

UNCERTAINTY CLASSIFICATION AND VISUALIZATION OF MOLECULAR INTERFACES

Aaron Knoll,^{1,*} Maria K.Y. Chan,³ Kah Chun Lau,² Bin Liu,³ Jeffrey Greeley,³ Larry Curtiss,² Mark Hereld,¹ & Michael E. Papka¹

¹Mathematics and Computer Science Division

²Materials Science Division

³Center for Nanoscale Materials

Argonne National Laboratory, Argonne, IL 60439 USA

Original Manuscript Submitted: 07/11/2012; Final Draft Received: 07/11/2012

Molecular surfaces at atomic and subatomic scales are inherently ill-defined. In many computational chemistry problems, boundaries are better represented as volumetric regions than as discrete surfaces. Molecular structure of a system at equilibrium is given by the self-consistent field, commonly interpreted as a scalar field of electron density. While experimental measurements such as chemical bond and Van der Waals radii do not spatially define the interface, they can serve as useful indicators of chemical and inert interactions, respectively. Rather than using these radial values to directly determine surface geometry, we use them to map an uncertainty interval in the electron density distribution, which then guides classification of volume data. This results in a new strategy for representing, analyzing and rendering molecular boundaries that is agnostic to the type of interaction.

KEY WORDS: *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” representing a diffusing atom or adjacent molecule. 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 by 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, and molecular dynamics (MD) simulations to study diffusion, and other temperature-dependent 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.

*Correspond to: Aaron Knoll, E-mail: knoll@mcs.anl.gov, URL: <http://www.mcs.anl.gov/~knoll>

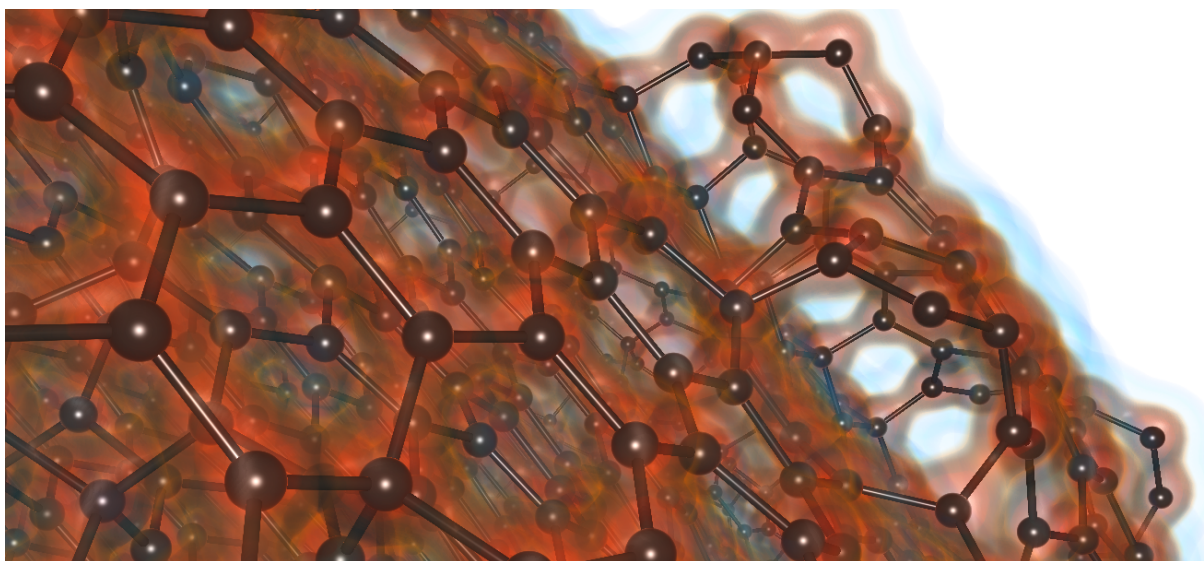


FIG. 1: Material interface of thermally annealed amorphous carbon nanosphere structures generated from molecular dynamics. The interface is derived from an approximate electron density field, using our classification of an interval between Van der Waals and chemical bond thresholds.

Rather than seek a single discrete surface to represent a material interface, we note that the desired surface in most cases exists between electronic bond and Van der Waals radii, forming an uncertainty interval. We then visualize and analyze the interface defined by this interval as a function of electronic structure, or charge density, as shown in Figure 1. Classification of the electron density expresses the molecular interface in ways that extrinsic molecular surfaces do not. We use volume rendering to illustrate that molecular interfaces are not discrete boundaries, but phenomena that vary with the type of interaction and are inherently difficult to quantify. In a sense, this 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 experimentally-acquired geometric radii to determine its boundaries. We employ this approach in rendering volume data from DFT computation, and density fields approximated for larger MD data modeled after DFT simulation on a bulk compound.

