JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS

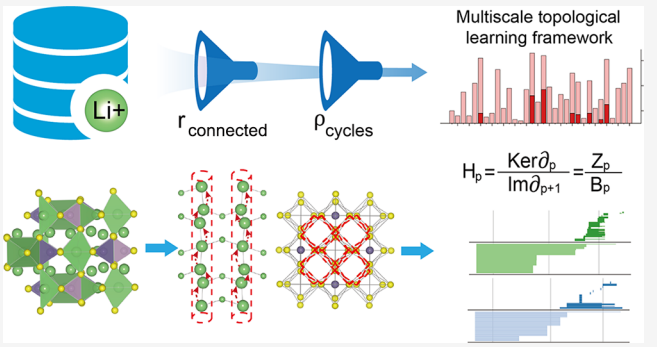
Article

Superionic Ionic Conductor Discovery via Multiscale Topological Learning

Dong Chen,[⊥] Bingxu Wang,[⊥] Shunning Li, Wentao Zhang, Kai Yang, Yongli Song, Guo-Wei Wei,* and Feng Pan*

Cite This: https://doi.org/10.1021/jacs.5c04828

ABSTRACT: Lithium superionic conductors (LSICs) are crucial for next-generation solid-state batteries, offering exceptional ionic conductivity and enhanced safety for renewable energy and electric vehicles. However, their discovery is extremely challenging due to the vast chemical space, limited labeled data, and understanding of complex structure−function relationships required for optimizing ion transport. This study introduces a multiscale topological learning (MTL) framework that integrates algebraic topology and unsupervised learning to efficiently tackle these challenges. By modeling lithium-only and lithium-free substructures, the framework extracts multiscale topological features and introduces two topological screening metrics, cycle density and minimum connectivity distance, to ensure structural connectivity and ion diffusion compatibility. Promising candidates are clustered via unsupervised algorithms to identify those that resemble known superionic conductors. For final refinement, candidates that pass chemical screening undergo ab initio molecular dynamics simulations for validation. This approach led to the discovery of 14 novel LSICs, four of which have been independently validated in recent experiments. This success accelerates the identification of LSICs and demonstrates broad adaptability, offering a scalable tool for addressing complex material discovery challenges.



1. INTRODUCTION

The discovery of superionic conductors—materials with exceptional ion transport properties—is crucial for advancing electrochemical energy storage and conversion technologies, including batteries,[1−3] fuel cells,[4] and ceramic membranes.[5,6] Among these, lithium superionic conductors (LSICs) are particularly promising alternatives to conventional organic liquid electrolytes due to their high ionic conductivity, broad electrochemical stability, and enhanced safety.[7] These attributes are vital for improving the performance, energy density, and lifespan of lithium-ion batteries. However, the discovery of LSICs remains a significant challenge. Only a limited number of lithium-based compounds, such as $Li_{10}GeP_2S_{12}$ (LGPS),[8] garnet $Li_7La_3Zr_2O_{12}$ (LLZO),[9,10] NASICON,[11] and $Li_{1.3}Al_{0.3}Ti_{1.7}(PO_4)_3$ (LATP),[12,13] exhibit room-temperature ionic conductivity comparable to liquid electrolytes. This limited number, coupled with insufficient ionic conductivity data, complicates the discovery of new LSICs. Furthermore, the experimental process to validate these materials is both expensive and time-consuming, and traditional computational methods, such as density functional theory (DFT) and ab initio molecular dynamics (AIMD) simulations, are extremely expensive and intractable for large-scale screening. Despite their potential, current LSICs do not meet the comprehensive requirements for widespread commercialization, underscoring the urgent need for new materials capable of overcoming these challenges.

Ion diffusion in solids, driven by lithium-ion migration through interconnected channels within the crystal structure, is central to the performance of LSICs. The framework of LSICs—comprising mobile lithium ions and immobile lithium-ion-free sublattices—determines the migration pathways and energy distributions.[14−16] While some LSICs, like LGPS and $Li_7P_3S_{11}$, feature bcc anionic sublattices that facilitate low-energy ion migration, others with nonbcc frameworks, such as garnet and NASICON, also demonstrate high conductivity.[17−19] These findings highlight the limitations of existing structural descriptors in capturing the diverse structural features that contribute to ion transport in LSICs. Prior efforts to identify LSICs have primarily relied on empirically defined or manually curated geometric descriptors, such as bond

Received: March 20, 2025
Revised: May 23, 2025
Accepted: May 23, 2025

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

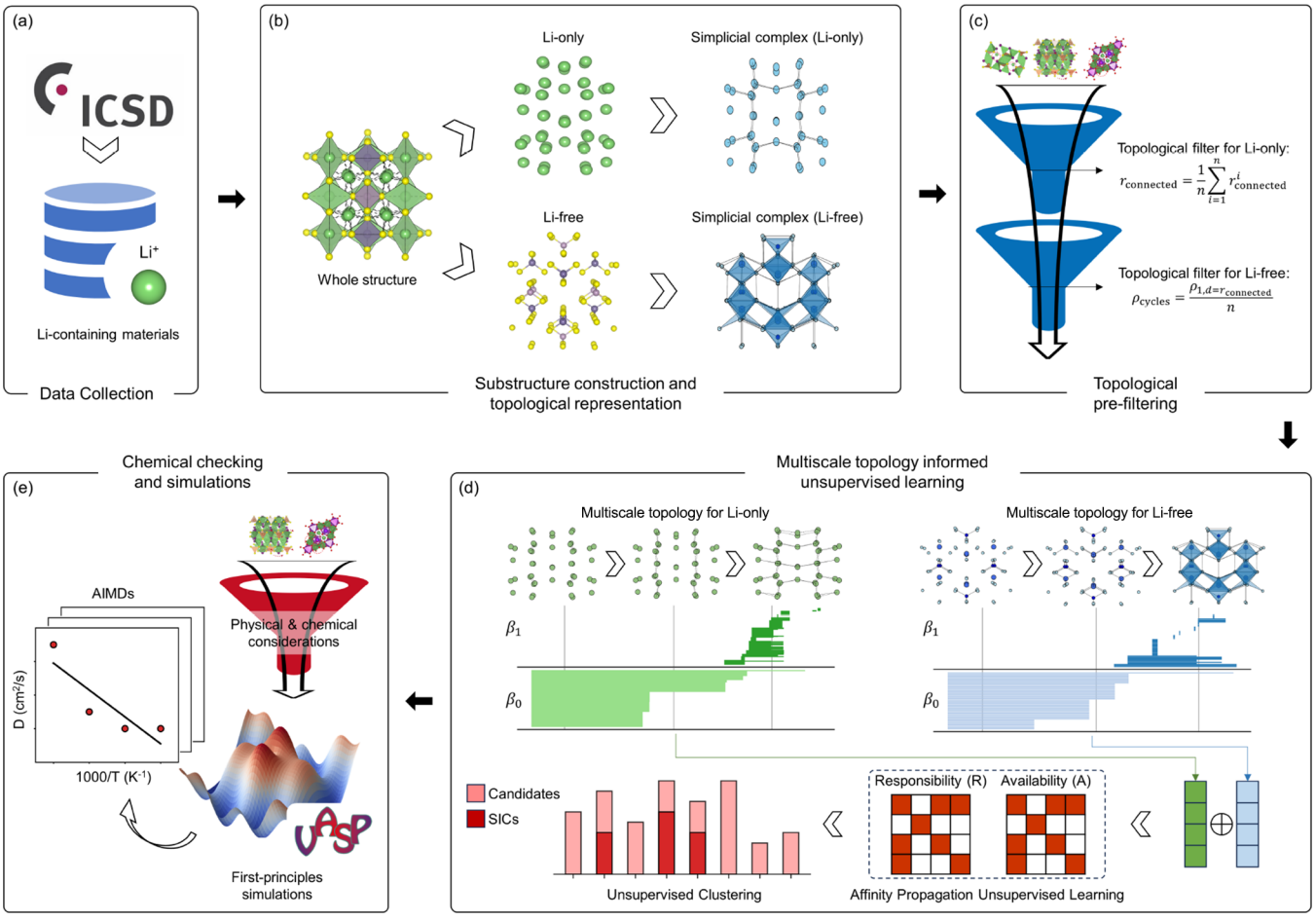


Figure 1. Workflow for a multiscale topological learning approach to discovering lithium superionic conductors. (a) Data collection phase filters materials containing lithium ions from the ICSD database to identify potential candidates. (b) Preliminary study of known LSIC structures, where lithium-only substructures (Li-only) and lithium-free frameworks (Li-free) are modeled as simplicial complexes and analyzed independently. (c) Topological representation of Li-only and Li-free substructures using simplicial complexes, capturing high-order interactions and deriving features like connectedness ($r_{\text{connected}}$) and cycle density (ρ_{cycles}) to narrow the search space. (d) Multiscale topological features derived via persistent homology and affinity propagation clustering, grouping materials based on topological similarity to highlight clusters with LSIC candidates. (e) Final physical and chemical validation, including first-principles analysis, was performed to identify the most promising LSIC candidates.

valence site energy, Li−Li coordination numbers, or statistical measures derived from Voronoi tessellations and radial distributions.[20,21] While such approaches have shown promise, they often require significant feature engineering and are typically limited to capturing local structural environments.

