Morse-Smale Analysis of Ion Diffusion in *Ab Initio* Battery Materials Simulations

Attila Gyulassy, Aaron Knoll, Kah Chun Lau, Bei Wang, Peer-Timo Bremer, Michael E. Papka, Larry A. Curtiss, Valerio Pascucci

Abstract

Ab initio molecular dynamics (AIMD) simulations are increasingly useful in modeling, optimizing and synthesizing materials in energy sciences. In solving Schrödinger's equation, they generate the electronic structure of the simulated atoms as a scalar field. However, methods for analyzing these volume data are not yet common in molecular visualization. The Morse-Smale complex is a proven, versatile tool for topological analysis of scalar fields. In this paper, we apply the discrete Morse-Smale complex to analysis of first-principles battery materials simulations. We consider a carbon nanosphere structure used in battery materials research, and employ Morse-Smale decomposition to determine the possible lithium ion diffusion paths within that structure. Our approach is novel in that it uses the wavefunction itself as opposed distance fields, and that we analyze the 1-skeleton of the Morse-Smale complex to reconstruct our diffusion paths. Furthermore, it is the first application where specific motifs in the graph structure of the complete 1-skeleton define features, namely carbon rings with specific valence. We compare our analysis of DFT data with that of a distance field approximation, and discuss implications on larger classical molecular dynamics simulations.

1 Introduction

First principles (*ab initio*) simulations of molecular structures, employing density functional theory (DFT) or Hartree-Fock (HF) methods, are increasingly common

Attila Gyulassy, Aaron Knoll, Bei Wang, Valerio Pascucci
SCI Institute, University of Utah, e-mail: jediati, knolla, beiwang, pascucci @sci.utah.edu

Kah Chun Lau, Michael E. Papka, Larry A. Curtiss
Argonne National Laboratory, e-mail: kclau, papka, curtiss @anl.gov

Peer-Timo Bremer
Lawrence Livermore National Laboratory, e-mail: bremer5@llnl.gov

in materials science applications. Unlike classical dynamics simulations, they solve the Schrödinger equation to more accurately determine molecular geometry. They compute the wavefunction, the electronic structure of the molecules they simulate, which consists of scalar fields of component molecular orbitals. Most analysis of such simulations is carried out on the resulting molecular geometry as opposed to the wavefunction. However, the wavefunction offers advantages: it is the closest we can come to "ground truth" concerning the structure of the molecule, and its volumetric representation can be used for topological analysis of scalar fields.

We consider an application in materials science: determining the charge capacity of simulated battery anode structures. The material in question is a sphere of carbon sculpted from a solid block of graphite, heated to a high temperature (2500 Kelvin) via molecular dynamics, and then annealed. The resulting "nanosphere" resembles ordinary graphitic soot, but possesses numerous channels that can accommodate lithium ion electrolyte. These channels form as a result of defects in the graphite structure, which in turn change the coordination number (valence) of their component carbon atoms. Higher average coordination number indicates more defect sites, and structures that are better-suited as battery anodes. While effective, simple statistical analysis does not fully quantify the charge capacity of these structures. To do that, we must understand the paths that lithium ions may diffuse through inside the carbon nanosphere structure.

The goal of this work is to use Morse-Smale analysis to tackle this problem, analyzing the scalar field of the wavefunction itself from DFT computation. We compute a 1-skeleton of the Morse-Smale complex to determine likely ion diffusion paths in the nanosphere, identifying defect sites in the carbon structure through which lithium ions may pass. We present a model for these features, and examine statistical properties of the topology to establish a methodology to analyze such data. We apply this technique to extract and compare results from the DFT-computed wavefunction to those of a distance field.

2 Background

Our scientific goal is to examine the structure of carbon nanospheres throughout the heating and annealing process, and analyze the suitability for the resulting structure as a battery anode material. To do this at a relatively small scale (hundreds of atoms), but ensure higher physical accuracy, we conduct an *ab initio* molecular dynamics (AIMD) simulation, specifically DFT computation using the VASP code [18].

In experiments, the monodispersed carbon nanospheres that are used for a sustainable lithium energy storage electrode can be synthesized effectively by autogenic reactions of hydrocarbon precursors (e.g. polyethylene from plastic waste, etc.) at high temperature and pressure, enabling synthesis of a battery anode from recycled materials [29]. From the reported studies, the unique carbon microstructures (*i.e.* layered graphitic motifs and sufficient carbon defects) are critical in promoting lithium diffusion into and out from the interior of carbon nanospheres for a practical lithium ion battery operations (*i.e.* capacity, voltages and charging rates).
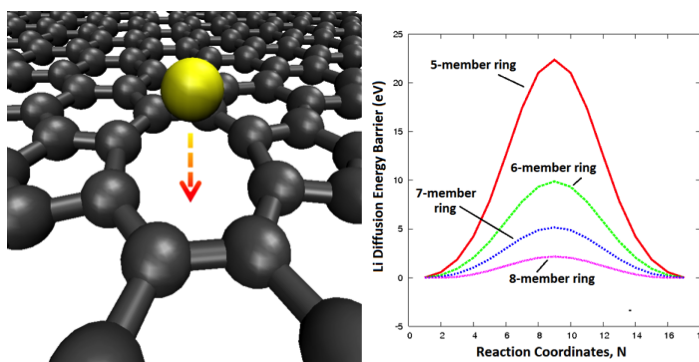
Fig. 1 *Left:* the symmetric feature of the energy barriers are due to the Li ion diffusion into and out of the center of n-membered ring at a single graphene sheet. *Right:* The Li ion diffusion energy barrier (in eV) obtained from DFT calculation.

