COMMAT: DATASETS AND BENCHMARKS FROM COMPLEX MATERIALS FOR GRAPH MACHINE LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent research has demonstrated the efficacy of graph learning over a wide spectrum of materials, including molecular graphs, crystals, mechanical metamaterials, and strongly disordered systems. In this work, we draw attention to the broad class of *complex materials*, which combine order and disorder, and fall outside the above categories, yet have shown superior properties throughout the materials science literature. We present a Complex Material Benchmark (ComMat), including three graph datasets of complex materials from experimental and computational research studies, unifying distinctly developed data-to-graph pipelines under a standardized graph-based representation. In particular, we provide the first publicly available 3D graph dataset of a nanoscale network derived from 3D tomography. We then quantitatively show that these graphs are fundamentally different from existing materials datasets. We design various predictive tasks to advance machine learning (ML) methods, including experimentally measured properties, simulated mechanical response, and structural awareness. Extensive benchmark experiments are conducted over popular graph learning models, revealing their limitations and the need for further development in handling complex materials. ComMat is openly released to accelerate ML research and innovation in complex material design.

1 INTRODUCTION

The discovery of high-performance materials is often the bottleneck in the development of tools in science and engineering applications. Such materials must often combine contradictory properties. Examples include coatings that must simultaneously be electrically conductive, flexible, and transparent, for applications in flexible electronics; or structural batteries that must be simultaneously load-bearing, ionically conductive, but electrically insulating. With advancements in synthesis and manufacturing technologies, our ability to create these materials increasingly "outpaces our ability to test them" (Sarkisov & Kim (2015)). This has driven the development of many computational tools for studying structure–property relationships in modern materials, including inverse design methods that tailor structures to achieve desired properties. However, even with these advances, progress remains constrained by the high cost of simulating systems at realistic scales and with diverse structural details. These limitations have motivated the development of machine learning (ML) for materials science (Choudhary et al. (2022)).

Graph representations offer a flexible framework for encoding material structures across scales, enabling deeper exploration of their structure–property relationships. Data from experiments (e.g., microscopy images) and simulations (e.g., molecular dynamics trajectories) can be readily mapped to graphs, enabling the application of network science and graph learning methods to uncover deeper physical insights. However, most studies are focused on crystalline materials (Yan et al. (2022)) and molecular graphs (Gilmer et al. (2017)). In each case, the material can be well defined by a relatively small amount of information, either by using unit-cells for crystals or molecular graphs for small molecules ($\sim$100 nodes). This greatly simplifies graph representations and subsequent ML tasks.

Our first claim is that there is a very broad category of practically relevant materials that do not fall into the above categories and are not currently investigated with ML. Specifically, we are concerned with *complex materials*. In this context, *complex* refers to the combined presence of *order and disorder* in a material's structure (Mao & Kotov (2024)). This structural characteristic has been shown to

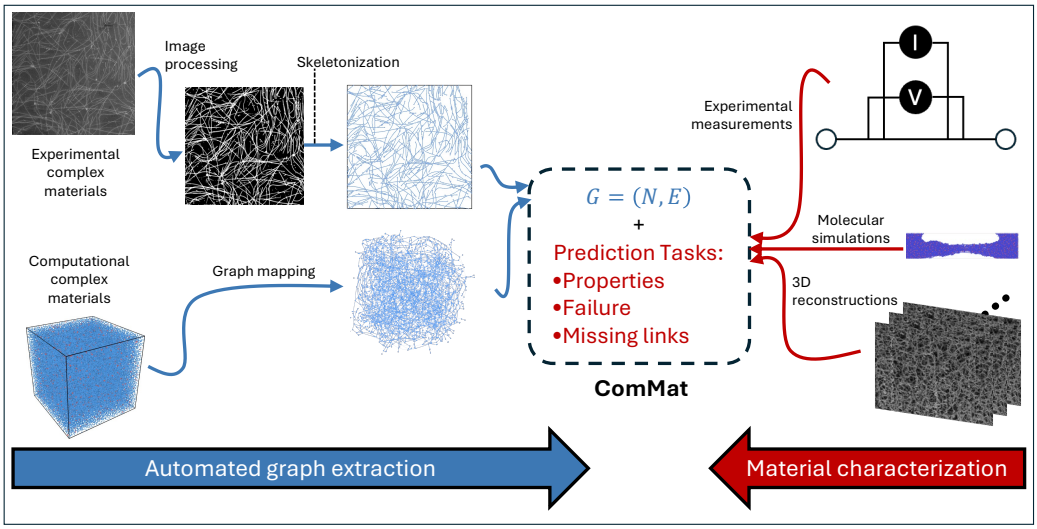

Figure 1: Left: Graph representations of complex materials derived from multi-modalities - experimental microscopy of a 2D percolating network of nanowires processed using *StructuralGT* (top) and molecular dynamics trajectories (bottom). Right: Information from complex material characterization, from top to bottom, experimental measurements of properties, molecular simulations of fracture, 3D imaging. Middle: Combination of complex material graphs and ground truth information to form ComMat samples. Each ComMat sample involves predicting complex material behavior from the associated graph, such as experimentally measured resistances, locations of failing polymer links in molecular simulations, or 3D network structure as imaged by microscopy. As a further example, we include an excerpt of a 3D graph in Appendix A.

impart many superior material properties of great importance in high-performance applications, including improved transport properties (Smart et al. (2007)), fracture toughness (Fulco et al. (2025)), anisotropy (Wu et al. (2024)), and optical activity (Kuznetsova et al. (2025)). Such materials are also synthetically more accessible, which makes them practically attractive for low-cost manufacturing and future technologies. Simultaneously, complex materials are a predominant form in Nature and so are also of fundamental interest (Jiang et al. (2020)).

However, the opportunities raised by the prevalence and utility of material complexity simultaneously raises challenges with regard to their ML-driven study. Specifically, the most efficient representations of complex materials remain challenging due to the combined order and disorder inherent in their structure (Mao & Kotov (2024)). The multi-scale nature of complex materials and the absence of a generalizable theory make their structure–property relationships difficult to explore (Ortiz-Tavárez et al. (2025)). The commonly used unit cell representations do not apply due to the lack of periodic order in complex materials. Given that complex materials are harder to generate from toy models, there is also a stronger reliance on generating them from either costly experiments or computationally expensive simulations. The resulting data scarcity motivates both centralizing the data that is already available, and subsequently developing one-shot learning models for their analysis.

One of the most promising frameworks for complex material study involves graphs, allowing for a universal representation while flexibly encoding structural details at different levels through adjustable graph properties. Recent advancements in graph-based methods have enabled quantitative study of complex materials for both experimental and computational researchers. Firstly, the development of *StructuralGT* (Vecchio et al. (2021)) - an image informed package for generating and analyzing graphs - has allowed experimental researchers to rapidly extract graphs from images. This process is depicted in the top-left quadrant of Figure 1. Meanwhile, the increasing availability of computing resources has pushed researchers to simulate increasingly complex systems. With such increasing complexity, some researchers have found benefits of graph-based representations for detecting phase transitions (Yang et al. (2024)), probing structural changes (Choi & Cho (2016)), and predicting properties (Zhang & Riggleman (2024)). In these cases, simulation trajectories are con-

verted to graphs directly from the system's bond information, as depicted in the bottom-left quadrant of Figure 1. We argue that a unified graph-based representation for these complex material - in computational and experimental structures - allows us to apply findings in the broadest possible manner. For example, the ability of Geodesic Edge Betweenness Centrality to predict failure locations in various materials has been confirmed by separate research groups in computational (Mangal et al. (2023); Zhang & Riggleman (2024)) and experimental studies (E. Berthier (2019)).

One of the greatest challenges associated with building a complex material ML benchmark is the distinctly high cost associated with collecting complex material data. Graphs from experimental samples must first be *synthesized*, requiring specialized laboratory expertise, equipment, and materials. Then they must be *imaged*, requiring expensive and difficult to operate microscopy equipment. Finally, *image-to-graph extraction* requires novel image processing including binarization, skeltonization, and node/edge detection. Python APIs with ML library compatibilities, such as *StructuralGT*, have only just been released. For graphs derived from molecular simulations, developing a physically reliable simulation pipeline, including critical steps such as force-field parameterization and validation of model assumptions, requires significant time and specialized expertise. Rigorous post-processing is also needed to ensure that the extracted networks are physically meaningful. In both experimental and computational cases, defining a practically relevant task associated with each material requires intimate knowledge of the material properties and industries that find them interesting. Combined with all the above necessitates highly interdisciplinary collaborations, which is an additional challenge.