As such, there is a pressing need for more comprehensive and quantitative methods to understand the structure−function relationships in these materials. In addition to ab initio methods, several alternative approaches have been developed to investigate ionic transport in solids. Geometrical analysis combined with the bond valence method has proven effective in predicting ion migration pathways across diverse material systems.[22−25] The bond valence−Ewald method further refines this approach by incorporating long-range electrostatic interactions.[26,27] Additionally, theoretical frameworks such as effective medium theory (EMT) and the random resistance model (RRM), as well as data-driven machine learning techniques, offer complementary perspectives for evaluating ionic conductivity, particularly in high-throughput screening contexts.[28−32] While traditional techniques and computational approaches such as graph-based modeling and deep learning have provided valuable insights,[19,33] they often overlook the higher-order interactions and topological relationships crucial for accurately predicting ion transport.

Mathematically, topology encompasses the study of space, connectivity, dimensionality, and transformations. By providing a high level of abstraction, topology serves as an effective tool for analyzing structured data in the physical world, particularly in high-dimensional contexts. However, while topology offers valuable insights, it often oversimplifies geometric information, leading to a loss of structural detail during feature extraction. Persistent homology (PH),[34,35] a burgeoning branch of algebraic topology, presents a promising avenue for reconciling geometry and topology by facilitating a more nuanced understanding of spatial structures in a multiscale topological manner. PH has found applications in predicting the stability of carbon isomers through the combination of simple linear regression models.[36] Additionally, the introduction of element-specific persistent homology has enabled the preservation of crucial structural information during topological abstraction, particularly beneficial for handling multielement structures.[37] This approach has been successfully employed in predicting the affinity and solubility of molecular proteins in biomedicine.[38,39] Furthermore, by restricting its scope of action, persistent homology has been

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

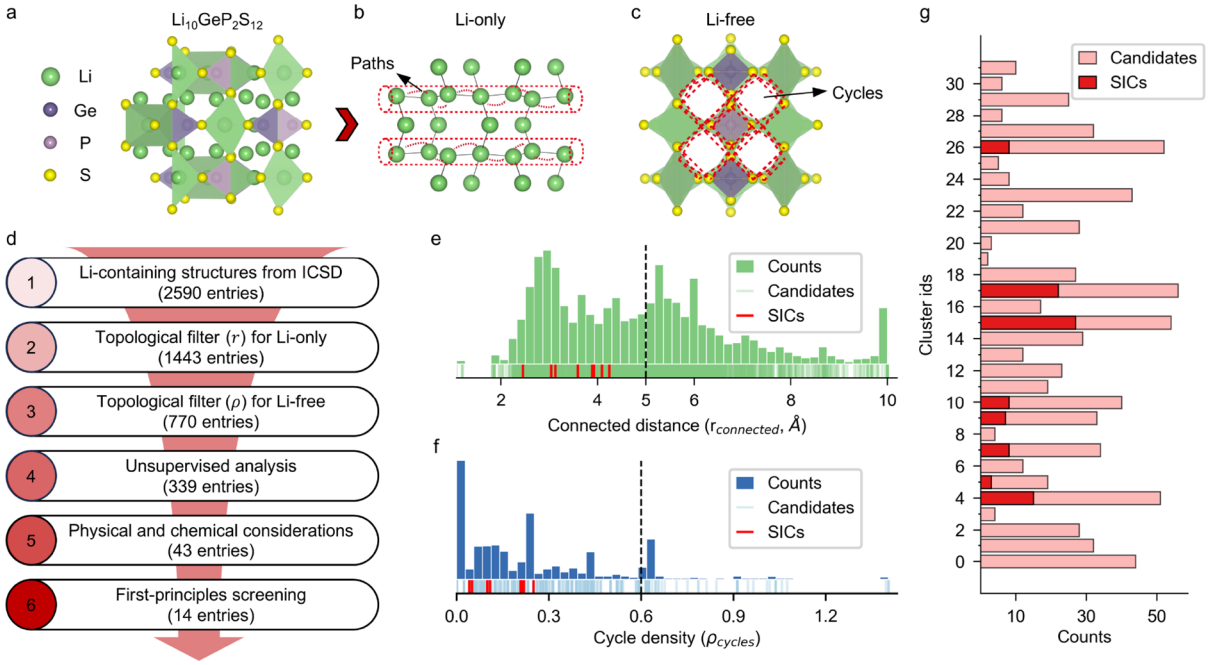


Figure 2. Results of the multiscale topology-driven workflow for LSIC discovery. (a) Crystal structure of the solid ionic conductor (LSIC) $Li_{10}GeP_2S_{12}$, used as an example. (b) Li-only substructure extracted from the LSIC. (c) Li-free substructure is derived from the same material. (d) Overview of the materials discovery workflow showing six stages with the corresponding number of materials filtered at each stage. (e) Distribution of the minimum connectivity distances ($r_{connected}$) for Li-only substructures. The red lines in the rug plot highlight the known LSICs, and the dashed line marks the threshold of $r_{connected}$ = 5 Å used in the filtering process. (f) Distribution of the pore occupancy index (ρ_{cycles}) for Li-free substructures. The red lines in the rug plot indicate known LSICs, and the dashed line denotes the filtering threshold ρ_{cycles} = 0.6. (g) Results from the affinity propagation algorithm. The lighter red bars represent all materials retained after topological prefiltering, while the darker red bars indicate the known LSICs. The horizontal axis corresponds to the number of structures in each cluster.

extended to the realm of inorganic crystalline materials exhibiting periodicity.[40−43] It has proven effective in predicting the formation energies of these materials, showcasing its versatility across different domains[44] and underscoring its versatility and potential in materials discovery. However, this approach has not been applied to the predictive discovery of new materials.

Building on these insights, this study introduces a multiscale topological learning (MTL) framework to accelerate the discovery of LSICs. Leveraging persistent homology, the framework extracts multiscale topological features from lithium-ion-only (Li-only) and lithium-ion-free (Li-free) substructures. These substructures are modeled as simplicial complexes to capture higher-order interactions, enabling a more nuanced representation of structural properties. This topological approach preserves critical structural information, offering valuable insights into the spatial organization and functional roles of these substructures in lithium-ion conduction. Next, the present framework introduces two key topological metrics: cycle density (ρ_{cycles}) and minimum connectivity distance ($r_{connected}$) for quantitative analysis. These metrics quantify the connectivity of Li-only substructures and assess the suitability of Li-free environments for ion diffusion, forming the basis for initial candidate filtering. The resulting materials are further scrutinized with an unsupervised machine learning model, which clusters materials based on similarities in terms of their multiscale topological features. The clustering results indicate that most known LSICs are concentrated within specific clusters, suggesting that other materials in these groups may also exhibit promising ionic conductivity. Finally, a chemical checking process filters out non-LSIC materials followed by AIMD simulations to validate the remaining candidates. While AIMD simulations are computationally intensive, they are applied exclusively to a small subset of candidates, thereby optimizing resource utilization. This integrated approach not only reduces both computational and experimental costs but also enhances the accuracy of the results, culminating in the identification of 14 novel LSICs and showcasing the efficacy of the proposed framework in accelerating material discovery.

2. RESULTS

2.1. Workflow and Conceptual Schematic. Figure 1 presents the workflow for a multiscale topology approach aimed at discovering lithium superionic conductors (LSICs). In the initial step (Figure 1a), the data collection phase filters materials containing lithium ions from the ICSD-2019 database, identifying promising candidates for analysis, which turns out to be 2590 unique materials. Preliminary mechanical filtering criteria applied during this stage are detailed in the Supporting Information. Figure 1b shows the second stage, where a preliminary study of well-known LSIC structures is conducted. Given that ionic conductivity is influenced by both the connectivity of lithium substructures and the stability of the surrounding framework, the Li-only and the Li-free are modeled as independent topological spaces using simplicial complexes and analyzed separately.

In the next stage, illustrated in Figure 1c, a topological approach is applied to each structure by representing Li-free and Li-only substructures with simplicial complexes. This topological representation captures high-order interactions within the material structure, with each *n*-simplex in the

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

complex representing different types of interactions: 0-simplices denote atoms, 1-simplices (edges) capture pairwise atomic interactions, and 2-simplices encode triplet interactions among three atoms. By capturing such high-order interactions, this topological approach provides a deeper structural characterization, essential for understanding ionic conductivity mechanisms in LSICs. Two key features—connectedness ($r_{connected}$) and cycle density (ρ_{cycles})—are derived through this analysis. These features serve as effective filters for narrowing the search space, with $r_{connected}$ encoding information about the Li-only substructure's conductivity and ρ_{cycles} reflecting the stability of the Li-free framework.