In order to understand how Li ion can be diffused into or from (or intercalated into or de-intercalated from) carbon microstructures, the Li ion diffusion energy barriers along the diffusion paths within the carbon microstructures of an electrode have to be determined at atomistic level. To accurately quantify the Li ion diffusion dynamics, DFT nevertheless remains the preferred choice in theoretical descriptions at the atomistic level. In addition to computing physically accurate atom geometry, solving the Schrödinger equation explicitly generates the electronic structure (electron or charge density cloud), volume data on which we can apply topological analysis.

From the reported studies [29], the interlayer Li diffusion paths are most probably determined by the presence of large n-membered ring (e.g. n > 6) defect sites due to the extensive thermal graphitization and significant carbon dislocations during the high temperature synthesis of carbon. To model the anticipated Li ion diffusion energy barrier through the n-membered rings, the atomistic simulation is carried out based on DFT calculations with plane wave basis sets as implemented in the VASP code [18]. All the DFT calculations were spin-polarized and carried out using the gradient corrected exchange-correlation functional of Perdew, Burke and Ernzerhof (PBE) [26] under the projector augmented wave (PAW) method, with plane wave basis sets up to a kinetic energy cutoff of 400 eV. For a Li-ion diffusion barrier, as the size of the n-membered ring increases, the barrier for a Li ion to diffuse through decreases significantly, as confirmed by the DFT study in this work (Figure 1) and previous studies [35]. For the larger 9-membered ring, the Li diffusion energy barrier is even smaller, *i.e.* 0.15 eV, slightly smaller than the reported value (*i.e.* 0.5 eV) in carbon nanotubes [24].

In addition to large n-membered rings of carbon defects, the Li diffusion is also facilitated through the intra-layer in-plane diffusion within the layered graphite present in carbon nanospheres [19]. From the computed free energy surface explored by our CPMD metadynamics simulation, we found the Li in-plane diffusion barrier at arbitrary direction in between the graphitic layers is comparatively small (i.e. 8 kcal/mol = 0.35 eV), which is consistent with the reported values (i.e. 0.26 – 0.50 eV) [30]. Thus, it is reasonable to assume that the diffusion dynamics of Li ions are

driven by inter-layer diffusion through n-membered rings (referred to as *defects* in the perfect C-6 carbon ring structure) as well as intra-layer movement.

We can pair these insights with topological analysis to identify Li-accessible tunnels inside the nanosphere, The approach of our work is to use the Morse-Smale complex to define an initial skeleton of minima and 1-saddles, count the size of carbon rings (number of maxima) adjacent to that skeleton to determine which tunnels an ion would likely pass through, and thus extract the diffusion path skeleton. Moreover, the 1-skeleton represents both inter and intra-layer diffusion paths, making it well-suited for the analysis of this particular structure.

2.1 Morse-Smale Complex

The following provides a brief introduction to the Morse-Smale complex, which we use to identify features in the DFT data, and topological simplification, used to study the function at multiple scales and reason about the stability of the identified features.

Morse functions and the Morse-Smale (MS) complex. Let f be a real-valued smooth map $f : \mathbb{M} \to \mathbb{R}$ defined over a compact d-manifold $\mathbb{M}$. A point $p \in \mathbb{M}$ is critical when $|\nabla f(p)| = 0$, *i.e.* the gradient is zero, and is non-degenerate when its Hessian (matrix of second partial derivatives) is non-singular. The function f is a *Morse function* if all its critical points are non-degenerate and no two critical points have the same function value. In this case the *Morse Lemma* states that there exists local coordinates around p such that f has the following *standard form*: $f_p = \pm x_1^2 \pm x_2^2 \cdots \pm x_d^2$. The number of minus signs in this equation gives the *index* of critical point p. In three-dimensional functions, minima are index-0, 1-saddles are index-1, 2-saddles are index-2, and maxima are index-3.