2. RELATED WORK

Molecular surface visualization relies on abstractions of geometry defined by atomic position, rather than interpretation of electronic structure. Comprehensive overviews of early molecular surface work are provided by Connolly [1] and O'Donnell [2]. The first and simplest molecular surface representation was the Van der Waals surface or Corey-Pauling-Koltun (CPK) model [3, 4], consisting of the union of Van der Waals spheres. The solvent-accessible surfaces of Richards [5] 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 [6], defined by a probe atom rolling continuously across the Van der Waals surface. These surfaces are implemented in most molecular graphics packages [7–9] using Varshney's SURF [10] or MSMS [11] algorithms. Akkiraju and Edelsbrunner [12] noted the equivalence of molecular surfaces and alpha shapes, and proposed a method for triangulating molecular surfaces. Other molecular surface formulations serve geometric or illustrative purposes. Minimal molecular surfaces [13] choose a surface that minimizes Gibbs free energy by minimizing mean curvature of the surface geometry. Other surface abstractions [14, 15] have been used 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 [16] using compact polynomial basis functions. Direct volume rendering, as a general visualization modality, was introduced by

Levoy [17]. Electronic structure data from computational chemistry codes such as DFT are inherently volumetric scalar fields. In general, the use of isosurfaces for representing electron density clouds is well accepted in the chemistry literature, though not as an outright replacement for molecular surfaces [18]. Goodsell et al. [19] demonstrated direct volume rendering of charge density fields. Similar work explored volume rendering of electron density from *ab initio* molecular dynamics [20]. Qiao et al. [21, 22] explore volume rendering of electron orbitals in quantum dot device simulations. Jang et al. [23] propose direct evaluation of wavefunctions for accurately volume rendering molecular orbitals.

To our knowledge, no paper has considered uncertainty in the definition of a molecular surface with respect to its electronic structure. Hu et al. [24] 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 rendering of multifield data with associated uncertainty was explored with 1D and 2D transfer functions [25]. Uncertainty-driven multidimensional classification of scalar volume data has also been explored [26]. Uncertainty pertaining to isosurfaces with computed error has also been visualized [27]. Point-based approaches allow for expression of spatial uncertainty in both polygonal and volumetric data [28].

Volume approximations of charge density and other atomic properties are implemented for fixed radial distributions in VMD [7] (e.g. the VolMap tool). We improve on this by incorporating density distributions corresponding to different interatomic bond lengths, from smaller-scale bulk DFT computation, into our model. While isosurfacing is ubiquitous, direct volume rendering remains uncommon in molecular visualization tools [7–9]. In our work, we have developed a prototype domain-specific visualization tool for classifying and visualizing molecular boundaries based on electronic structure.

3. BACKGROUND

Molecular structure can be described in both intramolecular and intermolecular regions by continuous scalar fields of electron density (or charge density). In quantum mechanics, electron density measures the probability of an electron being present at a given spatial location [29]. Solving Schrödinger’s equation using a linear combination of atomic orbitals (LCAO) results in wavefunctions with radial and angular momentum for each electron [30].

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

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 [31]) or inverse (density functional theory [32]) approaches. Molecular orbitals consist of the sum of all atomic orbitals.

$$\psi(r, \theta, \phi) = \sum_N \varphi_{nlm}(r, \theta, \phi) \quad (2)$$

The wavefunction amplitude ψ is complex-valued, and the electronic charge density ρ is given as

$$\rho = e\psi^2 \quad (3)$$

Where e is the electronic charge.

From computed wavefunction coefficients, one can then plot charge density as a structured grid of real numbers, and visualize or analyze this data directly. However, although this scalar field defines the overall structure of the molecule, it does not define boundaries. Picking different isovalues of charge density can lead to different interpretations of molecular shape and size. For molecules expecting to donate or accept electrons, the interface of interest is often given by the scalar field of the highest-occupied molecular orbital (HOMO) and lowest-unoccupied molecular orbital (LUMO). However, these orbitals are still distributed over relatively wide (multi-Angstrom) regions, and are

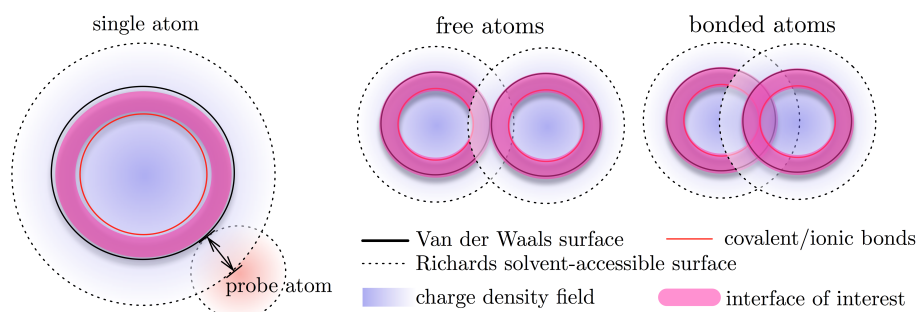


FIG. 2: Atomic radii. Regions between chemical bond radii and Van der Waals radius can be of interest in defining a molecular interface.

poorly suited for visualization as a single isosurface. Moreover, many interactions of interest do not involve interchange of electrons or formation of new bonds. Ideally, we would like a method of classifying charge density that gives an accurate indication of the actual shape and size of the molecule, and is agnostic to the type of interaction.

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 [33]. This is typically small compared to uncertainty associated with the type of interface interaction.

