is handled via particle mesh Ewald summation.<sup>66</sup> For comparison, models including the Halgren buffered 14-7 potential were also included. For these calculations, a van der Waals cutoff distance of 10 Å is used. To evaluate the cost in the context of a generalized AMOEBA-like model, timings are also presented that include polarization with the induced dipole convergence criteria set to  $10^{-5}$  D RMS.

## IV. RESULTS

### A. Noble gas dimers

To assess the validity of the orbital overlap model for Pauli repulsion, we first considered the case of Pauli repulsion between noble gas dimers. Because they are neutral and have spherical symmetry, noble gas dimers are a natural first testing ground for a Pauli repulsion model. Specifically, we set out to test the underlying assumption that the Pauli repulsion energy should be proportional to  $S^2/R$ . The results, plotted for a range of distances of the neon and argon dimers, are plotted in Fig. 3.

There are several features worth noting in Fig. 3. The first is that for these simple systems the SAPT exchange repulsion energy is clearly exponential. The plot reveals a near-linear relationship between the internuclear distance and the natural log of the SAPT exchange repulsion energy. The second noteworthy feature is the quality of the  $S^2/R$  model energy calculated from SCF monomer orbitals. This is not necessarily surprising, given that the model was derived from the *ab initio* Pauli repulsion of the helium dimer, but agreement is remarkable, given that this estimate is obtained without the need for dimer calculations. The third important item of note is the poor quality of the density overlap approximation for the Pauli repulsion energy. The distance-dependence problem addressed in Sec. II B is abundantly clear in this simple example. The density overlap is evidently not proportional to the Pauli repulsion energy for a meaningful range of distances for noble gas dimers. Although the rapid divergence of the density overlap model at a long range for the neon dimer is likely due to the use of the smaller aug-cc-pVDZ basis set, the same qualitatively different radial dependence is observed when the overlap is computed with a much larger (aug-cc-pV5Z) basis set.



**FIG. 3.** Comparison of QM monomer-based methods for estimating Pauli repulsion in noble gas dimers. The natural log of the repulsion energy is plotted against the dimer separation distance to illustrate the exponential relationship. For a range of distances for Ne–Ne and Ar–Ar dimers, the  $S^2$ ,  $S^2/R$ , and density overlap methods were tested against SAPT exchange repulsion energy. The proportionality constant for each method was arbitrarily fixed to reproduce the middle value of the distance range (3.0 ang for Ne–Ne and 4.0 ang for Ar–Ar). In both cases, the  $S^2/R$  method is virtually indistinguishable from the SAPT result.

The final feature to point out regarding the noble gas dimers is subtle but important. Also plotted in Fig. 3 are the results for assuming that  $S^2$ , as opposed to  $S^2/R$ , is proportional to the Pauli repulsion energy. The results show that although the agreement might be close over a small range of distances, the overall slope is slightly too small. This shows the importance of the  $1/R$  factor in the Pauli repulsion expression. There has been some discussion in the literature about what, if any, function of  $R$  should precede  $S^2$  for Pauli repulsion.<sup>41,62</sup> These results clearly indicate that  $1/R$  is the correct choice. As an empirical matter, of course, for more complicated molecules, other choices can be made in the context of a total energy model. However, these results show the  $S^2/R$  model to be the most natural fit to the fundamental Pauli repulsion phenomenon.

## B. S101x7 dataset

Having established the validity of the  $S^2/R$  model for noble gas systems, we set about determining whether the model is appropriate for a more complicated dataset. The S101x7 database was chosen to represent a range of biomolecular dimer interactions and three models were fit: multipolar Pauli repulsion, isotropic Pauli repulsion, and a buffered 14-7 model. As stated in Sec. III, the first two repulsion-only models were fit with natural log weighted least squares, while the buffered 14-7 model, termed vdW2017, was fit to unweighted SAPT dispersion and exchange repulsion data simultaneously. The results of these fits are given in Table II and illustrated in Fig. 4.

The results show the trade-off in accuracy that is taken for using an isotropic model. The isotropic Pauli repulsion and vdW2017 models exhibit similarly large errors for the S101x7 dataset of 2.37 kcal/mol and 2.66 kcal/mol, respectively. This error is driven by the closest contact points in the dataset but is still large for intermediate distances as well. Notably, the vdW2017 model has an error of close to 1 kcal/mol for points at equilibrium and beyond. To some extent, errors at this distance for the buffered 14-7 potential are compensating for the dispersion part of the function, but this comes at the detriment of having a separate and interpretable Pauli repulsion model. The isotropic Pauli repulsion function root mean square error, on the other hand, decays more rapidly with distance.

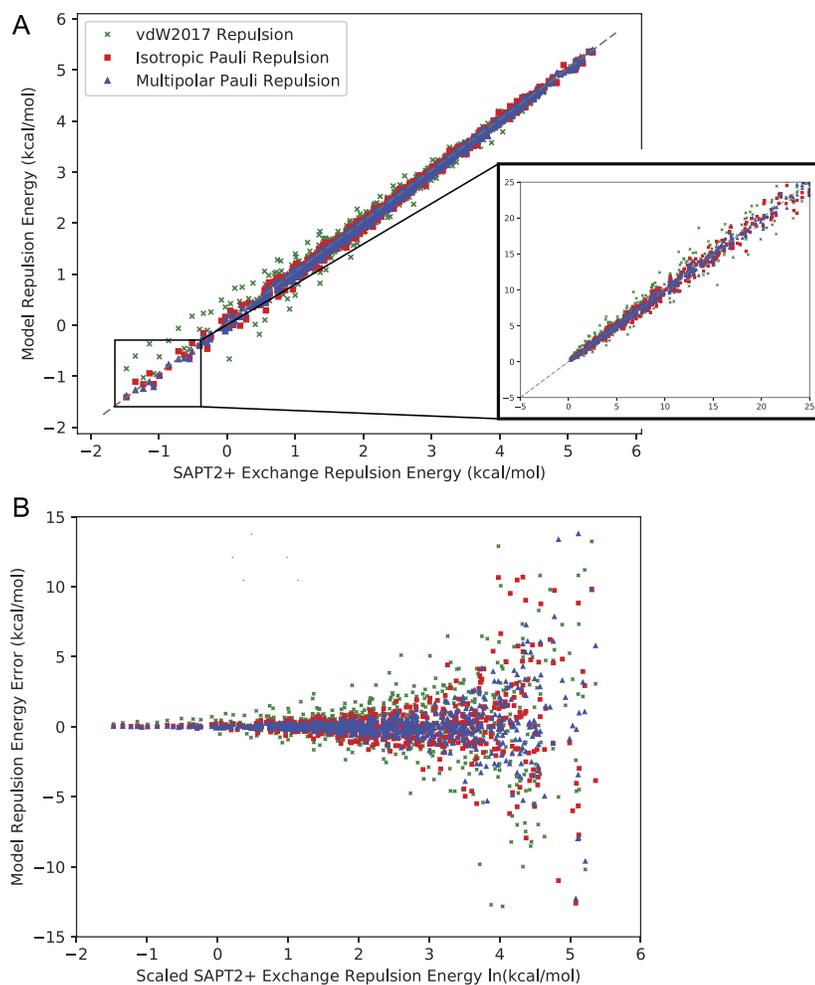
The quantitative benefit of using an anisotropic Pauli repulsion function is readily apparent from Table II and Fig. 4. The multipolar Pauli repulsion model requires more terms to compute, but it fits the S101x7 nearly twice as well as its isotropic counterpart. The total RMSE (Root Mean Square Error) is being driven almost entirely by the closest-range points ( $0.7\times$ ) in the dataset. For intermediate and near-equilibrium points, the multipolar Pauli repulsion model gives errors of well under 1 kcal/mol. Because the fitting was performed against log-transformed data, it is not surprising to see this behavior. Moreover, this behavior should be considered desirable since the  $0.7\times$  points of the dataset largely fall just outside of the realm accessed during molecular dynamics simulation under ambient conditions. Figure 4 illustrates the tighter fit to SAPT that is achieved with the anisotropic multipolar Pauli repulsion model. The results of these fits will be

**TABLE II.** Root mean square error on the S101x7 dataset. Shown are the errors relative to SAPT exchange-repulsion. “Short-Range” indicates data points at 0.7 times the dimer equilibrium distance. “Intermediate” indicates data points 0.8–0.95 times the dimer equilibrium distance. “Long-Range” indicates data points at or beyond the dimer equilibrium distance. Note that all values are absolute errors and not log-weighted.

	Total RMSE (kcal/mol)	Short-range RMSE (0.7) (kcal/mol)	Intermediate RMSE (0.8–0.95) (kcal/mol)	Long-range RMSE (1.0–1.1) (kcal/mol)
Multipolar Pauli repulsion	1.71	4.14	0.99	0.37
Isotropic Pauli repulsion	2.37	5.68	1.46	0.44
vdW2017 repulsion	2.66	5.94	2.02	0.83

used to evaluate the models for the remainder of the paper. Tables showing the fitted parameters for the isotropic and multipolar Pauli repulsion models, as well as the vdW2017 model are available in the [supplementary material](#). We will explore the factors contributing to the superior fit of the multipolar Pauli repulsion model in Sec. V.

