Uncertainty Classification of Molecular Interfaces

Aaron Knoll, Bin Liu, Kah Chun Lau, Maria K.Y. Chan, Aslihan Sumer
Jeffrey Greeley, Larry Curtiss, Julius Jellinek, Mark Hereld and Michael E. Papka

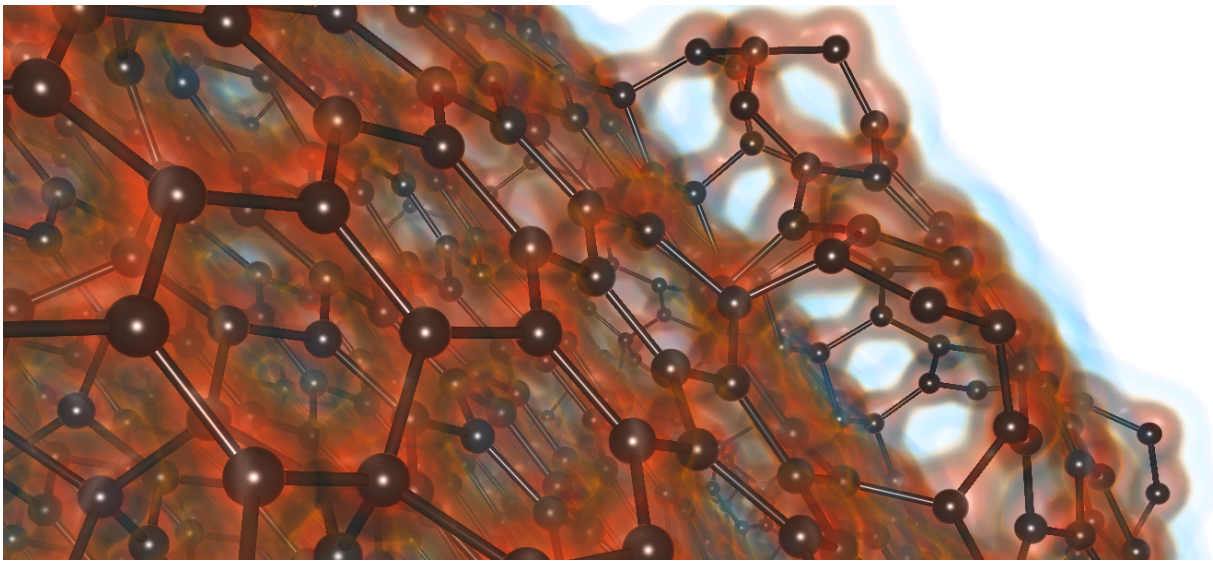

Fig. 1. Material interface of thermally annealed amorphous carbon "supersoot" structures generated from molecular dynamics. The interface is derived from an approximate electron density field, using our uncertainty-driven classification of Van der Waals and chemical bond thresholds.

Abstract— Molecular surfaces at atomic and subatomic scales are inherently ill-defined. In many computational chemistry problems, interfaces are better represented as volumetric regions than as discrete surfaces. The geometry of this interface is largely defined by electron density and electrostatic potential fields. While experimental measurements such as chemical bond and Van der Waals radii do not directly specify the interface, they are physically relevant in modeling molecular structure. Rather than use these radial values to directly determine surface geometry, we use them to define an uncertainty interval in an electron density distribution, which then guides classification of volume data. This results in a strategy for representing, analyzing and rendering molecular structure and interface.

Index Terms—uncertainty, volume rendering, molecular, atomic, quantum mechanics, electron density, charge density, kernel density estimation, radial basis functions

1 Introduction

Representation of surfaces in nanostructured materials data poses challenges for existing visualization techniques. At nanometer and smaller scales, surfaces are defined by the interactions between atoms and their electronic bonds, namely covalent and ionic bonds, and Van der Waals forces. The extremal surface of an individual atom is commonly defined by its Van der Waals radius. Commonly, molecular surfaces are defined by the union of Van der Waals spheres, covered by a probe atom. However, these models fail to accurately describe surface behavior in numerous systems. For amorphous structures in molecular dynamics (MD) simulations, Van der Waals radii are too large and overestimate the surface considerably. However, ionic and covalent bonds are too short, resulting in disjoint geometry.

In reality, molecular interfaces are dictated atomic and molecular interactions, which in turn are governed by multiple physical phenomena. No single first-principles model can accurately define all molecular interfaces, but different models are appropriate for applications with certain scales and given assumptions. Modern materials scientists use density functional theory (DFT) to determine electronic bonding behavior, followed by separate molecular dynamics (MD) simulation of diffusion, annealing, and heat-related macromolecular phenomena. In most cases, computational chemists care less about the underlying surface science than about the general structure and its characteristics (volume, surface area, and mass). For simple analysis, relative statistics based on size and position of atoms are sufficient, without surface classification. However, for visualization and validation, determining interface boundaries is ultimately desirable for understanding computational data.

Rather than seek a single discrete surface to represent a material interface, we note that the desired surface in most cases exists be-

- *Aaron Knoll, Mark Hereld and Michael E. Papka are with the Mathematics and Computer Science (MCS) Division at Argonne National Laboratory (ANL), E-mail: knoll|hereld|papka@mcs.anl.gov.*
- *Bin Liu, Maria K.Y. Chan, and Jeffrey Greeley are with the Center for Nanoscale Materials (CNM) at ANL, E-mail: bliu|mchan|greeley@anl.gov.*
- *Kah Chun Lau and Larry Curtiss are with the Materials Science Division (MSD) at ANL, E-mail: kclau|curtiss@anl.gov.*
- *Aslihan Sumer and Julius Jellinek are with Chemical Sciences and Engineering (CSE) at ANL, E-mail: asumer|jellinek@anl.gov.*

Manuscript received 31 March 2011; accepted 1 August 2011; posted online 23 October 2011; mailed on 14 October 2011.
For information on obtaining reprints of this article, please send email to: tvcg@computer.org.

tween electronic bond and Van der Waals radii, forming an uncertainty interval. Our goal is then to visualize and analyze the interface via this interval. Classification of the electron density field expresses the molecular interface in ways that extrinsic molecular surfaces do not. Volume rendering with this classification illustrates how molecular interfaces are not discrete boundaries, but probabilistic phenomena that vary with the type of interaction. In some ways, our approach is the reverse of conventional Connolly molecular surfaces: rather than defining a molecular surface geometrically and color-coding according to electron density, we define the surface as a level set of density and use geometric radii for uncertainty classification. We employ this approach in rendering volume data from DFT computation, and density fields approximated for larger MD data modeled after DFT from a bulk compound.