3.2 Molecular surfaces

Volume rendering of electron density has gained relatively little traction, and there are few guidelines for choosing isovalues of electron density. Common practice is to pick some minimal nonzero isovalue corresponding to an outer isosurface. While adequate for visualization, the resulting isosurfaces vary greatly, and this approach is not suitable for analysis.

Defining the surface of a molecule is difficult 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 elementary particles (e.g. from a scanning electron microscope). The simplest “CPK” molecular surface model consists of overlapping spheres [3, 4]. Generally, Van der Waals spheres form a convex hull of the molecule, but spheres consisting of ionic or covalent radii may be disjoint. In the biomolecular community, problems such as protein docking have motivated the use of solvent-excluded surfaces (SES) and solvent-accessible surfaces (SAS) [5, 6]. 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. While SES can represent boundaries inside the Van der Waals radius, they have limited relation to physical molecular structure. Another solution to forming continuous geometry is to model the surface implicitly using radial basis functions such as those of Blinn [16]. However, arbitrary basis functions have similarly limited 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 over a charge density distribution. In turn, this can be used to classify wavefunction or electron density fields.

We make two simplifying assumptions in our treatment of charge density fields. We work with the real-valued wavefunction amplitude ψ' , defined as clamp negative values of ρ to zero and ignore complex values of the wavefunction ψ . Since we assume we are operating on the same units as our computed DFT dataset, we also ignore the constant e from Equation 3. Then, from charge density ρ plotted in DFT, we compute real wavefunction amplitude:

$$\psi' = |\psi|\sqrt{e} = \sqrt{\rho} \quad (4)$$

This derived quantity ψ' has units $\sqrt{e/\text{\AA}^3}$, and is uniquely defined by the charge density ρ . Using ψ' is advantageous for classification since it is positive-definite with $\psi' > 0$ even outside the Van der Waals radius. Like ρ , $\psi'(r)$ (usually) varies monotonically with r outside the core region, and can be classified unambiguously. However, classifying ψ' is numerically cleaner, since ρ is frequently small, transfer functions use quantized ranges (often 8-bit), and DFT computations output spatially low-resolution volume data that is sensitive to interpolation. Lastly, as shown in Section 4.3, ψ' lends itself to radial basis function modeling of a positive-definite approximate scalar field, Ψ .

To perform our classification, we first determine the spatial interval

$$\bar{r} = [r, \bar{r}] \quad (5)$$

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

$$\overline{\psi'} = [\psi'(r), \psi'(\bar{r})] \quad (6)$$

Then, classification of wavefunction amplitude in $\overline{\psi'}$ corresponds 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 distributions with primarily radial (s orbital) behavior such as the O-O example in Figure 3, it suffices to plot wavefunction amplitude 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 discussed above does not hold. This is shown in the C-C bond in Figure 3 (right), which has non-monotonic ψ' (the $\theta = 0$ distribution). Moreover, ψ' exhibits low electron density at certain orientations, and significantly higher density at others. When ψ' is non-monotonic, $\psi'(r)$ is no longer a bijection, 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 4.3)
- Classify the minimum and maximum ψ' 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 ψ' for all θ, ϕ (or a subset thereof). While more ambiguous for inner orbitals, this choice has no adverse impact on classification of outer orbital geometry near the Van der Waals radius, which is a better indicator of chemical bond geometry.

We note that the C-C dimer example is shown for illustrating potentially non-monotonic orbitals; it is a poor choice for modeling the classification of larger carbon structures in this paper.