As can be seen from Fig. 4(a), there are some large errors in the fitted S101 dataset, particularly for the isotropic and vdW2017 models. Nearly all of these errors occur at the short-range 0.7 $\times$  points of the database where the absolute repulsion energy is very high. This is apparent from Fig. 4(b), which shows the error of each point in the fits plotted against the log



**FIG. 4.** S101x7 Pauli repulsion energy. (a) Model Pauli repulsion energy plotted against SAPT2+ exchange repulsion energy for all dimers. Both model and SAPT data are plotted on natural log weighted access for clarity. The dashed line indicates perfect agreement. (b) Model error plotted against absolute (log-weighted) SAPT repulsion energy.

magnitude of the repulsion energy. All of the errors greater than 5 kcal/mol occur when the SAPT repulsion energy is greater than 50 kcal/mol. Additionally, many of the dimers where the anisotropic multipole Pauli repulsion model produces the greatest reduction in the error fit with intuition. Some of the largest decreases in error are for the DMSO-DMSO dimer, dimers involving phosphate, and pi-pi stacking interactions. These are all interactions with significant electrostatic anisotropy (large dipole and quadrupole moments), and the multipole Pauli repulsion model fits these data more precisely.

As stated in the Introduction, it is important that a repulsion model be interpretable in addition to being accurate. One simple measure of interpretability is the reasonableness of the fitted model parameters. To assess the sensibility of the parameters for the multipolar Pauli repulsion model, we calculated an atomic “size” for each atom class defined in Table I. The metric for size, presented in Table III, is the atomic radius corresponding to an atom-atom homodimer internuclear distance at which the repulsive energy reaches 1.0 kcal/mol. Because the multipolar Pauli repulsion model is anisotropic, we only include the charge-charge portion of the energy to cleanly separate the size from the orientational dependence.

**TABLE III.** Atomic “size” for multipolar Pauli repulsion atom classes. The radius is calculated as half the distance at which an atom-atom homodimer experiences 1.0 kcal/mol of repulsion energy. Only the charge-charge component of the repulsion is included. This is equivalent to the repulsion energy of the homodimer averaged over all possible dimer orientations at the standard distance.

N	Pauli repulsion class	Radius (Å)
1	H (nonpolar)	1.12
2	H (nonpolar, alkane)	1.16
3	H (polar, N-H/N aromatic)	1.01
4	H (polar, O-H)	1.03
5	H (aromatic, C-H)	1.00
6	H (polar, S-H)	1.08
7	C (sp <sup>3</sup> )	1.48
8	C (sp <sup>2</sup> , alkene)	1.73
9	C (sp <sup>2</sup> , C=O)	1.50
10	C (aromatic, C-C)	1.64
11	C (aromatic, C-X)	1.64
12	N (sp <sup>3</sup> )	1.58
13	N (sp <sup>2</sup> )	1.63
14	N (aromatic)	1.58
15	O (sp <sup>3</sup> , hydroxyl, water)	1.50
16	O (sp <sup>2</sup> , carbonyl)	1.64
17	O (O <sup>-</sup> in AcO <sup>-</sup> )	1.71
18	O (O <sup>-</sup> in HPO <sub>4</sub> <sup>-2</sup> )	1.75
19	O (O <sup>-</sup> in H <sub>2</sub> PO <sub>4</sub> <sup>-</sup> )	1.68
20	O (O in H <sub>3</sub> PO <sub>4</sub> )	1.66
21	P (phosphate)	1.55
22	S (sulfide, R-SH)	1.95
23	S (sulfur IV, DMSO)	2.02
24	F (organofluorine)	1.40
25	Cl (organochloride)	1.87
26	Br (organobromine)	2.05

Broadly, the sizes in Table III show a chemically intuitive picture of atomic size. The sizes follow periodic trends, and the differences across classes of the same element are reasonable. We note that although similar to the “size” (radius) parameter of the Lennard-Jones 12-6 or Halgren buffered 14-7 potentials, the size metric here should not be quantitatively compared. The size parameters in those van der Waals functions implicitly include the dispersion contribution in addition to repulsion.

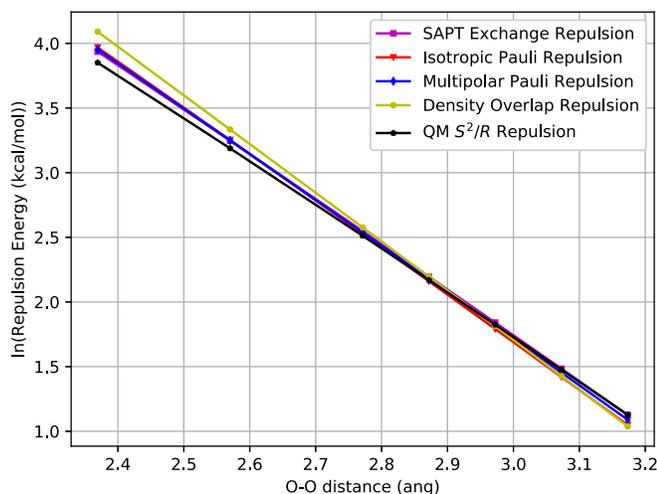
The S101x7 database provides extensive coverage for biomolecular chemical space and the radial dependence of interactions. The results of the fit show that using an exponential-based function matches this radial dependence better than the buffered 14-7 potential. The errors of the isotropic Pauli repulsion and vdW2017 models are similar for the closest (0.7×) points of the dataset. However, at distances just past equilibrium, the errors in the isotropic Pauli repulsion model become asymptotically smaller compared against the buffered 14-7 potential. This suggests radial scans of S101x7 are effective at determining the exponential parameter,  $\alpha$ , and the prefactor,  $K$ . The S101x7 database, however, contains relatively less orientational information. This requires us to carefully consider the charge,  $q$ , parameters for heavy atoms that are largely responsible for handling the angular dependence of the multipolar Pauli repulsion model. In the following test cases, we explore a variety of systems that specifically target the angular degrees of freedom that are less sampled by the S101 dataset.

### C. Water dimers

Water is an important case in force field development, for the reason that it is the solvent in which interesting biomolecular phenomena usually occur. In addition to being important for applications, water is also curious because of its anisotropic repulsive properties. To examine the performance of our model on this system, we shall consider three separate series of water dimers: one in which water dimer dissociation is considered, one in which the “flap angle” (defined in the inset of Fig. 6) of the water dimer is systematically varied, and one which consists of 10 well-studied independent stationary points on the water dimer surface.

Because water-water interactions are so important to the end goal of biomolecular simulations, the quality of the fit to the water dimer dissociation data of the S101x7 dataset is instructive. Figure 5 shows the performance of a number of repulsion models against reference SAPT2+ exchange repulsion energies plotted on a natural log scale.

The two Pauli repulsion models parameterized in this work compare quite well with the SAPT2+ result. Both the isotropic and multipolar Pauli repulsion models capture the magnitude of the interaction as well as the shape across the range of distances. This result is borne out by the S<sup>2</sup>/R comparison also shown in Fig. 5. Although this is less straightforward to compute for polyatomic molecules, we defined R as the O-O distance and computed S<sup>2</sup> according to Eq. (45) with each monomer’s MOs. This measure is also in very good agreement with the SAPT results, with the divergence

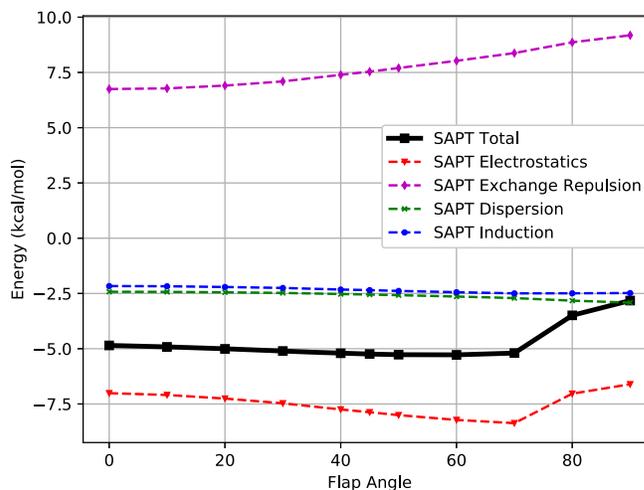


**FIG. 5.** Water dimer dissociation Pauli repulsion. Water monomer geometries were fixed at equilibrium and placed at distances from 0.7 to 1.1 times the equilibrium O–H distance. The SAPT exchange repulsion curve is almost entirely obscured by the multipole Pauli repulsion and isotropic Pauli repulsion curves. See text for definition of  $S^2/R$  and density overlap models.

at short range likely due to our neglect of the hydrogens in the definition of  $R$ .