An integral line in f is a path in $\mathbb{M}$ whose tangent vector agrees with the gradient of f at each point along the path. The integral line passing through a point p is the solution to $\frac{\partial}{\partial t} L(t) = \nabla f(L(t)), \forall t \in \mathbb{R}$, with initial value $L(0) = p$. Each integral line has an origin and destination at critical points of f, at $t = \pm\infty$. *Ascending* and *descending* manifolds are obtained as clusters of integral lines having common origin and destination respectively. The descending manifolds of f form a cell complex that partitions $\mathbb{M}$; this partition is called the *Morse* complex. Similarly, the ascending manifolds also partition $\mathbb{M}$ in a cell complex. A Morse function f is a *Morse-Smale function* if ascending and descending manifolds of its critical points only intersect transversally. An index-i critical point has an i-dimensional descending manifold and a $(d-i)$-dimensional ascending manifold. The simply-connected cells formed by the intersections of ascending and descending manifolds form the cells of the Morse-Smale (MS) complex. A three-dimensional MS complex is a cell complex where cells of dimension zero through three are called nodes, arcs, quads, and crystals, respectively. Each arc is a 1-manifold bounded by two nodes, 0-manifolds, each quad is a 2-manifold bounded by arcs, and finally, each crystal of the MS complex is bounded by quads. Cells of the MS complex satisfy several combinatorial properties: end points of arcs are critical points whose indices differ exactly by one; quads

contain exactly four arcs on their boundary (although some might be repeated); and the closure of the boundary of a crystal contains a collection of quads, arcs, saddles and exactly one minimum and one maximum. The *1-skeleton* of the MS complex is formed by the nodes and arcs, representing much of the connectivity information of the complex.

Topological simplification. A function f is simplified by repeated cancellation of pairs of critical points connected by an arc in the MS complex. The local change in the MS complex indicates a smoothing of the gradient vector field and hence of the function f. Forman [9] showed how a cancellation could be achieved in a discrete gradient field by reversing the gradient path between two critical cells. Gyulassy et al. [12] provided a full characterization of cancellation operations in terms of how they affect the connectivity of the complex and the geometry of the ascending/descending manifolds, operating solely on the combinatorial structure of the complex. Each cancellation operation removes a pair of critical points, reconnects arcs of the complex, and merges their ascending and descending manifolds with their neighbors geometry. Repeated application of cancellations in order of *persistence*, the absolute difference in function value of the canceled critical points, results in a hierarchy of MS complexes and a multi-resolution representation of features. Gyulassy et al. [13] described data structures and search algorithms to reconstruct the ascending and descending manifolds of any critical point at any stage of simplification, allowing rapid browsing of the extracted features at multiple scales.

3 Related Work

We review works that are most relevant to our proposed techniques. Most of the geometric and topological methods discussed here originate from molecular shape analysis, in particular, in the detection of protein cavities. A cavity is an empty space enclosed by the molecule, and it includes voids (without openings that allow access to the surrounding solvent), pockets and tunnels (with openings). A tunnel connects multiple surface sites through pathways; while a pocket connects a site in the interior with a surface site. We refer the reader to Lidow et al. [23, 22] for illustrations of these scenarios. In our case diffusion paths correspond to the pockets and tunnels of a given molecule. Most of the geometry-based cavity detection algorithms rely on the computation of Voronoi diagram, its dual graph Delaunay triangulation, weighted Delaunay triangulation or its close relative alpha shapes. Because of the large number of work in the field, we include a few significant ones.

Voronoi diagram-based techniques that have a special focus on the analysis and visualization of tunnels include works in [23, 22], tools MOLE [27] and CAVER [28], where the work in [22] detects structures from a molecular dynamics trajectory.

Alpha shapes are closely related to alpha complexes which are subcomplexes of the Delaunay triangulation of the point set. Alpha shape theory [6, 8] has been used for the detection of protein cavities [7, 20, 21] and has been featured in tools such as CAST [20, 5], Proshape [17], CAVER [28] and MolAxis [34]. A related concept that is similar to Alpha shape, but better in terms of remaining connected

for all resolutions, called Beta shape [15], has been proposed to define cavities [16] in molecules.

A few works exist that employ Morse complex or Morse-Smale complex in shape analysis. Sousbie used Morse complexes to compute filamentary structures in of the cosmic web [31]. Topological spines [4], used as visual representations that preserve the topological and geometric structure of a scalar field, have been developed based on the extraction of sparse subsets of the Morse-Smale complex. The works in [3, 25] decompose the protein surfaces into segmented features for the analysis of protein-protein interactions. However these works are restricted to study surface geometry of the proteins, while our proposed technique focuses on both surface and interior structures. On the other hand, Bajaj et. al. [2] model structural features of molecules by computing stable and unstable manifolds of the critical points of the distance function induced by the iso-surface. Their work is most relevant to our algorithm in a sense that both algorithms rely on properties associated with stable and unstable manifolds of critical points of distance functions to the surface of the molecule; and both detect pockets and tunnels simultaneously. The work in [2] forms pockets and tunnels by clustering and merging adjacent stable manifolds of critical points based on their scalar value; such an approach is not applicable to diffusion in carbon nanospheres, since, as we show in section 6, there is no scalar threshold (either of the DFT field or distance field) that distinguishes between ring structures of different valences. The full MS complex has been used to identify atomic structures on volume data, identifying the atoms and bonds in a C4H4 molecule and orbitals of a hydrogen atom [14]. Günther et al. [10] refined topological analysis of electron density, by using a derived gradient for identification of both covalent and non-covalent bonds.

In the chemistry literature, relatively little analysis is carried out on the wavefunction field itself, as opposed to atom geometry. Bader analysis [1] decomposes charge density into regions of uniform gradient each associated with one atom, for example using Voronoi partitioning. It is similar to Morse theory in that it uses gradient descent to partition the scalar field, but would not help in identifying tunnels between 1-saddles and minima.