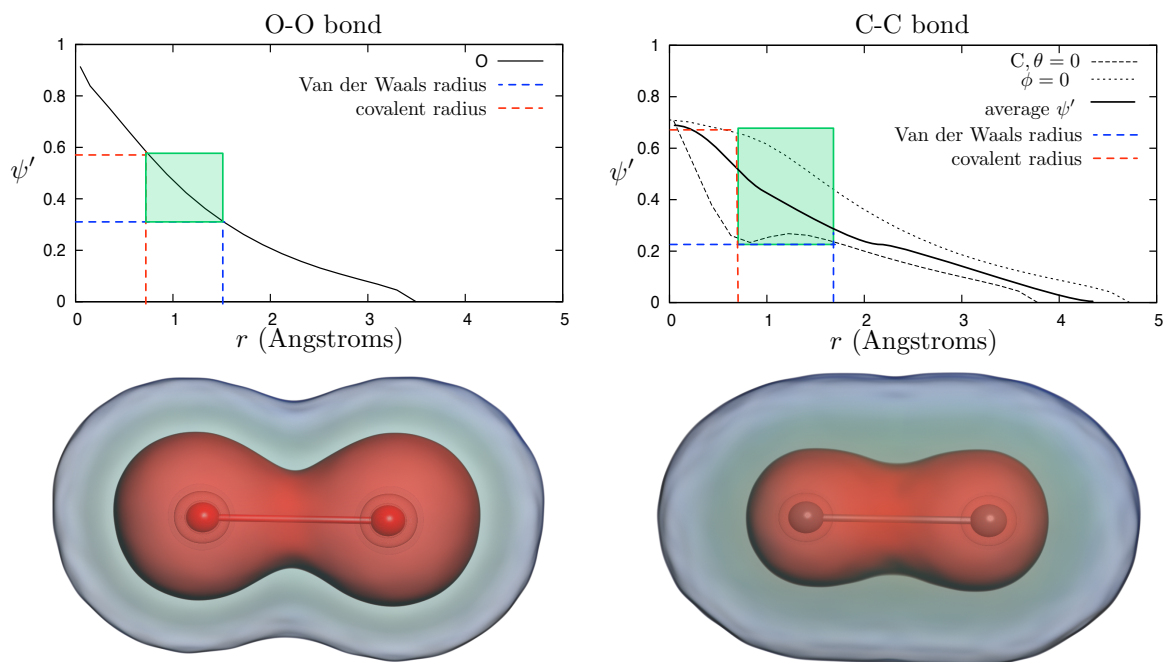


FIG. 3: Wavefunction amplitude ψ' , as a function of radius r in Angstroms, in O-O and C-C bonds. The interval between Van der Waals and covalent radii defines as interval in ψ' that we use in classification. For the anisotropic C-C orbitals, we show minimum and maximum distributions at angular coordinates of $\theta = 0$ and $\phi = 0$, respectively, and average ψ' over all θ, ϕ . The blue and red isosurfaces corresponding to the lower (outer) and upper (inner) extents of this interval, respectively. Most interactions of interest, and thus the material interface, will occur in between.

4.2 Varying bond profiles

Bonds between the same two atoms vary depending on their component orbitals. Even for atoms of the same type in a homogeneous structure, bond length and electron density profile can change significantly. An example is the carbon nanosphere shown in Figure 4. While this structure consists mostly of C6 graphene chains with sp^2 covalent bonds, imperfections can exhibit different behavior, resulting in different bond lengths and consequently a range of different distributions. Similarly to how we can build an interval $\overline{\psi'}$ for a distribution of profiles for anisotropic ψ' molecular orbitals, we can represent a distribution of multiple bond specimens of the same diatomic pair.

Chiefly, we are interested in density (and by extension ψ') along bond lines and exterior regions of the structure. For diatoms, this is trivial to determine based on minimum and maximum profiles. For systems with more than two atoms, the maximal profile can exhibit interference from other atoms in the system, obscuring the lower classification bound (the Van der Waals radius) toward the tail. As illustrated in Figure 4, this is also the case for the average ψ' distribution. To accurately approximate the tail, one could for example solve for a least-squares fit of radial basis functions that satisfy the average distribution, based on all or a subset of profiles in the system. A less costly approach is to use the minimum profile as a lower bound and the average profile as an upper bound. Additional isovalues, such as the minimum of *bond* profiles, can be of interest within this interval.

4.3 Approximate wavefunction fields

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

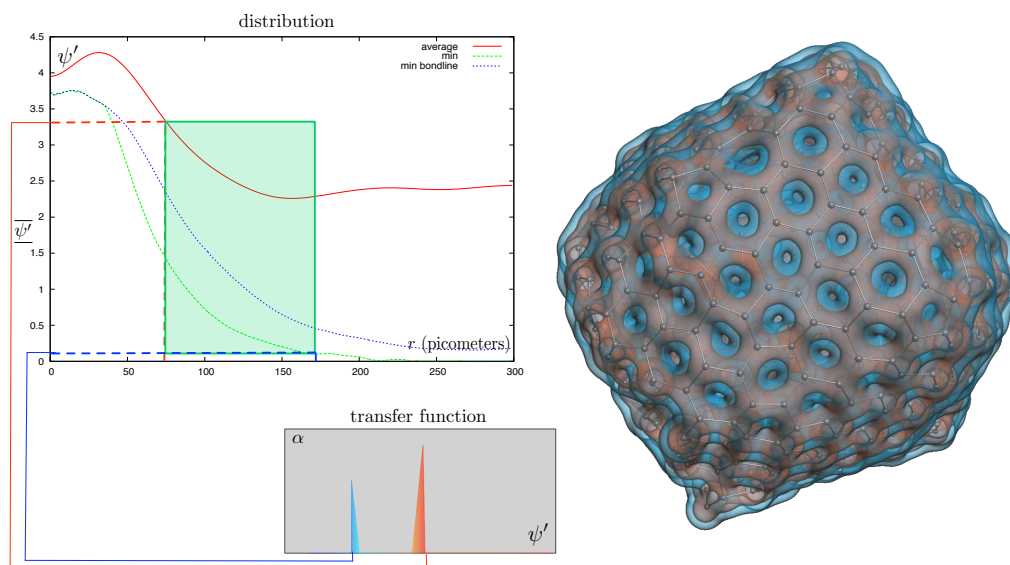


FIG. 4: Classification of a 732-atom carbon nanosphere from DFT computation in VASP. We show how the spatial interval between covalent and Van der Waals radii is mapped to an interval $\underline{\psi'}$ of wavefunction amplitude, which is then mapped to isosurfaces or features in a transfer function.