This further motivates the current cohesive, interdisciplinary, and open-source dataset for complex material ML-driven study. We take raw data from both the computational and experimental pipelines mentioned above, and unify them under the general $G = (N, E)$ representation, whilst maintaining geometric information via node position attributes. The complex materials included in our study are chosen to satisfy two criteria: (1) adherence to our definition of complexity, involving the coexistence of ordered and disordered structural features, and (2) practical significance in real-world applications. Our dataset comprises a nanofiber graph containing more than 330,000 nodes, samples extracted from experimental nanowire images, and molecular simulations of polymer networks. It enables multiple ML tasks aligned with practical material applications, including fracture prediction (edge-level), resistance prediction (graph-level), and link prediction, as depicted in the right half of Figure 1. By evalutaing fundamental structural parameters, we quantitatively show that each of our complex materials is fundamentally distinct from materials reported in previous databases.

Our contributions are as follows. 1) We introduce ComMat, a collection of three graph datasets from complex materials, integrating both 2D and 3D structures from experimental and computational sources. We provide complexity analysis to distinguish ComMat properties from existing materials datasets. To the best of our knowledge, this is the first open-source benchmark of complex materials at this scale. 2) We design a diverse set of predictive tasks, ranging from edge-level to graph-level, to drive ML advancements in materials science. 3) We conduct extensive experiments to benchmark a wide range of mainstream graph learning baselines on our proposed tasks, establishing a robust foundation for future research.

2 Background & Related Work

2.1 Background on Order, Disorder, & Complexity

When a material has translational order, it is said that its composition is identical at points that are apart by some given separation: $f(\mathbf{r}+\mathbf{R}) = f(\mathbf{r})$, where f returns the vector that defines the content of the material at a particular point, $\mathbf{r}$ is an arbitrary position, and $\mathbf{R}$ is some separation vector. This allows for an efficient representation of the structure because one only needs to define the unit cell and the manner in which it can be used to reconstruct the entire material. Such representations were used by Yan et al. (2022) who represent atomic crystalline unit cells with $\mathbf{A}, \mathbf{P}, \mathbf{L}$, where each of the matrices is a set of vectors identifying the contents, positions and manner of the unit cell repeated to form the bulk material. Subsequent work by Liu et al. (2022) included further geometric features, such as bond angles, while work by Yan et al. (2024) achieves SE(3) invariant and SO(3) equivariant representations. Similar methods have been applied to polymer crystalline materials (Zeng et al. (2018)), while extensions to weakly disordered materials involve training on their strictly ordered

counterparts, followed by extensions that accommodate the minor deviations from perfect order (Chen et al. (2021); Wang et al. (2022); Eremin et al. (2024)).

Just as is the case with the above ordered materials, when materials are strongly disordered, their behavior can be inferred from their local environment (Kim et al. (2023)). As a result, graph learning studies have allowed researchers to differentiate liquids from amorphours solids (Swanson et al. (2020)), predict their long-time evolution (Bapst et al. (2020)), and inversely design their structure for mechanical property improvements (Wang & Zhang (2021)). Local environment representations have also been shown to reproduce mechanical properties (Sak (2025)) and short-range order in high entropy alloys (Sheriff et al. (2024)).

When materials are neither ordered nor disordered, but instead complex, their local behavior is often dependent on structural features beyond just their local environment. This phenomenon has been demonstrated by Smart et al. (2007), who show that particles with high betweenness centrality values experience higher heat flow than those with high nodal degree. Similarly, concentration of stress in granular media and strut lattices is dependent on global connectivity patterns, as captured by betweenness centrality metrics (Kollmer & Daniels (2019); E. Berthier (2019); Reyes-Martinez et al. (2024)).

By examining the definition of betweenness centrality, we may learn the problems with applying popular unit-cell like graph representations of complex materials: Edge betweenness centrality, EBC_G, of an edge, e, is defined as

$$\mathrm{EBC}_G(e) = \frac{1}{N(N-1)} \sum_{s,t \in N} \frac{\sigma_{st}(e)}{\sigma_{st}} \tag{1}$$

where N is the number of nodes. σ_{st} is the number of shortest paths between a given source and target, s, t, respectively. $\sigma_{st}(e)$ is the number of shortest paths between s, t containing e. So given that the local betweenness value depends on the entire graph, because of the dependence on σ_{st}, learning betweenness centrality for complex materials would require inefficiently numerous iterations of the 1-hop neighbor information exchange scheme that is central to graph learning techniques that use the message-passing (MP) paradigm (Gilmer et al. (2017)). This is because each iteration of the MP paradigm only augments the representations of each node with its immediate neighbors. It is this defining feature of local behavior depending on global connectivity that often separates complex materials from their ordered and disordered counterparts and thus motivates our complex material dedicated dataset.

2.2 Related Work

In preparing this dataset, we surveyed the current literature on graph learning material datasets. We learnt that all previous studies use the message passing paradigm, which cannot easily learn centrality parameters, which have previously shown to impact complex material properties. We also learnt that many rely on unit-cell representations that are incompatible with our materials. Finally, we learnt that there are no existing benchmarks for complex materials. Below we review some existing datasets and benchmarks to highlight the gap in the literature that our work fills. Further related work is discussed in Appendix D.

Mechanical Metamaterials A popular class of materials for graph machine learning is metamaterials, defined as those with interesting properties arising from their precisely arranged geometry (Fang et al. (2022)). Having exploded in recent years due to advances in 3D printing and additive manufacturing, there is increasing demand for their computational generation and analysis. Works targeting metamaterial design often use experiments or finite-element simulations for establishing ground truth target properties and behavior, such as MetaMatBench (Chen et al. (2025a)), which standardizes 5 diverse metamaterial datasets into a unified representation $M = (L, U, y)$. They also provide a robust toolbox with 17 adapted ML models for property prediction and inverse design, alongside a novel evaluation suite featuring 12 metrics and finite-element analysis for physics-aware validation. Zhan et al. (2025) then extend this work to collectively consider all modalities associated with ordered mechanical metamaterial design.

Although mechanical metamaterials are not necessarily ordered, the above works focus on ordered materials, and thus, rely on unit-cell based graph representations. While Grega et al. (2024) include

nodal perturbations in their GNN trained dataset, the importance of tuning disorder beyond the mechanical metamaterial unit cell is highlighted by Fulco et al. (2025), who show that its presence enhances fracture toughness. In other words, identifying the best candidates must involve inclusion of complex materials, relying on graph representations incompatible with the unit-cell approach used in the above works.

Nanomaterials In a similar vein to metamaterials, there are numerous studies of crystalline nanomaterials whose assumption of order enables unit-cells for efficient graph-based representations. Instead of finite-element, the ground truth for most ML tasks comes from Density Functional Theory (DFT). ECD Chen et al. (2025b) proposes large-scale datasets for electronic charge densities prediction. It contains 140,646 crystal geometries with medium-precision calculations and a subset of 7,147 geometries with high-precision data. Friis-Jensen et al. (2024) present two novel datasets for nanomaterials: CHILI-3K, a medium-scale dataset of over 6 million nodes focused on monometallic oxides, and CHILI-100K, a large-scale dataset of over 183 million nodes derived from experimentally determined crystal structures with broad chemical diversity. They also define several property and structural prediction tasks, benchmarking a wide array of GNN models. Instead of studying bulk crystals, they study small crystallites whose entirety can be represented with a graph.

3 ComMat Datasets

In Sections 3.1–3.3, we detail three datasets from complex materials with broad real-world applications: polymer networks obtained from 3D molecular dynamics simulations, silver nanowire networks derived from 2D microscopy images, and an aramid nanofiber network reconstructed from 3D tomographic images. Because this contribution is concerned with graph machine learning benchmarks, all our materials are networks that have previously been shown to benefit from graph-based representations. In Section 3.4, we quantify the structural complexity of these materials.

3.1 Polymer Network

Polymer networks are formed by crosslinking polymer chains at multifunctional junctions and are widely used for tissue engineering, implantable devices and drug delivery (Li & Mooney (2016); Nonoyama & Gong (2021); Yuk et al. (2022)). A central challenge is understanding how these networks fracture, since performance is often limited by when and where failure occurs. The underlying network topology (the arrangement of chains and junctions) plays a critical role in fracture behavior (Dobrynin et al. (2023)). In molecular dynamics simulations, both generating these networks and simulating their fracture require very large systems (tens of thousands of particles) to capture macroscopic behavior, demanding thousands of CPU hours on high-performance computing resources.