Our proposed technique uses persistence simplification to separate noise from features, and similar simplification algorithms have been employed to simplify surface features based on MSCs [3, 25, 4, 33] or alpha shapes [32].

4 Li Diffusion Paths in DFT Data

In this section, we first discuss our approach to handling DFT data, then present the model for identifying features in such data, defining features according to the structure of the MS complex. Next, we present justification and validation for the model, analysing the stability of the extracted features. Finally, we discuss details of the methodology, corner cases, and remaining ambiguities.
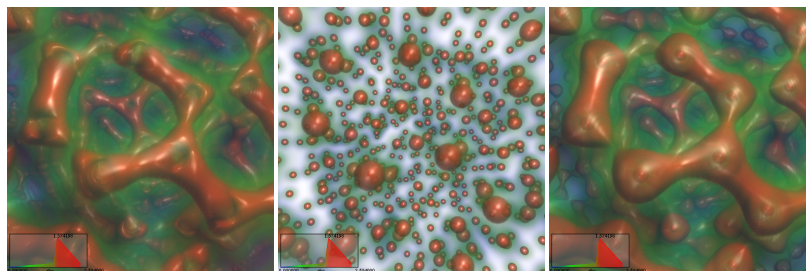
Fig. 2 Modification of DFT data to restore maxima at the nuclei. *Left:* original DFT data *F*. *Center:* RDF impulses *G*. *Right:* Augmented $F' = sup\{F, G\}$.

4.1 Preparation of DFT Data

DFT simulations employ a variety of different functionals and strategies to maximize accuracy and reduce computation time. As interior orbitals (1s, 2s) typically remain static throughout a simulation, it is common to ignore their computation entirely. A side-effect of this optimization is that electron density is ill-defined at the nuclei: whereas the true wavefunction (e.g., obtained through x-ray crystallography or computation of all orbitals) would exhibit maxima at the atomic nuclei, our data instead show empty space (minima) at the nuclei.

As our analysis relies entirely on pre-existing connections within the MS complex to define the skeleton, it is challenging to correct this phenomenon *post facto*. We found it simplest to modify the scalar field itself to restore maxima at the carbon nuclei. To do this, we created a field of summed radial distance functions $G = \sum g_i$, $g_i(d) = Z \, exp(-(d/r)^2)$, where $Z=6$ is the atomic number (maximum possible electron density) and $r = .36$ Ångstrom is half the covalent radius of sp^2 carbon (graphite). Then, given our DFT all-electron density field F, our augmented field is $F' = sup\{F, G\}$. This allows us to correctly identify maxima at the nuclei in the MS complex, and is shown in Figure 2.

We use standard techniques from discrete Morse theory to compute an initial MS complex [11]. Our analysis entails every component of the MS complex taken together, necessitating coherent simplification to remove low-persistence features and artifacts due to discretizing a function onto a mesh. One challenge in using DFT data is the exponential nature of the electron density as a function of distance from an atom. For instance, the features of the 732-atom nanosphere dataset we use in our results span 9 orders of magnitude; the difference between minima and 1-saddles, and 2-saddles and maxima occurs on the scales of 10^{-4} and 10^2, respectively. To obtain a coherent view of the structure, we apply a straightforward strategy of computing the MS complex on the DFT field, but then rescaling critical points using the *log* function before applying persistence simplification.

4.2 Theoretical Model for Analyzing Electron Density

We follow a standard model relating the critical points of electron density to various features of the carbon nanosphere. For each of the following features, we identify the component of the MS complex that corresponds to it, and illustrate the features in figure 3.

Fig. 3 The full MS complex (a) of the 732-atom nanosphere provides the structure for identifying features. Carbon atoms (red spheres) and covalent bonds (yellow arcs) are identified as maxima, high-valued 2-saddles, and the arcs between them (b). Rings are the boundary of ascending 2-manifolds emanating from 1-saddles (c), and rings with valence > 6 are identified as defect sites (d). Free Li ion diffusion paths (e) are the 1-skeleton connecting voids (blue spheres) with low density value.
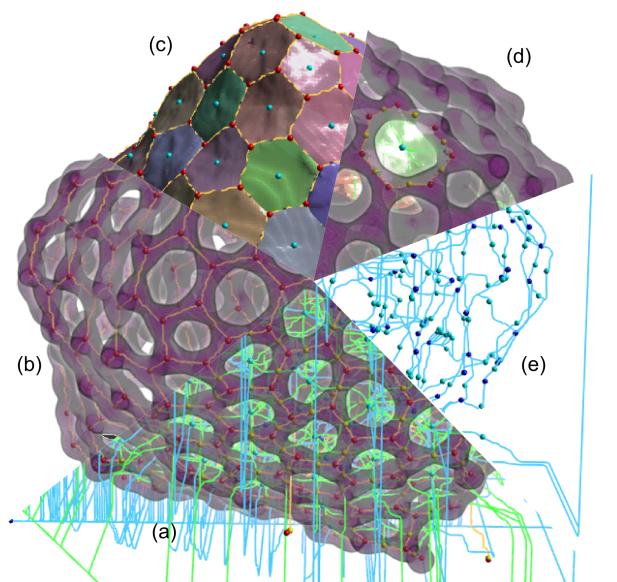
Carbon atoms: It is well-understood in chemistry that the electron density reaches a maximum at atom locations, and in our model, we use maxima of the electron density field as a proxy for carbon atoms.