for this kind of problem [34], and indeed our proposed solution is similar to the one first used by Blinn [16]. Our contribution is the use of the average wavefunction amplitude 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 electron distribution over radius. If we divide this by the number of electrons N , we have an average distribution

$$\Theta(r) = \frac{1}{N} \int \psi'(r, \theta, \phi) d\theta d\phi \quad (7)$$

This is shown by the solid black line in Figure 3. In practice, density distribution varies with the chosen DFT approach (plane-wave or LCAO) and whether or not core electrons are fully simulated. Although density in the outer valence regions is relatively low, these regions 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 with position p_i and $\vec{x} \in \mathfrak{R}^3$,

$$r_i(\vec{x}) = |\vec{x} - \vec{p}_i|$$

$$\Psi(\vec{x}) = \sum_i \Theta_i(r_i(\vec{x})).$$

The kernel density model is a per-atom approximation of per-electron phenomena, and is not a replacement for computation of orbitals from DFT. However, it suffices for large molecules with predictable, homogeneous structure. As with ψ' 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 [35] 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.

DFT computations on bulk compounds were performed in VASP [36]. Utilities for computing the radial distributions and generating the corresponding uncertainty interval and transfer function are written in C++, accessible either as a standalone program or a library.

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 most diatomic examples. However, DFT for larger systems consisting of thousands of atoms require greater computational resources.

5.1 Transfer Functions

Given the interval $\overline{\psi'}$, we can create a wide range of transfer functions that express the electron density volume with respect to the chemical bond and Van der Waals thresholds. In Figures 3 and 4 we use a simple RGB warm-cold colormap, with peaks at each threshold to suggest “inner” and “outer” surfaces. While it is straightforward to automatically generate such a transfer function, or indeed two isosurfaces, it is often useful to tweak the transfer function manually to highlight desired phenomena. In Figure 6, we use a low-opacity (purple-red) ramp to highlight the uncertainty region between these thresholds, which peaks at an opaque green feature corresponding to ionic bond radius. In the lower-right panel of Figure 8, we choose a ramp to show only features *outside* the $\overline{\psi'}$ interval, allowing us to visualize void space.

6. RESULTS AND USE CASES

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 nanosphere our renderer outperforms rasterization of isosurface and ball-and-stick models in VMD [7] by up to 10x, with no required isosurface precomputation. 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 computations

A straightforward application of our technique is classifying data directly from DFT computation. An example of this is shown on the small carbon nanosphere in Figure 4. While it is possible to volume render ρ directly using a linear ramp (e.g. opacity directly proportional to charge density), occlusion often obscures geometric features inside the scalar field. The main advantage of our semi-automatic approach over naive classification is that we have some physical gauging of appropriate isovalues based on empirical measurements.

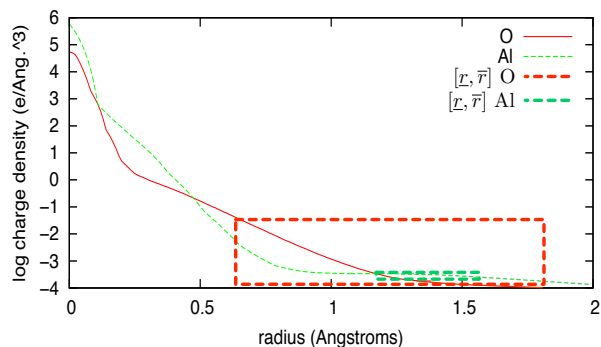


FIG. 5: Charge density ρ plotted as a function of radius for Al and O from bulk DFT computation, used to model radial basis functions $\psi'(r)$ of electron density in nanobowl MD data.

While our approach works well for bulk compound computations, we note that it is less appropriate for classification of DFT simulations involving complex chemical bonding. Nonetheless, even in these computations, a spatial gauge relating ρ or ψ can be useful, given the appropriate distributions.

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 4.3.

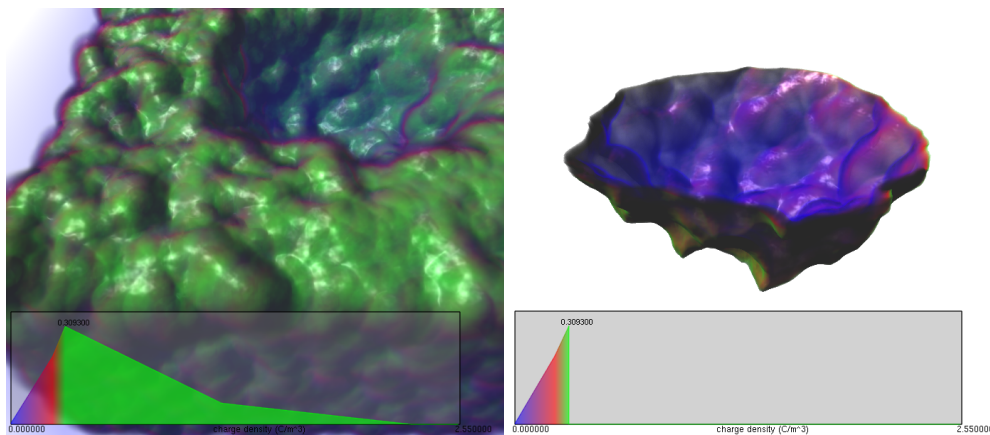


FIG. 6: Classification of charge density ρ applied to our approximate electron density volume $\psi' = \sqrt{\rho}$. We show uncertainty in the material interface between ionic and Van der Waals radii via a red-blue transition in these transfer functions. This provides a simple mechanism for segmenting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