We introduce a structured dataset derived from coarse-grained molecular dynamics (MD) simulations that frames a long-standing materials science problem as a machine learning benchmark. The MD simulation protocol follows established approaches (Barney et al. (2022)). Briefly, the simulations mimic experimental systems where linear polymer chains with reactive end groups crosslink with tetrafunctional junctions, and the resulting network structures are represented as graphs: junctions are modeled as nodes and polymer chains as edges. Graph representations naturally encode topological defects such as self-loops, parallel edges, and higher-order cycles. Prior work has shown that polymer chains with fewer local topological defects, higher-than-average geodesic edge betweenness centrality, and stronger alignment with the loading direction are more likely to break under uniaxial tensile deformation (Zhang & Riggleman (2024)). These insights highlight opportunities for the ML community to apply graph models to predict fracture behavior directly from network topology.

We provide ensemble datasets from 10 combinations of polymer mole fraction (ϕ) and chain length (N), which produce networks with varying topology and defect concentrations. Because fracture is not fully deterministic and thermal fluctuations influence which chains break, we performed isoconfigurational ensemble simulations, where the same network structure was fractured multiple times with randomized initial particle velocities. This approach yields per-chain breakage probabilities (P_{break}) and sequences of break events. The benchmark dataset supports two core tasks: 1) Link prediction: recovering missing connections in partially observed networks, providing a pathway to accelerate generation of large-scale polymer systems. 2) Fracture prediction: predicting where

and in what order polymer chains break under uniaxial deformation, directly linking microscopic network topology to macroscopic failure behaviors.

3.2 NANOWIRE NETWORKS

Nanowire networks show unique combinations of transparency, flexibility, and conductivity, making them ideal for flexible electronics and electrical shields for aerospace applications (Wu et al. (2024); Tan et al. (2020)). Surrogate random stick models have been used by theorists to study nanowire networks (Jagota & Scheinfeld (2020)). However, these were recently shown to differ from the real complex structures (Wu et al. (2024)), which was shown to have an impact on their properties. While the infeasibility of testing all possible real networks motivates their ML-driven study, these materials are distinctly challenging to collect data for. Specialized laboratory expertise and equipment for material synthesis, including a purpose-built novel flow system that allows the optimal air-solvent mixture required for film synthesis, is required. Property measurement for the resultant network then requires running four-point probe experiments specifically designed for assessment of conductive film resistances. Advanced scanning electron microscopy is then required for resolving nanowire network structure (costing \$100k-\$1mn).

In order to open the possibility for data-scare model development, we present this dataset of 33 graphs across four sets of scanning electron microscopy images of silver nanowires. Each set corresponds to a sample with a particular electrical sheet resistance that was measured experimentally, using the well-established four-point probe technique, with experimental measurement error less than 10 $\%$. Each image has a graph associated with it, which was extracted with *StructuralGT*, as depicted in the top left quadrant of Figure 1. For this dataset, we test the ability for graph learning to predict the electrical resistance of the networks from the extracted graphs.

The challenge in developing ML models for this dataset thus far comes from the combined difficulties of the above mentioned experimental synthesis, imaging, and characterization, leading to data-scarcity.

3.3 ARAMID NANOFIBERS

Synthesized from well-known Kevlar, aramid nanofiber networks show enhanced thermal and mechanical properties (Yang et al. (2011)), and have since shown application in membranes (Wang et al. (2022)), structural batteries (Wang et al. (2020)), and protective coatings (Wang et al. (2025b)). Aramid nanofiber networks take even longer to synthesize than nanowire networks ($\sim$ 10 days). To confidently predict their complex, multidirectional behavior from their structure, accurate 3D imaging is required. Our nanofiber graph is the first public example from a 3D tomographic image, processed by *StructuralGT* and yielding a graph of over 330,000 nodes. For this dataset, we test a series of models' abilities to predict the presence of links when the graph is incomplete.

3.4 DATASET COMPLEXITY

While our material graphs come from a diverse range of experimental and computational sources, they share a common feature: structural complexity. Although quantitative complexity measures remain an open question in materials science, here we adopt the general heuristic of a mix of order and disorder/defects as a hallmark of complexity, commensurate with information theory and thermodynamics (Mao & Kotov (2024)). Leveraging a wide range of graph theory metrics, we contrast the highly ordered and disordered structures with our own structures in Fig. 2. The clustering and square clustering distributions for all networks are shown in Fig.2 a-l. We also calculate generalized fractal dimensions (GFD) and node-based multifractal analysis (NMFA) for topological complexity Xiao et al. (2021). In Fig.2o we show standard parameters quantifying local environments, such as degree, clustering, square clustering, and degree assortativity. We also include graph Ollivier-Ricci curvature (ORC), which is a crucial feature in identifying the community structure of complex networks (Sia et al. (2019)), and asymmetry from the analyses in (Xiao et al. (2021)). Separate from graph theory, nematic order parameter (Wu et al. (2024)) is used to measure the correlation between orientations of geometric entities, and increases with increasing orientational order. Full definitions are given in Appendix C.

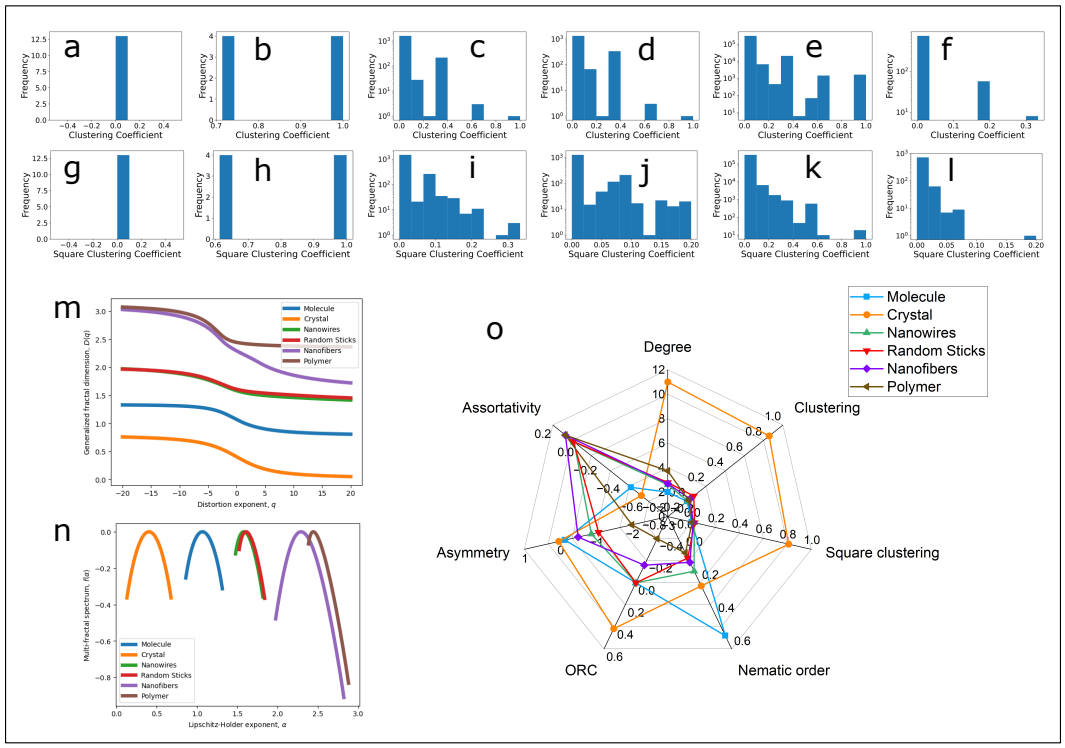

Figure 2: Clustering coefficient distributions for graphs from a variety of material. (a) Molecular graph of aspirin, from PubChem; (b) Unit-cell for zincblende, from CHILI (Friis-Jensen et al. (2024)); (c) A graph from an experimental image of a nanowire network; (d) A random stick models created for this work; (e) A graph from a 3D tomographic reconstruction of aramid nanofibers; (f) A graph from the molecular simulations of polymer networks. (g-l) is the same but for square clustering coefficients. (m) Generalized fractal dimension of each graph; (n) Multi-fractal spectrum of each graph; (o) a radar plot depicting the values of some graph theory metrics and order parameters or each of the graphs (average degree, average clustering coefficient, average square clustering coefficient, nematic order parameter, average Ollivier-Ricci curvature (ORC), asymmetry, degree assortativity. The nanowire and random stick values have been normalized w.r.t. their ordered disordered limiting values for 2D structures (see Appendix C).