Covalent bonds: The covalent bonds between carbon atoms also appear as relatively high values, and typically a high-valued 2-saddle connecting two maxima in the 1-skeleton of the MS complex indicates a bond.

Voids: Voids are minima of the electron density field, and represent locally minimal energy configurations. In our model we identify voids as low-valued minima in the MS complex.

Free Li diffusion path: The intra-layer in-plane regions are characterized by regions where Li ions have a comparatively small energy barrier to diffusion, corresponding to low electron density values. We model this as the network connecting voids in the 1-skeleton corresponding to very low electron densities.

N-member rings: The bonds and atoms in a Carbon ring form a circular ridge-line in space, that forms the boundary of a disk. The disk itself separates low-valued voids, and corresponds to the ascending 2-manifold emanating from a 1-saddle in the MS complex. To account for possible pinching of the boundary of a disk, we

compute the valence of a ring as the number of carbon atoms in the largest simple cycle on the boundary.

Defect sites: Defect sites, allowing Li to diffuse across layers, are a carbon rings with valence > 6, and are identified as 1-saddles that are not part of free Li diffusion paths and are also connected to > 6 bond 2-saddles in the MS complex.

4.3 Model Validation: Feature Stability

The model presented in section 4.2 is missing some key information needed for practical feature extraction: the MS complex must first be simplified to a user-supplied threshold, a second threshold is needed to identify the 2-saddles corresponding to covalent bonds, and finally, a third threshold determines free Li ion diffusion paths. Ideally, one can arrive at stable thresholds for each of these without *a priori* knowledge, by simply examining the distribution and persistence of features of the computed MS complex. The goal is to be able to identify such features and then relate them back to what we know about the data, as a strong validation of the approach. In the following, we describe this methodology as it is applied to the 732-atom DFT data.

We begin with identification of Carbon atoms; in our model, they are maxima of the MS complex. Figure 4(a) shows the count of critical points of the MS complex as a function of persistence simplification threshold. A large stable threshold is identified in the range $[0.03 : 0.6]$. In our example data, we find exactly 732 high-valued maxima in this threshold range, corresponding exactly to the carbon atoms in the atomistic simulation. For simplification thresholds less than 0.03, low-valued maxima appear on some of the covalent bonds, and for thresholds greater than 0.6, high-valued maxima begin to merge together.

Next, we validate our model for high-valued 2-saddles representing covalent bonds by finding a stable threshold that distinguishes between 2-saddle-maximum arcs of the complex. In figure 4(b), we show each arc of the MS complex as a point, with coordinates given by the lower and upper critical points of the arc. A clear structure is apparent from this figure, notably the separation of high-valued 2-saddle-maximum arcs from the other arcs in the dataset. Selecting the 2-saddle-maximum arcs in the box 2-saddle$= [-0.05 : 2.0]$ provides a stable threshold, that corresponds to the covalent bonds in the data. This is furthermore shown in figure 4(c) to be stable for the same simplification threshold range as the carbon atoms, yielding exactly 1110 covalent bonds in these threshold ranges.

The rings that do not allow Li ions to diffuse, those with valence ≤ 6, can be identified by the count of 2-saddles corresponding to covalent bonds. Intuitively, these are given by 1-saddles at some simplification threshold. The count of barrier rings as a function of persistence is shown in figure 4(d). This plot indicates that a smaller stable threshold, as 1-saddles are lost after a persistence simplification threshold of 0.11.
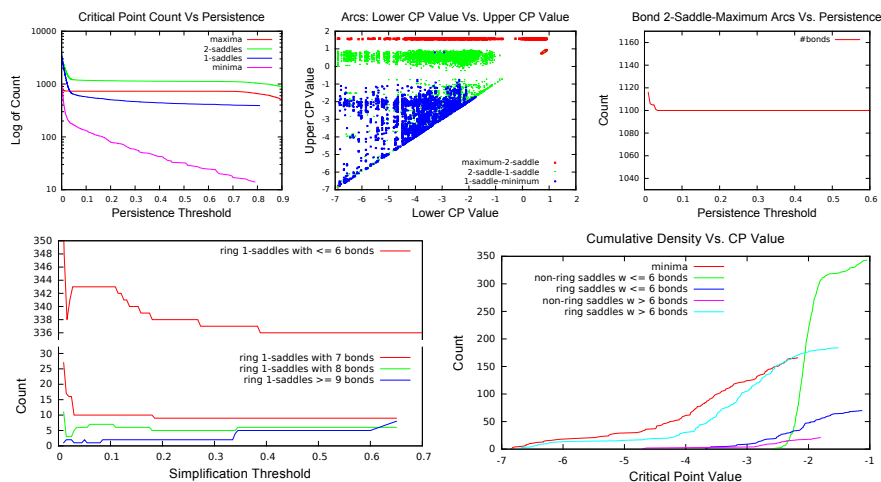
Fig. 4 Analysis of the feature set encoded by the MS complex leads to insights about the location and stability of features. Counting critical points as a function of persistence simplification threshold (a) shows that carbon atoms can be extracted in a stable manner. Plotting each arc of the complex with a point whose coordinates are the values of the critical points (b) reveals that the 2-saddle-maximum connections corresponding to covalent bonds are well separated from other features, and furthermore stable with respect to simplification (c). Counting the number of n-valence rings reveals that barrier 1-saddles are stable in a smaller range of simplification (d). Finally, we identify a conservative estimate for threshold free diffusion paths by identifying flatter regions in the cumulative density function of critical points (e) while avoiding any barrier rings.