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 [36]. We used these distributions of ρ to determine amplitude $\psi' = \sqrt{\rho}$, which we then used to model our approximate wavefunction volume Ψ . Classification is performed in units of ρ . 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 $\bar{\Psi} = \sqrt{\bar{\rho}}$ 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 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 \AA^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: Nanospheres

Our last application involves visualization of large amorphous carbon nanosphere structures from molecular dynamics. Appearing as ordinary soot to the human eye, 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 [37] for various-size structures, the largest of which is 740,000 atoms. Structurally, the nanosphere is a mix of diamond (sp^3) and graphite (sp^2) carbon configurations. Due to its large size, neither ball-and-stick nor Connolly surfaces are helpful in visualizing the full nanosphere 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 allow lithium-ion transport. Clearly, DFT simulation is prohibitively costly for structures of this size.

To visualize this structure, we employ our kernel density model using the minimum bond line wavefunction distribution shown in Figure 4, followed by uncertainty classification. 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 ions may be able to pass through. 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. Volume classification using the model shown in Figure 4 helps us flexibly understand the extents of the material interface, and may aid in correlation 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.

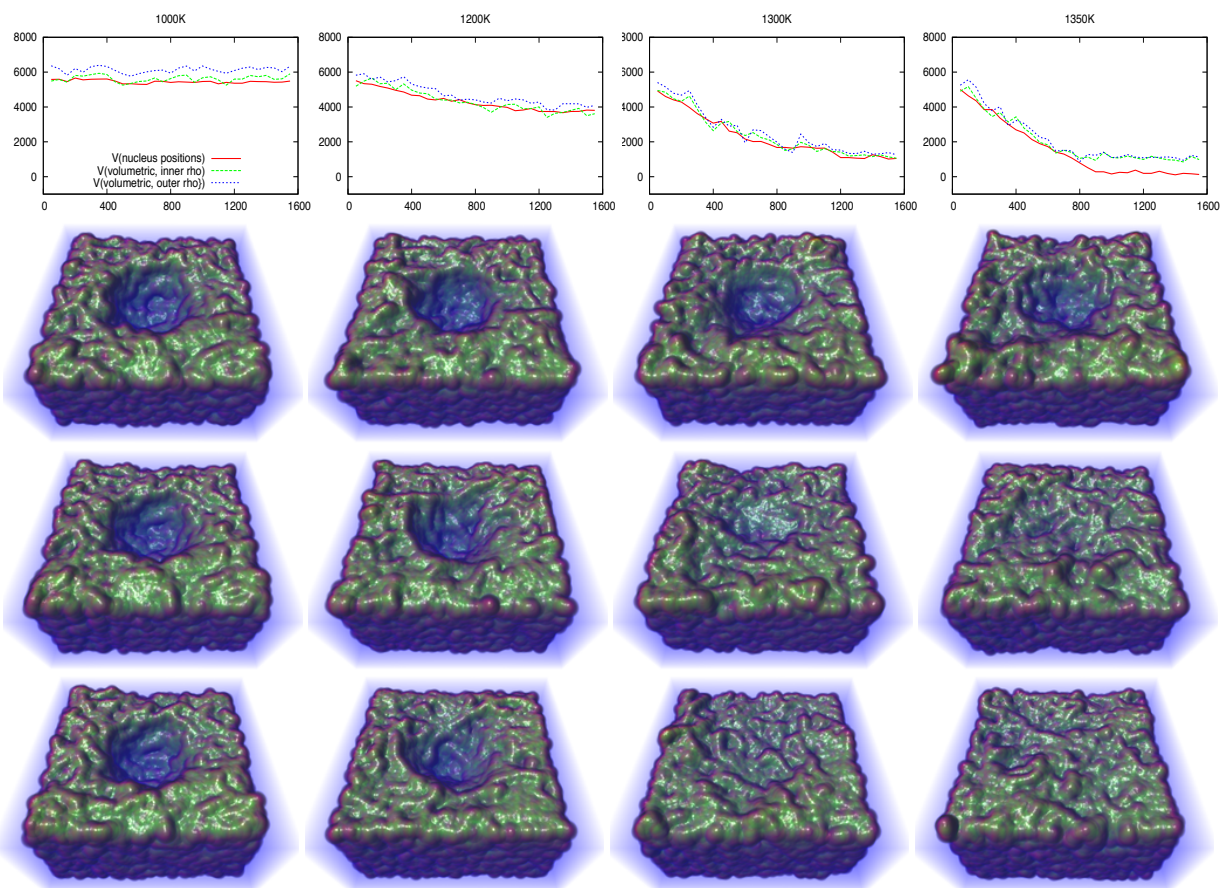


FIG. 7: Top: analysis from our classification, plotting bowl volume (\AA^3) over the simulation time (0 - 1550 picoseconds) for separate MD nanobowl runs at 1000K, 1200K, 1300K and 1350K temperatures, shown in the four columns. Bottom rows show visualizations at 260, 800 and 1500 ps, respectively, at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

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). Additional support was provided from an Early Career Award (J.G.). Liu, Lau, Chan, Greeley and Curtiss were supported by the Center for Electrical Energy Storage: Tailored Interfaces, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences. Use of the Center for Nanoscale Materials was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract No. DE-AC02-06CH11357. The authors also thank Aslihan Sumer and Julius Jellinek in the ANL Chemical Sciences and Engineering Division for their insights.