Also shown in Fig. 5 are the results from the QM density overlap calculation. The proportionality constant,  $K_{ij}$ , for this model was chosen (as with the QM  $S^2/R$  model) to reproduce the SAPT repulsion energy at the equilibrium distance. One can see that the quantitative agreement is comparably good for this model. However, the density overlap model does show the same characteristic distance dependence problem illustrated for the noble gas dimers; it is slightly too repulsive at a short range and not repulsive enough at a long range. This erroneous distance dependence arises due to the unit consistency issue identified in Sec. II. It is worth noting that for practical simulation purposes, this error in the radial dependence may be tolerable, given other sources of error in a force field.

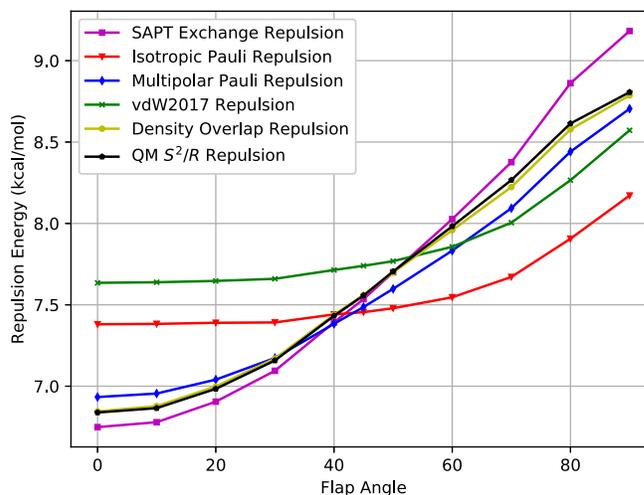
Another important slice of the water potential energy surface is the “flap angle” energy dependence of the water dimer.<sup>74,75</sup> High level *ab initio* calculations predict this angle ( $\theta$  in the inset of Fig. 6) to be  $57^\circ$ .<sup>76,77</sup> Typical 3- and 4-site point charge force fields for water such as TIP3P, SPC, and TIP4P generally predict a flap angle to be too flat<sup>78</sup> (less than  $57^\circ$ ) due to their inability to reproduce the molecular quadrupole moment of water. The opposite behavior was observed by Ren and Ponder when developing the original AMOEBA water model. Prior to their decision to scale down the quadrupole moments, the AMOEBA water model reproduced the molecular quadrupole moment very well but predicted a flap angle of  $70^\circ$ .<sup>31</sup> A scaling of the quadrupole moments by 70% served to correct the angle. Electrostatics, however, are only half of the story of the water dimer flap angle. Figure 6 shows that, in fact, *ab initio* electrostatics do strongly favor a flap angle of about  $70^\circ$ .



**FIG. 6.** Water dimer “flap angle” SAPT energy decomposition analysis. While the dispersion and induction components of the total energy are relatively flat across this slice of the potential energy landscape, the electrostatic and exchange repulsion change substantially and in opposite directions. The two trends largely cancel each other out in the total energy.

However, this does not correspond to the behavior of the total energy surface, which is basically flat for angles from  $45^\circ$  to  $70^\circ$ . To get this flat surface requires a compensating contribution from Pauli repulsion, and indeed Fig. 6 shows that as the flap angle is increased through this range, while the electrostatic energy consistently becomes more negative, the Pauli repulsion energy steadily trends more positive.

Which, if any, molecular mechanics models are capable of capturing this kind of phenomenon? Shown in Fig. 7 are several water dimer Pauli repulsion models evaluated for a range of flap angle values.



**FIG. 7.** Water dimer “flap angle” Pauli repulsion. The isotropic models (isotropic Pauli repulsion and vdW2017) clearly miss the sensitivity to angle change, while the anisotropic models mirror the shape of the SAPT exchange repulsion.

The multipolar Pauli repulsion model as well as the QM-based  $S^2/R$  and density overlap methods all reproduce the shape of the angular dependence of the water dimer repulsion well. The multipolar Pauli repulsion model, despite not being fit to any water angular dependence data (there are no angular scans in S101x7), reproduces the SAPT data quite well. This is due entirely to the natural description of anisotropy that comes through the atomic multipoles. The fact that the electrostatic and exchange repulsion curves in Fig. 6 are nearly mirror images for a fixed distance is not a coincidence. The interpretation of this result is that while the quadrupole interactions become more attractive as the flap angle is increased, this same overlap causes the repulsion to increase as well. It is the same underlying change in overlap that is driving both trends. This is again borne out by the *ab initio*  $S^2/R$  calculation (evaluated in the same way as before) which also mirrors the SAPT result. Interestingly, the density overlap model also does very well in describing this angular dependence. This result makes sense because, in contrast to the dissociation case, the distance between atoms for this slice of the surface is largely unchanged throughout the scan. This means that the density overlap distance dependence problem is hidden, while the accounting of anisotropy (in this case implicitly through the density) gives the correct angular trend.

What is apparent from Fig. 7, however, is that isotropic Pauli repulsion models cannot capture the flap angle dependence of the water dimer. Neither the isotropic Pauli repulsion model nor the vdW2017 model experiences any change in the repulsion energy until the flap angle becomes large enough that the hydrogen atoms of the acceptor molecule swing around to feel the repulsion of the donor oxygen. These models completely miss the quadrupole repulsion effect responsible for the shape of the flap angle repulsion curve.

The water dimer flap angle provides an excellent test case for cancellation of errors in advanced force fields. As force fields work to reproduce energy components individually, care must be taken to advance the physical models of each part in concert. If the electrostatic model is advanced to include anisotropy without including any such anisotropy in the repulsion, the electrostatics part of the force field will accurately capture that component of the energy. However, for the flap angle degree of freedom, this will incur an error in the total energy of over 2 kcal/mol over an area in which the total energy should be essentially flat. Figure 7 shows that no cancellation of error scheme for an isotropic repulsion model is sensitive to this degree of freedom; the only way to correct it is to include anisotropy in the repulsion as well. This is the reason why the original AMOEBA water model deviated from the physically derived electrostatic model and scaled down the quadrupoles. Having a fully anisotropic Pauli repulsion function means that these components sit at the same level of theory, and this in turn allows us to regain sensitivity to cancellation of errors in the angular degrees of freedom.

The angular dependence of Pauli repulsion is not only apparent in the minimum energy water dimer. We also

considered the ten water dimer structures introduced by Tschumper *et al.*<sup>77</sup> Figure 8 shows the error, relative to the SAPT2+ exchange repulsion energy for the multipolar Pauli repulsion, isotropic Pauli repulsion, and vdW2017 models.

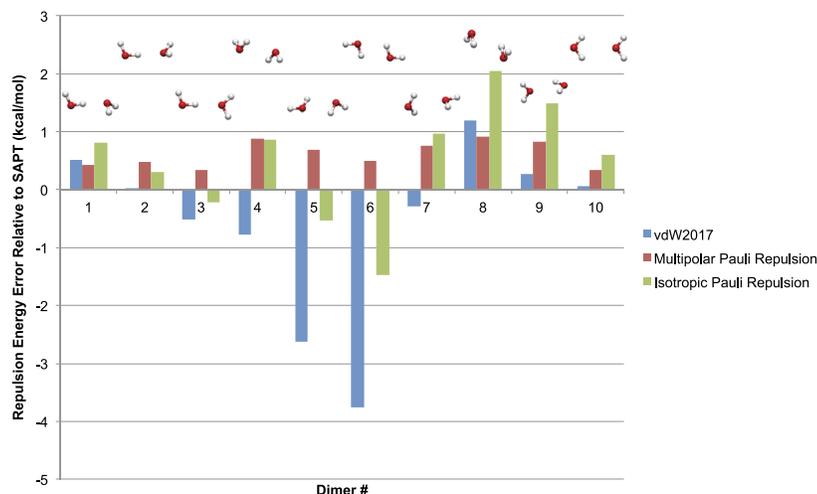
The multipolar Pauli repulsion model displays errors of less than 1 kcal/mol for every dimer configuration. The two isotropic models, however, suffer from large, nonrandom errors on several of the dimers. Both the vdW2017 and isotropic Pauli repulsion models feature large and opposing errors on dimers 6 and 8. This indicates an angular repulsion dependence that an isotropic model is incapable of capturing. In fact, we attempted to fit the isotropic Pauli repulsion model directly to exchange repulsion energies of the ten water dimers and found the same opposing errors for dimers 6 and 8.