So far, we have established that carbon atoms, covalent bonds, and barrier rings can be extracted in a very stable manner, with simplification threshold of 0.06 ± 0.03. This stability aligns with expectation, due to the regular structure imposed by the chemistry and physics of the annealed nanosphere. We make the logical step that any 2-saddles that are *not* covalent bonds must be part of an intra-layer region, where Li ions can diffuse relatively easily. Furthermore, since the ascending 2-manifolds of 1-saddles form topological disks that fill in the interior of covalently bonded carbon rings, any 1-saddle adjacent to a non-bond 2-saddle (and any minimum attached to that) must also be part of the free diffusion region. Finally, we consider the case where based on the electron density value alone we can determine if the Li ions may diffuse. In figure 4(e) we show cumulative density functions of minima and 1-saddles, categorized by whether they are ring, non-ring, defect, or non-defect points. To pick the threshold that guarantees that Li may diffuse freely, we must pick a threshold below the lowest ring saddle values with ≤ 6 bonds, which occurs at value -3.7. The most permissive, yet relatively stable threshold occurs around -4.

<div style="text-align:center;">(a)   (b)   (c)   (d)</div>

<div style="text-align:center;">(e)   (f)   (g)   (h)</div>
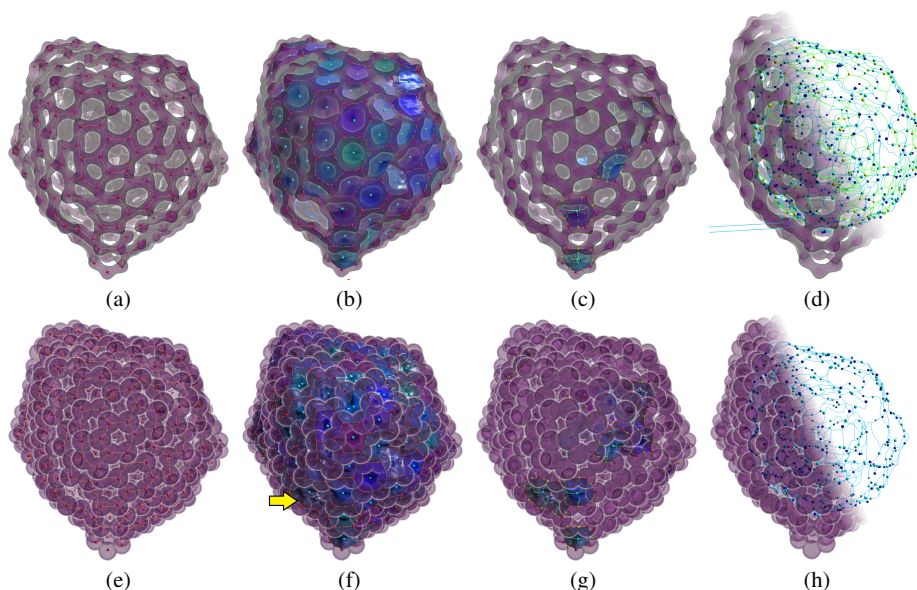
Fig. 5 A visual comparison of features of the DFT data (top row) with features of the distance field (bottom row). The exact same set of carbon atoms and bonds are found in each (a,e). While the majority of rings are the same, the DFT data classifies some defect rings as intra-layer free diffusion paths. For instance (f) displays a ring (yellow arrow) not present in (b). This difference can be accounted for with (c,d,g,h), where (g) shows a defect detected in the distance field , whereas (d) shows the same site as part of an intra-layer free path. Other defects sites match (c,g), and the free paths in both (d) and (g) completely connect every void in the interior.