First, because of stoichiometric constraints not present in most complex materials, the molecular graph for aspirin has (square) clustering coefficient of 0 for all atoms (Figure 2 a,g). Due to its simplicity and small size, it has near-zero asymmetry, and high nematic order (Figure 2o). For the crystal unit cell, high order leads to (square) clustering coefficients taking only one of two values (Figure 2 b,h). High density of connections leads to high clustering values, degree and ORC, shown in Figure 2o. Additionally, it can be shown that there are only 19 unique orientations of the bonds, which results in the high nematic order parameter. As with molecular graph, the asymmetry is near-zero, suggesting no dominant structure. The exponent α_0 from NMFA - the α value at maximum $f(\alpha)$ - is crucial to quantify the complexity of network structure. The higher α_0 indicates the higher degree of complexity in network structure (Xiao et al. (2021)). From Figure 2 n, both crystalline unit cells and small molecule graphs - common targets of graph learning studies - manifest low exponent α_0, therefore do not exhibit complex structure.

Meanwhile, the NMFA spectra show significantly higher α_0 in nanofibers and polymer datasets, demonstrating higher degree of complexity in their networks (Fig.2 n). The negative asymmetry implies thorn-like structures dominate in the network. To highlight how hidden geometric order may result in unique complex material properties relevant to their application, we can compare complex nanowire network results to results from the commonly used random stick surrogate model. Comparisons show that, despite having very similar clustering distributions, GFDs and NMFA (Fig.2 c,d,i,j), the real networks have significantly higher hidden nematic order (Figure 2o). This is because

Table 1: Statistics of datasets.

Dataset	#Graph	#Node (per graph)	#Edge (per graph)	Diameter
Nanofibers	1	336968	901066	154
Nanowire	33	712 - 3899	1670 - 9162	50 - 128
Polymer	10	800 - 8000	1493 - 15642	12 - 44

Table 2: Evaluation of the Link Prediction on ComMat datasets. The metric used is AUC ↑. Best is bolded and runner-up is underlined.

Model	Nanofibers			Nanowire			Polymer		
	5%	10%	20%	5%	10%	20%	5%	10%	20%
GCN	63.60±1.10	63.02±1.03	59.14±0.80	75.00±1.84	73.04±1.74	71.90±3.70	80.68±0.27	79.47±0.25	78.36 ± 1.11
GAT	56.40±0.55	55.23±0.78	53.10±0.65	66.80±1.30	67.52±1.65	64.46±2.71	76.18±0.69	77.04±0.51	77.10±0.44
GraphSAGE	**77.19±0.44**	**74.95±0.17**	**68.54±0.92**	66.13±1.86	62.70±1.34	56.64±1.99	78.36±1.34	78.14±0.39	77.41±0.48
Centrality-GCN	68.78 ± 2.93	65.63±2.36	61.56 ± 1.80	**77.62 ± 1.18**	**74.45±1.93**	**72.13 ± 3.70**	81.45 ± 0.16	79.24±1.16	77.89 ± 0.85
Centrality-SAGE	76.80 ± 0.44	74.84±0.46	67.53 ± 1.23	66.43 ± 2.35	63.85±1.90	57.30 ± 2.31	**83.26 ± 2.88**	80.48±2.63	76.97 ± 0.48
EGNN	73.02±11.9	73.44±10.6	67.98±12.9	62.73±11.7	55.10±8.33	55.61±9.73	72.68±0.71	75.43±1.12	75.50±0.28
Equiformer	76.14±10.3	63.51±11.5	61.52±10.1	59.54±1.28	53.92±1.80	54.63±1.58	82.91±0.09	**82.57±0.48**	**80.24±0.94**

random stick networks are generated from i.i.d. sampling of orientations and positions, while real nanowire networks experience crowding effects, resulting in alignment of neighboring edges, and hence correlations in the orientations. Because it is a geometric effect, the discrepancy is not detected by most topological analyses (Fig.2m-o). Although subtle, this structural difference has been shown to impart unexpected directional dependencies of properties (Wu et al. (2024)).

Thus, both graph theory and materials metrics confirm the presence of complexity in our proposed datasets. Additionally, the presence of defects in the polymer network is enumerated by a considerable number of nodes with non-zero clustering coefficients (Figure 2 f), which should be zero for all nodes in ideal networks (Zhang & Riggleman (2024)).

4 Experiment and Benchmarking

In this section, we elaborate on the experimental setup for our proposed datasets, where the overview of statistics is provided in Table 1. We perform experiments on comprehensive standard graph learning models, evaluating their effectiveness and discussing potentials for future research.

4.1 Evaluation Tasks

Link prediction Link prediction aims to predict missing links from incomplete graph, which is a fundamental task to evaluate model capability to capture the network structure. Link prediction is crucial for graphs derived from both experimental imaging and MD simulations because microscopy-based skeletonization often introduces missing or spurious edges, and generating networks via MD simulations can be computationally expensive. A learned link prediction model can denoise experimental graphs by distinguishing true physical connections from imaging artifacts, and can also approximate MD-generated connectivity patterns at a fraction of the computational cost. Robust link prediction performance enables more accurate, scalable, and physically meaningful graph representations for studying structure–property relationships in complex materials.

We conduct the link prediction as a binary classification task, predicting whether there exists a link between two nodes. For Nanowire and Polymer datasets, we use the largest graph for this task. For each graph, we retain a portion of total links for validating and testing, with the ratio of 5% / 10% / 20% sequentially. The models are trained on the subgraph constructed from the remaining links. The ratio of positive and sampled negative edges is set as 1:1. The performance is measured by the Area Under Curve (AUC).

Resistance prediction We formulate the property prediction task on the Nanowire dataset as a graph regression problem, which aims to predict the electrical resistance for each graph. For each set of Nanowire, we split the total number of graphs into 60% / 20% / 20% for training, validation,

Table 3: Evaluation of resistance prediction for Nanowire. Best is bolded and runner-up is underlined.

Metrics	GCN	GAT	GraphSAGE	EGNN	Equiformer
MAE ↓	61.48 ± 0.38	60.98 ± 0.29	62.83 ± 1.68	<u>59.89 ± 0.12</u>	**57.46 ± 2.76**
RE ↓	0.49 ± 0.01	0.49 ± 0.01	<u>0.48 ± 0.01</u>	0.51 ± 0.01	**0.39 ± 0.07**

Table 4: Evaluation of the link breaking probability for polymer. Best is bolded and runner-up is underlined.

	GCN	GAT	GraphSAGE	EGNN	Equiformer
mMAE ↓	<u>0.30</u> ± 0.03	**0.29** ± 0.03	0.32 ± 0.05	0.31 ± 0.03	0.50 ± 0.01

and testing, respectively. We use the mean absolute error (MAE) and relative error (RE) to measure the deviation of predictions from ground truth.

Fracture prediction Predicting fracture in polymer networks is crucial for evaluating material performance across a wide range of applications. We design this task as link regression problem, where the model predicts the breaking probability P_{break} for each target link. For each graph, the total links are partitioned into training (60%), validation (20%), and testing (20%) sets. To maintain the balance in ground-truth distribution, we sample a maximum of 100 links for each label. MAE is used to measure the error for each graph, and we report the mean MAE (mMAE) for the overall performance of all graphs.

4.2 Graph Learning Models

Classic GNNs Classic GNNs adopt the MP paradigm to learn from graph-structured data, where each node iteratively updates its feature representation by aggregating information from its neighbors. Among that, GCN Kipf (2016) aggregates information by operating convolution, which is approximately equivalent to mean pooling over the neighbor nodes. GAT Veličković et al. (2017) applies a self-attention mechanism to assign weights to each neighbor, thereby modeling their levels of importance. SAGE Hamilton et al. (2017) concatenates information from its neighbors with the target node representation during aggregation, ensuring that the node's original identity is preserved and plays a key role in its new embedding.

Centrality-aware GNNs We adopt centrality encoding Ying et al. (2021), which embeds node degree centrality together with features vector, to popular GNN methods such as GCN and GraphSAGE, denoted as Centrality-GCN and Centrality-SAGE.

Geometric GNNs These methods leverage spatial information to learn equivariant graph representations under geometric transformations, such as translation, rotation, and reflection. E-GNN Satorras et al. (2021) employs equivariance by incorporating relative squared distances into the MP function and updating node coordinates based on relative position vectors. Equiformer Liao et al. (2023) leverages irreducible representations to integrate SE(3)-equivariant features into the network's channels, preserving geometric information without altering the graph structure.

4.3 Results

Link prediction Results from Table.2 show that existing baselines can achieve promising performance on this task. Overall, centrality-aware GNNs can match or outperform their classic counterparts, suggesting that incorporating centrality could be beneficial and promising for future development. Geometric GNNs often show higher fluctuation performance while being less computationally efficient. This might indicate that the geometric equivariance properties are not as beneficial for this particular link prediction task, and these models may require careful hyperparameter tuning. Overall, the performance degrades as more of the graph structures are hidden, reflecting challenges in structural missing.