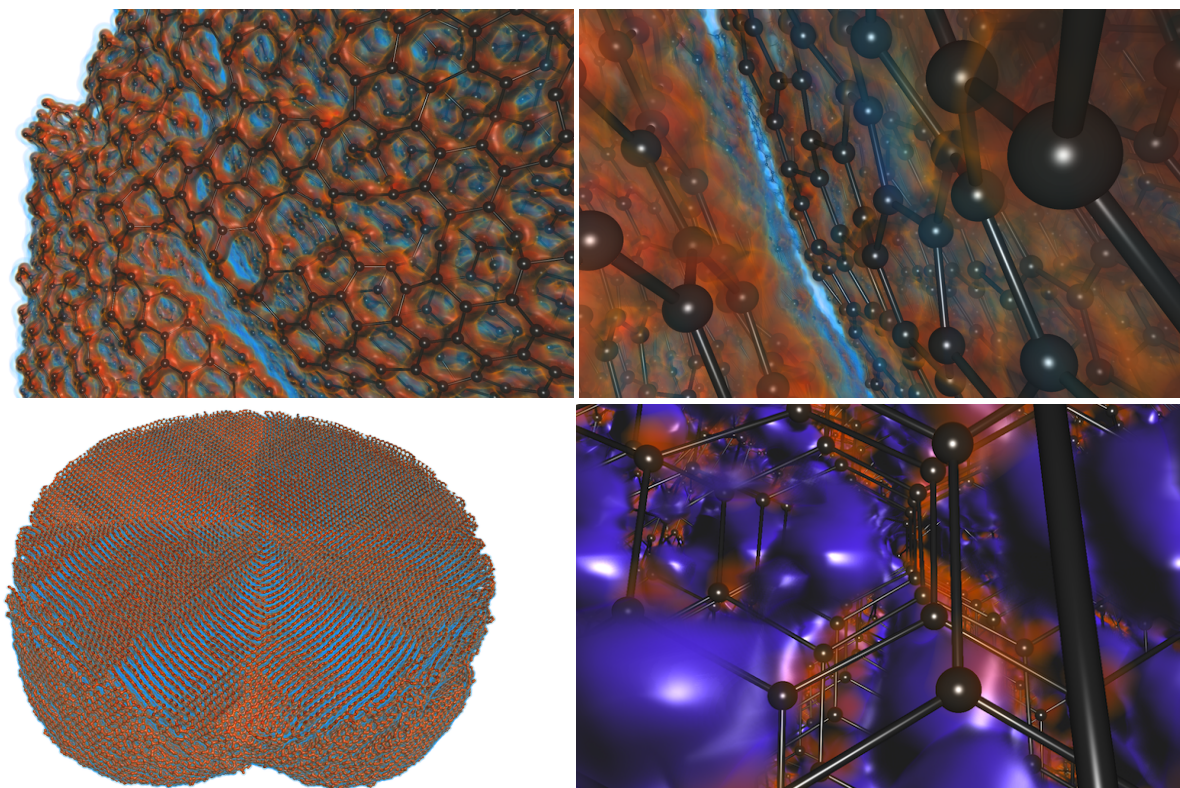


FIG. 8: Amorphous carbon nanosphere structures resulting from thermal annealing. Uncertainty classification lets us better explore material interfaces, including channels where Lithium ions may diffuse.

REFERENCES

1. Connolly, M., Molecular Surfaces: A Review, *Network Science Online*, 1996, <http://www.netsci.org/Science/Compchem/feature14.html>.
2. Odonnell, T., The Scientific and Artistic Uses of Molecular Surfaces, *Network Science Online*, 1996, <http://www.netsci.org/Science/Compchem/feature15.html>.
3. Corey, R. B. and Pauling, L., Molecular Models of Amino Acids, Peptides and Proteins, *Applied Physics Letters*, 24:621–627, 1953.
4. Loltun, W. L. Space Filling Atomic Units and Connectors for Molecular Models. US Patent 3170246, 1965.
5. Lee, B. and Richards, F., The Interpretation of Protein Structures: Estimation of Static Accessibility, *Journal of Molecular Biology*, 55:379–400, 1971.
6. Connolly, M., Analytical Molecular Surface Calculation, *Journal of Applied Crystallography*, 16:548–558, 1983.
7. Humphrey, W., Dalke, A., and Schulten, K., VMD: visual molecular dynamics, *Journal of molecular graphics*, 14(1):33–38, 1996.
8. Avogadro: an open-source molecular builder and visualization tool, <http://avogadro.openmolecules.net/>.
9. Schrödinger, LLC, . The PyMOL molecular graphics system, version 1.3r1. August 2010.
10. Varshney, A., Brooks Jr, F., and Wright, W., Computing smooth molecular surfaces, *Computer Graphics and Applications, IEEE*, 14(5):19–25, 1994.
11. Sanner, M., Olson, A., and Spehner, J., Reduced surface: an efficient way to compute molecular surfaces,