5 Comparison with Distance Field

Production-scale nanosphere simulations are expected to produce results on the order of one million atoms, making DFT data impossibly expensive to compute. One alternative we explore in this section is to use the atom positions to create a distance field. In particular, we generate a signed distance field, that is the negative of the minimum distance from a grid vertex to a Carbon atom, preserving the convention of identifying carbon atoms as maxima. The same data-driven approach is applied from section 4.3 to obtain stable threshold of 0.1 for persistence simplification, -0.9 for identification of covalent bonds, and -5.0 to identify free paths. Figure 5 shows a direct comparison of features extracted from the DFT data and the distance field. There is a one-to-one correspondence between Carbon atoms and covalent bonds, between the two fields (a,e). The first differences appear in the identification of ring saddles (b,f), where certain 1-saddles appear as free paths in the DFT field, and instead appear as defect sites in the distance field. This difference can be noted in (c,g), where a defect present in the distance field (g) is instead identified as a free path connecting the exterior to the interior of the nanosphere (d). In both fields, ev-
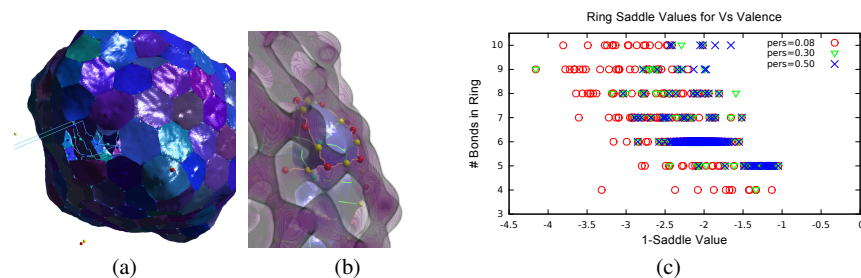
Fig. 6 In the DFT field, free paths were identified between all layers of the nanosphere (a). A corner case illustrates the need to compute the length of simple cycles when counting the valence of a ring (b). The distribution of values for rings of varying valence shows that there are no scalar thresholds that separate features (c).

ery void in the interior of the nanosphere is completely connected by free diffusion paths.

6 Discussions

We apply our data-driven approach to the 732-atom DFT field, and found a stable persistence simplification threshold of 0.06, a stable selection threshold for bonds of > -0.05, and a conservative estimate of free paths with value lower than -4. With these thresholds, we found that the entire interior of the 732-atom DFT field is accessible to Li ion diffusion. Figure 6(a) illustrates that according to the definitions we presented the free paths directly connect the exterior of the nanosphere with every layer. Figure 6(b) demonstrates a scenario where the ascending 2-manifold of a 1-saddle does not have a simple circle as its boundary. To avoid incorrectly classifying this as a fault, we perform an additional check on the ring, only counting the atoms that form a simple cycle. Such configurations arise when carbon atoms are bonded in the third dimension outside of the plane of the carbon ring. Finally, we justify using the ring valence as opposed to selecting a threshold for diffusion (as was done in [2]). Figure 6(c) shows the function values for various valence rings. Even for large simplification thresholds, rings identified as valence 5 or 6 may have *lower* function value than higher valence rings.

7 Conclusion/Future Work

We have presented a new technique for analyzing Li ion diffusion paths in carbon nanospheres that utilizes identifying motifs (rings) in the 1-skeleton of the MS complex. We have further shown that feature analysis is stable, and appropriate thresholds can be identified from the topology of the data itself without *a priori*

knowledge. Finally, we compared the analysis of the DFT data to a distance field, showing similarities, but also differences in the classification of free path vs. defect sites. Further study is needed to understand the implications on qualitative measurements of Li storage capacity of nanospheres. We plan to use this methodology to understand the energy storage properties of large-scale nanospheres.

References

1. R. F. Bader. *Atoms in molecules*. Wiley Online Library, 1990.
2. C. Bajaj, A. Gillette, and S. Goswami. Topology based selection and curation of level sets. In H.-C. Hege, K. Polthier, and G. Scheuermann, editors, *Topology-Based Methods in Visualization II, Mathematics and Visualization*, pages 45–58. Springer Berlin Heidelberg, 2009.
3. F. Cazals, F. Chazal, and T. Lewiner. Molecular shape analysis based upon the morse-smale complex and the connolly function. *Proceedings 19th Annual Symposium on Computational Geometry*, pages 351–360, 2003.
4. C. D. Correaa, P. Lindstrom, and P.-T. Bremer. Topological spines: A structure-preserving visual representation of scalar fields. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1842–1851, 2011.
5. J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, and J. Liang. CASTp: computed atlas of surface topography of proteins with structural and topographical map- ping of functionally annotated residues. *Nucleic acids research*, 34(2):W116–W118, 2006.
6. H. Edelsbrunner and P. Fu. Measuring space filling diagrams and voids. Technical Report UIUC-BI-MB-94-01, University of Illinois at Urbana-Champaign, 1994.
7. H. Edelsbrunner and P. Koehl. The geometry of biomolecular solvation. *Combinatorial & Computational Geometry*, 52:243–275, 2005.
8. H. Edelsbrunner and E. P. Mucke. Three-dimensional alpha shapes. *ACM Trans. Graphics*, 13(43-72), 1994.
9. R. Forman. A user's guide to discrete Morse theory. In *Séminare Lotharinen de Combinatore*, volume 48, 2002.
10. D. Günther, R. A. Boto, J. Contreras-Garcia, J.-P. Piquemal, and J. Tierny. Characterizing molecular interactions in chemical systems. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Scientific Visualization 2014)*, 20(12):2476–2485, December 2014.
11. A. Gyulassy, P. Bremer, and V. Pascucci. Computing Morse-Smale complexes with accurate geometry. *IEEE Transactions on Visualization and Computer Graphics*, 18:2014 – 2022, 12 2012.
12. A. Gyulassy, P.-T. Bremer, V. Pascucci, and B. Hamann. A practical approach to Morse-Smale complex computation: Scalability and generality. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1619–1626, 2008.
13. A. Gyulassy, N. Kotava, M. Kim, C. D. Hansen, H. Hagen, and V. Pascucci. Direct feature visualization using Morse-Smale complexes. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1549–1562, 2012.
14. A. Gyulassy, V. Natarajan, V. Pascucci, P.-T. Bremer, and B. Hamann. Topology-based simplification for feature extraction from 3D scalar fields. In *Proceedings of the IEEE Visualization 2005 (VIS'05)*, pages 535–542. IEEE Computer Society, 2005.