In the following stage (Figure 1d), multiscale topological features (persistent homology) are computed through both Li-only and Li-free frameworks, and affinity propagation clustering groups the remaining candidates based on topological feature similarity. This unsupervised clustering reveals internal structural patterns, placing similar materials in proximity within the topological space. Known LSICs tend to cluster within specific groups, highlighting clusters that are likely to contain additional LSIC candidates. Finally, as shown in Figure 1e, physical and chemical validation, including first-principles-based analysis, is applied to materials within promising clusters. This final evaluation identifies the most viable LSIC candidates, demonstrating the effectiveness of this multiscale topology-based unsupervised learning approach for LSIC discovery.

2.2. Topological Screening. Given the limited number of identified LSICs, understanding their internal structural characteristics is essential for advancing materials discovery in this field. In classical LSICs, lithium ions migrate in a cooperative manner characterized by co-diffusion rather than isolated jumping, which is typical of nonsuper lithium-ion conductors.[18,45] This cooperative migration, facilitated by lower energy barriers, indicates that both lithium−lithium interactions and the surrounding framework's structure strongly influence lithium-ion mobility. Additionally, Coulomb interactions among lithium ions affect the migration energy barrier.[33] When fractionally or integrally occupied lithium sites are close by (less than 2 Å apart), these interactions produce a continuous lithium-ion probability density within the structure. To fully capture these interactions and effectively identify potential LSICs, it is necessary to analyze both the Li-only and Li-free substructures. Figure 2a−c depicts the construction of the Li-only and Li-free substructures from the original material, exemplified by $Li_{10}GeP_2S_{12}$. This process establishes the foundation for identifying LSIC candidates. In the Li-only substructure (Figure 2b), the red channels represent the conductive paths of lithium ions. In the Li-free substructure (Figure 2c), the red cycles illustrate the structural environment surrounding the lithium paths. To streamline the search for suitable LSIC materials among Li-containing compounds, a preliminary filtering process was applied. This filtering process is based on two key topological features—$r_{connected}$ and ρ_{cycles}—that were derived, using a topology-informed approach, for the analysis of Li-only and Li-free frameworks.

To characterize the distinct structural roles in lithium-ion conduction, each material is decomposed into Li-only and Li-free substructures. The Li-only network captures the connectivity of mobile ions, quantified by the minimum connectivity radius ($r_{connected}$), which reflects percolation behavior relevant to concerted Li migration. The Li-free framework defines the geometry of diffusion channels, with its cycle density (ρ_{cycles}), derived from persistent homology, indicating the openness of the host lattice. Initially, each Li-only and Li-free substructure was represented as a simplicial complex, an advanced extension of graphs capable of encoding high-order interactions via n-simplices in a multidimensional topological space. This complex structure provides a means of analyzing high-order properties that extend beyond pairwise interactions, capturing more intricate atomic configurations through higher-dimensional simplices. By applying algebraic topology techniques, specifically homology and persistent homology, to these simplicial complexes, we extracted topological invariants, known as Betti numbers (β), to describe structural features across different dimensions. Here, β_0 denotes the count of independent components, while β_1 represents the number of independent cycles, both of which are essential for assessing material connectivity. Persistent homology was further employed to track changes in these topological invariants across a range of spatial scales. Through a distance-based filtration parameter, the evolution of topological invariants as a function of atomic connectivity was visualized with barcodes, producing unique, scale-dependent topological fingerprints for each structure. An example of topological invariants in the 0- and 1-dimensions is shown in Section 3. This approach enables the extraction of key topological and geometric characteristics for both Li-only and Li-free substructures, leading to the development of two essential metrics for filtering materials.

For the Li-only structure, the metric $r_{connected}$ was calculated as the minimum connectivity radius, signifying the critical distance at which all lithium ions in the structure become interconnected. A 10 Å truncation radius is chosen to calculate the connectivity radius ($r_{connected}$) for lithium-ion frameworks, as it effectively captures the local environment around lithium ions without including irrelevant long-range interactions, balancing computational efficiency and the key features of lithium-ion migration. This was determined by taking each lithium ion within the crystal cell as a center and calculating the connectivity within a spherical region of 10 Å. The atoms within this range are connected to form a simplicial complex. The connectivity of lithium ions reflects their transport efficiency. The connectivity radius for each lithium ion was averaged as follows:

$$r_{connected} = \frac{1}{n}\sum_{i=1}^{n} r_{connected}^{i} \quad (1)$$

The parameter n represents the number of lithium ions within this cutoff radius of 10 Å, reflecting the number of lithium ions interacting within this region. The 'connected' lithium ions are defined purely based on distance, with two lithium ions considered 'connected' if their distance is within the 10 Å cutoff radius, thereby capturing the direct interactions that influence ion migration. This value provides insight into the minimum connectivity distance required for ion mobility in the Li-only substructure.

Figure 2e illustrates the distribution of $r_{connected}$, a measure of lithium connectivity, calculated for the Li-only substructures of all Li-containing materials in the data set. The distribution is presented as a histogram, with a rug plot shown at the bottom of the figure. Green lines on the rug plot represent the distribution of all materials, while red lines indicate the $r_{connected}$ values for known superionic conductors (LSICs), including $Li_7P_3S_{11}$, NASICON, and LLZO. A detailed list of these LSICs

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

is provided in Table S2. The threshold at 5 Å represents the midpoint of the observed connected distance range (0−10 Å), chosen as a lenient cutoff to balance inclusivity and specificity. Interestingly, all known LSICs exhibit $r_{connected}$ values below 5 Å, signifying strong lithium connectivity. This observation highlights a critical characteristic of superionic conductors: the lithium ions are closely paired, ensuring good ionic conductivity. Consequently, a threshold of 5 Å was chosen to screen materials with poor lithium connectivity, effectively narrowing down the data set from 2,590 to 1,443 materials for further analysis.

In the Li-free framework, the topological feature ρ_{cycles} was derived from the value of β_1 in the topological fingerprint, representing the number of independent "holes" or cycles in the structure. These cycles, or voids, within the framework are essential for facilitating lithium-ion migration. To ensure ionic conductivity, an appropriate number of cycles is required; too many cycles could destabilize the framework, while too few could hinder lithium-ion movement. Here, ρ_{cycles} was calculated as

$$\rho_{cycles} = \frac{\beta_{1,d=r_{connected}}}{n} \quad (2)$$

where $\beta_{1,d=r_{connected}}$ is the value of β_1 at $d = r_{connected}$, and n denotes the number of lithium sites. This metric captures the balance of voids necessary for ion migration, providing a measure of the Li-free framework's suitability for LSIC functionality.

Figure 2f presents the distribution of ρ_{cycles}, a measure of cycle density, for the Li-free frameworks of the remaining structures after filtering based on $r_{connected}$. The heights of the histogram bars represent the counts of ρ_{cycles} values across all Li-free frameworks. At the bottom, a rug plot is shown, where the blue lines indicate the distribution of ρ_{cycles} for all materials, and the red lines mark the corresponding values for known LSICs. The analysis reveals that effective Li-free frameworks exhibit relatively low cycle density. This finding suggests that a balance is required: the framework must have a sufficient cycle ratio to stabilize the environment surrounding the Li pathways but should not possess excessively high cycle density, which could lead to structural instability or collapse. The threshold value of 0.6 corresponds to the midpoint of the full range of cyclic density values (0−1.2) across all screened structures. Based on this observation, a threshold of 0.6 was set for ρ_{cycles}, filtering out unconsolidated Li-free frameworks and refining the selection of candidate materials. The threshold values for both metrics were established based on known LSICs, enabling high-throughput screening of the material database to expedite the identification of potential LSIC candidates.

2.3. Multiscale Topological Clustering. In this study, we combined persistent homology, a promising algebraic topology tool, with an unsupervised learning approach to identify potential LSICs among lithium-based materials. Persistent homology offers a robust means of characterizing the structures of both Li-only and Li-free sublattices, providing a comprehensive, multiscale topological fingerprint for each material. The preliminary step used two key topological features derived from the barcodes. The full breadth of features, capturing a more complete spectrum of multiscale topological interactions, was subsequently applied to enhance the clustering process and identify LSIC candidates with greater accuracy.

To systematically compare materials, we construct fixed-length feature vectors from topology-derived barcodes. For the β_0 of Li-only, since the starting segments of β_0 barcodes are all 0, we extract seven statistical features from their terminating values: minimum, maximum, mean, sum, standard deviation, median, and $r_{connected}$. This choice reflects the fact that the Li-only substructure captures the connectivity of the lithium diffusion network; therefore, greater emphasis is placed on the zero-dimensional topological information (β_0) that encodes the multiscale connectivity of Li sites. For the β_1 of Li-free, we compute 15 statistics (five for each of the start, end, and persistence of one-dimensional barcodes): maximum, minimum, sum, mean, and standard deviation. This is motivated by the fact that the Li-free substructure imposes energetic and geometric constraints on Li-ion movement. The rigidity, symmetry, and spatial arrangement of the host framework define the size and alignment of bottlenecks and interstitial sites, making one-dimensional cycle features (β_1) particularly relevant. In total, 22 standardized topological features are generated and used as inputs for an unsupervised learning model to detect potential LSIC candidates. This approach, unlike supervised learning, is well-suited for LSIC discovery, where the scarcity of known LSICs makes supervised training impractical.