2 Related Work

Comprehensive overviews of early molecular surface work are provided by Connolly [11] and O'Donnell [30]. The first and simplest molecular surface representation was the Van der Waals surface or Corey-Pauling-Koltun (CPK) model [12, 27], consisting of the union of Van der Waals spheres. The solvent-accessible surfaces of Richards [24] expand the Van der Waals radius by an additional radius of a probe atom. The most common implementation of molecular surfaces employ the method of Connolly [10], defined by a probe atom rolling continuously across the Van der Waals surface. Connolly surfaces are implemented in most molecular graphics packages [19, 1, 38], commonly using Varshney's SURF [44] or MSMS [37] algorithms. Other molecular surface formulations serve geometric or illustrative purposes. Minimal molecular surfaces [3] minimize mean curvature and thus Gibbs free energy of the surface. Surface abstractions [9, 43] are useful for visualization of larger molecules, particularly proteins, with subsurface structures.

Volumetric renderings of molecular data were first used in the illustrative animations of Blinn [4] using compact polynomial basis functions. Direct volume rendering, as a general visualization modality, was introduced by Levoy [26]. Many computational chemistry data are inherently scalar fields, including charge density from Hartree-Fock or DFT, and electrostatic potential from molecular dynamics. In general, the use of volume isosurfaces for representing electron density clouds is well accepted in the chemistry literature, though not as an outright replacement for molecular surfaces [39]. Goodsell et al. [15] demonstrated direct volume rendering of charge density fields. Similar work explored volume rendering of electron density from *ab initio* molecular dynamics [13]. Qiao et al. [33, 34] explore volume rendering of electron orbitals in quantum dot device simulations. Jang et al. [20] propose direct evaluation of density functionals of both atoms and molecular orbitals for interactive volume rendering.

In visualizing uncertainty, volume rendering of multifield data with associated uncertainty was explored with 1D and 2D transfer functions [14]. Uncertainty-driven multidimensional classification of scalar volume data has also been explored [21]. Uncertainty pertaining to isosurfaces with computed error has also been visualized [36]. Point-based approaches allow for expression of spatial uncertainty in both polygonal and volumetric data-[16]. Hu et al. [18] considered atomic movement in formulating and visualizing uncertainty of molecular surfaces. Their approach is to model the movement of each atom as a distribution, and propagate that uncertainty to the Connolly molecular surface model.

Volume approximations of charge density and other atomic properties are implemented in VMD [19], but not modeled after bulk DFT computation. For the most part, volume rendering is uncommon in molecular visualization tools [19, 1, 38]. General-purpose software such as Visit [8] offer volume rendering and molecular visualization capabilities, but leave classification in the hands of the users. We propose a domain-specific tool for classifying and visualizing molecular volume data.

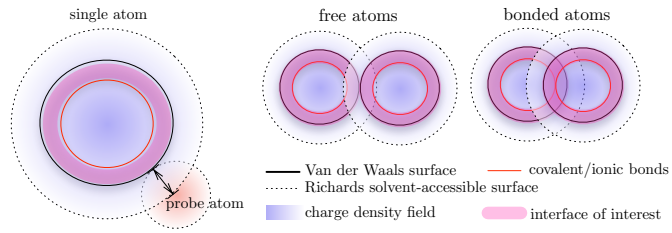

Fig. 2. Depending on the type of interaction, regions between chemical bonds and Van der Waals radius can be of interest in defining a molecular interface.

3 Background

Molecular structure is determined by intramolecular and intermolecular forces. These are represented as continuous scalar fields of electron density (or charge density) and potential energy, respectively.

In quantum mechanics, electron density measures the probability of an electron being present at a given spatial location [31]. Solving for Schrödinger's equation using a linear combination of atomic orbitals (LCAO) results in wave functions with radial and angular momentum for each electron [35].

$$\theta_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_{lm}(\theta,\phi) \tag{1}$$

The simplest case, the 1s orbital of the hydrogen atom has $Y_{lm} = 1$ and a Gaussian $R_{nl} = Ae^{-Br^2}$. For a molecule with electronic bonds, bond geometry can be determined by minimizing the global energy of the system through coefficients of all molecular orbitals (e.g. A and B). This can be computed using either forward (variational quantum Monte Carlo [28]) or inverse (density functional theory [17]) approaches. Molecular orbitals consist of the sum of all atomic orbitals.

$$\psi(r,\theta,\phi) = \sum_e \theta_{nlm}(r,\theta,\phi) \tag{2}$$

Electron density ψ has positive or negative sign associated with independent orbitals; *charge density* normalizes this:

$$\rho(r,\theta,\phi) = \psi(r,\theta,\phi)^2 \tag{3}$$

Classifications of electron and charge densities are semantically equivalent. We use unsigned electron density due to its lower dynamic range, and because it can be better modeled with Gaussian bases approximating LCAO DFT. For the purposes of this paper, we are interested in absolute value of electron density:

$$|\psi| = \sqrt{\rho} \tag{4}$$

From the spherical basis, it is then trivial to plot $|\psi(x,y,z)|$ as a structured grid in R^3. Typically, the surface of the density field is suggested by outer valence electrons (the highest-occupied molecular orbital, or HOMO). However, all-electron density is also a suitable indicator. In any case, electron density of the equilibriated system defines chemical bonds, thus the "geometry" of the molecule.

Weak intermolecular forces are responsible for other structural aspects. These include Pauli repulsion between atoms, and electrostatic attractive and repulsive forces between dipoles and induced dipoles, including polarization and London dispersion forces. Collectively these are known as Van der Waals forces. For most problems, these are paired with Newtonian physics and modeled using molecular dynamics (MD) [2]. In classical MD, these forces are assumed to operate independently of chemical bonds according to the Born-Oppenheimer approximation [6], and a ground state potential energy surface (from separate DFT computation) guides MD computation. In more costly *ab initio* MD, intramolecular and intermolecular forces both contribute to energy potentials. Energy potentials define a separate scalar field manifesting larger-scale intermolecular structure. In general, electron density is more indicative of surface and interface structure.