Resistance prediction The results in Table 3 highlight two findings: first, all baselines obtain suboptimal performance on this task, as shown by the high MAE and RE values. It shows that the task remains challenging for current mainstream GNN models, and more effective methods should be investigated. Second, geometric GNNs significantly outperform classic counterparts on this task. This suggests the geometric information is essential for physics-related problems, where target properties could be intrinsically governed by the material's physical shape.

Fracture prediction We also witness the underperformance of benchmark graph methods on this task, as shown by the high MAE values in Table.4. The classic GNNs tends to be more robust than geometric class, suggesting the local connectivity patterns could be more effective for predicting the breaking probability.

Discussion Overall, there is no single model that dominates for all tasks. We find that benchmark models are unable to predict behavior that has previously been shown to be deterministic and predictable from complex material graph structure (Wu et al. (2024); Zhang & Riggleman (2024)). This may arise from a combination of complex material data scarcity and from the MP's inability to efficiently learn features uniquely crucial to predicting complex material performance. Specifically, betweenness centrality parameters and edge-flows (as calculated from networked linear transport methods) are known to be strong predictors of complex material performance. However, those long-range connectivity patterns are computationally expensive, limiting their utilization on large-scale graphs. Experimental results show that integrating simple, local centrality can yield performance gain, suggesting the potential for further centrality-aware advances. Future research could investigate more effective, efficient approaches to incorporate global indicators with local aggregation to enhance performance.

While we acknowledge the limited data for the resistance prediction task, such data scarcity is an inevitable reality in complex materials as well as frontier research. We position this benchmark as a rigorous challenge, urging for the development of models capable of learning on extremely sparse data. We also invite the community to explore broader utilization of this dataset. A possible scenario is to leverage it as an independent, zero-shot benchmark for assessing materials foundation models Miret & Krishnan (2025); Yan et al. (2025).

5 Conclusion

In this work, we present ComMat, the first open graph benchmark of complex materials to foster ML research for advanced materials science. ComMat consists of 3 complex materials datasets unified under graph representations from multi-modality sources, including previously unreported 3D nanonetwork graphs. We show how the complex material structures at this intersection exhibit structural features distinct from previous material datasets. By providing a range of material scales, properties, and predictive tasks, ComMat enables a thorough and challenging evaluation of machine learning models. Our benchmark experiments highlight a significant performance gap, showing that popular graph ML models currently struggle with the challenges posed by these complex materials. These limitations raise the urgent need for novel graph learning algorithms that can accelerate the design and discovery workflows for advanced materials.

REFERENCES

Graph representation of local environments for learning high-entropy alloy properties. *Machine Learning: Science and Technology*, 6:025005, 4 2025. ISSN 2632-2153. doi: 10.1088/2632-2153/ADC0E1. URL https://iopscience.iop.org/article/10.1088/2632-2153/adc0e1https://iopscience.iop.org/article/10.1088/2632-2153/adc0e1/meta.

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J. Ballard, Joshua Bambrick, Sebastian W. Bodenstein, David A. Evans, Chia-Chun Hung, Michael O'Neill, David Reiman, Kathryn Tunyasuvunakool, Zachary Wu, Akvilė Žemgulytė, Eirini Arvaniti, Charles Beattie, Ottavia Bertolli, Alex Bridgland, Alexey Cherepanov, Miles Congreve, Alexander I. Cowen-Rivers, Andrew Cowie, Michael Figurnov, Fabian B. Fuchs, Hannah Gladman, Rishub Jain, Yousuf A. Khan, Caroline M. R. Low, Kuba Perlin, Anna Potapenko, Pascal Savy, Sukhdeep Singh, Adrian Stecula, Ashok Thillaisundaram, Catherine Tong, Sergei Yakneen, Ellen D. Zhong, Michal Zielinski, Augustin Žídek, Victor Bapst, Pushmeet Kohli, Max Jaderberg, and Demis Hassabis John M. Jumper. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630:493–500, 6 2024. ISSN 14764687. doi: 10.1038/S41586-024-07487-W. URL https://www.nature.com/articles/s41586-024-07487-w.

Timothy Atkinson, Thomas D. Barrett, Scott Cameron, Bora Guloglu, Matthew Greenig, Charlie B. Tan, Louis Robinson, Alex Graves, Liviu Copoiu, and Alexandre Laterre. Protein sequence modelling with bayesian flow networks. *Nature Communications*, 16:1–14, 12 2025. ISSN 20411723. doi: 10.1038/S41467-025-58250-2. URL https://www.nature.com/articles/s41467-025-58250-2.

V. Bapst, T. Keck, A. Grabska-Barwińska, C. Donner, E. D. Cubuk, S. S. Schoenholz, A. Obika, A. W.R. Nelson, T. Back, D. Hassabis, and P. Kohli. Unveiling the predictive power of static structure in glassy systems. *Nature Physics*, 16:448–454, 4 2020. ISSN 17452481. doi: 10.1038/S41567-020-0842-8. URL https://www.nature.com/articles/s41567-020-0842-8.

Albert László Barabási. Scale-free networks: A decade and beyond. *Science*, 325:412–413, 7 2009. ISSN 00368075. doi: 10.1126/SCIENCE.1173299/SUPPL_FILE/1173299S3.MOV. URL /doi/pdf/10.1126/science.1173299?download=true.

Christopher W. Barney, Ziyu Ye, Ipek Sacligil, Kelly R. McLeod, Han Zhang, Gregory N. Tew, Robert A. Riggleman, and Alfred J. Crosby. Fracture of model end-linked networks. *Proceedings of the National Academy of Sciences*, 119(7):e2112389119, February 2022. doi: 10.1073/pnas.2112389119.

Ilyes Batatia, David Peter Kovacs, Gregor N. C. Simm, Christoph Ortner, and Gabor Csanyi. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (eds.), *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=YPpSngE-ZU.

Chi Chen, Yunxing Zuo, Weike Ye, Xiangguo Li, and Shyue Ping Ong. Learning properties of ordered and disordered materials from multi-fidelity data. *Nature Computational Science*, 1:46–53, 1 2021. ISSN 26628457. doi: 10.1038/S43588-020-00002-X. URL https://www.nature.com/articles/s43588-020-00002-x.

Jianpeng Chen, Wangzhi Zhan, Haohui Wang, Zian Jia, Jingru Gan, Junkai Zhang, Jingyuan Qi, Tingwei Chen, Lifu Huang, Muhao Chen, et al. Metamatbench: Integrating heterogeneous data, computational tools, and visual interface for metamaterial discovery. *arXiv preprint arXiv:2505.20299*, 2025a.

Pin Chen, Zexin Xu, Qing Mo, Hongjin Zhong, Fengyang Xu, and Yutong Lu. Ecd: A machine learning benchmark for predicting enhanced-precision electronic charge density in crystalline inorganic materials. In *The Thirteenth International Conference on Learning Representations*, 2025b.

Jun-Ho Choi and Minhaeng Cho. Ion aggregation in high salt solutions. V. Graph entropy analyses of ion aggregate structure and water hydrogen bonding network. *The Journal of Chemical Physics*, 144(20):204126, May 2016. ISSN 0021-9606. doi: 10.1063/1.4952648.

Kamal Choudhary, Brian DeCost, Chi Chen, Anubhav Jain, Francesca Tavazza, Ryan Cohn, Cheol Woo Park, Alok Choudhary, Ankit Agrawal, Simon J. L. Billinge, Elizabeth Holm, Shyue Ping Ong, and Chris Wolverton. Recent advances and applications of deep learning methods in materials science. *npj Computational Materials*, 8(1):59, April 2022. ISSN 2057-3960. doi: 10.1038/s41524-022-00734-6.

Andrey V. Dobrynin, Yuan Tian, Michael Jacobs, Evgeniia A. Nikitina, Dimitri A. Ivanov, Mitchell Maw, Foad Vashahi, and Sergei S. Sheiko. Forensics of polymer networks. *Nature Materials*, 22 (11):1394–1400, November 2023. ISSN 1476-4660. doi: 10.1038/s41563-023-01663-5.

K.E. Daniels E. Berthier, M.A. Porter. Forecasting failure locations in 2-dimensional disordered lattices. *Proceedings of the National Academy of Sciences of the United States of America*, 2019. doi: 10.1073/pnas.1900272116. URL www.pnas.org/cgi/doi/10.1073/pnas.1900272116. Failure points of disordered lattice materials may be predicted from their GT parameters.

Roman A. Eremin, Innokentiy S. Humonen, Alexey A. Kazakov, Vladimir D. Lazarev, Anatoly P. Pushkarev, and Semen A. Budennyy. Graph neural networks for predicting structural stability of cd- and zn-doped -cspbi3. *Computational Materials Science*, 232:112672, 1 2024. ISSN 0927-0256. doi: 10.1016/J.COMMATSCI.2023.112672. URL https://www.sciencedirect.com/science/article/abs/pii/S0927025623006663?via%3Dihub.