15. D.-S. Kim, Y. Cho, K. Sugihara, J. Ryu, and D. Kim. Three-dimensional beta-shapes and beta-complexes via quasi-triangulation. *Computer-Aided Design*, 42(10):911–929, 2010.

16. D.-S. Kim and K. Sugihara. Tunnels and voids in molecules via voronoi diagram. *Proc. Symp. Voronoi Diagrams in Science and Engineering (ISVD)*, pages 138–143, 2012.

17. P. Koehl, M. Levitt, and H. Edelsbrunner. Proshape: understanding the shape of protein structures. *Software at biogeometry*, 2004.

18. G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.

19. K. C. Lau, J. Lu, J. Low, D. Peng, H. Wu, H. M. Albishri, D. A. Al-Hady, L. A. Curtiss, and K. Amine. Investigation of the decomposition mechanism of lithium bis (oxalate) borate (libob) salt in the electrolyte of an aprotic li–o2 battery. *Energy Technology*, 2(4):348–354, 2014.

20. J. Liang, H. Edelsbrunner, P. Fu, S.-H. P., and S. Subramaniam. Analytical shape computation of macromolecules: Ii. inaccessible cavities in proteins. *Proteins Structure Function and Genetics*, 33(1):18–29, 1998.

21. J. Liang, H. Edelsbrunner, and C. WoodWard. Anatomy of protein pockets and cavities. *Protein Science*, 7(9):1884–1897, 1998.

22. N. Lindow, D. Baum, A. Bondar, and H. Hege. Dynamic channels in biomolecular systems: Path analysis and visualization. *Proc. IEEE Symposium on Biological Data Visualization (BioVis)*, pages 99–106, 2012.

23. N. Lindow, D. Baum, and H. Hege. Voronoi-based extraction and visualization of molecular paths. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2025–2034, 2011.

24. V. Meunier, J. Kephart, C. Roland, and J. Bernholc. Ab initio investigations of lithium diffusion in carbon nanotube systems. *Physical review letters*, 88(7):075506, 2002.

25. V. Natarajan, Y. Wang, P.-T. Bremer, V. Pascucci, and B. Hamann. Segmenting molecular surfaces. *Computer Aided Geometric Design*, 23(6):495–509, 2006.

26. J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

27. M. Petřek, P. Košinová, J. Koča, and M. Otyepka. MOLE: A voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure*, 15(11):1357–1363, 2007.

28. M. Petřek, M. Otyepka, P. Banáš, P. Košinová, J. Koča, and J. Damborský. CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics*, 7:316, 2006.

29. V. G. Pol, J. Wen, K. C. Lau, S. Callear, D. T. Bowron, C.-K. Lin, S. A. Deshmukh, S. Sankaranarayanan, L. A. Curtiss, W. I. David, et al. Probing the evolution and morphology of hard carbon spheres. *Carbon*, 68:104–111, 2014.

30. J. Song, B. Ouyang, and N. V. Medhekar. Energetics and kinetics of li intercalation in irradiated graphene scaffolds. *ACS applied materials & interfaces*, 5(24):12968–12974, 2013.

31. T. Sousbie. The persistent cosmic web and its filamentary structure–i. theory and implementation. *Monthly Notices of the Royal Astronomical Society*, 414(1):350–383, 2011.

32. R. Sridharamurthy, H. Doraiswamy, S. Patel, R. Varadarajan, and V. Natarajan. Extraction of robust voids and pockets in proteins. *Eurographics Conference on Visualization (short paper)*, 2013.

33. K. Weiss, F. Iuricich, R. Fellegara, and L. De Floriani. A primal/dual representation for discrete morse complexes on tetrahedral meshes. *Computer Graphics Forum*, 32(3pt3):361–370, 2013.

34. E. Yaffe, D. Fishelovitch, H. J. Wolfson, D. Halperin, and R. Nussinov. Molaxis: efficient and accurate identification of channels in macromolecules. *Proteins*, 73:72–86, 2008.

35. F. Yao, F. Gunes, H. Q. Ta, S. M. Lee, S. J. Chae, K. Y. Sheem, C. S. Cojocaru, S. S. Xie, and Y. H. Lee. Diffusion mechanism of lithium ion through basal plane of layered graphene. *Journal of the American Chemical Society*, 134(20):8646–8654, 2012.