3.1 Atomic radii

Conventional definitions of molecular surfaces rely on empirically-measured radii of their component atoms from X-ray crystallography. The radius of an atom depends highly on the molecule it resides in, and the interactions of interest. Illustrated in Figure 2, the most common atomic radii are:

- *Bond radii* of ionic, covalent or metallic bonds, determined directly from X-ray crystallography.
- *Van der Waals radius*, or half the distance between the nuclei of two free atoms of the same type, which relates to the London dispersion component of Van der Waals forces.

Van der Waals radius itself can be determined in multiple ways, resulting in 5-10 pm error for most atoms [5]. This is typically small compared to uncertainty associated with the type of interface interaction.

3.2 Molecular surfaces

As described in Section 2, most molecular surfaces consist of Van der Waals spheres [12, 27], those surfaces offset to accommodate a probe [24, 10], or abstracted simplifications of either [3, 9]. In general, volume rendering these scalar fields has not gained traction in chemistry visualization communities. Electron and charge densities are typically visualized using the minimal nonzero isovalue corresponding to an outer isosurface. Electrostatic potential is commonly represented as the zero-level surface.

Defining the surface of a molecule is impossible without knowledge of the "probe" interacting with it. As shown in Figure 2, the probe can be any type of free atom or molecule, as well as subatomic particles including photons and electrons (e.g. from a scanning electron microscope). Inert free atoms and large polymers such as proteins motivate the use of solvent-accessible surfaces [24, 10]. However, in the case of smaller particles and free ions, one often desires representation of features within the Van der Waals radius corresponding to chemical bonds. A logical solution would be to compute Connolly surfaces at both Van der Waals radius and an inner chemical bond radius. However, chemical bond surfaces can be disjoint. Another solution to forming continuous geometry is to model the surface implicitly using radial basis functions such as those of Blinn [4]. However, these basis functions have no direct physical meaning.

4 Molecular Interface Uncertainty Classification

To flexibly represent material interfaces without assuming a specific probe atom, we posit that many interactions of interest in fact occur *in between* inner chemical bonds and outer Van der Waals radii. Our main insight is that, for any given type of chemical bond, empirically-measured radii correspond to isovalues in the associated electron/charge density distribution. In this way, the measured values of chemical bond radii and Van der Waals radius define an uncertainty interval in the density distribution. In turn, this can be used to classify all-electron density in direct volume rendering.

We choose to classify absolute value of electron density $\Psi = |\psi| = \sqrt{\rho}$ since:

1. $|\psi|$ is a continuous scalar field
2. For all r greater than the Van der Waals radius r_v, $|\psi(r)|$ exists and $|\psi(r)| > 0$. Moreover, $\Psi(r)$ is positive-definite and surjective, implying it can be classified even outside the Van der Waals radius.
3. Linear combinations of $|\psi|$ form a better model of LCAO DFT than using ρ
4. In most (though not all) cases, $|\psi(r)|$ varies monotonically with r, in these cases, Ψ is a bijection and classification is unambiguous.

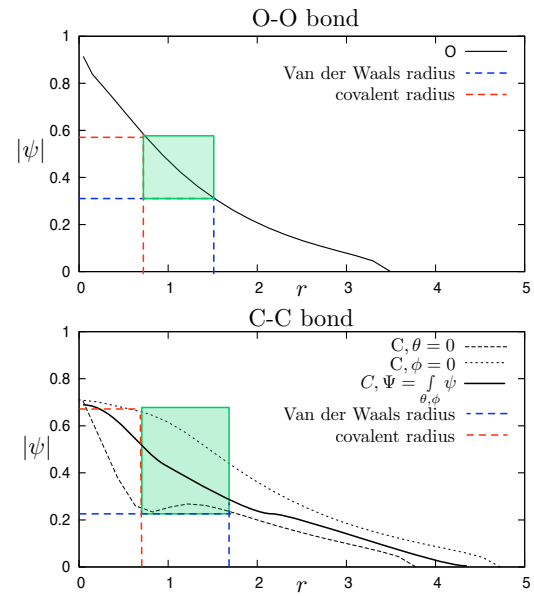

Fig. 3. Electron density $|\psi| = \sqrt{\rho}$, in $e/\AA^3$, as a function of radius r in Angstroms, in O-O and C-C bonds. The interval between Van der Waals and covalent radii defines as interval in Ψ that we use in classification. For the anisotropic C-C orbitals, we show minimum and maximum distributions at angular momentum values of $\theta = 0$ and $\phi = 0$, respectively, and the average cumulative distribution Ψ over all θ, ϕ. These distributions map electron density as a function of radius, and drive our uncertainty classification.

Our technique then proceeds as follows. From DFT computation on a bulk compound, or chemical Monte-Carlo, we plot electron density Ψ with respect to radius for each atom in our system. We consider the interval

$$\bar{r} = [\underline{r}, \bar{r}] \tag{5}$$

where $\underline{r}$ is the smallest measured chemical (covalent, ionic or metallic) radius, and $\bar{r}$ is the upper bound on the measured of the Van der Waals radius. These are shown via the dotted lines bracketing range and domain in Figure 3. When $|\psi|$ is monotonic, this corresponds to

$$\overline{\underline{\Psi}} = [|\psi(\underline{r})|, |\psi(\bar{r})|] \tag{6}$$

Then, classification of electron density in $\overline{\underline{\Psi}}$ corresponds directly to the extents of the inner chemical bonds and outer Van der Waals interactions, i.e. the molecular interface.

We use this classification in forming a transfer function, mapping scalar value to opacity and color. We then employ direct volume rendering to visualize our data. The classification can equally be used to measure volume, surface area, and other geometric properties.

4.1 Anisotropic Orbitals

For electron density distributions with primarily radial (s orbital) behavior such as the O-O example in Figure 3, it suffices to plot electron density of free atoms along a single line (e.g. $\theta = \phi = 0$, or along the X axis through an atom).