Affinity propagation clustering[46] was selected due to its ability to automatically determine the optimal number of clusters based on pairwise data similarity, without requiring prior specification or labeled training data. This property is particularly advantageous for unsupervised topological analysis in materials discovery, where the intrinsic structure of the feature space is unknown, and conventional methods such as k-means may impose rigid assumptions on cluster quantity and geometry. The adaptive clustering process enables AP to determine high-quality clusters based on the data's intrinsic structure, avoiding the need for predefined cluster numbers or centroids.

To enable systematic comparison across materials, we constructed 22 fixed-length topological descriptors from persistence barcodes, which served as inputs for unsupervised clustering. The clustering result is illustrated in Figure S1. As shown in Figure 2g, the known LSIC materials, represented in dark colors, are notably concentrated within a limited number of clusters (8 out of 32), while unclassified materials are shown in lighter shades. The presence of unknown materials within clusters containing known LSICs suggests that these unclassified materials may also exhibit superionic conductivity based on their topological similarity. This multiscale topology-informed unsupervised model enables efficient, label-free identification of LSIC candidates without reliance on predefined hyperparameters or conductivity labels. Ultimately, our approach identified 339 materials clustered alongside known LSICs, providing a refined pool of candidates for further investigation based on their similarity to established LSICs. Since unsupervised learning lacks labeled data, conventional accuracy metrics are inapplicable; instead, model effectiveness was indirectly validated via AIMD simulations of representative candidates.

2.4. Chemical Validation and First-Principles Verification. To further validate the LSIC candidates filtered through the unsupervised learning model, a rigorous chemical screening process was applied to ensure both structural and chemical suitability for practical applications. Several criteria were established for this stage of validation: (1) compounds

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

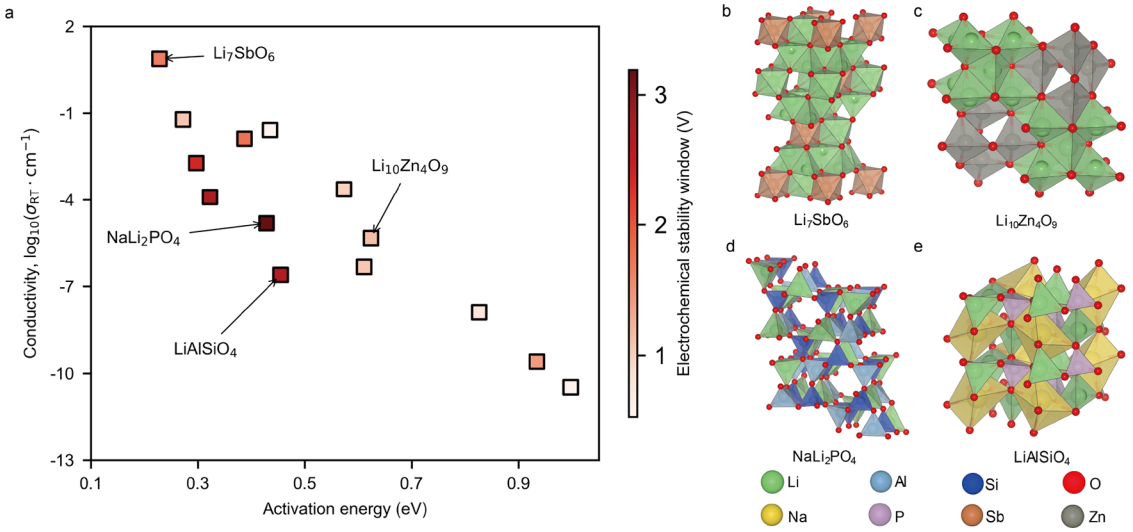


Figure 3. (a) Ionic conductivity as a function of the lithium-ion diffusion activation barrier for 14 potential LSICs was calculated by using AIMD at room temperature (300 K). The color bar indicates the electrochemical stability window of these materials. (b), (c), (e), and (d) demonstrate the structures of Li_7SbO_6, $Li_{10}Zn_4O_9$, $LiAlSiO_4$, and $NaLi_2PO_4$, respectively.

with two or fewer elements were excluded. Binary compounds were excluded because, despite their high conductivity and extensive study, they offer limited opportunities for uncovering novel structural motifs;[47] (2) compounds with radioactive elements or water molecules were eliminated; (3) alloys were excluded; (4) compounds with elements in abnormal oxidation states, which could compromise stability, were removed; (5) specific classes of compounds, such as all Li-X-O ternary systems where X is S, I, Si, C, P, Al, Ge, Se, B, or Cl, and Li−P−S systems, were excluded due to their extensive prior study and practical limitations such as poor chemical stability, allowing the focus to shift toward discovering novel and more stable candidates in less-explored chemical spaces;[12,48] (6) to ensure chemical and electrochemical stability, compounds with transition metals (e.g., Fe, Mn, Ni, Ti, Mo, V, and Co) were not considered, as their multiple oxidation states and partially filled d orbitals may induce undesirable redox activity and electronic conduction[48−51] and oxides containing elements such as N, Re, Ho, Hf, Ru, Eu, and Lu were also omitted; (7) materials containing more than 500 atoms were removed due to computational limitations and challenges in experimental validation; and (8) compounds in which lithium shared atomic sites with other elements were excluded to avoid hindrance of lithium-ion diffusion channels. These postlearning screening steps represent a combination of chemical and physical considerations: chemical criteria aim to ensure intrinsic stability, practical applicability, and novelty by excluding unstable, metallic, or well-studied compounds; physical criteria focus on structural feasibility and ion transport compatibility, such as limiting system size and avoiding site-sharing between lithium and other atoms. A total of 339 alternative materials were subjected to this screening process, as detailed in Table S3, ultimately narrowing the pool to 45 candidates (Table S4).

Following the chemical screening phase, AIMD simulations were employed to evaluate the ionic conductivity, lithium-ion diffusion activation barriers, and electrochemical stability of the 45 selected materials. AIMD simulations were conducted at elevated temperatures (800−1400 K) to ensure sufficient lithium-ion mobility within tractable simulation times, as direct simulations at room temperature are computationally prohibitive due to limited diffusion events. Ionic conductivity at ambient conditions was then estimated via Arrhenius extrapolation, a widely accepted approach in solid-state ionics.[18,19] These simulations were conducted at elevated temperatures (800, 1000, 1200, and 1400 K) to accurately capture lithium-ion diffusion behavior and calculate activation barriers, as detailed in Table S5. The last column of Table S5 presents SE_{E_a}, representing the standard error associated with activation energy estimation; higher values imply reduced reliability of the computed E_a, primarily due to artifacts stemming from AIMD simulations under complex and fluctuating diffusion environments. Such anomalies arise when cooperative or multibody ion migration mechanisms dominate, leading to deviations from simple Arrhenius behavior—an issue previously documented in molecular dynamics studies and reflective of methodological limitations rather than physical reality.[52,53] Unlike prior studies that rely on empirical descriptors to preselect AIMD-friendly structures with well-defined ion pathways, our multiscale topological learning framework identifies candidates based on high-dimensional topological similarity. This enables the discovery of materials with unconventional and complex diffusion networks, where multiple ion migration pathways may coexist and overlap. Such structural complexity presents challenges for conventional AIMD simulations and highlights the unique perspective offered by our topological approach in capturing diverse ion transport features beyond traditional methods. By integration of these results with electrochemical stability window (ESW) calculations, the analysis provided a comprehensive assessment of the structural and dynamic properties of the candidates.