Taken as a whole, all of the data presented for water repulsion interactions tell a consistent story that not including repulsion anisotropy on the water dimer potential energy surface will incur errors of 1-2 kcal/mol for accessible dimer configurations. It also shows that the multipolar Pauli repulsion model is capable of bringing those errors down to ~0.5 kcal/mol. The ability of the model to predict angular dependence that is not in the fitting set, such as the flap angle and dimers 2-10, suggests that the electrostatic multipole description is a natural fit for Pauli repulsion anisotropy.

#### D. The “sigma hole” effect

As stated in the Introduction, halogen bonding *vis a vis* the so-called “sigma hole” effect is of particular interest to biomolecular force fields. Over 35% of drugs in clinical phase III trials contain at least one halogen atom.<sup>10</sup> The “sigma hole” terminology refers to the area of positive charge found at the distal tip of the halogen atom in a halogen-containing compound. It has long been accepted that this feature suggests that the linear halogen B...X-Y bonding geometry characteristic of the “sigma hole” effect is driven by electrostatics.<sup>79</sup> Anthony Stone showed that this assumption is only partially correct. Using simple model systems, Stone showed that while electrostatics is indeed responsible for the overall attraction that causes halogen bonds to form, Pauli repulsion is largely responsible for the characteristic, often linear, geometry of these bonds.<sup>4</sup> Given the importance of halogen bond interactions, we chose to consider a range of test systems to assess the quality of the multipolar Pauli repulsion model. We consider a pair of representative examples from Stone’s work, a halobenzene example proposed by Nohad Gresh and co-workers,<sup>13</sup> an acetone-bromobenzene dimer suggested by Hobza and co-workers,<sup>80</sup> and a drug-like dimer system from Alzate-Morales and co-workers.<sup>12</sup>

From Stone’s work, we consider two representative halogen bonding configurations: the “head-on” ammonia-ClF dimer and the “from the side” ethene-ClF dimer. Figures 9 and 10 show the results for the models on ammonia-ClF and ethene-ClF dimer, respectively.

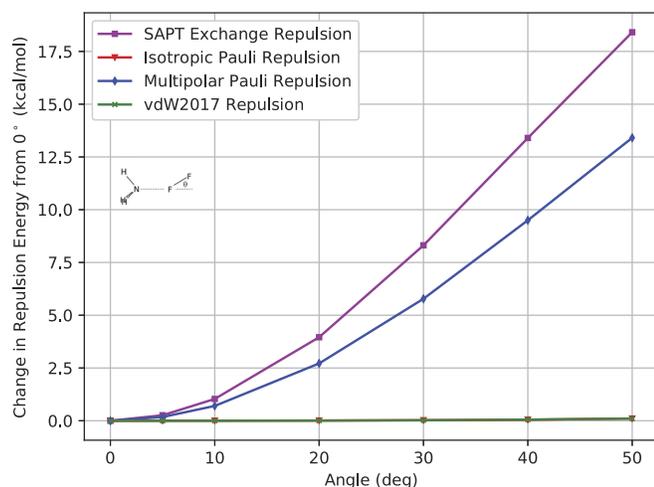


**FIG. 8.** Pauli repulsion energy error for ten water dimers. The error (SAPT minus model) is plotted for each configuration. The isotropic models exhibit errors of opposing sign for dimers 5 and 6 vs. 8 and 9. The multipolar Pauli repulsion model does not suffer from this constraint.

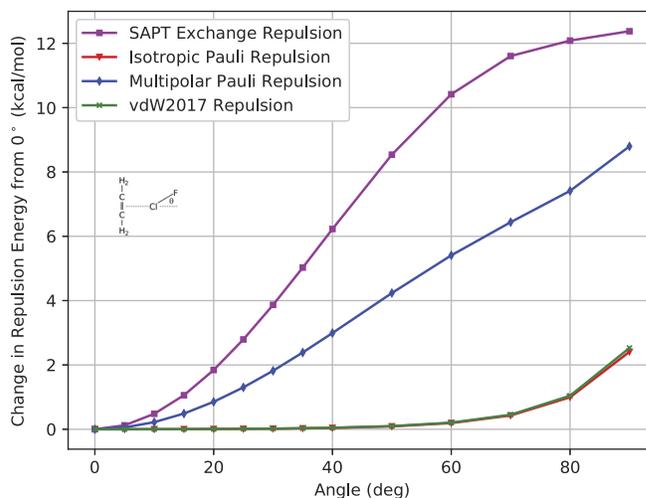
As the  $B \cdots X-Y$  angle varies away from linear for both systems, the *ab initio* Pauli repulsion rises sharply. The isotropic models miss this entirely. For ammonia-ClF, the repulsion energy of the isotropic Pauli repulsion and vdW2017 models is virtually flat throughout the scan, and for ethene-ClF, these models only start to vary once the fluorine swings around far enough to interact directly with the ethene molecule. This indicates that isotropic Pauli repulsion models will miss the strong linear preference of these halogen-bonded complexes. The multipolar Pauli repulsion model, however, does not miss this angular effect. In both cases, the anisotropic model correctly captures the immediate increase in Pauli repulsion that occurs as the halogen bond deviates from linearity. Although the multipole Pauli repulsion model

underestimates the anisotropy of repulsion in both examples, this is most likely a consequence of the DMA-based protocol used for determining multipole moments. For Cl-F, the DMA Cl dipole and quadrupole moments differ from those of similar halogens in the S101 database. This is likely responsible for the difference since the angular dependence of repulsion is driven by dipole and quadrupole interactions. However, the general qualitative agreement between the SAPT and the multipolar Pauli repulsion model shows that a multipole-based description of electrostatics works well to describe the angular dependence of halogen bond repulsion in these cases.

The concept that force field anisotropy is necessary to accurately model halogen bonding is not new. Recent studies



**FIG. 9.** Variation of the Pauli repulsion energy with respect to tilt angle ( $B \cdots X-Y$ ) for the ammonia-ClF dimer. The N to Cl distance is fixed at 2.376 Å. The isotropic Pauli repulsion and vdW2017 lines are indistinguishable.



**FIG. 10.** Variation of the Pauli repulsion energy with respect to tilt angle ( $B \cdots X-Y$ ) for the ethene-ClF dimer. The C=C bond midpoint to Cl distance is fixed at 2.766 Å.

using the AMOEBA force field<sup>81</sup> as well as the SIBFA force field<sup>13</sup> have explored this idea. Both studies stressed the importance of anisotropic electrostatics but largely neglected a discussion about anisotropic repulsion. In particular, the work of Gresh and co-workers studied the interactions of halobenzene–water complexes. To examine the repulsive contribution to these interactions, we chose the chlorobenzene–water dimer as a test system. Figure 11 shows the halogen bond angle scan for this system.

Although the energy variation is smaller for this system since the O to Cl contact distance is long, at 3.33 Å, the model trends are clear. The multipolar Pauli repulsion model mirrors the immediate change in repulsion energy that is felt when the complex deviates from linear. The two isotropic models, however, do not sense this effect. The energy of these models is flat until a rotation of  $\sim 60^\circ$ , where steric repulsion begins.

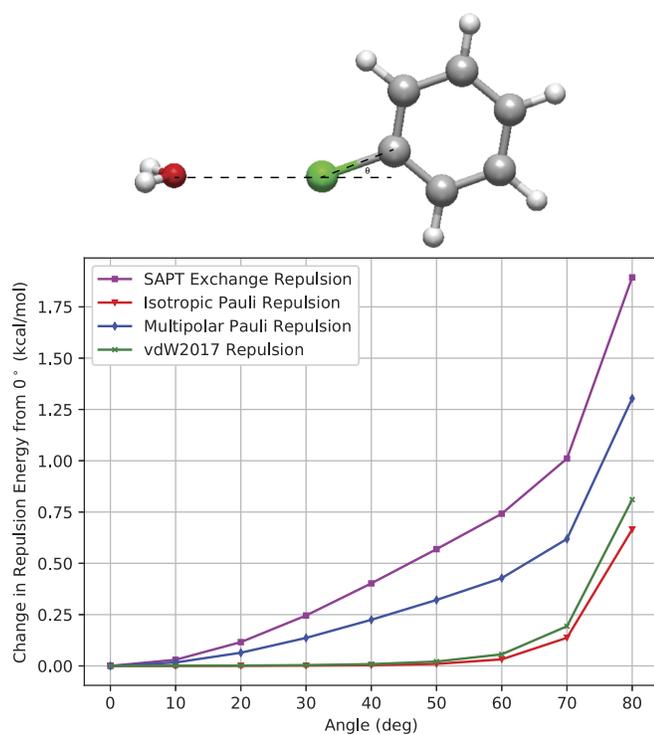
Another useful test system for halogen bonding is bromobenzenes interacting with acetone. A crystallographic survey by Auffinger and co-workers showed that of the halogen bonded structures in the PDB, 70% involved a protein backbone carbonyl oxygen and that of those structures, 94% involved a halogen atom bonded to an aromatic or heterocyclic aromatic ring.<sup>14</sup> Hobza and co-workers proposed the bromobenzene–acetone complex as a simple probe for examining this kind of important halogen bonding. We used this probe to assess the quality of the carbonyl oxygen containing

halogen bond behavior of the multipolar Pauli repulsion model. Shown in Fig. 12 is an angular scan of the acetone–bromobenzene Pauli repulsion energy surface.