Xin Fang, Jihong Wen, Li Cheng, Dianlong Yu, Hongjia Zhang, and Peter Gumbsch. Programmable gear-based mechanical metamaterials. *Nature Materials*, 21:869–876, 8 2022. ISSN 14764660. doi: 10.1038/S41563-022-01269-3. URL https://www.nature.com/articles/s41563-022-01269-3.

Ulrik Friis-Jensen, Frederik L Johansen, Andy S Anker, Erik B Dam, Kirsten MØ Jensen, and Raghavendra Selvan. Chili: Ch emically-i nformed l arge-scale i norganic nanomaterials dataset for advancing graph machine learning. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4962–4973, 2024.

Sage Fulco, Michal K. Budzik, Hongyi Xiao, Douglas J. Durian, and Kevin T. Turner. Disorder enhances the fracture toughness of 2d mechanical metamaterials. *PNAS Nexus*, 4, 2 2025. ISSN 27526542. doi: 10.1093/PNASNEXUS/PGAF023. URL https://dx.doi.org/10.1093/pnasnexus/pgaf023.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. *CoRR*, abs/1704.01212, 2017. URL http://arxiv.org/abs/1704.01212.

Alex Graves, Rupesh Kumar Srivastava, Timothy Atkinson, and Faustino Gomez. Bayesian flow networks. 8 2023. URL https://arxiv.org/pdf/2308.07037.

Ivan Grega, Ilyes Batatia, Gabor Csanyi, Sri Karlapati, and Vikram Deshpande. Energy-conserving equivariant GNN for elasticity of lattice architected metamaterials. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=smy4DsUbBo.

Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.

Milind Jagota and Isaac Scheinfeld. Analytical modeling of orientation effects in random nanowire networks. *Physical Review E*, 101:12304, 2020. doi: 10.1103/PhysRevE.101.012304. Method for computing expected value of adjacentcy matrix (and hence conductivity) for a nanowire network, given a distribution of positions and orientations.

Wenfeng Jiang, Zhi-bei Qu, Prashant Kumar, Drew Vecchio, Yuefei Wang, Yu Ma, Joong Hwan Bahng, Kalil Bernardino, Weverson R. Gomes, Felippe M. Colombari, Asdrubal Lozada-Blanco, Michael Veksler, Emanuele Marino, Alex Simon, Christopher Murray, Sérgio Ricardo Muniz, André F. de Moura, and Nicholas A. Kotov. Emergence of complexity in hierarchically organized chiral particles. *Science*, 368:642–648, 2020. ISSN 10959203. doi: 10.1126/science.aaz7949. URL https://www.science.org/doi/10.1126/science.aaz7949.

Eun Cheol Kim, Dong Jae Chun, Chung Bin Park, and Bong June Sung. Machine learning predicts the glass transition of two-dimensional colloids besides medium-range crystalline order. *Phys. Rev. E*, 108:044602, Oct 2023. doi: 10.1103/PhysRevE.108.044602. URL https://link.aps.org/doi/10.1103/PhysRevE.108.044602.

TN Kipf. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Jonathan E Kollmer and Karen E Daniels. Betweenness centrality as predictor for forces in granular packings. *Soft Matter*, 2019. doi: 10.1039/c8sm01372a.

Vera Kuznetsova, Alain Kadar, Anita Gaenko, Engin Er, Tao Ma, Kody G. Whisnant, Jessica Ma, Bing Ni, Natasha Mehta, Ji Young Kim, Yurii K. Gun'ko, and Nicholas A. Kotov. Graph-property relationships for complex chiral nanodendrimers. *ACS Nano*, 19:6095–6106, 2 2025. ISSN 1936086X. doi: 10.1021/ACSNANO.4C12964/ASSET/IMAGES/LARGE/NN4C12964_0006.JPEG. URL /doi/pdf/10.1021/acsnano.4c12964?ref=article_openPDF.

Jianyu Li and David J. Mooney. Designing hydrogels for controlled drug delivery. *Nature Reviews Materials*, 1(12):1–17, October 2016. ISSN 2058-8437. doi: 10.1038/natrevmats.2016.71.

Yi-Lun Liao, Brandon Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. *arXiv preprint arXiv:2306.12059*, 2023.

Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=givsRXsOt9r.

Deepak Mangal, Mohammad Nabizadeh, and Safa Jamali. Topological origins of yielding in short-ranged weakly attractive colloidal gels. *The Journal of Chemical Physics*, 158(1):014903, January 2023. ISSN 0021-9606. doi: 10.1063/5.0123096.

Xiaoming Mao and Nicholas Kotov. Complexity, disorder, and functionality of nanoscale materials. *MRS Bulletin*, 49:352–364, 4 2024. ISSN 08837694. doi: 10.1557/S43577-024-00698-6/METRICS. URL https://link.springer.com/article/10.1557/s43577-024-00698-6.

Santiago Miret and NM Anoop Krishnan. Enabling large language models for real-world materials discovery. *Nature Machine Intelligence*, 7(7):991–998, 2025.

M. E. J. Newman. *Networks: an introduction*. Oxford University Press, Oxford; New York, 2010.

Takayuki Nonoyama and Jian Ping Gong. Tough Double Network Hydrogel and Its Biomedical Applications. *Annual Review of Chemical and Biomolecular Engineering*, 12 (Volume 12, 2021):393–410, June 2021. ISSN 1947-5438, 1947-5446. doi: 10.1146/annurev-chembioeng-101220-080338.

José M. Ortiz-Tavárez, William Stephenson, and Xiaoming Mao. A unified model for linear responses of physical networks, August 2025.

Joshua Owoyemi and Nazim Medzhidov. Smilesformer: Language model for molecular design. In *ICLR 2023 - Machine Learning for Drug Discovery workshop*, 2023. URL https://openreview.net/forum?id=SHybourYZsv.

PubChem. Aspirin — c9h8o4 — cid 2244 - pubchem. URL https://pubchem.ncbi.nlm.nih.gov/compound/2244#section=3D-Conformer.

Marcos A. Reyes-Martinez, Alain Kadar, Steven Dunne, Sharon C. Glotzer, Christopher L. Soles, and Nicholas A. Kotov. Graph-theoretical description and continuity problems for stress propagation through complex strut lattices. 12 2024. URL https://arxiv.org/pdf/2412.15344.

Lev Sarkisov and Jihan Kim. Computational structure characterization tools for the era of material informatics. *Chemical Engineering Science*, 121:322–330, 1 2015. ISSN 0009-2509. doi: 10.1016/J.CES.2014.07.022. URL https://www.sciencedirect.com/science/article/abs/pii/S0009250914003674.

Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Killian Sheriff, Yifan Cao, Tess Smidt, and Rodrigo Freitas. Quantifying chemical short-range order in metallic alloys. *Proceedings of the National Academy of Sciences*, 121:e2322962121, 6 2024. ISSN 10916490. doi: 10.1073/PNAS.2322962121. URL /doi/pdf/10.1073/pnas.2322962121?download=true.

Jayson Sia, Edmond Jonckheere, and Paul Bogdan. Ollivier-ricci curvature-based method to community detection in complex networks. *Scientific reports*, 9(1):9800, 2019.

A. Smart, P. Umbanhowar, and J. Ottino. Effects of self-organization on transport in granular matter: A network-based approach. *Europhysics Letters*, 79:24002, 7 2007. ISSN 0295-5075. doi: 10.1209/0295-5075/79/24002. URL https://iopscience.iop.org/article/10.1209/0295-5075/79/24002https://iopscience.iop.org/article/10.1209/0295-5075/79/24002/meta.

Kirk Swanson, Shubhendu Trivedi, Joshua Lequieu, Kyle Swanson, and Risi Kondor. Deep learning for automated classification and characterization of amorphous materials. *Soft Matter*, 16:435–446, 1 2020. ISSN 1744-6848. doi: 10.1039/C9SM01903K. URL https://pubs.rsc.org/en/content/articlehtml/2020/sm/c9sm01903khttps://pubs.rsc.org/en/content/articlelanding/2020/sm/c9sm01903k.

Dongchen Tan, Chengming Jiang, Qikun Li, Sheng Bi, and Jinhui Song. Silver nanowire networks with preparations and applications: a review. *Journal of Materials Science: Materials in Electronics*, 31:15669–15696, 9 2020. ISSN 1573482X. doi: 10.1007/S10854-020-04131-X/FIGURES/25. URL https://link.springer.com/article/10.1007/s10854-020-04131-x.

Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in 't Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications*, 271:108171, February 2022. ISSN 0010-4655. doi: 10.1016/j.cpc.2021.108171.