To balance conductivity and stability, thresholds were established based on experimental and computational guidelines. Lithium-ion diffusion activation barriers were constrained between 0.1 and 1.0 eV, ensuring the exclusion of materials with impractically low barriers, which may indicate structural instability, while allowing for sufficient ionic mobility. Candidates who pass the threshold of activation barriers are shown in Table S6. Similarly, an ESW threshold of 0.5 V was applied to ensure chemical stability under slightly

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Table 1. Potential LSICs Filtered through AIMD Simulations, Including ICSD IDs and Corresponding Calculated Properties

ICSD-IDs	compositions	structure type	activation barrier (eV)	ionic conductivity (mS/cm)	electrochemical stability window (V)
9987	$Li_6Ga_2(BO_3)_4$	$Li_3AlB_2O_6$	0.826	1.291e-8	1.654
15631	Li_7SbO_6[54]		0.228	7.634	1.137
23634	$Li_{10}Zn_4O_9$[55]		0.624	4.694e-6	1.405
35250	$K_2Li_{14}Pb_3O_{14}$	$K_2Li_{14}Pb_3O_4$	0.998	3.291e-11	0.553
40245	Li_3BiO_3		0.573	2.32e-4	0.977
59640	$Li_4Zn(PO_4)_2$	$Li_4O_8P_2Zn$	0.387	1.32e-2	1.772
69967	$NaLi_2PO_4$	Li_3PO_4[57,58]	0.429	1.52e-5	3.19
71035	KLi_6BiO_6	KLi_6IrO_6	0.272	6.16e-2	1.064
72840	Li_6KBiO_6		0.611	4.792e-7	1.063
74864	$CsKNa_2Li_8(Li(SiO_4))_4$	$CsKNa_2Li_8(LiSiO_4)_4$	0.296	1.89e-3	2.382
78819	$Li_{10}N_3Br$		0.435	2.60e-2	0.530
92708	$LiAlSiO_4$[56]	$LiGaSiO_4$	0.455	2.504e-7	2.667
95972	Li_2MgSiO_4	Li_2ZnSiO_4	0.323	1.25e-4	2.739
262642	$In_2Li_2SiS_6$	Cd_4GeS_6	0.934	2.546e-10	0.755

reducing conditions, such as those encountered during cycling, as detailed in Table S6. These thresholds prioritize materials that achieve an optimal balance between high ionic conductivity, structural stability, and compatibility with lithium−metal anodes or other battery components.

From this comprehensive analysis, 14 materials were identified that satisfied the desired criteria. Figure 3a illustrates the ionic conductivity as a function of the diffusion activation barriers for these final candidates, many of which demonstrate excellent ionic conductivity in the order of 10^{-2} S/cm at room temperature (300 K). Detailed results for these candidates, including Arrhenius plots of lithium-ion diffusion coefficients, structural representations, and isosurfaces of lithium-ion probability densities obtained from AIMD simulations, are provided in Figures S2−S15 and Table 1.

This multistage validation process highlights the importance of integrating chemical screening with structural and dynamic assessments to identify high-potential LSICs. Notably, several materials identified in this study—such as Li_7SbO_6, $Li_{10}Zn_4O_9$, and $LiAlSiO_4$—have been independently validated experimentally or patented, further substantiating the approach's predictive power.

Li_7SbO_6 (Figure 3b), synthesized via solid-state reaction, demonstrated high cycling stability and rate performance as a lithium-ion battery anode. It retained >50% of its capacity at 5 and 10 mA cm^{-2} with nearly 100% Coulombic efficiency.[54] $Li_{10}Zn_4O_9$ (Figure 3c) was observed as a nanoscale phase that significantly enhances ionic conductivity in complex oxide matrices,[55] consistent with our predicted low diffusion barrier (0.2 eV). $LiAlSiO_4$ (Figure 3e), prepared as a sputtered thin film, exhibited a room-temperature ionic conductivity of 2.7×10^{-5} S/cm and strong stability in all-solid-state electrochromic devices.[56] Finally, $NaLi_2PO_4$ (Figure 3d), due to its structural similarity to Li_3PO_4, a well-known solid-state electrolyte with ionic conductivities in the 10^{-7}−10^{-6} S/cm range,[57,58] is predicted to possess similar potential for Li-ion conduction.

These findings validate the proposed model's efficacy in identifying promising LSICs and emphasize its potential to accelerate the discovery of advanced materials for next-generation lithium-ion batteries. Moreover, the identified candidates that have yet to be experimentally tested present exciting opportunities for future research, demonstrating the robustness and scalability of the methodology.

2.5. Discussions. This study highlights the effectiveness of a multiscale topology analysis approach integrated with unsupervised learning for quantitatively characterizing lithium-ion diffusion channels and their surrounding frameworks within crystal structures. Conventional persistent homology methods often treat materials as undifferentiated point clouds, lacking chemical context. The proposed multiscale topological framework introduces chemical specificity by decomposing structures into Li-only and Li-free sublattices, capturing features relevant to lithium-ion conduction across multiple spatial scales and yielding interpretable, physically grounded descriptors. The workflow and filtered structures at each stage are summarized in Figure 2d. This approach allows us to capture both local and global structural characteristics of materials in a physically meaningful and data-driven manner. A high-throughput topological analysis of lithium-containing compounds provided quantitative insights into their crystal structures and significantly narrowed the pool of potential LSIC candidates.

In materials discovery via machine learning, data quality critically influences model performance, especially given limited data sets and high-dimensional feature spaces. As noted by Liu et al.,[59] maintaining an appropriate balance among sample size, feature dimensionality, and model complexity is essential to mitigate the effects of the "curse of dimensionality." This study integrates domain knowledge to guide data curation that robustly characterizes potential lithium-ion conduction pathways. Through domain-informed data governance, the resulting models utilize physically meaningful descriptors, thereby improving predictive accuracy and expediting the identification of lithium superionic conductors. The initial phase of the strategy reduces the search space by analyzing two critical factors: cycle density (ρ_{cycles}) for lithium-free substructures (Li-free) and minimum connectivity distance ($r_{connected}$) for lithium-only substructures (Li-only). This dual-filtering approach ensures the retention of structures meeting the essential criteria for lithium-ion diffusion and stable frameworks. Specifically, all identified LSICs exhibit Li-free sublattices with ρ_{cycles} below 0.6, ensuring a balanced cycle density conducive to stability, and Li-only sublattices with $r_{connected}$ below 5 Å, enabling efficient ionic conduction.

Although connectivity radius and cycle density provide an intuitive basis, the multiscale topological framework integrates diverse persistent homology features across dimensions and scales. Combined with affinity propagation clustering, this enables a comprehensive and interpretable identification of

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

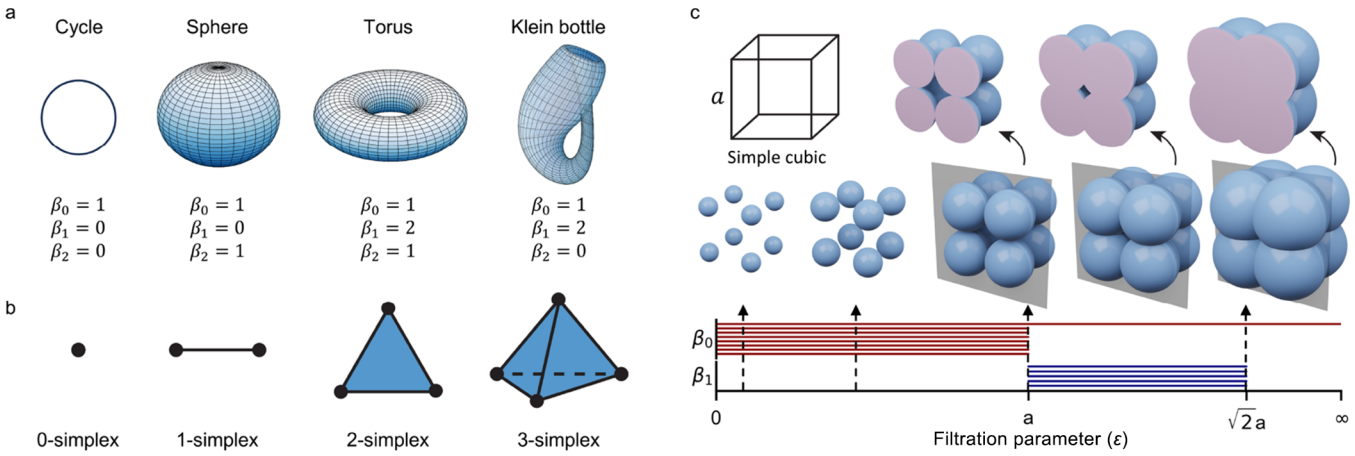


Figure 4. (a) Examples of topological spaces and their Betti numbers. A cycle has $\beta_0 = 1$, $\beta_1 = 1$, $\beta_2 = 0$; a sphere has $\beta_0 = 1$, $\beta_1 = 0$, $\beta_2 = 1$; a torus has $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1$; and a Klein bottle exhibits nontrivial Betti numbers with $\beta_0 = 1$, $\beta_1 = 1$, and $\beta_2 = 0$. (b) Building blocks of simplicial complexes, represented by simplices of increasing dimensions: vertices (0-simplices), edges (1-simplices), triangles (2-simplices), and tetrahedra (3-simplices). (c) Workflow of persistent homology is illustrated using a Vietoris–Rips complex. A simple cubic structure is analyzed by progressively increasing the filtration parameter d, which expands balls around each vertex. As d grows, topological features such as connected components (β_0) and loops (β_1) emerge and persist. The persistence of cycles in each phase of the cubic structure is visualized through barcodes corresponding to β_1.

LSIC candidates. These features capture both the global structural characteristics, using a multiscale filtration process, and the inherent properties of the structures, encompassing ionic transition pathways and their environmental frameworks. By comparison of these refined candidates with known LSIC structures through affinity propagation clustering, the method effectively identifies potential LSICs. This unsupervised learning step highlights materials structurally similar to known LSICs while uncovering novel, previously unstudied materials. This approach successfully identified all known LSIC structures and revealed 45 additional potential LSIC candidates.