However, for bonds with anisotropic behavior resulting from p or d orbitals, the monotonicity property (4) above does not hold. This is shown in the C-C bond in Figure 3 (bottom), which has non-monotonic $|\psi(r)|$ the line along $\psi = 0$. Moreover, $|\psi|$ exhibits low electron density at certain orientations, and significantly higher density at others. When $|\psi|$ is non-monotonic, $|\psi(r)|$ is no longer a bijection (i.e. $|\psi(r_1)| = |\psi(r_2)|$ for some $r_1 \neq r_2$), resulting in ambiguous classification. Principally, we are left with two choices:

- Integrate over θ, ϕ to find an average distribution, and classify using this (see Section 4.2)

- Classify the minimum and maximum $|\psi|$ corresponding to all θ, ϕ.

For DFT density data we recommend the second option, illustrated in the classification of Figure 3 (bottom). Expanding the uncertainty interval is a better choice when inner orbitals and individual bonds are of interest, but electron density at these features may vary. This can be accomplished by plotting $|\psi|$ for all θ, ϕ (or a subset thereof). While more ambiguous for inner orbitals, this choice has no adverse impact on classification of outer (valence electron) orbital geometry near the Van der Waals radius, which is a better indicator of chemical bond geometry.

4.2 Approximate electron density model

For molecular dynamics data, we are given atom geometry and sometimes electrostatic potential. However, the potential field is not immediately useful in defining molecular interfaces, particularly near chemical bonds. In many cases, MD data is too large for electron density to be computed directly using DFT methods. Thus, we propose using density distributions from DFT to define atomic radial basis functions. Kernel density methods have been well studied for this kind of problem [40], and indeed our proposed solution is similar to the one first used by Blinn [4]. Our contribution is the use of the average electron density distribution from bulk DFT as the basis for our kernel density model, which allow for uncertainty classification with associated physical meaning.

Integrating ψ over its angular components results in an all-electron distribution over radius. If we divide this by the number of electrons E, we have an average distribution

$$\Theta(r) = \frac{1}{E} \int \psi(r, \theta, \phi)\, d\theta\, d\phi \tag{7}$$

This is shown by the solid black line in Figure 3. The density distribution varies somewhat with the chosen DFT approach (planewave or LCAO) and whether or not inner atomic orbitals are fully simulated. Although density in the outer valence regions is relatively low, these features often have the greatest impact on molecular surface geometry. This must be taken into account when normalizing Θ for the number of electrons.

The distribution is computed for each atom in the system, based on the types of bonds present in the molecule. For example, in amorphous aluminum oxide, we would compute Ψ_{Al} based on Al-Al and Al-O bonds, and Ψ_O based on Al-O and O-O bonds from DFT computation on a bulk compound. Then, for all atoms i,

$$r_i(\vec{x}) = |\vec{x} - \vec{p}_i|$$
$$\Psi(\vec{x}) = \sum_i \Theta_i(r_i(\vec{x})).$$

The kernel density model is a per-atom approximation of per-electron phenomena, and is not meant as a replacement for computation of orbitals from DFT or quantum Monte Carlo. However, it suffices for large molecules with predictable, homogeneous structure, and is arguably better than Connolly surfaces at showing bond structure as well as materials interface. As with $|\psi|$ computed from DFT, we use our classification in both volume rendering and analysis.

5 Implementation

Our automated classification has been integrated into Nanovol, a GPU volume renderer designed for materials visualization and analysis. Nanovol employs peak finding classification and distance-based sampling [22] to achieve roughly 3x better performance than brute-force DVR for similar quality. It employs a uniform grid for space-skipping and efficient raycasting of ball-and-stick geometry. It also provides higher-order interpolation options, which can be useful in smoothing low-resolution volume data from DFT computation.

Electron density distributions were acquired from DFT computation of bulk compounds in GPAW [29], a planewave density functional theory code that computes outer orbitals and omits inner electrons as necessary. Charge density distributions for amorphous aluminum distribution were generated using VASP [23]. From these DFT computations, we compute the electron or charge density distribution and corresponding transfer function using utilities written in C++, accessible either as standalone programs or as a simple library.

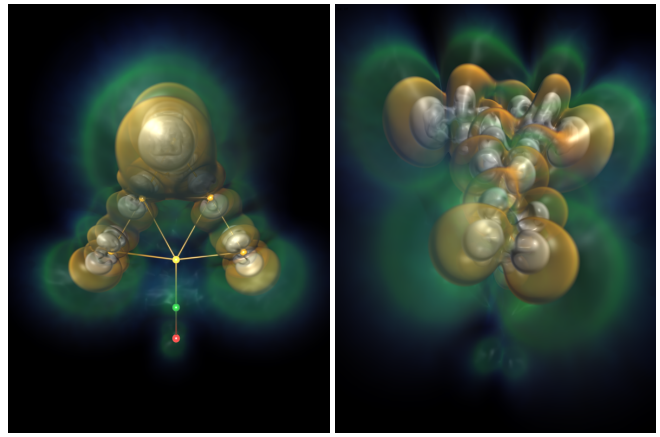

Fig. 4. Partially automatic classification of $|\psi|$ of the HOMO, in a CO-Pt_7 system computed using LCAO DFT in NWChem [42]. Our uncertainty classification uses the density distribution of the C-O bond to classify outer valence electron regions. Inner orbital geometry is classified *ad hoc* as isosurface-like features.

The proposed technique is a general approach for automatic classification (and if necessary, estimation) of electron density fields for representing molecular structure and interface. To implement this more generally, a database could store electron density distributions and distribution intervals for every atom and bond of interest. When considering a new molecule, one would determine component atoms and use the appropriate distributions for uncertainty classification and the per-atom kernel density model. So far, we have applied our technique to several DFT and MD molecular structures, and we consider two in particular in Section 6. However, computation of such a distribution database for every atom and bond of interest would require significant effort. Bulk DFT computation for two atoms is not generally costly, requiring several seconds for small examples (C, O). DFT for larger atoms such as aluminum require greater computational resources.

6 Use Cases and Results

We apply our classification approach to electron density data generated from DFT, and to MD atomic datasets with an approximate electron density volume generated with our kernel density model. Although the appropriate electron density classification varies greatly with the application, we find that our physically-based uncertainty classification is an improvement over *ad hoc* classification, and that properly classified volume rendering is a good modality for molecular visualization.