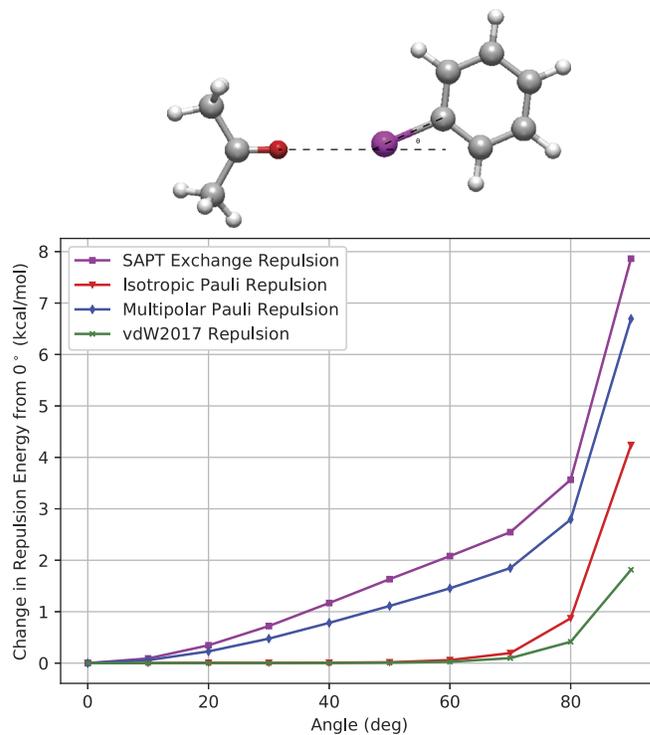
Clearly, while the multipolar repulsion model is not in perfect agreement with SAPT, it is the only model that captures the angular dependence trend. The preference the acetone–bromobenzene system shows for the linear configuration is being driven in no small part by this angular dependence in Pauli repulsion. The multipolar Pauli repulsion model gets this dependence qualitatively right. As expected, the isotropic models miss this variation in the rotational degrees of freedom.

The final example of halogen bonding we surveyed was a model “drug binding” system of *N*-methylacetamide (NMA) and chlorobenzene proposed by Alzate-Morales and co-workers.<sup>12</sup> This system was chosen to be a close approximation of a drug-like molecule interacting with a peptide backbone. Again, we performed SAPT calculations of the exchange repulsion energy at a range of interaction angles and compared these to the results from each model.

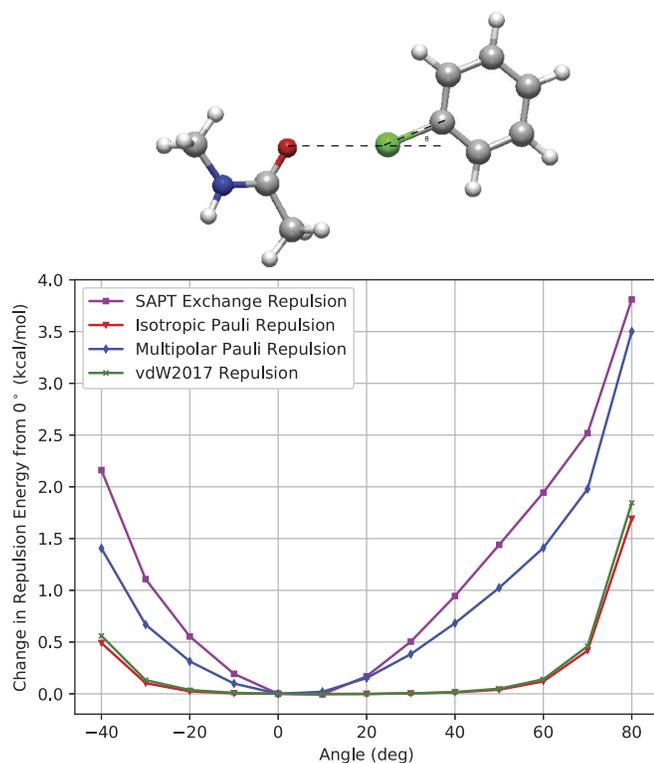
The results in Fig. 13 confirm the trend of the other test cases. The multipolar Pauli repulsion model correctly picks up the trends in both angular directions for the repulsion energy. Since NMA is an asymmetric molecule, we performed a scan from 80 to  $-40^\circ$  and found that the multipolar Pauli repulsion model reproduces the increase in repulsion energy for both positive and negative angular deviations. The isotropic



**FIG. 11.** Variation of the Pauli repulsion energy with respect to the tilt angle ( $B \cdots X-Y$ ) for the water–chlorobenzene dimer. The O to Cl distance is fixed at 3.33 Å.



**FIG. 12.** Variation of the Pauli repulsion energy with respect to the tilt angle ( $B \cdots X-Y$ ) for the acetone–bromobenzene dimer. The O to Br distance is fixed at 3.15 Å.



**FIG. 13.** Variation of the Pauli repulsion energy with respect to the tilt angle ( $B \cdots X-Y$ ) for the NMA-chlorobenzene dimer. The O to Cl distance is fixed at 3.0 Å.

models both capture the beginning of steric repulsion that occurs at each end of the angular scan but are not sensitive to the anisotropic change in repulsion in the middle. The anisotropy in the multipolar Pauli repulsion model is picking up not only the increase in energy associated with rotating away from linear but also the asymmetry about the  $O \cdots Cl-C$  angle.

Much like hydrogen bonds, halogen bonds can be useful tools for molecular design because they are strong and exhibit a marked geometric preference. As has been shown in previous work, much of this strength and some of the

geometry preference are expressed through the electrostatic, dispersion, and polarization components of the intermolecular energy. However, the component most responsible for enforcing the generally linear geometry of halogen bonds is Pauli repulsion. The tests presented in this section show that no isotropic model (without employing off-atom repulsion sites) is capable of reproducing this effect. Moreover, this section shows that the electrostatic description of monomers via atomic multipole expansions is sufficient to capture the signature anisotropy of these interactions. It bears noting that despite the repetitive feel of the angular halogen bonding results shown, there are no anisotropic parameters being fit. The differences in anisotropy, including the asymmetric shape of the NMA-chlorobenzene well, are entirely determined by the *ab initio* derived atomic multipole moments of the molecules. The results here show that for “sigma hole” type interactions, the multipole moments can simultaneously provide a suitable description of both electrostatic and repulsion anisotropy.

### E. Computational cost

For any molecular mechanics model that aims to be useful for biomolecular simulation, the computational cost must be considered. In particular, for the multipolar Pauli repulsion model, this is a matter of concern because multipole calculations are known to be computationally expensive. This model is intended to be used in tandem with a multipole electrostatics model; in particular, it is parameterized against the AMOEBA multipole model, so we tested the computational efficiency in that context. The results in Table IV show that the additional cost for this model is minimal.

When the multipolar Pauli repulsion model is paired with our previously published damped dispersion model, the resulting calculations are around 20% slower than the current standard AMOEBA buffered 14-7 van der Waals function. Furthermore, when this cost is put into the context of the entire AMOEBA energy function, including polarization, the extra cost becomes nearly negligible. The multipolar Pauli repulsion and overlap damped dispersion combination yields a model that is 5% slower than its buffered 14-7 counterpart. Polarization is the costliest component of the AMOEBA force field, so adding a slightly more expensive Pauli repulsion

**TABLE IV.** Computational cost of the multipolar Pauli repulsion model. Timings are for 100 energy and force evaluations in a standard Open-MP parallel implementation in Tinker.

	Time for 100 energy and force evaluations (s)
AMOEBA electrostatics with charge penetration + multipolar Pauli repulsion + dispersion	1.7
AMOEBA electrostatics with charge penetration + vdW2017	1.4
AMOEBA electrostatics with charge penetration + polarization + multipolar Pauli repulsion + dispersion	4.6
AMOEBA electrostatics with charge penetration + polarization + vdW2017	4.4

function makes very little difference to the overall computational efficiency of the model.

We note here that this kind of computational efficiency is predicated upon two important factors. First, it relies upon using a multipole model for electrostatics. If one were to deploy the multipolar Pauli repulsion model as a stand-alone energy term, it would be an order of magnitude more expensive than a standard van der Waals function. When it is used with a multipole model, and given that the multipole moments are constrained to be identical for the electrostatics and Pauli repulsion models, a large amount of the algebra to compute dipole and quadrupole forces is shared between the two models. Second, the speed of the model benefits greatly from employing cutoffs. The standard van der Waals cutoff for the AMOEBA force field is 9–12 Å. Because the multipolar Pauli repulsion model separates the repulsive and dispersive contributions to the van der Waals energy, it is free to use a much shorter cutoff. These tests were performed with a conservative truncated cutoff of 5 Å, but even better performance can be achieved by shortening this distance and employing a polynomial switching function.