- Biopolymers*, 38(3):305–320, 1996.
12. Akkiraju, N. and Edelsbrunner, H., Triangulating the surface of a molecule, *Discrete Applied Mathematics*, 71(1):5–22, 1996.
 13. Bates, P., Wei, G., and Zhao, S., Minimal molecular surfaces and their applications, *Journal of Computational Chemistry*, 29(3):380–391, 2008.
 14. Cipriano, G. and Gleicher, M., Molecular surface abstraction, *IEEE Transactions on Visualization and Computer Graphics*, pp. 1608–1615, 2007.
 15. Zwan, M.v. d., Lueks, W., Bekker, H., and Isenberg, T., Illustrative molecular visualization with continuous abstraction, In *Computer Graphics Forum*, Vol. 30, pp. 683–690. Wiley Online Library, 2011.
 16. Blinn, J., A Generalization of Algebraic Surface Drawing, *ACM Transactions on Graphics*, 1(3):235–256, July 1982.
 17. Levoy, M., Efficient Ray Tracing for Volume Data, *ACM Transactions on Graphics*, 9(3):245–261, July 1990.
 18. Shusterman, A. J. and Shusterman, G. P., Teaching chemistry with electron density models, *Journal of Chemical Education*, 74(7):771, 1997.
 19. Goodsell, D. S., Mian, I. S., and Olson, A. J., Rendering volumetric data in molecular systems, *Journal of Molecular Graphics*, 7:41–47, March 1989.
 20. Covick, L. and Sando, K., Volume-rendering representations of ab initio electron densities, *Journal of Molecular Graphics*, 11(1):37 – 39, 1993.
 21. Qiao, W., Ebert, D., Entezari, A., Korkusinski, M., and Klimeck, G., Volqd: Direct volume rendering of multi-million atom quantum dot simulations, In *Visualization, 2005. VIS 05. IEEE*, pp. 319–326. IEEE, 2005.
 22. Qiao, W., McLennan, M., Kennell, R., Ebert, D., and Klimeck, G., Hub-based simulation and graphics hardware accelerated visualization for nanotechnology applications, *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):1061–1068, 2006.
 23. Jang, Y. and Varetto, U., Interactive volume rendering of functional representations in quantum chemistry, *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1579–5186, 2009.
 24. Hu, M., Chen, W., Zhang, T., and Peng, Q., Direct volume rendering of volumetric protein data, *Advances in Computer Graphics*, pp. 397–403, 2006.
 25. Djurcilov, S., Kim, K., and Pang, A., 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*, p. 243. Springer Verlag Wien, 2001.
 26. Kniss, J., Uitert, R., Stephens, A., Li, G., Tasdizen, T., and Hansen, C., Statistically quantitative volume visualization, In *Proceedings IEEE Visualization*, pp. 287–294, 2005.
 27. Rhodes, P., Laramée, R., Bergeron, R., and Sparr, T., Uncertainty visualization methods in isosurface rendering, In *Eurographics*, pp. 83–88. Citeseer, 2003.
 28. Grigoryan, G. and Rheingans, P., Point-based probabilistic surfaces to show surface uncertainty, *IEEE Transactions on Visualization and Computer Graphics*, 10(5):564 –573, sept.-oct. 2004.
 29. Pauling, L. and Wilson, E. B., *Introduction to quantum mechanics: with applications to chemistry*, Courier Dover Publications, 1985.
 30. Quinn, C., *Computational quantum chemistry: an interactive guide to basis set theory*, 2002.
 31. McMillan, W. L., Ground state of liquid he⁴, *Phys. Rev.*, 138:A442–A451, Apr 1965.
 32. Hohenberg, P. and Kohn, W., Inhomogeneous electron gas, *Phys. Rev.*, 136:B864–B871, Nov 1964.
 33. Bondi, A., Van der Waals Volumes and Radii, *The Journal of Physical Chemistry*, 68(3):441–451, 1964.
 34. Silverman, B., *Density Estimation for Statistics and Data Analysis*, Vol. 26, Chapman & Hall/CRC, 1986.
 35. Knoll, A., Hijazi, Y., Westerteiger, R., Schott, M., Hansen, C., and Hagen, H., Volume Ray Casting with Peak

- Finding and Differential Sampling, *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1571–1578, Nov-Dec 2009.
36. Kresse, G. and Hafner, J., Ab initio molecular dynamics for liquid metals, *Physical Review B*, 47(1):558, 1993.
 37. Plimpton, S., Fast parallel algorithms for short-range molecular dynamics, *Journal of Computational Physics*, 117(1):1–19, 1995.

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.