In Nanovol, we achieve interactive performance (5-30 fps at 1 gigapixel resolution) when volume rendering most data sets on an 8-core Intel 2.8 GHz Core i7 and an NVIDIA 580 GTX with 1.2 GB RAM. While volume raycasting performs more slowly than rasterization-based approaches for small data, for large data such as the 740k atom supersoot our renderer outperforms rasterization of isosurface and ball-and-stick models in VMD [19] by up to 10x. Nonetheless, our classification approach can be generalized to any volume rendering software, or employed in choosing isovalues corresponding to the minimum and maximum electron densities in our uncertainty classification.

6.1 Density Functional Theory: CO Adsorption

Our classification strategy can most directly be applied to data from density functional theory computation. In Figure 4, we consider a carbon monoxide adsorption in a Platinum catalyst cluster, with the goal of understanding the bonding energetics of CO as it attaches to the catalyst. Bond formation was simulated using LCAO DFT in

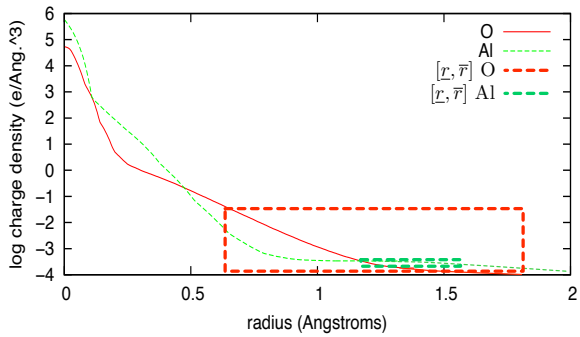

Fig. 5. Charge density ρ plotted as a function of radius for Al and O from bulk DFT computation, used to model radial basis functions $|\psi(r)|$ of electron density in nanobowl MD data.

NWChem [42]. Here, we examine $|\psi|$ of the highest-occupied molecular orbital (HOMO), which corresponds roughly to density of the outer (valence) electron.

We are primarily interested in the behavior of the bond between CO and the cluster. While electron density is low near the C-Pt bond, this is nonetheless the main feature of interest. To classify this, we used the distributions of C and O from Figure 3, and found density values corresponding to covalent (green) and Van der Waals (blue) radii. Values of Ψ corresponding to regions outside the Van der Waals radius are not chemical bonds, and are classified as zero opacity. $|\psi|$ in the CO bond is relatively low compared to density of inner electron orbitals. The high-density, high-gradient regions near the platinum nuclei were classified *ad hoc* using sharp spikes in the transfer function, similar to isosurfaces. Overall, this gives a good sense of the behavior of both inner and outer orbital geometry. The main advantage of our semi-automatic approach over naive classification is that we have some physical gauge of these extents, based on empirical measurements.

6.2 Molecular dynamics: Nanobowls

While DFT computation is useful in examining the behavior of chemical bonds, molecular dynamics are used to model larger-scale properties such as heat exchange, thermal annealing, diffusion and dispersion. For MD data, an electron density field must be approximated from its underlying atomic geometry, e.g. using the model described in Section 4.2.

We have applied our technique in MD computation of amorphous aluminum oxide (Al_2O_3) "nanobowl" structures, which are designed as multipurpose catalyst support structures [7, 41, 25]. One goal of these simulations is to understand the stability of these structures at varying temperatures. Preservation of bowl volume is a good measure of stability.

In preliminary analysis of the nanobowl data, all-electron charge density distributions ρ were computed for Al-Al, Al-O and O-O bonds in VASP [23]. We used these distributions of ρ to determine electron density $|\psi| = \sqrt{\rho}$, and used $\psi(r)$ as radial bases in our kernel density model. We perform classification based on charge density ρ, even though our kernel density model uses electron density $|\psi|$. The semilog plot of these distributions is shown in Figure 5. Since Al_2O_3 is largely ionic, we used ionic bond radius instead of covalent radius for the inner radius, and Van der Waals for the outer radius as usual. We chose to perform classification in ρ-space. This is functionally the same as using ψ, but with sharper transition between the interface endpoints in the transfer function. We find that the uncertainty interval $\overline{\Psi} = \sqrt{\overline{\rho}}$ is dominated by oxygen; to highlight the contributions of aluminum we chose a red key color to classify density corresponding to that atom's ionic radius. By classifying only this interface, we can

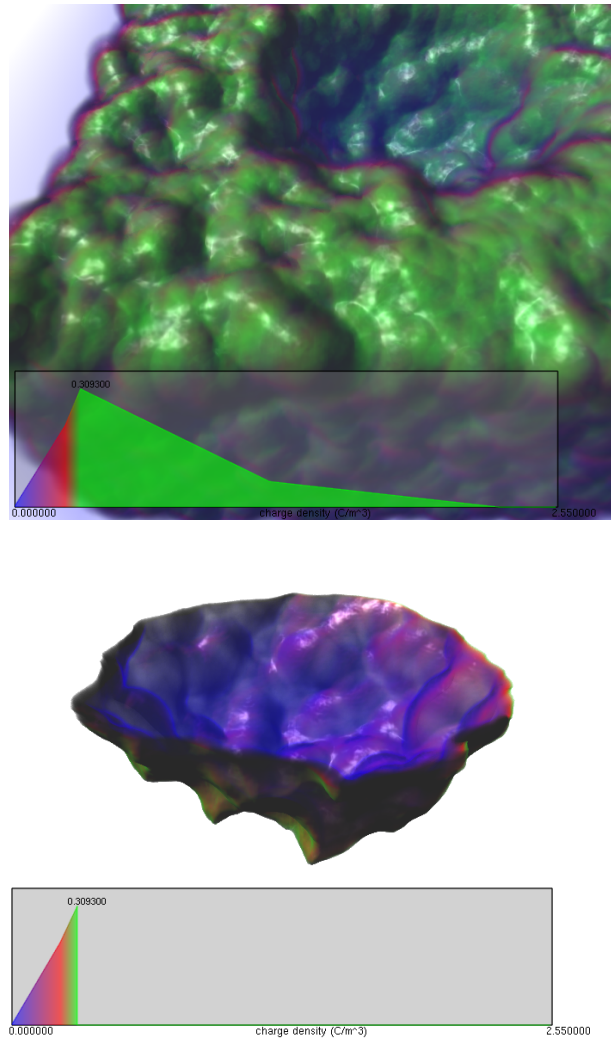

Fig. 6. Classification of charge density ρ applied to our approximate electron density volume $|\psi| = \sqrt{\rho}$. We show uncertainty in the material interface between ionic and Van der Waals radii via a red-blue transition in these transfer functions. This provides a simple mechanism for segmenting the bowl volume.

effectively segment the bowl from the rest of the structure. Examples of these classifications are shown in Figure 6.