## V. DISCUSSION AND CONCLUSIONS

Pauli repulsion is one of the most important parts of any classical intermolecular potential energy model. In most energy decomposition analyses, it is the only component of the total energy that is always positive. This means that for condensed phase systems, total models rely heavily on the Pauli repulsion term to reproduce bulk phase data. For this reason, it is important for a good Pauli repulsion model to be both accurate and physically interpretable. We have presented here the multipolar Pauli repulsion model as an option that fulfills both of these aims.

Despite the terms in which it is often discussed, there is nothing mystical about the phenomenon of Pauli repulsion. It is true that the effect arises from the enforcement of the laws of quantum mechanics, but the result can be fully understood within a classical physics interpretation. The Pauli exclusion principle demands the total wavefunction of an electronic system be antisymmetric. If we take as our reference the unperturbed monomer wavefunctions of two interacting molecules, then upon overlap, the enforcement of antisymmetry will lead to a loss in electron density in the overlap region. That loss of electron density, relative to the unperturbed reference state, causes a straightforward Coulombic repulsion between the two nuclei. The multipolar Pauli repulsion model presented here follows this explanation and thus gives a classical physical interpretation of this quantum mechanical effect. The electrostatic nature of this effect is borne out by the success of the model in reproducing the anisotropy of repulsion using electrostatic multipole moments.

In addition to being physically interpretable, the multipolar Pauli repulsion model is shown to yield good quantitative fits to *ab initio* data. The model fits the S101x7 dataset to an accuracy of 1.7 kcal/mol and fits the near-equilibrium points of that dataset to an error of much less than 1 kcal/mol. This

fit spans a large range of chemical space, from hydrogen bonds and halogen bonds to pi-pi interactions and charged species. The results are shown to be transferable to systems outside of the S101 set as well. Particularly, we have shown that the multipolar Pauli repulsion model captures the angular dependence of the repulsion energy associated with halogen bonding at a range of contact distances. These test systems not only show the transferability of the exponential parameters that were fit but also justify the claim to atomic multipole parameters as a description of anisotropic electron distribution overlap.

This work is certainly not the first to acknowledge the importance of the anisotropy of repulsion in intermolecular interactions, and the multipolar Pauli repulsion model is not the first model to include such effects. What makes this model noteworthy is that it circumvents two obstacles that have traditionally stood in the way of adopting anisotropic models: parameter underdetermination and computational cost. Any atomic anisotropic model (repulsive or otherwise) requires a local frame and a set of parameters that obey the symmetries of that frame. In the absence of any richer set of data, these requirements mean that if one wishes to fit to intermolecular energies, there will be a large number of parameters to fit to a (usually small, depending on computational resources) set of scalar values. This is a recipe for overfitting, and it is the reason that anisotropic repulsive models have largely been limited to specific systems for which large amounts of dimer data can be generated. The multipolar Pauli repulsion model evades this problem by constraining the parameters responsible for conferring anisotropy on the model to those that are derived from a much richer data source: the molecular density. The atomic multipolar parameters that are derived from *ab initio* monomer calculations not only constrain the parameter space of the model to avoid overfitting, but moreover, as shown in Sec. II, they do so through a series of theoretically justified approximations. The DMA multipoles come from the same monomer wavefunction that is required to calculate  $S^2$ , and this model uses that information to its advantage. This symbiosis not only stands the multipolar Pauli repulsion model on solid theoretical ground but also ameliorates the concern of computational cost typically associated with anisotropic models. Because the local frames and atomic multipoles are identical between the electrostatics and repulsion models, there is very little additional overhead incurred when using the multipolar Pauli repulsion model with a multipolar electrostatics model. Timings show that the cost of implementing a multipole-based repulsion model in an AMOEBA-like force field is minimal. By avoiding the overfitting and cost problems that have proved prohibitive, the multipolar Pauli repulsion model presented here provides a blueprint of one tractable way to include anisotropic repulsion in biomolecular force fields.

Not every force field needs the level of detail presented in this model. There are undoubtedly applications for which isotropic, point-based force fields are adequate for predicting quantities of interest. In fact, even for calculations that will require more advanced force fields, the multipolar Pauli repulsion model, as presented here, is probably

insufficient without additional tuning—several of the test cases presented, despite qualitative agreement with SAPT, fall short of accurately reproducing a truly *ab initio* potential energy surface. All force fields, advanced or not, will require some measure of error cancellation.

A final point of this paper is that in order to benefit from cancellation of errors the level of detail across different parts of the model must match. The water dimer data presented here show this point nicely. If one uses an anisotropic multipolar description of electrostatics that, inevitably, has some error in the associated intermolecular angular degrees of freedom, the only way to cancel that error is by having the other components of the force field be sensitive to those same degrees of freedom. The water dimer example shows that for this important case (and likely many others) it is largely the Pauli repulsion that provides the balancing force. This is specific evidence of the broader truism in molecular modeling that a theoretically “consistent” model is a good model. For a point charge force field, it is possible that including an anisotropic repulsion model might make the model worse by introducing error that cannot be canceled by other components of the force field. Likewise, for force fields based on multipolar electrostatics models, we suggest that the multipolar Pauli repulsion model will not just be more accurate with respect to energy decomposition analysis, and it will also confer the ability to achieve favorable cancellation of error across the total model.

While the theoretical framework presented here is applicable to any multipolar force field, the specifics of the parameterization and testing of the model are aimed at a particular goal. The development of the next generation of the AMOEBA force field is underway, and the multipolar Pauli repulsion model has been constructed explicitly for that purpose. It is intended to be used with a multipolar description of electrostatics that includes our earlier work on charge penetration and in conjunction with our previously published overlap damped dispersion model. The code that implements all of these components in the Tinker molecular mechanics software package is freely available on the web.<sup>70</sup> Work combining all of these components into a next-generation water model and full biomolecular force field will be reported in due course. The multipolar Pauli repulsion model provides a cheap, intuitive, and interpretable way to put this important component of the future force field on an equal footing with its counterparts.

## SUPPLEMENTARY MATERIAL

See [supplementary material](#) for tables containing the isotropic and multipolar model parameters fit to the S101 database.

## ACKNOWLEDGMENTS

J.W.P. would like to thank the National Institutes of Health for support of the development and extension of the AMOEBA model via Grant Nos. R01 GM106137 and R01 GM114237.