The proposed strategy demonstrates a highly efficient method for LSIC discovery by integrating advanced mathematical frameworks and machine learning techniques. The initial focus on two key topological features, combined with a comprehensive multiscale topological analysis, efficiently narrows a vast data set while maintaining high predictive accuracy. Moreover, the approach's generalizable framework can be extended to other materials of interest, offering a scalable and innovative pathway for materials discovery.

To validate these refined candidates, more precise AIMD simulations were conducted to assess their ionic conductivity, lithium diffusion activation barriers, and electrochemical stability. Among the candidates, 14 materials met stringent criteria, including a lithium-ion diffusion activation barrier below 1.0 eV and an electrochemical stability window greater than 0.5 V. Several of these materials have been experimentally validated as excellent LSICs, further confirming the model's predictive capability. The remaining candidates offer promising avenues for future experimental evaluation. Overall, this robust and efficient workflow ensures the discovery of materials with the desired properties, even when only limited verified knowledge is available.

This work demonstrates the potential of combining advanced topological methods with unsupervised learning for efficient material discovery. The proposed methodology is not limited to LSICs. It can be adapted to discover other materials with the desired properties, providing a versatile and generalizable strategy for addressing complex challenges in materials science.

Some candidate materials, such as $K_2Li_{14}Pb_3O_{14}$ in our data set, contain multiple alkali metal cations (e.g., both Li^+ and K^+), which raises important questions about their combined influence on ionic transport. In such systems, the well-known Mixed Alkali Effect (MAE)[60] can emerge, where the presence of multiple mobile ion species leads to nonlinear suppression of each ion's mobility. In this study, we restricted our analysis to lithium-ion conduction pathways and did not explicitly address the effects of mixed cationic species. Nonetheless, we recognize the importance of this topic and suggest that future work could extend our methodology to include mixed alkali systems, thereby enabling a more comprehensive understanding of fast-ion conductors involving Na^+, K^+, and other cations.

3. METHODS

3.1. Multiscale Topology Data Analysis. *3.1.1. Simplicial Complex Representation.* In this work, both Li-free and Li-only structures are modeled by using simplicial complexes, which extend graphs to higher dimensions, providing richer structural and topological insights. A simplex, the building block of a simplicial complex, generalizes geometric shapes like points (0-simplices), line segments (1-simplices), triangles (2-simplices), and tetrahedra (3-simplices) to arbitrary dimensions, as shown in Figure 4b. For material representation, atoms are treated as 0-simplices (vertices), and atomic interactions are captured by higher-dimensional simplices, reflecting structural hierarchy and connectivity. A k-simplex, defined as $\sigma^k = \{v | v = \sum_{i=0}^{k} \lambda_i v_i, \sum_{i=0}^{k} \lambda_i = 1, 0 \le \lambda_i \le 1\}$, is the convex hull of $k + 1$ affinely independent points. A simplicial complex K is a collection of simplices satisfying: (1) every face of a simplex in K is also in K; (2) the intersection of any two simplices is either empty or a common face.

3.1.2. Homology and Persistent Homology. Homology provides an algebraic framework to analyze simplicial complexes, revealing topological features, such as connectedness, holes, and voids across dimensions. Central to this framework are chains, chain groups, chain complexes, and boundary operators. A k-chain is a formal sum of k-simplices with coefficients in a chosen field (e.g., $\mathbb{Z}_2$), and the

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

collection of all k-chains forms the k-chain group C_k. The boundary operator ∂_k maps k-chains to $(k-1)$-chains:

$$\partial_k \sigma^k = \sum_{i=0}^{k} (-1)^i [v_0, \ldots, \hat{v_i}, \ldots, v_k] \tag{3}$$

where $\hat{v}_i$ omits the i-th vertex. This operator defines cycles ($\mathrm{Ker}(\partial_k)$: chains with no boundary) and boundaries ($\mathrm{Im}(\partial_{k+1})$: chains that are boundaries of higher-dimensional simplices). These relationships form a chain complex:

$$\cdots \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \xrightarrow{\partial_{k-1}} \cdots \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0 \tag{4}$$

where $\partial_{k-1} \bigcirc \partial_k = 0$. The k-th homology group H_k is defined as $H_k = \mathrm{Ker}(\partial_k)/\mathrm{Im}(\partial_{k+1})$, and measures k-dimensional holes in the simplicial complex. The Betti numbers $\beta_k = \mathrm{rank}(H_k)$ quantify the number of independent k-dimensional features, such as connected components (β_0), tunnels (β_1), and cavities (β_2). Figure 4a shows examples of topological spaces and their Betti numbers, a cycle has $\beta_0 = 1$, $\beta_1 = 0$, and $\beta_2 = 0$, while more complex shapes such as the torus and Klein bottle have nontrivial higher-dimensional Betti numbers.

Persistent homology extends homology to multiscale analysis, capturing the persistence of topological features as a parameter (e.g., a scale parameter ϵ) varies.[34,35] This is achieved through filtration, a sequence of nested simplicial complexes $\{K_i\}$ such that $K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n$. This work uses the Vietoris–Rips filtration, where simplices are added based on a distance threshold ϵ. Persistent homology tracks the evolution of homological features through filtration steps:

$$\emptyset = H(K_0) \to H(K_1) \to \cdots \to H(K_n) = H(K) \tag{5}$$

The p-persistent k-th homology group describes features persisting across filtration steps i to $i+p$: $H_k^{i,p} = Z_k^i/(B_k^{i+p} \cap Z_k^i)$, where Z_k^i and B_k^{i+p} are the cycles and boundaries at steps i and $i+p$, respectively. Persistent homology is often visualized using barcodes, where each bar represents a topological feature's birth and death as ϵ increases. Figure 4c illustrates a simple cubic at varying thresholds ϵ and their corresponding persistent patterns.

3.2. Clustering. The Affinity Propagation (AP) algorithm is a clustering technique designed to identify a set of exemplars among data points and assign each point to its nearest exemplar, forming distinct clusters.[46] Unlike traditional clustering methods like K-Means, AP does not require prespecifying the number of clusters. Instead, it dynamically determines the clusters based on the similarities among data points.

The algorithm begins by calculating the pairwise similarity between the data points. For data points $\mathbf{x}_i$ and $\mathbf{x}_k$, the similarity is defined as $s(i,k) = -\|\mathbf{x}_i - \mathbf{x}_k\|^2$, which measures how well $\mathbf{x}_k$ can serve as the exemplar for $\mathbf{x}_i$. Two key matrices, the responsibility matrix ($\mathbf{R}$) and the availability matrix ($\mathbf{A}$), are then iteratively updated to identify exemplars. These updates continue until the algorithm converges, producing exemplars that maximize cluster similarity. Each data point is assigned to the cluster corresponding to its most suitable exemplar, defined by the combination of responsibility and availability scores. This iterative process ensures robust cluster formation without requiring predefined parameters such as the number of clusters.

For this study, the implementation of AP from the scikit-learn library was employed.[61] This method's ability to dynamically identify cluster centers makes it particularly suitable for analyzing the complex, high-dimensional feature space generated by the multiscale topological method. It facilitated the identification of clusters representing structurally and chemically similar materials, enabling effective material categorization and candidate screening.

3.3. First-Principles Simulation. In this work, all density functional theory (DFT) calculations were performed using the Vienna Ab Initio Simulation Package (VASP), utilizing the Projector Augmented Wave (PAW) method in conjunction with the Perdew–Burke–Ernzerhof (PBE) exchange-correlation functional.[62,63] The plane wave basis set employed a cutoff energy of 520 eV to ensure computational accuracy and efficiency. For structural optimization, k-point meshes centered on the Γ-point were generated with a minimum spacing of 0.4 Å between k-points. A finer k-point spacing of 0.25 Å was used for accurate energy calculations.

Ab initio molecular dynamics (AIMD) simulations were conducted to assess lithium-ion diffusion. The systems were first relaxed and then heated to 1200 K over 10 ps, followed by equilibration at 800, 1000, 1200, and 1400 K for 20 ps, excluding the initial 2 ps of each trajectory. A time step of 2 fs was used for the AIMD simulations with Γ-point and k-point sampling.

Ionic diffusivity (D) was calculated using the mean square displacement (MSD) formula:

$$D = \frac{1}{2dN\Delta t} \sum_{i=1}^{N} \langle [\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)]^2 \rangle_t \tag{6}$$

where d is the dimensionality of diffusion, N is the number of ions, and $\mathbf{r}_i(t)$ is the displacement of the i-th ion.

The ionic conductivity (σ) was then derived using the Nernst–Einstein relation:

$$\sigma = \frac{nq^2}{k_B T} D \tag{7}$$

where n is the ion density, q is the ion charge, k_B is the Boltzmann constant, and T is the temperature. These calculations provided key insights into the ionic transport properties of the materials, including diffusion coefficients, lithium-ion diffusion activation barriers, and electrochemical stability windows.