In Figure 7, we apply our uncertainty classification to simulation runs at four different temperatures. We use our classification to measure volume in $\mathring{A}^3$ corresponding to both lower and upper density thresholds (ionic bonds and Van der Waals). We compare this to raw positional analysis, which used a Gaussian-smoothed heightfield of interpolated nucleus positions to estimate volume. Since volume statistics are relative, positional analysis is sufficient to determine whether a structure is stable or not. However, our uncertainty classification is advantageous for visualization and subsequent validation. With our uncertainty analysis, we note slightly more fluctuation in volume, though the same overall trends remain. Since our volumetric model accounts for bond length, not simply the convex hull of atomic nuclei, we expect it would correlate more strongly with experimental measurements. Overall, uncertainty classification suggests physical upper and lower estimates for a theoretical molecular surface, which could be useful in correlating with experimental measurements from electron microscopy.

6.3 Molecular dynamics: Supersoot

Our last application involves visualization of large amorphous carbon "supersoot" structures. Appearing as ordinary soot to the human eye,

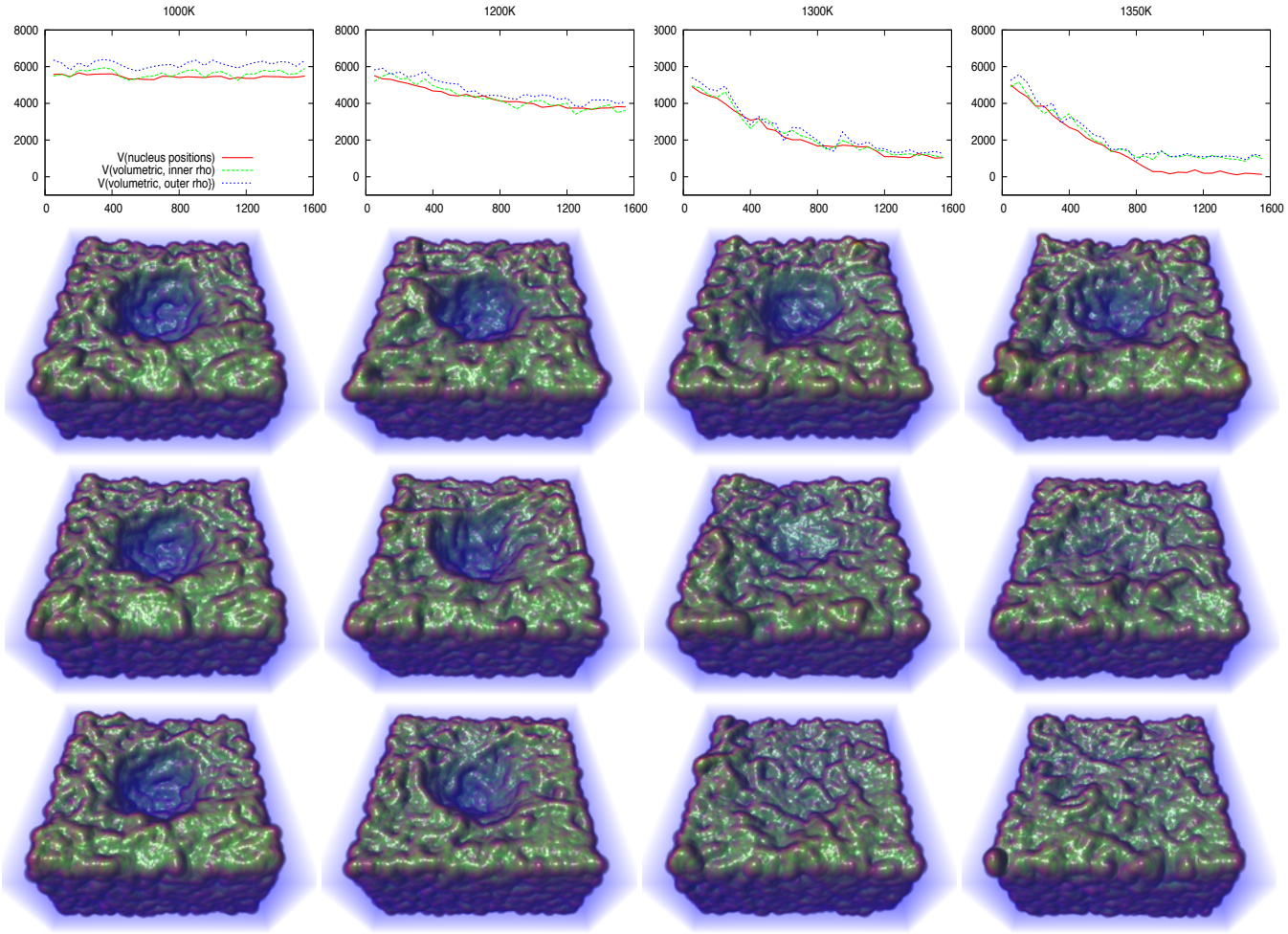

Fig. 7. Top: analysis from our classification, plotting bowl volume ($\mathring{A}^3$) over the simulation time (0 - 1550 picoseconds) for separate MD nanobowl runs at 1000K, 1200K, 1300K and 1350K temperatures, shown in the four columns. Bottom rows show visualizations at 260, 800 and 1500 ps, respectively, at the corresponding temperature of that column. Our uncertainty classification defines upper and lower bounds for bowl volume, which can be useful in validation.