## REFERENCES

- 1R. F. W. Bader, “Pauli repulsions exist only in the eye of the beholder,” *Chem. - Eur. J.* **12**, 2896–2901 (2006).
- 2P. Politzer, J. S. Murray, and T. Clark, “Mathematical modeling and physical reality in noncovalent interactions,” *J. Mol. Model.* **21**, 52 (2015).
- 3B. M. Deb, *The Force Concept in Chemistry* (Van Nostrand Reinhold, 1981).
- 4A. J. Stone, “Are halogen bonded structures electrostatically driven?,” *J. Am. Chem. Soc.* **135**, 7005–7009 (2013).
- 5M. L. Laury, L.-P. Wang, V. S. Pande, T. Head-Gordon, and J. W. Ponder, “Revised parameters for the AMOEBA polarizable atomic multipole water model,” *J. Phys. Chem. B* **119**, 9423–9437 (2015).
- 6Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, “Polarizable atomic multipole-based AMOEBA force field for proteins,” *J. Chem. Theory Comput.* **9**, 4046–4063 (2013).
- 7C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder, and P. Ren, “AMOEBA polarizable atomic multipole force field for nucleic acids,” *J. Chem. Theory Comput.* **14**, 2084–2108 (2018).
- 8M. L. Laury, Z. Wang, A. S. Gordon, and J. W. Ponder, “Absolute binding free energies of the SAMPL6 cucurbit[8]uril host-guest challenge via the AMOEBA polarizable force field,” *J. Comput.-Aided Mol. Des.* **32**, 1087–1095 (2018).
- 9S. Rybak, B. Jeziorski, and K. Szalewicz, “Many-body symmetry-adapted perturbation theory of intermolecular interactions. H<sub>2</sub>O and HF dimers,” *J. Chem. Phys.* **95**, 6576–6601 (1991).
- 10Z. Xu, Z. Yang, Y. Liu, Y. Lu, K. Chen, and W. Zhu, “Halogen bond: Its role beyond drug-target binding affinity for drug discovery and development,” *J. Chem. Inf. Model.* **54**, 69–78 (2014).
- 11J. Grant Hill and A. C. Legon, “On the directionality and non-linearity of halogen and hydrogen bonds,” *Phys. Chem. Chem. Phys.* **17**, 858–867 (2015).
- 12F. Adasme-Carreño, C. Muñoz-Gutierrez, and J. H. Alzate-Morales, “Halogen bonding in drug-like molecules: A computational and systematic study of the substituent effect,” *RSC Adv.* **6**, 61837–61847 (2016).
- 13K. El Hage, J.-P. Piquemal, Z. Hobaika, R. G. Maroun, and N. Gresh, “Could an anisotropic molecular mechanics/dynamics potential account for sigma hole effects in the complexes of halogenated compounds?,” *J. Comput. Chem.* **34**, 1125–1135 (2013).
- 14P. Auffinger, F. A. Hays, E. Westhof, and P. S. Ho, “Halogen bonds in biological molecules,” *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16789–16794 (2004).
- 15J. A. Rackers, Q. Wang, C. Liu, J.-P. Piquemal, P. Ren, and J. W. Ponder, “An optimized charge penetration model for use with the AMOEBA force field,” *Phys. Chem. Chem. Phys.* **19**, 276–291 (2017).
- 16J. D. Van der Waals, *Over de Continuïteit* (van den Gas-en Vloeistofoestand, Leiden, 1873).
- 17W. Pauli, “Über den zusammenhang des abschlusses der elektronengruppen im atom mit der komplexstruktur der spektren,” *Z. Phys.* **31**, 765–783 (1925).
- 18J. E. Jones, “On the determination of molecular fields. II. From the equation of state of a gas,” *Proc. R. Soc. London, Ser. A* **106**, 463–477 (1924).
- 19J. E. Lennard-Jones, “Cohesion,” *Proc. Phys. Soc.* **43**, 461–482 (1931).
- 20F. London, “Zur theorie und systematik der molekularkräfte,” *Z. Phys.* **63**, 245–279 (1930).
- 21J. A. McCammon, B. R. Gelin, and M. Karplus, “Dynamics of folded proteins,” *Nature* **267**, 585–590 (1977).
- 22M. Born and J. E. Mayer, “Zur gittertheorie der ionenkristalle,” *Z. Phys.* **75**, 1–18 (1932).
- 23R. A. Buckingham, “The classical equation of state of gaseous helium, neon and argon,” *Proc. R. Soc. London, Ser. A* **168**, 264–283 (1938).
- 24J. C. Slater, “The normal state of helium,” *Phys. Rev.* **32**, 349–360 (1928).
- 25N. L. Allinger, “Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms,” *J. Am. Chem. Soc.* **99**, 8127–8134 (1977).
- 26N. L. Allinger, Y. H. Yuh, and J.-H. Lii, “Molecular mechanics. The MM3 force field for hydrocarbons. I,” *J. Am. Chem. Soc.* **111**, 8551–8566 (1989).

- <sup>27</sup>N. L. Allinger, K. Chen, and J.-H. Lii, "An improved force field (MM4) for saturated hydrocarbons," *J. Comput. Chem.* **17**, 642–668 (1996).
- <sup>28</sup>T. A. Halgren, "The representation of van der Waals (vdW) interactions in molecular mechanics force fields: Potential form, combination rules, and vdW parameters," *J. Am. Chem. Soc.* **114**, 7827–7843 (1992).
- <sup>29</sup>T. A. Halgren, "Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94," *J. Comput. Chem.* **17**, 490–519 (1996).
- <sup>30</sup>T. A. Halgren, "Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions," *J. Comput. Chem.* **17**, 520–552 (1996).
- <sup>31</sup>P. Ren and J. W. Ponder, "Polarizable atomic multipole water model for molecular mechanics simulation," *J. Phys. Chem. B* **107**, 5933–5947 (2003).
- <sup>32</sup>C. Millot and A. J. Stone, "Towards an accurate intermolecular potential for water," *Mol. Phys.* **77**, 439–462 (1992).
- <sup>33</sup>A. J. Misquitta and A. J. Stone, "Ab initio atom–atom potentials using Cam-CASP: Theory and application to many-body models for the pyridine dimer," *J. Chem. Theory Comput.* **12**, 4184–4208 (2016).
- <sup>34</sup>E. M. Mas, R. Bukowski, K. Szalewicz, G. C. Groenenboom, P. E. S. Wormer, and A. van der Avoird, "Water pair potential of near spectroscopic accuracy. I. Analysis of potential surface and virial coefficients," *J. Chem. Phys.* **113**, 6687–6701 (2000).
- <sup>35</sup>T. S. Totton, A. J. Misquitta, and M. Kraft, "A first principles development of a general anisotropic potential for polycyclic aromatic hydrocarbons," *J. Chem. Theory Comput.* **6**, 683–695 (2010).
- <sup>36</sup>L. Salem, "The forces between polyatomic molecules. II. Short-range repulsive forces," *Proc. R. Soc. London, Ser. A* **264**, 379–391 (1961).
- <sup>37</sup>J. I. Musher and L. Salem, "Energy of interaction between two molecules," *J. Chem. Phys.* **44**, 2943–2946 (1966).
- <sup>38</sup>J. N. Murrell, M. Randić, and D. R. Williams, "The theory of intermolecular forces in the region of small orbital overlap," *Proc. R. Soc. London, Ser. A* **284**, 566–581 (1965).
- <sup>39</sup>J. N. Murrell and G. Shaw, "Intermolecular forces in the region of small orbital overlap," *J. Chem. Phys.* **46**, 1768–1772 (1967).
- <sup>40</sup>N. Gresh, "Energetics of Zn<sup>2+</sup> binding to a series of biologically relevant ligands: A molecular mechanics investigation grounded on *ab initio* self-consistent field supermolecular computations," *J. Comput. Chem.* **16**, 856–882 (1995).
- <sup>41</sup>J.-P. Piquemal, H. Chevreau, and N. Gresh, "Toward a separate reproduction of the contributions to the Hartree-Fock and DFT intermolecular interaction energies by polarizable molecular mechanics with the SIBFA potential," *J. Chem. Theory Comput.* **3**, 824–837 (2007).
- <sup>42</sup>J. H. Jensen and M. S. Gordon, "An approximate formula for the intermolecular Pauli repulsion between closed shell molecules," *Mol. Phys.* **89**, 1313–1325 (1996).
- <sup>43</sup>J. H. Jensen and M. S. Gordon, "An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. II. Application to the effective fragment potential method," *J. Chem. Phys.* **108**, 4772–4782 (1998).
- <sup>44</sup>S. Kita, K. Noda, and H. Inouye, "Repulsive potentials for Cl–R and Br–R (R = He, Ne, and Ar) derived from beam experiments," *J. Chem. Phys.* **64**, 3446–3449 (1976).
- <sup>45</sup>Y. S. Kim, S. K. Kim, and W. D. Lee, "Dependence of the closed-shell repulsive interaction on the overlap of the electron densities," *Chem. Phys. Lett.* **80**, 574–575 (1981).
- <sup>46</sup>R. J. Wheatley and S. L. Price, "An overlap model for estimating the anisotropy of repulsion," *Mol. Phys.* **69**, 507–533 (1990).
- <sup>47</sup>J.-P. Piquemal, G. A. Cisneros, P. Reinhardt, N. Gresh, and T. A. Darden, "Towards a force field based on density fitting," *J. Chem. Phys.* **124**, 104101 (2006).
- <sup>48</sup>R. E. Duke, O. N. Starovoytov, J.-P. Piquemal, and G. A. Cisneros, "GEM\*: A molecular electronic density-based force field for molecular dynamics simulations," *J. Chem. Theory Comput.* **10**, 1361–1365 (2014).
- <sup>49</sup>H. Gokcan, E. G. Kratz, T. A. Darden, J.-P. Piquemal, and G. A. Cisneros, "QM/MM simulations with the Gaussian electrostatic model, a density-based polarizable potential," *J. Phys. Chem. Lett.* **9**, 3062–3067 (2018).
- <sup>50</sup>G. A. Cisneros, "Application of Gaussian electrostatic model (GEM) distributed multipoles in the AMOEBA force field," *J. Chem. Theory Comput.* **8**, 5072–5080 (2012).
- <sup>51</sup>M. J. Van Vleet, A. J. Misquitta, and J. R. Schmidt, "New angles on standard force fields: Toward a general approach for treating atomic-level anisotropy," *J. Chem. Theory Comput.* **14**, 739–758 (2018).
- <sup>52</sup>M. J. Van Vleet, A. J. Misquitta, A. J. Stone, and J. R. Schmidt, "Beyond Born-Mayer: Improved models for short-range repulsion in *ab initio* force fields," *J. Chem. Theory Comput.* **12**, 3851–3870 (2016).
- <sup>53</sup>I. Nobeli, S. L. Price, and R. J. Wheatley, "Use of molecular overlap to predict intermolecular repulsion in N···H–O hydrogen bonds," *Mol. Phys.* **95**, 525–537 (1998).
- <sup>54</sup>J. B. O. Mitchell, J. M. Thornton, J. Singh, and S. L. Price, "Towards an understanding of the arginine-aspartate interaction," *J. Mol. Biol.* **226**, 251–262 (1992).
- <sup>55</sup>J. B. O. Mitchell and S. L. Price, "The nature of the N–H···O=C hydrogen bond: An intermolecular perturbation theory study of the formamide/formaldehyde complex," *J. Comput. Chem.* **11**, 1217–1233 (1990).
- <sup>56</sup>G. M. Day and S. L. Price, "A nonempirical anisotropic atom–atom model potential for chlorobenzene crystals," *J. Am. Chem. Soc.* **125**, 16434–16443 (2003).
- <sup>57</sup>M. Tafipolsky and K. Ansgor, "Toward a physically motivated force field: Hydrogen bond directionality from a symmetry-adapted perturbation theory perspective," *J. Chem. Theory Comput.* **12**, 1267–1279 (2016).
- <sup>58</sup>C. Domene, P. W. Fowler, M. Wilson, P. A. Madden, and R. J. Wheatley, "Overlap-model and *ab initio* cluster calculations of ion properties in distorted environments," *Chem. Phys. Lett.* **333**, 403–412 (2001).
- <sup>59</sup>L. V. Slipchenko and M. S. Gordon, "Electrostatic energy in the effective fragment potential method: Theory and application to benzene dimer," *J. Comput. Chem.* **28**, 276–291 (2007).
- <sup>60</sup>C. A. Coulson, "Two-centre integrals occurring in the theory of molecular structure," *Math. Proc. Cambridge* **38**, 210–223 (1942).
- <sup>61</sup>H. Inouye and S. Kita, "Experimental determination of the repulsive potentials between K<sup>+</sup> ions and rare-gas atoms," *J. Chem. Phys.* **56**, 4877–4882 (1972).
- <sup>62</sup>P. Söderhjelm, G. Karlström, and U. Ryde, "Comparison of overlap-based models for approximating the exchange-repulsion energy," *J. Chem. Phys.* **124**, 244101 (2006).
- <sup>63</sup>C. Nyeland and J. P. Toennies, "Modelling of repulsive potentials from atom charge density distributions: Interactions of inert gas atoms," *Chem. Phys. Lett.* **127**, 172–177 (1986).
- <sup>64</sup>E. Andreev, "On asymptotic calculation of the exchange interaction," *Theor. Chim. Acta* **28**, 235–239 (1973).
- <sup>65</sup>Q. Wang, J. A. Rackers, C. He, R. Qi, C. Narth, L. Lagardere, N. Gresh, J. W. Ponder, J.-P. Piquemal, and P. Ren, "General model for treating short-range electrostatic penetration in a molecular mechanics force field," *J. Chem. Theory Comput.* **11**, 2609–2618 (2015).
- <sup>66</sup>J. A. Rackers, C. Liu, P. Ren, and J. W. Ponder, "A physically grounded damped dispersion model with particle mesh Ewald summation," *J. Chem. Phys.* **149**, 084115 (2018).
- <sup>67</sup>T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, and C. D. Sherill, "Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies," *J. Chem. Phys.* **140**, 094106 (2014).
- <sup>68</sup>R. Qi, Q. Wang, and P. Ren, "General van der Waals potential for common organic molecules," *Bioorg. Med. Chem.* **24**, 4911–4919 (2016).
- <sup>69</sup>P. Ren, C. Wu, and J. W. Ponder, "Polarizable atomic multipole-based molecular mechanics for organic molecules," *J. Chem. Theory Comput.* **7**, 3143–3161 (2011).