Drew A. Vecchio, Samuel H. Mahler, Mark D. Hammig, and Nicholas A. Kotov. Structural Analysis of Nanoscale Network Materials Using Graph Theory. *ACS Nano*, 15(8):12847–12859, August 2021. ISSN 1936-0851. doi: 10.1021/acsnano.1c04711.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Conghao Wang, Gaurav Asok Kumar, and Jagath C. Rajapakse. Drug discovery and mechanism prediction with explainable graph neural networks. *Scientific Reports*, 15:1–14, 12 2025a. ISSN 20452322. doi: 10.1038/S41598-024-83090-3Y. URL https://www.nature.com/articles/s41598-024-83090-3.

Mingqiang Wang, Drew Vecchio, Chunyan Wang, Ahmet Emre, Xiongye Xiao, Zaixing Jiang, Paul Bogdan, Yudong Huang, and Nicholas A. Kotov. Biomorphic structural batteries for robotics. *Science Robotics*, 5:1912, 8 2020. ISSN 24709476. doi: 10.1126/SCIROBOTICS.ABA1912/SUPPL_FILE/ABA1912_SM.PDF. URL https://www.science.org/doi/full/10.1126/scirobotics.aba1912.

Mingqiang Wang, Shuhan Hu, Sangwok Bae, Volkan Cecen, Ahmet E. Emre, Zhengxiang Zhong, Li Liu, Carl Heltzel, Yudong Huang, and Nicholas A. Kotov. Aramid nanofiber aerogels: Versatile high complexity components for multifunctional composites. *Advanced Materials*, pp. 2502508, 2025b. ISSN 1521-4095. doi: 10.1002/ADMA.202502508. URL /doi/pdf/10.1002/adma.202502508https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.202502508https://advanced.onlinelibrary.wiley.com/doi/10.1002/adma.202502508.

Qi Wang and Longfei Zhang. Inverse design of glass structure with deep graph neural networks. *Nature Communications 2021 12:1*, 12:1–11, 9 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-25490-x. URL https://www.nature.com/articles/s41467-021-25490-x.

Zhilong Wang, Yanqiang Han, Junfei Cai, Sicheng Wu, and Jinjin Li. Deeptmc: A deep learning platform to targeted design doped transition metal compounds. *Energy Storage Materials*, 45: 1201–1211, 3 2022. ISSN 2405-8297. doi: 10.1016/J.ENSM.2021.11.020. URL https://www.sciencedirect.com/science/article/abs/pii/S2405829721005377.

Joseph L. Watson, David Juergens, Nathaniel R. Bennett, Brian L. Trippe, Jason Yim, Helen E. Eisenach, Woody Ahern, Andrew J. Borst, Robert J. Ragotte, Lukas F. Milles, Basile I.M. Wicky, Nikita Hanikel, Samuel J. Pellock, Alexis Courbet, William Sheffler, Jue Wang, Preetham Venkatesh, Isaac Sappington, Susana Vázquez Torres, Anna Lauko, Valentin De Bortoli, Emile Mathieu, Sergey Ovchinnikov, Regina Barzilay, Tommi S. Jaakkola, Frank DiMaio, Minkyung Baek, and David Baker. De novo design of protein structure and function with rfdiffusion. *Nature 2023 620:7976*, 620:1089–1100, 7 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06415-8. URL https://www.nature.com/articles/s41586-023-06415-8.

Wenbing Wu, Alain Kadar, Sang Hyun Lee, Sharon C Glotzer, Valerie Goss, and Nicholas A Kotov. Layer-by-layer assembled nanowire networks enable graph-theoretical design of multifunctional coatings. *Matter*, 2024. doi: 10.1016/j.matt.2024.09.014. URL https://doi.org/10.1016/j.matt.2024.09.014.

Xiongye Xiao, Hanlong Chen, and Paul Bogdan. Deciphering the generating rules and functionalities of complex networks. *Scientific Reports*, 11:1–15, 12 2021. ISSN 20452322. doi: 10.1038/S41598-021-02203-4. URL https://www.nature.com/articles/s41598-021-02203-4.

Keqiang Yan, Yi Liu, Yuchao Lin, and Shuiwang Ji. Periodic graph transformers for crystal material property prediction. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (eds.), *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=pqCT3L-BU9T.

Keqiang Yan, Cong Fu, Xiaofeng Qian, Xiaoning Qian, and Shuiwang Ji. Complete and efficient graph transformers for crystal material property prediction. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=BnQY9XiRAS.

Keqiang Yan, Montgomery Bohde, Andrii Kryvenko, Ziyu Xiang, Kaiji Zhao, Siya Zhu, Saagar Kolachina, Doğuhan Sarıtürk, Jianwen Xie, Raymundo Arróyave, et al. A materials foundation model via hybrid invariant-equivariant architectures. *arXiv preprint arXiv:2503.05771*, 2025.

Ming Yang, Keqin Cao, Lang Sui, Ying Qi, Jian Zhu, Anthony Waas, Ellen M. Arruda, John Kieffer, M. D. Thouless, and Nicholas A. Kotov. Dispersions of aramid nanofibers: A new nanoscale building block. *ACS Nano*, 5:6945–6954, 9 2011. ISSN 19360851. doi: 10.1021/NN2014003. URL /doi/pdf/10.1021/nn2014003?ref=article_openPDF.

Ruochen Yang, Kalil Bernardino, Xiongye Xiao, Weverson R. Gomes, Davi A. Mattoso, Nicholas A. Kotov, Paul Bogdan, and André F. de Moura. Graph Theoretical Description of Phase Transitions in Complex Multiscale Phases with Supramolecular Assemblies. *Advanced Science*, 11 (33):2402464, 2024. ISSN 2198-3844. doi: 10.1002/advs.202402464.

Chengxuan Ying, Tianle Cai, Shengjie Luo, Shuxin Zheng, Guolin Ke, Di He, Yanming Shen, and Tie-Yan Liu. Do transformers really perform badly for graph representation? *Advances in neural information processing systems*, 34:28877–28888, 2021.

Hyunwoo Yuk, Jingjing Wu, and Xuanhe Zhao. Hydrogel interfaces for merging humans and machines. *Nature Reviews Materials*, 7(12):935–952, December 2022. ISSN 2058-8437. doi: 10.1038/s41578-022-00483-4.

Minggang Zeng, Jatin Nitin Kumar, Zeng Zeng, Ramasamy Savitha, Vijay Ramaseshan Chandrasekhar, and Kedar Hippalgaonkar. Graph convolutional neural networks for polymers property prediction, 2018. URL https://arxiv.org/abs/1811.06231.

Wangzhi Zhan, Jianpeng Chen, Dongqi Fu, and Dawei Zhou. Unimate: A unified model for mechanical metamaterial generation, property prediction, and condition confirmation. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=X7uVxeFS9A.

Han Zhang and Robert A Riggleman. Predicting failure locations in model end-linked polymer networks. *Physical Review Materials*, 8(3):035604, 2024.

Kangjie Zheng, Siyue Liang, Junwei Yang, Bin Feng, Zequn Liu, Wei Ju, Zhiping Xiao, and Ming Zhang. SMI-editor: Edit-based SMILES language model with fragment-level supervision. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=M29nUGozPa.

APPENDIX

A 3D NANOFIBER NETWORK

In Figure 3, we show a few snapshots of a very small subset of the nanofiber graph.

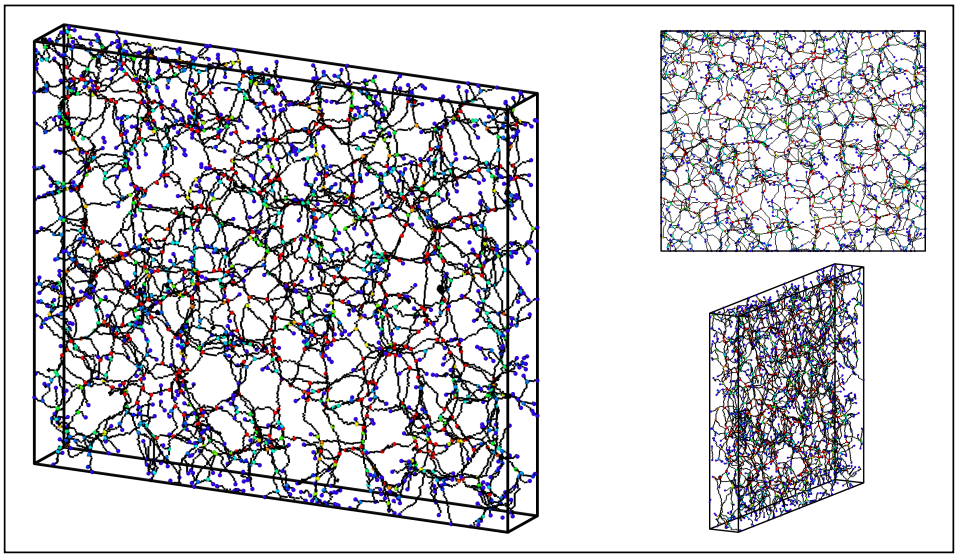

Figure 3: Snapshots of a subset of the Aramid Nanofiber graph used in this work. The nodes are colored by node betweenness centrality.