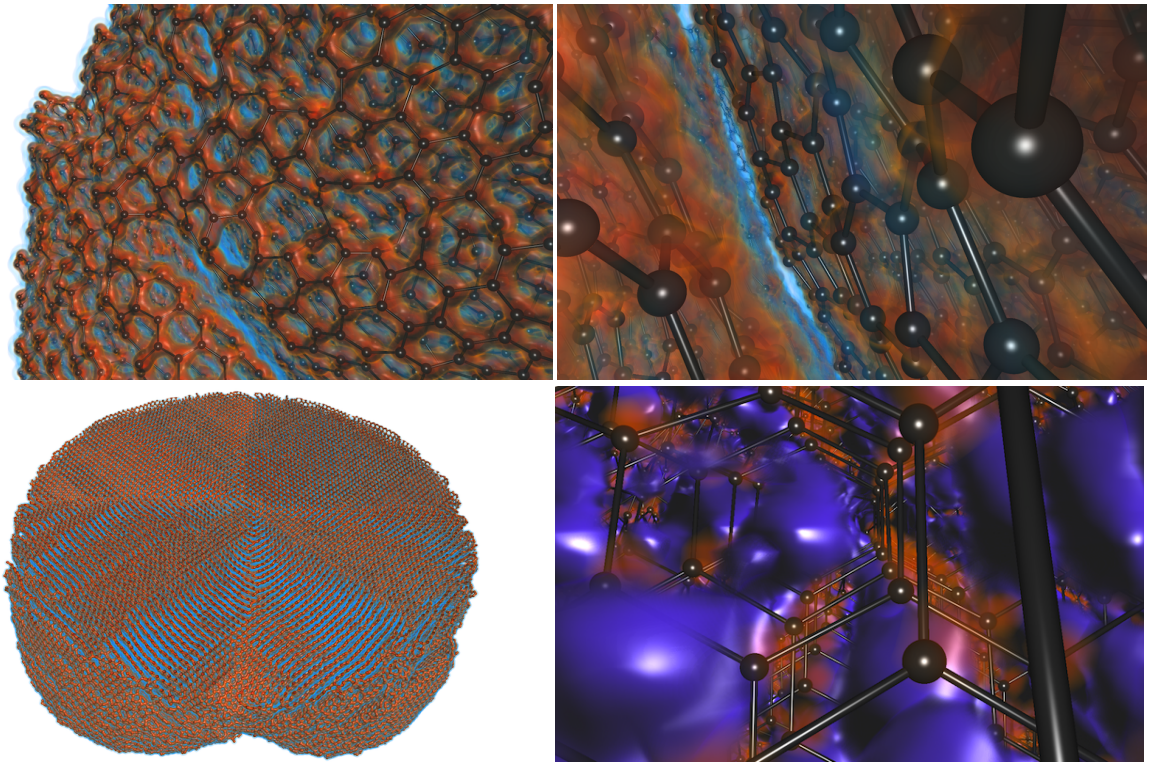

Fig. 8. Amorphous carbon “supersoot” structures resulting from thermal annealing, built in . Uncertainty classification lets us better explore material interfaces, including channels where Lithium electrolyte may be stored.

these materials are obtained through recycling of plastics and other organic compounds. At high temperature (2500K) they undergo thermal annealing, forming folds and channels that are well-suited for anode material in Lithium-ion batteries. This process is modeled in the LAMMPS MD code [32] for various-size structures, the largest of which is 740,000 atoms. Structurally, supersoot is a mix of diamond (sp_3) and graphite (sp_2) carbon configurations, with one or multiple diamond cores and graphite sheaths on the exterior. Due to its large size, neither ball-and-stick nor Connolly surfaces are helpful in visualizing the full supersoot model. Scientists are primarily interested in the macromolecular folds and channel structures, as distinct from standard diamond or graphite geometry. They are also interested in classifying space that can potentially contain Lithium electrolyte. Clearly, DFT is prohibitively costly for structures of this size.

To accomplish both goals, we employ our kernel density model and uncertainty classification. Because the C-C dimer (Figure 3) differs from both diamond or graphite configurations, we use DFT computation on bulk graphite and diamond as our distribution. Then, our kernel density model at 4 voxels per Angstrom generates a moderately large (roughly 1024^3) volume. Figure 1 and Figure 8 show ball-and-stick and volume rendering of this structure in Nanovol. In particular, our uncertainty classification allows for visualization of channels in between covalent and Van der Waals radii where Lithium ion electrolyte may be able to pass through and potentially be stored. This void space could consist of irregularities in graphitic C_6 (sp_2) geometry, the space in between graphite sheaths, or folding features resulting from the annealing process. Uncertainty classification and volume visualization help us better understand the material interface and correlate with experimental results.

7 Conclusions

We have presented an uncertainty classification method for molecular interfaces based on electron density distributions computed from bulk DFT or other first principles methods. This approach is useful in identifying interfaces based on measured chemical bond and Van der Waals radii, and in modeling continuous electron density fields from MD computation.

We are not the first to propose volume rendering of molecular data, or radial basis functions for approximation of electron or charge density fields. However, our contributions aim to associate physical meaning with these techniques, which remain largely unused by the molecular visualization community over two decades since they were first advocated. While similar visualizations can easily be achieved without our semi-automated uncertainty classification, we argue that it is scientifically useful to have this frame of reference when performing volume visualization.

In all, our approach argues for more widespread adoption of volume rendering modalities in molecular visualization software, and engaged discourse between theoretical, experimental and computational chemists, and visualization staff. Future work in this direction could better formalize approximation and classification of electron density for molecular interfaces, as well as explore the impact of electrostatic potential on surface science.

Acknowledgments

This work was supported by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract DE-AC02-06CH11357, and the Computational Postdoctoral Fellowship at Argonne National Laboratory under the American Reinvestment and Recovery Act (ARRA). Data CSE, MSD and CNM were supported as part of the Institute for Atom-efficient Chemical Transformations (IACT), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

References

[1] Avogadro: an open-source molecular builder and visualization tool, http://avogadro.openmolecules.net/.

[2] B. J. Alder and T. E. Wainwright. Studies in Molecular Dynamics. I. General Method. *Journal of Computational Physics*, 31:459–466, Aug. 1959.

[3] P. Bates, G. Wei, and S. Zhao. Minimal molecular surfaces and their applications. *Journal of Computational Chemistry*, 29(3):380–391, 2008.

[4] J. Blinn. A Generalization of Algebraic Surface Drawing. *ACM Transactions on Graphics*, 1(3):235–256, July 1982.

[5] A. Bondi. Van der Waals Volumes and Radii. *The Journal of Physical Chemistry*, 68(3):441–451, 1964.