- <sup>70</sup>J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardere, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, "Tinker 8: Software tools for molecular design," *J. Chem. Theory Comput.* **14**, 5273–5289 (2018).
- <sup>71</sup>J. A. Rackers, "GitHub branch for tinkertools development," [updated 2018; cited], Available from: <https://github.com/JoshRackers/tinker/tree/amoeba2>, 2018.
- <sup>72</sup>J. W. Ponder, P. Ren, and J.-P. Piquemal, "GitHub site for tinkertools," [updated 2018; cited], Available from: <https://github.com/TinkerTools>, 2018.
- <sup>73</sup>D. G. A. Smith, L. A. Burns, D. A. Sirianni, D. R. Nascimento, A. Kumar, A. M. James, J. B. Schriber, T. Zhang, B. Zhang, and A. S. Abbott, "Psi4NumPy: An interactive quantum chemistry programming environment for reference implementations and rapid development," *J. Chem. Theory Comput.* **14**, 3504–3511 (2018).
- <sup>74</sup>M. W. Mahoney and W. L. Jorgensen, "A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions," *J. Chem. Phys.* **112**, 8910–8922 (2000).
- <sup>75</sup>P. Ren and J. W. Ponder, "Temperature and pressure dependence of the AMOEBA water model," *J. Phys. Chem. B* **108**, 13427–13437 (2004).
- <sup>76</sup>W. Klopper, J. G. C. M. van Duijneveldt-van de Rijdt, and F. B. van Duijneveldt, "Computational determination of equilibrium geometry and dissociation energy of the water dimer," *Phys. Chem. Chem. Phys.* **2**, 2227–2234 (2000).
- <sup>77</sup>G. S. Tschumper, M. L. Leininger, B. C. Hoffman, E. F. Valeev, H. F. Schaefer III, and M. Quack, "Anchoring the water dimer potential energy surface with explicitly correlated computations and focal point analyses," *J. Chem. Phys.* **116**, 690–701 (2002).
- <sup>78</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.* **79**, 926–935 (1983).
- <sup>79</sup>K. E. Riley, J. S. Murray, J. Fanfrlík, J. Řezáč, R. J. Solá, M. C. Concha, F. M. Ramos, and P. Politzer, "Halogen bond tunability. II. The varying roles of electrostatic and dispersion contributions to attraction in halogen bonds," *J. Mol. Model.* **19**, 4651–4659 (2013).
- <sup>80</sup>K. E. Riley, J. S. Murray, P. Politzer, M. C. Concha, and P. Hobza, "Br...O complexes as probes of factors affecting halogen bonding: Interactions of bromobenzenes and bromopyrimidines with acetone," *J. Chem. Theory Comput.* **5**, 155–163 (2008).
- <sup>81</sup>X. Mu, Q. Wang, L.-P. Wang, S. D. Fried, J.-P. Piquemal, K. N. Dalby, and P. Ren, "Modeling organochlorine compounds and the  $\sigma$ -hole effect using a polarizable multipole force field," *J. Phys. Chem. B* **118**, 6456–6465 (2014).

Supplementary Information for:

*Classical Pauli Repulsion:*

*An Anisotropic, Atomic Multipole Model*

*Joshua A. Rackers and Jay W. Ponder*

**A. Damping equation for the case when  $\alpha_A = \alpha_B$ :**

$$S = \int \phi_A \phi_B dv = \frac{8\sqrt{q_A q_B}}{\alpha R} f_{damp}(R)$$

$$f_{damp}(R) = \left(1 + \alpha R + \frac{1}{3}(\alpha R)^2\right) e^{-\alpha R}$$

**B. Parameters for models fit to S101x7 database:**

**1. Multipole Pauli Repulsion**

Class	$K$	$\alpha$	$q$
1	2.25	4.63	1.00
2	1.82	4.23	1.00
3	1.11	4.21	1.00
4	1.18	4.20	1.00
5	1.24	4.43	1.00
6	1.20	4.01	1.00
7	17.87	8.82	4.00
8	13.65	6.08	2.00
9	1.42	3.56	3.92
10	84.37	12.57	2.00
12	1.37	3.77	4.00
13	1.37	3.70	3.70
14	3.61	4.16	2.00
15	4.62	4.27	2.15
16	3.93	4.40	2.48
17	3.57	4.74	3.00
18	1.43	4.14	6.00
19	1.19	3.77	5.87
20	1.25	3.73	5.82
21	1.47	4.02	5.75
22	1.63	4.15	5.78
23	1.74	4.40	4.98
24	3.40	3.62	3.39
25	1.52	3.33	6.00
26	1.38	4.72	5.05
27	1.91	3.76	5.91
28	2.02	3.52	6.63

## 2. Isotropic Pauli Repulsion

<i>Class</i>	<i>K</i>	$\alpha$
1	2.40	4.73
2	1.96	4.32
3	0.82	4.07
4	1.08	4.27
5	1.11	4.47
6	1.16	3.95
7	13.57	8.47
8	9.00	7.77
9	1.64	3.69
10	1.23	3.96
12	1.41	3.74
13	0.97	3.46
14	2.20	4.27
15	2.05	4.24
16	2.92	4.89
17	1.81	4.97
18	1.32	4.43
19	0.93	3.75
20	0.97	3.76
21	1.16	4.12
22	0.91	4.16
23	1.23	3.25
24	2.31	3.57
25	1.82	3.36
26	0.85	4.55
27	1.71	3.73
28	2.11	3.55