■ ASSOCIATED CONTENT

Data Availability Statement

The data set used in this study is from the ICSD database, and all data can be downloaded from the official ICSD Web site. Additionally, we have provided a list of ICSD numbers for the data at each filtering step on https://github.com/PKUsam2023/MTUL-LSIC/tree/main/filter_data. The related codes have been released as an open resource in the Github repository: https://github.com/PKUsam2023/MTUL-LSIC/tree/main.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.5c04828.

Data filtering workflow, compound tables, and topological and AIMD analyses (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Guo-Wei Wei − *Department of Mathematics, Department of Electrical and Computer Engineering, and Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States;* orcid.org/0000-0002-5781-2937; Email: weig@msu.edu

Feng Pan − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0002-8216-1339; Email: panfeng@pkusz.edu.cn

Authors

Dong Chen − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China; Department of Mathematics, Michigan State University, East Lansing, Michigan 48824, United States*

Bingxu Wang − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China*

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Shunning Li − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0002-5381-6025

Wentao Zhang − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China*

Kai Yang − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0001-7391-1495

Yongli Song − *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, China;* orcid.org/0000-0003-2231-2438

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.5c04828

Author Contributions

[⊥]D.C. and B.W. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is financially supported by the Shenzhen Science and Technology Research Grant (No. ZDSYS201707281026184), the Soft Science Research Project of Guangdong Province (No. 2017B030301013), and the Major Science and Technology Infrastructure Project of Material Genome Big-Science Facilities Platform supported by the Municipal Development and Reform Commission of Shenzhen. The work of Chen and Wei was supported in part by NIH grants R01GM126189, R01AI164266, and R35GM148196, NSF grants DMS-2052983, MSU Research Foundation, and Bristol-Myers Squibb 65109.

■ REFERENCES

(1) Dunn, B.; Kamath, H.; Tarascon, J.-M. Electrical energy storage for the grid: a battery of choices. *Science* **2011**, *334* (6058), 928−935.

(2) Seino, Y.; Ota, T.; Takada, K.; Hayashi, A.; Tatsumisago, M. A sulphide lithium super ion conductor is superior to liquid ion conductors for use in rechargeable batteries. *Energy Environ. Sci.* **2014**, *7* (2), 627−631.

(3) Thangadurai, V.; Narayanan, S.; Pinzaru, D. Garnet-type solid-state fast li ion conductors for li batteries: critical review. *Chem. Soc. Rev.* **2014**, *43* (13), 4714−4727.

(4) Wachsman, E. D.; Lee, K. T. Lowering the temperature of solid oxide fuel cells. *Science* **2011**, *334* (6058), 935−939.

(5) Tongwen, X. Ion exchange membranes: State of their development and perspective. *J. Membr. Sci.* **2005**, *263* (1−2), 1−29.

(6) Randau, S.; Weber, D. A.; Kötz, O.; Koerver, R.; Braun, P.; Weber, A.; Ivers-Tiffée, E.; Adermann, T.; Kulisch, J.; Zeier, W. G.; et al. Benchmarking the performance of all-solid-state lithium batteries. *Nature Energy* **2020**, *5* (3), 259−270.

(7) Janek, J.; Zeier, W. G. A solid future for battery development. *Nat. Energy* **2016**, *1* (9), 16141.

(8) Zhao, Y.; Wu, C.; Peng, G.; Chen, X.; Yao, X.; Bai, Y.; Wu, F.; Chen, S.; Xu, X. A new solid polymer electrolyte incorporating $Li_{10}GeP_2S_{12}$ into a polyethylene oxide matrix for all-solid-state lithium batteries. *J. Power Sources* **2016**, *301*, 47−53.

(9) Thangadurai, V.; Kaack, H.; Weppner, W. J. F. Novel fast lithium ion conduction in garnet-type $Li_5La_3M_2O_{12}$ (m= nb, ta). *J. Am. Ceram. Soc.* **2003**, *86* (3), 437−440.

(10) Murugan, R.; Thangadurai, V.; Weppner, W.; et al. Fast lithium ion conduction in garnet-type $Li_7La_3Zr_2O_{12}$. *Angew. Chem., Int. Ed.* **2007**, *46* (41), 7778.

(11) Arbi, K.; Rojo, J. M.; Sanz, J. Lithium mobility in titanium based nasicon $Li_1 + xTi_2{-}xAlx(PO_4)_3$ and $LiTi_2{-}xZrx(PO_4)_3$ materials followed by nmr and impedance spectroscopy. *Journal of the European Ceramic Society* **2007**, *27* (13−15), 4215−4218.

(12) Kamaya, N.; Homma, K.; Yamakawa, Y.; Hirayama, M.; Kanno, R.; Yonemura, M.; Kamiyama, T.; Kato, Y.; Hama, S.; Kawamoto, K.; et al. A lithium superionic conductor. *Nature materials* **2011**, *10* (9), 682−686.

(13) Aono, H.; Sugimoto, E.; Sadaoka, Y.; Imanaka, N.; Adachi, G.-y. Ionic conductivity of solid electrolytes based on lithium titanium phosphate. *J. Electrochem. Soc.* **1990**, *137* (4), 1023.

(14) Bachman, J. C.; Muy, S.; Grimaud, A.; Chang, H.-H.; Pour, N.; Lux, S. F.; Paschos, O.; Maglia, F.; Lupart, S.; Lamp, P.; et al. Inorganic solid-state electrolytes for lithium batteries: mechanisms and properties governing ion conduction. *Chem. Rev.* **2016**, *116* (1), 140−162.

(15) Wang, Y.; Richards, W. D.; Ong, S. P.; Miara, L. J.; Kim, J. C.; Mo, Y.; Ceder, G. Design principles for solid-state lithium superionic conductors. *Nature materials* **2015**, *14* (10), 1026−1031.

(16) Hull, S. Superionics: crystal structures and conduction processes. *Rep. Prog. Phys.* **2004**, *67* (7), 1233.

(17) Jun, K. J.; Sun, Y.; Xiao, Y.; Zeng, Y.; Kim, R.; Kim, H.; Miara, L. J.; Im, D.; Wang, Y.; Ceder, G. Lithium superionic conductors with corner-sharing frameworks. *Nature materials* **2022**, *21* (8), 924−931.

(18) He, X.; Zhu, Y.; Mo, Y. Origin of fast ion diffusion in super-ionic conductors. *Nat. Commun.* **2017**, *8* (1), 15893.

(19) Zhang, Y.; He, X.; Chen, Z.; Bai, Q.; Nolan, A. M.; Roberts, C. A.; Banerjee, D.; Matsunaga, T.; Mo, Y.; Ling, C. Unsupervised discovery of solid-state lithium ion conductors. *Nat. Commun.* **2019**, *10* (1), 5260.

(20) Sendek, A. D.; Cubuk, E. D.; Antoniuk, E. R.; Cheon, G.; Cui, Y.; Reed, E. J. Machine learning-assisted discovery of solid li-ion conducting materials. *Chem. Mater.* **2019**, *31* (2), 342−352.

(21) Laskowski, F. A. L.; McHaffie, D. B.; See, K. A. Identification of potential solid-state li-ion conductors with semi-supervised learning. *Energy Environ. Sci.* **2023**, *16* (3), 1264−1276.

(22) Pan, L.; Zhang, L.; Ye, A.; Chi, S.; Zou, Z.; He, B.; Chen, L.; Zhao, Q.; Wang, D.; Shi, S. Revisiting the ionic diffusion mechanism in Li_3PS_4 via the joint usage of geometrical analysis and bond valence method. *Journal of Materiomics* **2019**, *5* (4), 688−695.

(23) He, B.; Mi, P.; Ye, A.; Chi, S.; Jiao, Y.; Zhang, L.; Pu, B.; Zou, Z.; Zhang, W.; Avdeev, M.; et al. A highly efficient and informative method to identify ion transport networks in fast ion conductors. *Acta Mater.* **2021**, *203*, No. 116490.

(24) He, B.; Chi, S.; Ye, A.; Mi, P.; Zhang, L.; Pu, B.; Zou, Z.; Ran, Y.; Zhao, Q.; Wang, D.; et al. High-throughput screening platform for solid electrolytes combining hierarchical ion-transport prediction algorithms. *Sci. Data* **2020**, *7* (1), 151.

(25) He, B.; Ye, A.; Chi, S.; Mi, P.; Ran, Y.; Zhang, L.; Zou, X.; Pu, B.; Zhao, Q.; Zou, Z.; et al. Cavd, towards better characterization of void space for ionic transport analysis. *Sci. Data* **2020**, *7* (1), 153.

(26) Shi, W.; He, B.; Pu, B.; Ren, Y.; Avdeev, M.; Shi, S. Software for evaluating long-range electrostatic interactions based on the ewald summation and its application to electrochemical energy storage materials. *J. Phys. Chem. A* **2022**, *126* (31), 5222−5230.

(27) Ren, Y.; Liu, B.; He, B.; Lin, S.; Shi, W.; Luo, Y.; Wang, D.; Shi, S. Portraying the ionic transport and stability window of solid electrolytes by incorporating bond valence-ewald with dynamically determined decomposition methods. *Appl. Phys. Lett.* **2022**, *121* (17), 173904.