[6] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[7] R. Chen, H. Kim, P. McIntyre, D. Porter, and S. Bent. Achieving area-selective atomic layer deposition on patterned substrates by selective surface modification. *Applied Physics Letters*, 86:191910, 2005.

[8] H. Childs, E. Brugger, K. Bonnell, J. Meredith, M. Miller, B. Whitlock, and N. Max. A contract based system for large data visualization. In *Proceedings of IEEE Visualization*, volume 2005, pages 190–198. IEEE, 2005.

[9] G. Cipriano and M. Gleicher. Molecular surface abstraction. *IEEE Transactions on Visualization and Computer Graphics*, pages 1608–1615, 2007.

[10] M. Connolly. Analytical Molecular Surface Calculation. *Journal of Applied Crystallography*, 16:548–558, 1983.

[11] M. Connolly. Molecular Surfaces: A Review. *Network Science Online*, 1996. http://www.netsci.org/Science/Compchem/feature14.html.

[12] R. B. Corey and L. Pauling. Molecular Models of Amino Acids, Peptides and Proteins. *Applied Physics Letters*, 24:621–627, 1953.

[13] L. Covick and K. Sando. Volume-rendering representations of ab initio electron densities. *Journal of Molecular Graphics*, 11(1):37 – 39, 1993.

[14] S. Djurcilov, K. Kim, and A. Pang. Volume rendering data with uncertainty information. In *Data Visualization 2001: proceedings of the Joint Eurographics-IEEE TCVG Symposium on Visualization in Ascona, Switzerland, May 28-30, 2001*, page 243. Springer Verlag Wien, 2001.

[15] D. S. Goodsell, I. S. Mian, and A. J. Olson. Rendering volumetric data in molecular systems. *Journal of Molecular Graphics*, 7:41–47, March 1989.

[16] G. Grigoryan and P. Rheingans. Point-based probabilistic surfaces to show surface uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 10(5):564 –573, sept.-oct. 2004.

[17] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[18] M. Hu, W. Chen, T. Zhang, and Q. Peng. Direct volume rendering of volumetric protein data. *Advances in Computer Graphics*, pages 397–403, 2006.

[19] W. Humphrey, A. Dalke, and K. Schulten. VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.

[20] Y. Jang and U. Varetto. Interactive volume rendering of functional representations in quantum chemistry. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1579–5186, 2009.

[21] J. Kniss, R. Uitert, A. Stephens, G. Li, T. Tasdizen, and C. Hansen. Statistically quantitative volume visualization. In *Proceedings IEEE Visualization*, pages 287–294, 2005.

[22] A. Knoll, Y. Hijazi, R. Westerteiger, M. Schott, C. Hansen, and H. Hagen. Volume Ray Casting with Peak Finding and Differential Sampling. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1571–1578, Nov-Dec 2009.

[23] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558, 1993.

[24] B. Lee and F. Richards. The Interpretation of Protein Structures: Estimation of Static Accessibility. *Journal of Molecular Biology*, 55:379–400, 1971.

[25] S. Lee, L. M. Moline, M. J. Lopez, J. A. Alonsa, B. Hammer, B. Lee, S. Seifert, R. Winans, J. Elam, M. Pellin, and S. Vajda. Selective propene epoxidation on immobilized au6-10 clusters: The effect of hydrogen and water on activity and selectivity. *Langmuir*, 25(19):11216–11220, 2009.

[26] M. Levoy. Efficient Ray Tracing for Volume Data. *ACM Transactions on Graphics*, 9(3):245–261, July 1990.

[27] W. L. Loltun. Space Filling Atomic Units and Connectors for Molecular Models. US Patent 3170246, 1965.

[28] W. L. McMillan. Ground state of liquid he^4. *Phys. Rev.*, 138:A442–A451, Apr 1965.

[29] J. Mortensen, L. Hansen, and K. Jacobsen. Real-space grid implementation of the projector augmented wave method. *Physical Review B*,

71(3):035109, 2005.

[30] T. Odonnell. The Scientific and Artistic Uses of Molecular Surfaces. *Network Science Online*, 1996. http://www.netsci.org/Science/Compchem/feature15.html.

[31] L. Pauling and E. B. Wilson. *Introduction to quantum mechanics: with applications to chemistry*. Courier Dover Publications, 1985.

[32] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.

[33] W. Qiao, D. Ebert, A. Entezari, M. Korkusinski, and G. Klimeck. Volqd: Direct volume rendering of multi-million atom quantum dot simulations. In *Visualization, 2005. VIS 05. IEEE*, pages 319–326. IEEE, 2005.

[34] W. Qiao, M. McLennan, R. Kennell, D. Ebert, and G. Klimeck. Hub-based simulation and graphics hardware accelerated visualization for nanotechnology applications. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):1061–1068, 2006.

[35] C. Quinn. *Computational quantum chemistry: an interactive guide to basis set theory*. Academic Press, 2002.

[36] P. Rhodes, R. Laramee, R. Bergeron, and T. Sparr. Uncertainty visualization methods in isosurface rendering. In *Eurographics*, pages 83–88. Citeseer, 2003.

[37] M. Sanner, A. Olson, and J. Spehner. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, 38(3):305–320, 1996.

[38] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.3r1. August 2010.

[39] A. J. Shusterman and G. P. Shusterman. Teaching chemistry with electron density models. *Journal of Chemical Education*, 74(7):771, 1997.

[40] B. Silverman. *Density Estimation for Statistics and Data Analysis*, volume 26. Chapman & Hall/CRC, 1986.

[41] M. Tada, T. Sasaki, and Y. Iwasawa. Novel sio2-attached molecular-imprinting rh-monomer catalysts for shape-selective hydrogenation of alkenes; preparation, characterization and performance. *Phys. Chem. Chem. Phys.*, 4(18):4561–4574, 2002.

[42] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, et al. Nwchem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.

[43] M. van der Zwan, W. Lueks, H. Bekker, and T. Isenberg. Illustrative molecular visualization with continuous abstraction. In *Computer Graphics Forum*, volume 30, pages 683–690. Wiley Online Library, 2011.

[44] A. Varshney, F. Brooks Jr, and W. Wright. Computing smooth molecular surfaces. *Computer Graphics and Applications, IEEE*, 14(5):19–25, 1994.

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.