B MODEL SETTINGS

We provide the details of model parameters for each experiment tasks:

Link tasks On link prediction and fracture prediction task, we employ a GNN encoder of 2 layers and the decoder as a multilayer perceptron (MLP) for all baselines. Each GNN layer is followed by a layer normalization, and the hidden dimension is set as 32. For GAT, we use 4 heads of attention. For GraphSAGE, we use mean pooling for neighborhood aggregation. For EquiformerV2, we set the maximum degree of spherical harmonics as 2. the dropout ratio is set as 0.5 for classic GNNs and 0.1 for geometric GNNs.

Graph task For all baselines, we use a GNN encoder of 3 layers with the hidden dimension set as 32, followed by layer normalization. Finally, a graph mean pooling layer is applied to obtain a graph embeddings from all element nodes. The regressor is implemented as 2-layer MLP.

C GRAPH PARAMETER DEFINITIONS AND CALCULATION METHODS

C.1 AVERAGE DEGREE AND AVERAGE CLUSTERING COEFFICIENT

Average degree and average clustering coefficeint, Δ, were calculated using the **structural** module from *StructuralGT*, and follow the standard definitions given in the documentation:

$$\Delta = \frac{\sum_i \delta}{n},$$
$$\delta_i = \frac{2 * T_i}{k_i(k_i - 1)}$$

where T_i is the number of connected triples (visually triangles) on node i and k_i is the degree of node i.

C.2 Graph Ollivier-Ricci Curvature

Ollivier-Ricci Curvature Sia et al. (2019) measures the local geometric property in the network. An edge with positive curvature suggests it resides in a tight cluster, while a negative curvature edge tends to be a "bridge" between clusters of the network.

$$\kappa(x, y) = 1 - \frac{W_1(\mu_x, \mu_y)}{d(x, y)},$$

where $d(x, y)$ is the distance between two points x and y, $W_1(\mu_x, \mu_y)$ is the 1-Wasserstein distance (or Earth Mover's distance) between two probability distributions.

C.3 Nematic order parameter

The nematic order parameter was calculated using the **geometric** module of *StructuralGT*. The nematic order parameter S is defined via the eigenvalues of the nematic tensor, $\mathbf{Q}$. The nematic tensor is defined as

$$\mathbf{Q} = \mathbf{M} - \frac{1}{3}\mathbf{I}$$

where $\mathbf{M}$ is computed from a sum of multiplications of the vectors, $\mathbf{m}^{(i)}$, describing particle orientations:

$$\mathbf{M}_{\alpha\beta} = \sum_{i=1}^{N} \mathbf{m}_{\alpha}^{(i)} \mathbf{m}_{\beta}^{(i)}.$$

The eigenvalues of Q are $\frac{2}{3}S, -\frac{1}{3}S, -\frac{1}{3}S$. The closer a graph's value of S is to 1, the greater the degree of corelation between edge orientations. When S approaches 0, the graph's edges are uncorrelated.

For 2D networks, edges are confined to a plane, and so the lower limit for the nematic order parameter is 0.25, as shown in Wu et al. (2024). The values of nematic order parameter for the 2D networks in 2g (nanowires and random sticks) are normalized w.r.t. to this lower limit.

D Further Related Work

Biological materials Material datasets for non-ordered structures are highly prevalent throughout the biological communities. Owing to the small number of nodes in most of the molecular graphs from drug discovery studies ($\sim$100 nodes), such structures are well-suited to MP models (Batatia et al. (2022); Wang et al. (2025a)), or transformer architectures that treat molecular sequences as linguistic entities Owoyemi & Medzhidov (2023); Zheng et al. (2025).

Non-ordered materials that extend beyond this size include non-graph-based representations such as protein sequences. Most dataset-like studies involve applying new models to existing datasets and the most recent studies have acknowledged that the behavior of these structures, like complex materials, depends on more than just aggregations over local environments (Atkinson et al. (2025)). As such, there has been a recent push to extend beyond the transformer and message-passing architectures, especially towards diffusion models (Watson et al. (2023); Abramson et al. (2024)). Bottom-up development of diffusion models suited to the discreteness of protein sequences include Bayesian Flow Networks (Graves et al. (2023)). Extension of these models to graph-based representations is a critical missing piece in ML-driven complex material design and our dataset offers the best benchmark to test them.

E Further Network Analysis

Figure 4 shows degree distributions for each network. Unsuprisingly, the molecular and crystal graphs have simple degree distributions (Figure 4a,b). Random stick model and nanowire networks

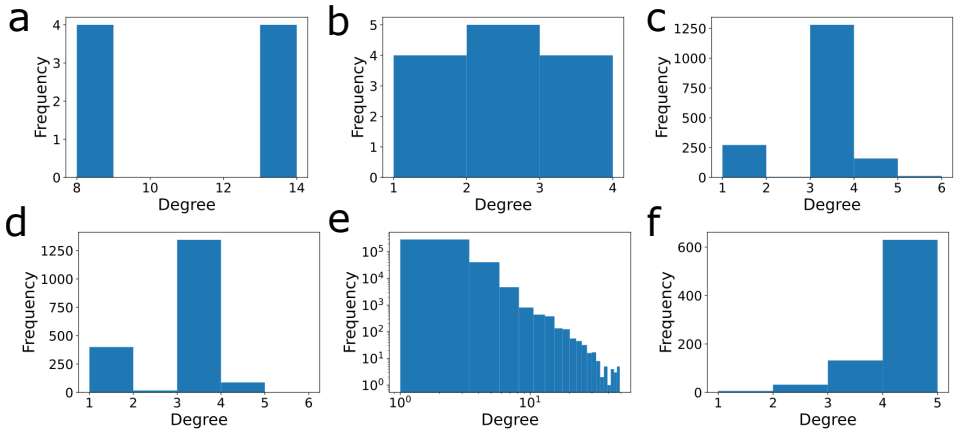

Figure 4: Degree distributions for graphs from a variety of material. (a) Unit-cell for zincblende, from CHILI (Friis-Jensen et al. (2024)); (b) Molecular graph of aspirin, from PubChem; (c) A graph from an experimental image of a nanowire network; (d) A random stick models created for this work; (e) A graph from a 3D tomographic reconstruction of aramid nanofibers; (f) A graph from the molecular simulations of polymer networks.

have very similar degree distributions (Figure 4c,d), which highlights the importance of geometric effects, as cpatured by the nematic order parameter in Figure 2o. Nanofiber network shows log-log distributions (Figure 4e), which is a commonly observed departure from the Poissonian distribution of random graphs, (and is often referred to scale-free (Barabási (2009))) While often attributed to the preferential attachment mechanism (Newman (2010)), geometric constraints mean that this doesn't apply here. The true source of this complex structure remains an open question, likely answerable by the ML community. Finally, polymer networks show defects, indicated by considerable number of nodes with degree less than four, which is the ideal network value (Zhang & Riggleman (2024)).

F Future Work

Acceleration of polymer network generation, which in current MD protocols requires costly simulations of chains and crosslinking junctions diffusing to form bonds, can be naturally framed as a link prediction task. While current models can capture some structural patterns, more fine-tuned or domain-informed approaches are needed to generate physically plausible networks with defect concentrations comparable to real materials. Limitations of this and similar directions include the data scarcity associated with the high computational cost of ground-truth data. However, this data scarcity also further motivates the development of models that generalize well on limited training data.

G Environment

For the polymer network dataset, molecular dynamics simulations were performed using the LAMMPS package (versions 09 Jan 2020 and 23 Jun 2022) Thompson et al. (2022) to generate model end-linked polymer networks and conduct fracture tests. Computational resources were provided by high-performance computing facilities under national and institutional programs.

Graph extraction and analysis for aramid nanofibers and silver nanowire network datasets were carried out on an Apple Mac mini 2020, with an Apple M1 CPU and 16 GB RAM, using a conda-forge distribution of StructuralGT 0.1.6 and Python 3.13.

The experiments for benchmark on graph ML tasks are conducted on a workstation including a 24-core CPU, 64GB RAM, and an RTX A6000 GPU. We implement the benchmarks using Python 3.10 and PyTorch 2.1 on Ubuntu-22.04.