(28) Li, Y.; Zhao, Y.; Cui, Y.; Zou, Z.; Wang, D.; Shi, S. Screening polyethylene oxide-based composite polymer electrolytes via combining effective medium theory and halpin-tsai model. *Comput. Mater. Sci.* **2018**, *144*, 338−344.

(29) Zou, Z.; Li, Y.; Lu, Z.; Wang, D.; Cui, Y.; Guo, B.; Li, Y.; Liang, X.; Feng, J.; Li, H.; et al. Mobile ions in composite solids. *Chem. Rev.* **2020**, *120* (9), 4169−4221.

(30) Ding, Y.; He, B.; Wang, D.; Avdeev, M.; Li, Y.; Shi, S. Software for evaluating ionic conductivity of inorganic−polymer composite solid electrolytes. *Energy Mater. Adv.* **2023**, *4*, No. 0041.

(31) Zhao, Q.; Avdeev, M.; Chen, L.; Shi, S. Machine learning prediction of activation energy in cubic li-argyrodites with hierarchically encoding crystal structure-based (hecs) descriptors. *Science Bulletin* **2021**, *66* (14), 1401−1408.
(32) Zhao, Q.; Zhang, L.; He, B.; Ye, A.; Avdeev, M.; Chen, L.; Shi, S. Identifying descriptors for li+ conduction in cubic li-argyrodites via hierarchically encoding crystal structure and inferring causality. *Energy Storage Materials* **2021**, *40*, 386−393.
(33) He, X.; Bai, Q.; Liu, Y.; Nolan, A. M.; Ling, C.; Mo, Y. Crystal structural framework of lithium super-ionic conductors. *Adv. Energy Mater.* **2019**, *9* (43), No. 1902078.
(34) Edelsbrunner, H.; Letscher, D.; Zomorodian, A. In Topological persistence and simplification, Proceedings 41st annual symposium on foundations of computer science; IEEE, 2000; pp 454−463.
(35) Zomorodian, A.; Carlsson, G. Computing persistent homology. *Discrete & Computational Geometry* **2005**, *33* (2), 249−274.
(36) Xia, K.; Feng, X.; Tong, Y.; Wei, G. W. Persistent homology for the quantitative prediction of fullerene stability. *Journal of computational chemistry* **2015**, *36* (6), 408−422.
(37) Cang, Z.; Wei, G.-W. Integration of element specific persistent homology and machine learning for protein-ligand binding affinity prediction. *Int. J. Numer. Methods Biomed. Eng.* **2018**, *34* (2), No. e2914.
(38) Xia, K.; Wei, G.-W. Persistent homology analysis of protein structure, flexibility, and folding. *International journal for numerical methods in biomedical engineering* **2014**, *30* (8), 814−844.
(39) Kedi, W.; Zhao, Z.; Wang, R.; Wei, G.-W. TopP−S: Persistent homology-based multi-task deep neural networks for simultaneous predictions of partition coefficient and aqueous solubility. *J. Comput. Chem.* **2018**, *39* (20), 1444−1454.
(40) Hiraoka, Y.; Nakamura, T.; Hirata, A.; Escolar, E. G.; Matsue, K.; Nishiura, Y. Hierarchical structures of amorphous solids characterized by persistent homology. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (26), 7035−7040.
(41) Lee, Y.; Barthel, S. D.; Dłotko, P.; Moosavi, S. M.; Hess, K.; Smit, B. Quantifying similarity of pore-geometry in nanoporous materials. *Nat. Commun.* **2017**, *8* (1), 15396.
(42) Vijay Anand, D.; Xu, Q.; Wee, J. J.; Xia, K.; Sum, T. C. Topological feature engineering for machine learning based halide perovskite materials design. *npj Comput. Mater.* **2022**, *8* (1), 203.
(43) Chen, D.; Chen, C.-L.; Wei, G.-W. Category-specific topological learning of metal−organic frameworks. *Journal of Materials Chemistry A* **2025**, *13* (13), 9292−9303.
(44) Aono, H.; Sugimoto, E.; Sadaoka, Y.; Imanaka, N.; Adachi, G.-y. Ionic conductivity and sinterability of lithium titanium phosphate system. *Solid State Ionics* **1990**, *40-41*, 38−42.
(45) Jalem, R.; Yamamoto, Y.; Shiiba, H.; Nakayama, M.; Munakata, H.; Kasuga, T.; Kanamura, K. Concerted migration mechanism in the li ion dynamics of garnet-type $Li_7La_3Zr_2O_{12}$. *Chem. Mater.* **2013**, *25* (3), 425−430.
(46) Frey, B. J.; Dueck, D. Clustering by passing messages between data points. *science* **2007**, *315* (5814), 972−976.
(47) Boukamp, B. A.; Huggins, R. A. Lithium ion conductivity in lithium nitride. *Phys. Lett. A* **1976**, *58* (4), 231−233.
(48) Zhang, Z.; Shao, Y.; Lotsch, B.; Hu, Y.-S.; Li, H.; Janek, J.; Nazar, L. F.; Nan, C.-W.; Maier, J.; Armand, M.; et al. New horizons for inorganic solid state ion conductors. *Energy Environ. Sci.* **2018**, *11* (8), 1945−1976.
(49) Zhang, R.; Strauss, F.; Jiang, L.; Casalena, L.; Li, L.; Janek, J.; Kondrakov, A.; Brezesinski, T. Transition-metal interdiffusion and solid electrolyte poisoning in all-solid-state batteries revealed by cryo-tem. *Chem. Commun.* **2023**, *59* (31), 4600−4603.
(50) Ong, S. P.; Mo, Y.; Richards, W. D.; Miara, L.; Lee, H. S.; Ceder, G. Phase stability, electrochemical stability and ionic conductivity of the li 10 ± 1 mp 2 × 12 (m= ge, si, sn, al or p, and x= o, s or se) family of superionic conductors. *Energy Environ. Sci.* **2013**, *6* (1), 148−156.
(51) Banerjee, A.; Wang, X.; Fang, C.; Wu, E. A.; Meng, Y. S. Interfaces and interphases in all-solid-state batteries with inorganic solid electrolytes. *Chem. Rev.* **2020**, *120* (14), 6878−6933.
(52) Morgan, B. J.; Madden, P. A. Relationships between atomic diffusion mechanisms and ensemble transport coefficients in crystalline polymorphs. *Phys. Rev. Lett.* **2014**, *112* (14), No. 145901.
(53) Wei, J. Adsorption and cracking of n-alkanes over zsm-5: negative activation energy of reaction. *Chemical engineering science* **1996**, *51* (11), 2995−2999.
(54) Kundu, M.; Mahanty, S.; Basu, R. N. Lithium hexaoxo antimonate as an anode for lithium-ion battery. *Nanomaterials and Energy* **2012**, *1* (1), 51−56.
(55) Sengupta, A.; Chamuah, A.; Ram, R.; Ghosh, C. K.; Diyali, S.; Biswas, B.; Ali, M. S.; Bhattacharya, S. Formation of $Li_{10}Zn_4O_9$, Li_2MoO_3, and $ZnSeO_3$ nanophases: roles in electrical conductivity and electrochemical stability in lithium ion conductors and their crystalline counterparts. *ECS Journal of Solid State. Science and Technology* **2022**, *11* (11), 113008.
(56) Zhao, Y.; Zhang, X.; Chen, X.; Li, W.; Li, Z.; Chen, M.; Sun, W.; Zhao, J.; Li, Y. All-solid-state electrochromic devices based on the lialsio4 electrolyte. *Mater. Lett.* **2021**, *292*, No. 129592.
(57) Prayogi, L. D.; Faisal, M.; Kartini, E.; Honggowiranto, W.; Supardi. Morphology and conductivity study of solid electrolyte Li_3PO_4. *AIP Conf. Proc.* **2016**, *1710*, No. 030047.
(58) Xu, Y.-N.; Ching, W. Y.; Chiang, Y.-M. Comparative studies of the electronic structure of $LiFePO_4$, $FePO_4$, Li_3PO_4, $LiMnPO_4$, $LiCoPO_4$, and $LiNiPO_4$. *J. Appl. Phys.* **2004**, *95* (11), 6583−6585.
(59) Liu, Y.; Yang, Z.; Zou, X.; Ma, S.; Liu, D.; Avdeev, M.; Shi, S. Data quantity governance for machine learning in materials science. *Natl. Sci. Rev.* **2023**, *10* (7), nwad125.
(60) Gedeon, O.; Jurek, K.; Drbohlav, I. Mixed-alkali effect in sodium−potassium glasses irradiated with electrons. *Journal of non-crystalline solids* **2010**, *356* (9−10), 456−460.
(61) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(62) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational materials science* **1996**, *6* (1), 15−50.
(63) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **1996**, *105* (22), 9982−9985.

https://doi.org/10.1021/jacs.5c04828
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX