Position: Topological Deep Learning is the New Frontier for Relational Learning

Theodore Papamarkou [1] **Tolga Birdal** [2] **Michael Bronstein** [3] **Gunnar Carlsson** [4 5] **Justin Curry** [6] **Yue Gao** [7] **Mustafa Hajij** [8] **Roland Kwitt** [9] **Pietro Liò** [10] **Paolo Di Lorenzo** [11] **Vasileios Maroulas** [12] **Nina Miolane** [13] **Farzana Nasrin** [14] **Karthikeyan Natesan Ramamurthy** [15] **Bastian Rieck** [16 17] **Simone Scardapane** [11] **Michael T. Schaub** [18] **Petar Veličković** [19 10] **Bei Wang** [20] **Yusu Wang** [21] **Guo-Wei Wei** [22] **Ghada Zamzmi** [23]

Abstract

Topological deep learning (TDL) is a rapidly evolving field that uses topological features to understand and design deep learning models. This paper posits that TDL is the new frontier for relational learning. TDL may complement graph representation learning and geometric deep learning by incorporating topological concepts, and can thus provide a natural choice for various machine learning settings. To this end, this paper discusses open problems in TDL, ranging from practical benefits to theoretical foundations. For each problem, it outlines potential solutions and future research opportunities. At the same time, this paper serves as an invitation to the scientific community to actively participate in TDL research to unlock the potential of this emerging field.

[1]Department of Mathematics, The University of Manchester, Manchester, UK. [2]Department of Computing, Imperial College London, London, UK. [3]Department of Computer Science, University of Oxford, Oxford, UK. [4]Department of Mathematics, Stanford University, Stanford, USA. [5]BlueLightAI Inc, USA. [6]University at Albany, New York, USA. [7]School of Software, Tsinghua University, Beijing, China. [8]University of San Francisco, San Francisco, USA. [9]Department of Artificial Intelligence and Human Interfaces, University of Salzburg, Austria. [10]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK. [11]Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Rome, Italy. [12]Department of Mathematics, University of Tennessee, Knoxville, USA. [13]Department of Electrical and Computer Engineering, UC Santa Barbara, Santa Barbara, USA. [14]Department of Mathematics, University of Hawai'i at Mānoa, Hawai'i, USA. [15]IBM Corporation New York, USA. [16]Helmholtz Munich, Munich Germany. [17]Technical University of Munich, Munich Germany. [18]RWTH Aachen University, Aachen, Germany. [19]Google DeepMind. [20]School of Computing, University of Utah, Utah, USA. [21]Computer Science and Engineering Department, University of California San Diego, San Diego, USA. [22]Department of Mathematics, Michigan State University, East Lansing, Michigan, USA. [23]University of South Florida, Florida, USA. Correspondence to: Theodore Papamarkou <theo.papamarkou@manchester.ac.uk>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

1. Introduction

Traditional machine learning often assumes that the observed data of interest are supported on a linear vector space and can be described by a set of feature vectors. However, there is growing awareness that, in many cases, this viewpoint is insufficient to describe several data within the real world. For example, molecules may be described more appropriately by graphs than feature vectors. Other examples include three-dimensional objects represented by meshes, as encountered in computer graphics and geometry processing, or data supported on top of a complex social network of interrelated actors. Hence, there has been an increased interest in importing concepts from geometry and topology into the usual machine learning pipelines to gain further insights into such types of data in a systematic way.

One of the most prominent examples of this research direction is geometric deep learning (GDL); see (Bronstein et al., 2021; Zhou et al., 2020; Wu et al., 2020; Nguyen & Wei, 2019). GDL aims to generalize neural networks to non-Euclidean domains, including manifolds and graphs. Graph neural networks (GNNs) form a central pillar within GDL (Wu et al., 2020; Zhou et al., 2020).

Topology is concerned with the study of properties that remain invariant under continuous deformations, and affords a powerful lens through which the global structure of data can be discerned. By characterizing topological features (including connected components, loops, and voids across multiple scales), topological tools such as persistent homology (Carlsson, 2009; Edelsbrunner & Harer, 2010) have become powerful methods to capture essential structures and patterns that elude conventional methods. The ability to quantify the topological signatures of data can enhance the robustness of traditional machine learning models, enabling them to discern a greater variety of meaningful structures in diverse and complex datasets. This use of topological ideas, which often runs under the umbrella of topological data analysis (TDA), is by now a well-established field of research. For example, persistent homology has led to victories in the D3R Grand Challenges, a worldwide annual competition series in computer-aided drug design (Nguyen

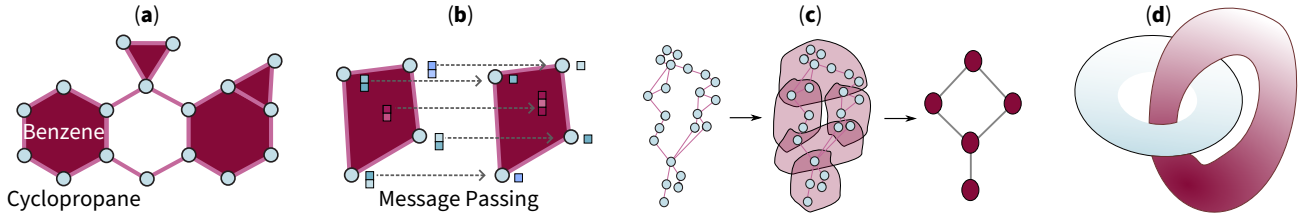

Figure 1. Topological spaces enrich deep learning methods from a variety of perspectives. (a): A topological space, modeled as a cell complex, enables a flexible molecular representation. This representation can improve the performance of a deep learning model supported on this space (Bodnar et al., 2021a). (b): Topological neural networks enable the processing of data, for example via higher-order message-passing schemes on a topological space. These networks find a wide range of applications, from computer graphics to drug discovery (Hajij et al., 2023b). (c): Topological spaces allow hierarchical representations of the underlying data that naturally correspond to pooling operations in deep learning (Hajij et al., 2023b). (d): The topological characteristics of the underlying data are crucial when selecting a neural network architecture. Data from a knotted structure in $\mathbb{R}^3$, such as the one shown, cannot be embedded in $\mathbb{R}^2$ with a single layer multilayer perceptron (MLP) from $\mathbb{R}^3$ to $\mathbb{R}^2$ (Olah, 2014).

et al., 2019; 2020). Moreover, topological descriptors such as contour trees, merge trees, Reeb graphs, mapper, Morse and Morse-Smale complexes are established tools in scientific data analysis and visualization (Banerjee et al., 2020; Hu et al., 2021b).

Various topological structures have been used as models to describe data in terms of its higher order relations, including cell complexes (Hajij et al., 2020), simplicial complexes (Schaub et al., 2022), sheaves (Hansen & Ghrist, 2019), and combinatorial complexes (Hajij et al., 2023b). Machine learning models have recently been designed to learn from data that are supported on these topological domains (Billings et al., 2019; Bunch et al., 2020; Hajij et al., 2020; Schaub et al., 2020; Roddenberry et al., 2021; 2022; Schaub et al., 2021; Giusti et al., 2023; Yang & Isufi, 2023).

Beyond using topological features to describe data, it is also possible to leverage topological features to understand and control computational models. In other words, one may use topological methods to understand the flow of information within computational models, such as neural networks, to gain insight about the models' functional behavior. Furthermore, one may design deep learning architectures according to certain topological notions to enhance or constrain the type of computations that can be performed, for example, by enforcing message-passing schemes that are compliant with certain topological invariants or promoting desirable topological properties of learned representations.

Although there is an overlap between the three ways described above to incorporate topological notions into machine learning, this paper focuses primarily on the latter two points of view, which hold great potential for machine learning applications and are not yet well established, in contrast to 'traditional' TDA (Edelsbrunner & Harer, 2010; Carlsson, 2009; Dey & Wang, 2022; Ghrist, 2014). The discussion is particularly concerned with deep learning models, due to their broad real-world adoption. So, the scope of this paper is dedicated to the nascent field of topological deep learning (TDL), highlighting open problems and identifying opportunities for future research. In this way, the paper invites the scientific community to contribute to research in TDL.

Why TDL? To articulate why TDL plays a critical role in the encoding, modeling and analysis of relational data, four practical advantages of TDL are presented. First, the topology of the underlying data space determines the choice of possible neural network architectures. Second, topological domains enable the modeling of data containing multi-way interactions (also known as higher-order relations). Third, TDL captures regularities inherent to manifolds, such as 'remeshing symmetry'. Fourth, TDL captures topological equivariances in the data. In summary, TDL takes into account topological characteristics that appear in relational data, and therefore is a natural choice for various machine learning problems. Figure 1 shows a range of examples in which topological approaches to deep learning are useful and sometimes crucial.

Choice of neural network architecture. Olah (2014) contributed one of the first works on TDL, showing that the topology of the underlying space must be considered when choosing a neural network architecture. In the early stages, the term 'TDL' was often used to refer to the incorporation of features generated by persistent homology within the input pipeline of a deep neural network (DNN; Cang & Wei, 2017; Hofer et al., 2017). However, this paper uses the term 'TDL' to refer to the collection of ideas and methods related to the use of topological concepts in deep learning. Following Hensel et al. (2021), TDL methods can be used in an observational fashion, improving the understanding of existing deep learning models and their topological formulations, or in an interventional fashion, allowing deep learning architectures to treat data supported on higher-order domains, such as cell complexes (Hajij et al., 2020; Bodnar et al., 2022; Giusti et al., 2023), simplicial complexes (Bodnar, 2022; Schaub et al., 2022; Mitchell et al., 2024), sheaves (Hansen

& Ghrist, 2019; Bodnar et al., 2021b), combinatorial complexes (Hajij et al., 2023b), and hypergraphs (Feng et al., 2019; Kim et al., 2020; Bai et al., 2021b; Behrouz et al., 2023).

Multi-way interactions. In TDL, multi-way interactions between entities constitute useful features capable of embedding topological structure via deep learning algorithms. Higher-order relations capture long-distance or seemingly disparate connections in a system, providing scope for effective or robust message-passing schemes (Hajij et al., 2020; Roddenberry et al., 2021; Bodnar, 2022; Hajij et al., 2023b; Yang & Isufi, 2023). While graph-based equivariant neural networks have been developed to incorporate certain multi-way interactions (Batatia et al., 2022; Musaelian et al., 2023), many types of multi-way interactions can arise in practice. TDL operates on various topological domains, such as simplicial and cellular complexes, and therefore provides a framework to model a multitude of possible types of multi-way interactions appearing in relational data.

Regularities inherent to manifolds. Beyond higher-order relations, the topological view may also capture the regularities inherent to manifolds. Examples include 'remeshing symmetry' over manifolds, such as being invariant to different triangulations of a sphere or inducing similar behaviors at different meshing resolutions. These regularities would be perhaps not impossible, but quite challenging, to express using purely combinatorial constructs offered by graph representation learning (GRL), while TDL is a natural domain to think about such effects.

Topological equivariances. TDL is a natural approach to capture 'topological equivariances'. For example, if a classification algorithm is meant to identify different knots, then it is useful to understand the stabilizer group involving isotopies of the complement. In general, GNNs and GDL are based on 'standard groups'. For instance, GNNs adopt the permutation group, and applications of GDL in molecular modeling use subgroups of the Euclidean group, such as the special Euclidean group $\mathrm{SE}(n)$ or the special orthogonal group $\mathrm{SO}(n)$. TDL incoporates more complex homeomorphism groups that act on a space, while trying to preserve some embedding information. This embedding information can include, for example, the arrangement of critical points in Morse-Smale complexes (Catanzaro et al., 2020) or the nesting of circles which leads to the concept of 'tree of shapes' in image processing (Caselles & Monasse, 2009). In particular, image segmentation provides an example of a topological structure that needs to be preserved under homeomorphisms of the domain.

Position. This position paper argues that TDL is the new frontier for relational learning. Relational data constitute a main modality of data emerging from natural and artificial systems (Veličković, 2023). In such systems, sets of of objects are interconnected via binary or higher-order relations, and relational data encode features of these interconnections. TDL provides a framework for learning from binary or higher-order relational data.

While TDL is pivotal in relational learning, it is more generally a valuable framework with many functions in the contemporary AI landscape. TDL informs the choice of neural network architecture subject to the topology of the underlying data space, captures regularities inherent to manifolds, and captures topological equivariances.

Topological concepts are expected to play a crucial role in articulating the position of TDL in relation to overlapping fields, such as GRL or GDL. Finding the conditions under which TDL becomes the go-to machine learning tool for practitioners is one of the main research problems in TDL. This paper builds on the engineering and mathematical knowledge of the community to identify possible research directions to solve this problem.

Paper structure. The remainder of the paper is organized into several sections, each of which raises specific open problems and research directions for TDL. The ordering of the sections starts with more concrete application-centric topics and ends with issues related to the theoretical foundations of TDL, before concluding with some final remarks. An extensive literature review of TDL is provided for the interested reader in appendix A.

2. Examples, Datasets and Benchmarks

This section highlights compelling applications and success stories of TDL in machine learning challenges to corroborate the advantages of TDL in various settings. In addition, it proposes ways to enrich the range of datasets and benchmarks that can facilitate the evaluation of TDL methods.

2.1. Examples

A natural use case of TDL involves attributed graphs, which combine structural information with feature information. Such graphs arise in numerous domains, for example, involving protein structures (Xia & Wei, 2014; Sverrisson et al., 2021), drug design (Cang & Wei, 2017), virus analysis (Chen et al., 2022), or structural representations of molecules (Jiang et al., 2021) and materials (Reiser et al., 2022; Townsend et al., 2020). However, applications that require higher-order topological structures have been limited to intrinsically complex data, such as those arising in biological sciences (Cang et al., 2018). Perhaps the most compelling examples and applications of TDL which consistently demonstrate the relevant advantages of TDL over existing methods are the victories of TDL in the D3R Grand Challenges (Nguyen et al., 2019; 2020), the discovery of SARS-CoV-2 evolution mechanisms (Chen et al.,

2020; Wang et al., 2021), and the successful forecasting of SARS-CoV-2 variants BA.2 (Chen & Wei, 2022), BA.4 and BA.5 (Chen et al., 2022) about two months in advance. More broadly, beyond graph-centric applications, other applications have benefited from a topological perspective of machine learning, including 2D shape analysis, unsupervised learning (Hofer et al., 2019), and classification (Chen et al., 2019; Hofer et al., 2020a; Curry et al., 2023).

Open problem 1. *Compelling applications of TDL have been developed, in which the topological properties of the data empirically demonstrate a competitive edge. However, a broader adaptation of TDL, which can revolutionize relational learning, has not yet taken place in real-world applications.*

Several application areas are plausible candidates for TDL to shine, since their underlying domains give rise to topological structures. To encourage the development and adaptation of TDL in practice, the presence or emergence of TDL applications in numerous scientific disciplines is elaborated further in appendix B.

Research direction 1. *Topological structures emerge in several scientific areas, including data compression, natural language processing (NLP), computer vision and computer graphics, chemistry, biological imaging, virus evolution, drug design, neuroscience, protein engineering, chip design, semantic communications, satellite imagery, and materials science. Synergies with researchers from these scientific disciplines are encouraged to develop real-world or impactful applications of TDL.*

2.2. Datasets

In addition to showcasing the practical benefits of TDL, applications can play an important role in the development and deployment of TDL models. In particular, standardized datasets derived from applications are instrumental in driving TDL research. The Open Graph Benchmark, a set of benchmark datasets, has been developed to facilitate reproducible graph machine learning research (Hu et al., 2020; 2021a). However, the community has not yet invested in resources to construct higher-order data and associated benchmarks because many graph methods do not take advantage of this information.

Open problem 2. *There is a scarcity of higher-order data. One of the major milestones for the advancement of TDL is the prolific generation of public higher-order datasets.*

Higher-order data can be collected or synthesized. The TDL application areas of subsection 2.1 are potential candidates to collect higher-order data. Conversely, a graph can be lifted to a higher-order domain (Papillon et al., 2023a). Lifting procedures for graphs supply mechanisms for the generation of synthetic topological data. This is why a survey of lifting procedures and associated message-passing schemes would be useful for TDL. Moreover, a generalization of graph-rewiring approaches and analyses to higher-order topological structures would be a fruitful direction.

Research direction 2. *Developing applications of TDL can produce higher-order datasets that naturally arise from the underlying domain. A systematic assessment and generalization of graph-lifting and rewiring algorithms is a plausible path towards synthetic higher-order datasets.*

2.3. Benchmarks

The curation of a collection of higher-order datasets can pave the way for TDL benchmarks. Benchmarking is necessary for the comparative assessment and development of novel TDL architectures and associated learning algorithms.

Open problem 3. *Benchmark suites are needed to enable efficient and objective evaluation of novel TDL research.*

The design of open source and reproducible benchmark suites for TDL requires a minimal collection of higher-order benchmark datasets, as well as implementations of graph-lifting algorithms for generating synthetic datasets in higher-order domains. To ease user experience, a taxonomy of higher-order datasets is a recommended feature, organizing benchmarks, for example, by dataset size and type of learning task. Benchmark suites for TDL are expected to have a comprehensive set of performance metrics that extend beyond predictive performance. For example, stability metrics are relevant to TDL, since it is anticipated that exploitation of topological structure in data can improve stability comparably to GNNs.

Research direction 3. *The basic components of TDL benchmark suites include higher-order datasets, graph-lifting algorithms, and predictive and stability metrics.*

3. Software

There are several graph-based learning software packages, such as `NetworkX` (Hagberg et al., 2008), `KarateClub` (Rozemberczki et al., 2020), `PyG` (Fey & Lenssen, 2019), and `DGL` (Wang et al., 2019). `NetworkX` facilitates computations on graphs, and `KarateClub` implements algorithms for unsupervised learning on graph-structured data. `PyG` and `DGL` are two geometric deep learning packages for graphs.

To the best of the authors' knowledge, four software packages provide functionality on higher-order structures, namely `HyperNetX` (Liu et al., 2021), `XGI` (Landry et al., 2023), `DHG` (Feng et al., 2019), and `TopoX` (Hajij et al., 2024). `HyperNetX` enables computations on hypergraphs, while `XGI` provides similar functionality on hypegraphs and simplicial complexes. `DHG` is a deep learning package for

graphs and hypergraphs. `TopoX` is a suite of Python packages designed to compute and learn with topological neural networks. The suite consists of three packages, `TopoNetX`, `TopoEmbedX` and `TopoModelX`. `TopoNetX` supports computations on graphs and higher-order domains, including colored hypergraphs, simplicial complexes, cell complexes, path complexes and combinatorial complexes. `TopoEmbedX` provides methods to embed higher-order domains into Euclidean domains. `TopoModelX` implements the majority of topological neural networks surveyed in Papillon et al. (2023a).

`pytorch-topological` combines several state-of-the-art packages for TDA, including `giotto-tda` (Tauzin et al., 2021) and `Ripser` (Bauer, 2021), thus enabling the creation of topology-driven algorithms that work on point clouds or structured data such as images. Similarly, `torch-ph` (Pérez et al., 2021) and `TopologyLayer` (Gabrielsson et al., 2020) support persistent homology computations, as well as differentiating through these computations to facilitate the development of topology-informed loss functions. `GUDHI` (Maria et al., 2014) provides a wider range of methods, including the optimization of functions based on persistent homology.

However, despite numerous theoretical advances in TDL, practical implementations are scarce due to the limited availability of easy-to-use software packages for deep learning on higher-order structures. TDL is a broad research area, so the contemporary TDL software landscape needs to be enriched with more functionality to cover additional higher-order domains (including domains related to higher-order multiplex networks, and dynamic versions of existing higher-order domains), graph-lifting and rewiring algorithms, TDL models, and learning algorithms (Alain et al., 2023).

Open problem 4. *Research experimentation and deployment of TDL models are often hindered by the limited availability of software. Software development for TDL is one of the most pressing open problems in accelerating the progress of theoretical and engineering research in TDL.*

To accelerate the development of TDL software, more human capital and financial investment are required. Since the currently available TDL software packages are open source, there are software engineers and researchers who volunteer to improve these packages. Coding challenges (Papillon et al., 2023a) help in this direction. However, more programmers are needed. Academic research grants are one funding avenue for computer scientists to contribute to the development of TDL software. To attract financial resources from investors, it is expected to demonstrate the value of TDL. In other words, solving open problem 1 (compelling examples of TDL) creates conditions to address open problem 4 (resources for the development of TDL software).

Transformers are popular due to their easy implementation in graphics processing units. As existing TDL methods might be more challenging to implement than transformers, software users might resort to transformers equipped with some topological encoding. This latter possibility provides an alternative direction for TDL research. It becomes apparent that understanding computational trade-offs and hardware friendliness of TDL algorithms is essential for the development and broad adaptation of TDL software. Advances towards hardware-friendly TDL implementations may include, for example, graph rewiring (Topping et al., 2022; Nguyen et al., 2023) and positional encoding (Wang et al., 2022).

Research direction 4. *The acquisition of more resources is required to advance TDL software development. Furthermore, research on the computational trade-offs and hardware friendliness of TDL algorithms might yield more accessible TDL software implementations.*

4. Complexity and Scalability

While the development of software for TDL is one problem, another imminent question is whether the benefits of TDL outweigh the resulting costs. The innovative design of TDL architectures can make a difference in practical applications. In the context of such applications, computational complexity and scalability are major challenges that become particularly prominent in TDL due to the inclusion of higher-order data. This section discusses the complexity and scalability implications of TDL.

4.1. Complexity

TDL models increase space and time complexity in comparison to GNNs, since the former operate on domains with supplementary higher-order structure. Thus, the question becomes whether the increased complexity of TDL is worthwhile.

Open problem 5. *A cost-benefit analysis framework for TDL has not been formalized. The increased time and space complexity of TDL relatively to GDL and GRL (due to the inclusion of higher-order data) must be systematically factored in to enable informed decision-making when it comes to choosing between TDL, GDL or GRL.*

An approach to understanding the role of complexity in TDL is to analyze the trade-off between complexity and performance as a function of higher-order structure. For example, the efficiency of a TDL algorithm may be evaluated by connecting the increase in predictive performance through the inclusion of higher-order features with the increase in the computational cost of neural network training.

Another approach to dissecting the notion of complexity in TDL is to study the propagation of topological information through neural network layers. For example, Naitzat et al.

(2020) have demonstrated that neural networks transform topologically complicated input data into topologically simpler forms as they pass through the layers. Petri & Leitão (2020) have expanded the work of Naitzat et al. (2020) by showing that it is not necessarily the topology of the data that is simplified, but the topology of the decision boundary. Wheeler et al. (2021) have studied how the persistence homology of data transforms as the data pass through successive layers of a deep neural network. While Naitzat et al. (2020), Petri & Leitão (2020) and Wheeler et al. (2021) do not focus on input data supported on higher-order structures, such works provide plausible ways of characterizing topological complexity, for instance, via layer-wise transformations of the underlying topology of the data.

Research direction 5. *Complexity in TDL may be formalized in terms of performance-complexity trade-offs dependent on higher-order (algebraic) structure or from the angle of propagation of topological information through neural network layers.*

4.2. Scalability

Limited software availability (open problem 4) is not the only challenge in the development of TDL. Lack of scalability impedes progress in TDL. Besides, the multitude of higher-order domains confines the interoperability of scalable TDL models and learning algorithms. For instance, if a scalable TDL model is developed for simplicial complexes, it would not be applicable to other higher-order domains, such as cell complexes, hypergraphs or path complexes.

Open problem 6. *Building scalable TDL models, which work across multiple higher-order domains, remains an open problem.*

One way of building resource-efficient TDL models might be to leverage more information from fewer data points or to develop optimization algorithms that converge provably faster on higher-order domains. Another direction, which has seen improved performance in TDA applications, involves subsampling (Moor et al., 2020; Wagner et al., 2021).

Distributed training emerges as another plausible path to scalable TDL. Message-passing schemes on graphs naturally lead to distributed learning algorithms by letting 'messages' travel through separate nodes. For example, Scardapane et al. (2021) have introduced a framework for distributed GNN training. One possibility is to extend the work of Scardapane et al. (2021) toward distributed TDL.

Research direction 6. *Extending existing results in distributed GNN training may pave the way to scalable training of TDL models.*

5. Explainability, Generalization and Fairness

This section focuses on the problems of explainability, generalization, and fairness. Such problems are common in machine learning and have been explored in other contexts, but how these can get translated in the context of TDL is largely unexplored. This section discusses how TDL can play an important role in solving these problems.

5.1. Explainability and Generalization

Understanding deep learning is a remarkably difficult task and is one of the main concerns of the community. One question is how topology can contribute to the understanding of deep learning, including explainability for neural network weights, decisions, training mechanisms, and characterization of generalization error.

Open problem 7. *Harnessing topology to address questions related to explainability and generalization in deep learning delineates a family of open problems.*

Recent results have demonstrated that the topology of data, training trajectories, neural network weights and internal representations are related to generalization error in deep learning (Rieck et al., 2019b; Hofer et al., 2020a; Birdal et al., 2021; Andreeva et al., 2023; Dupuis et al., 2023). Along these lines, a feasible research endeavor to predict the generalization error may involve training of TDL models on the computational graph, training trajectories of neural networks, or promoting desirable topological properties of internal representations that are linked to generalization. The upshot of any advances along these directions would be to draw greater attention to TDL, given the potential impact of TDL on fundamental deep learning research.

Research direction 7. *Training TDL models on the computational graph, controlling topological properties of weight trajectories traced out during stochastic optimization of neural networks, or appropriately enforcing desirable topological properties of a network's internal representation of the data may hold the keys to understand the puzzling generalization behavior in deep learning and to move a step closer to explainable AI.*

5.2. Fairness

In addition to explainability and generalization, there are other topics in GDL that have not yet been extended to TDL, such as model fairness (Spinelli et al., 2022) and adversarial attacks (Zheng et al., 2021; Zhou et al., 2023). To this end, at least two issues need to be addressed, namely the definition of valid fairness metrics and the development of graph sampling or rewiring algorithms.

Open problem 8. *TDL model fairness has not been adequately studied.*

In a graph, fairness can be quantified, for example, using homophily, which counts the ratio of inter-class edges versus intra-class edges. Telyatnikov et al. (2023) have proposed a way to extend homophily measures to hypergraphs. By adapting the work of Telyatnikov et al. (2023), the notion of homophily can be extended to higher-order features other than hyperedges. Bias in higher-order structures can be assessed by generalizing the assortative mixing coefficient (Newman, 2003) beyond graphs. Furthermore, dyadic fairness metrics for graphs, such as disparate impact (Laclau et al., 2021) and statistical parity (Rahman et al., 2019), may be extended to polyadic fairness metrics for higher-order structures.

Research direction 8. *To build fair TDL models, existing notions of fairness for graphs, such as homophily, assortative mixing and dyadic fairness, can be generalized for higher-order structures.*

6. Theoretical Foundations

This section is dedicated to the theoretical foundations of TDL. It focuses on open problems and associated research directions regarding the advantages of TDL, topological representation learning (TRL), and transformers in TDL.

6.1. Advantages of Topological Deep Learning

Why is topology relevant in deep learning? A first step in answering this question has been taken by developing empirically verifiable examples in which TDL outperforms traditional learning methods in some respect (see also open problem 1). Validating such possible empirical findings requires the establishment of theoretical answers. How can theory justify the relevance of topology in deep learning? Can theory prove (or disprove) that higher-order domains provide a richer representation than vector spaces?

When do higher-order relations become useful? Higher-order relations may not benefit every application. Is it possible to discover the context in which higher-order relations become useful (Battiston et al., 2020; Benson et al., 2021; Bick et al., 2023)? Is it possible to pinpoint higher-order relations that lead to improved downstream performance? Once discovered, can these relations be certified as semantically meaningful? Can physical or biological relevance be drawn from higher-order relations with visual tools or semantic metrics? Can higher-order relations help scientists understand phenomena that are currently obscure?

What are the advantages of TDL over GRL? One question is whether in certain cases higher-order relations are not required and the same performance might be achieved by stacking more layers in regular GNNs or by using graph transformer networks (Yun et al., 2019). This can potentially happen, but it is known that neural architecture design matters (especially in the low-data regime) with respect to capturing higher-order relations. For example, the work of Sanford et al. (2023) implies that existing transformers are efficient in modeling pairwise relations, but not three-way relations. Furthermore, TDL may help avoid some of the common pathologies that arise with deeper GNNs, such as vanishing gradients and degradation.

Open problem 9. *Theoretical foundations have not yet been adequately laid to consolidate the relative advantages of TDL. More theoretical research is needed to shed light on the relevance of topology in deep learning and contextual understanding of downstream performance gained by higher-order relations.*

Expressivity proofs for TDL. For some GNNs and datasets, it is known both theoretically and empirically that persistent homology outperforms GNNs (Horn et al., 2022). There are graphs with two connected components or two cycles that cannot be distinguished by the ordinary Weisfeiler-Lehman test for graph isomorphism and, thus, also not by message-passing GNNs. Furthermore, the increased expressivity of higher-order message passing has been proven (Bodnar et al., 2021a). Expressivity proofs provide a mechanism to unravel the relevance of topology in deep learning. For example, conditions may be sought under which DNNs or GNNs can or cannot learn certain topological invariants.

Studying properties of TDL. One way to evaluate the relevance of the topology in TDL is to study the properties of TDL. This raises the question of which properties of TDL to focus on. For instance, investigating over-smoothing and over-squashing in deep TDL models may provide useful insights for the relative advantages of TDL over GNNs. Interpreting the spectral properties of Hodge Laplacians may illuminate how message passing benefits from topological structures. It is also worth examining whether neural network parameter priors based on higher-order structures yield any gains in reliability or uncertainty quantification.

Generative TDL. Generative modeling aims to learn the distribution of training samples by generating new instances. While image generation is now well-studied (Rombach et al., 2022), generating real-world data supported on non-Euclidean or mixed domains presents an unresolved challenge. This has led to research on the generation of 3D point clouds (Zhao et al., 2019; Zeng et al., 2022), meshes (Liu et al., 2023b; Poole et al., 2022), and graphs (Guo & Zhao, 2022; Vignac et al., 2023). Yet, there is no unifying representation that can generate all these 3D representations. This is one of the premises of TDL. In the field of structural biology, the generation of molecules and proteins is currently approached as graph generation (Vignac et al., 2023; Hoogeboom et al., 2022; Ingraham et al., 2019; Zhang et al., 2023) or by generative modeling on 3D rigid bodies (Yim et al., 2023). Generative modeling for graphs generally

involves creating new graphs from scratch. However, for many applications where data have inherent geometric structures, relying solely on graphs, as done in Veličković (2023), can potentially omit useful higher-order information. Recent studies have shown that considering higher-order relations yields more accurate representations of both scene graphs and molecular structures (Zhan et al., 2022; Watson et al., 2023). In addition to 2D and 3D computer vision, diffusion models have become the state-of-the-art in graph-based molecular generation (Xu et al., 2022; Vignac et al., 2023; Haefeli et al., 2022; Jing et al., 2023). In particular, Jing et al. (2023) have introduced torsional diffusion, a hybrid method that uses topological principles to diffuse in a toroidal space, narrowing the search space and improving both model performance and inference speed. This approach demonstrates how topology can enhance existing generative methods. Despite these advances, the development of topological diffusion and flow-matching models remains an uncharted territory. Data characterized by higher-order structures align naturally with TDL. In some cases, such structures are not inherently present, but can be inferred (Tang et al., 2023a;b), raising questions about the nature and implications of generating higher-order features absent from the original data. The assignment of specific semantic meanings and structures to these generated forms is an ongoing research challenge, paving the way for more interpretable TDL. This area of research, both recent and progressive, holds the potential to advance generative TDL.

Sheaf theory for TDL. There has been some work on categorical machine learning (Fong et al., 2021), as well as TDA to study activation layers (Naitzat et al., 2020), but there has not been much work connecting the two. Category theory offers rich structures for knowledge representation, but these structures are typically quite rigid. Persistent homology is less committal and allows algebraic structure to come into existence as a function of some metrized parameter space. Sheaf theory can serve as a bridge between category theory and persistent homology in the context of TDL. For example, a combination of applied category theory and TDA may lead to algorithms that automatically summarize how neural networks produce their output, making deep learning more understandable to humans. Moreover, Fong et al. (2021) have introduced a theoretical sheaf approach to detect local merging relations in digital images, and Bodnar et al. (2022) have developed neural sheaf diffusions to identify topological data features that are important for real-world problems. Persistent sheaf Laplacians allow the incorporation of physical laws in the topological representation of data (Wei & Wei, 2021). The neural sheaf diffusion models of Bodnar et al. (2022) address some limitations of classical graph diffusion equations and the corresponding GNNs, providing ideas that can be extended to push the limits of TDL through algebraic topology.

Research direction 9. *Expressivity proofs may offer a comparative characterization of topological invariance in GNNs and TDL models. Properties of TDL models, such as over-smoothing and over-squashing, and spectral properties of Hodge Laplacians may help demystify the relevance of topology in TDL. Sheaf theory for generative TDL is a plausible research avenue to demonstrate how twisting of feature spaces in a GNN affects expressivity and detects inconsistent logic present in large language models.*

6.2. Topological Representation Learning

There are several possible interpretations of the umbrella term TRL. As a high-level definition, TRL is concerned with automatically learning and exploiting multiscale topological descriptions of the data and of the neural network's internal embeddings, which scale efficiently to larger dimensions and which provide benefits in efficiency and transparency compared to the equivalent graph-based representations.

One question in TRL is how to develop training strategies that enable a neural network to elicit higher-order information from the data for the task at hand. An immediate follow-up question in this context is whether this uncovered higher-order structure is meaningful in any semantically valid sense. In the context of GNNs, it is, for example, unclear whether such an uncovered higher-order structure might not be equivalent to the message-propagation phase in the spirit of Veličković (2022). However, even if this is true, it might still be possible to analyze these structures to improve the explainability and transparency of a neural network.

Another question in TRL is how to develop methods that promote or penalize certain topological properties of a neural network's internal representation of the data or its output. On the one hand, such properties could be informed by a priori knowledge or corresponding properties in the input domain, as successfully demonstrated in image segmentation problems (Hu et al., 2021b; Gupta et al., 2023), learning representations to synthesize new shapes (Waibel et al., 2022), using random walks on complexes to obtain cell embeddings (Billings et al., 2019; Schaub et al., 2020), or preserving topological properties while training autoencoders (Moor et al., 2020; Hajij et al., 2020; Trofimov et al., 2023). On the other hand, topological characteristics of representations can be guided by establishing provably beneficial properties for generalization (Hofer et al., 2019; 2020a) or can be used for comparing neural network representations (Barannikov et al., 2022). Another facet of this interpretation of representation learning involves the analysis and modification of the decision boundaries of classifiers. Here, Chen et al. (2019) have shown that regularizers based on persistent homology lead to better generalization and higher predictive performance.

Open problem 10. *It remains unclear how to automate the learning of topological representations in TDL, that is, how to obtain semantically meaningful topological information from the data during neural network training. It is an open problem as to what properties of representations to promote or penalize when interested in generalization questions in TDL. Further, it is often far from straightforward how to appropriately encode desirable properties of the input space, not least due to the complexity and possibly high-dimensionality of the data.*

The explainability of predictions made by GNNs (Luo et al., 2020; Lucic et al., 2022) is complicated by the fact that even a small subgraph is difficult for a human to visualize and grasp quickly. Such complications are amplified in TDL, since visualization and human apprehension of higher-order domains is virtually nonexistent to date. Nevertheless, information in a higher-order domain can be summarized more flexibly by a smaller set of cells or hyperedges, leading to potentially interpretable structure. In other words, the richer structure attained by higher-order relations may enable more effective representation learning. For example, the work of Battiloro et al. (2023b) on latent topological inference (LTI) has introduced a trainable method to identify higher-order interactions between entities, combining advances in TDL and differentiable sampling. Building upon LTI might lead to a general framework for TRL on arbitrary higher-order domains. Building on the graph-based representation learning approach of Hamilton et al. (2017), the work of Hajij et al. (2022b; 2023b) has introduced a TRL framework for neighbourhood-based methods, which provides scope for topological neural network architecture search.

Research direction 10. *Higher-order relations, encoded by cells or hyperdeges, offer candidate components for effective TRL. Latent topological inference outlines a tentative methodological path toward generalized representation learning for TDL. Existing work on TRL provides a means to develop topological neural network architecture search algorithms.*

6.3. Transformers in Topological Deep Learning

Looking for a 'transformer architecture' for graphs has led to graph transformer models (Dwivedi & Bresson, 2020; Min et al., 2022), along with the first 'foundation models' for graph datasets (Liu et al., 2023a), which can be transferred (either in a zero-shot fashion or with fine-tuning) to smaller tasks. It is not yet known what the 'transformer architecture' of TDL models is, if it exists. It is expected that over the next years new transformer models for topological domains may be defined, along with algorithms such as diffusion generative algorithms or structured state-space models for cell and simplicial complexes.

Conversely, it is interesting to explore whether higher-order variants of transformers can be used to improve tasks in mainstream domains, if computational bottlenecks can be solved. For example, is there any meaningful latent information embedded in text or images that can be learned and leveraged by a higher-order transformer?

Open problem 11. *Developing a transformer architecture for TDL models would lay a unified foundation for TDL across different higher-order domains. A follow-up question would be the applicability of such a 'topological transformer' to text, audio, or imaging data.*

Advective diffusion transformers have been introduced by Wu et al. (2023) to generalize GNNs in the presence of varying graph topologies. A first step in extending the work of Wu et al. (2023) might be to consider diffusion transformers inspired by diffusions on hypergraphs (Prokopchik et al., 2022; Wang et al., 2023) or heat diffusion on simplicial complexes (Aktas & Akbas, 2021).

Research direction 11. *A first milestone for a transformer architecture for TDL might be a transformer architecture that is specific to a given higher-order domain. For instance, a transformer might be inspired by existing diffusions on hypergraphs or simplicial complexes.*

7. Final Remarks

To accelerate the transition of TDL from theoretical constructs to real-world applications, the research community is invited to adopt a multidisciplinary approach rooted in innovation, collaboration, and open science. Such an approach necessitates the convergence of diverse fields, bringing together experts in mathematics, computer science, and machine learning, to foster a rich cross-pollinating environment. By encouraging open dialogue and the exchange of ideas, TDL can evolve rapidly, breaking through current barriers, and sparking groundbreaking advancements.

One way to facilitate this interdisciplinary approach is through the organization of TDL workshops and conferences. These gatherings are the key to creating platforms for knowledge sharing, discussion, and the initiation of collaborative ventures for TDL. In these events, it is important to adopt open science principles.

This paper can help guide collaborative efforts and research discussions on TDL. It highlights key challenges and outlines actionable research strategies that can pave the way for concerted and effective efforts in the field. By openly sharing ideas and research directions in this paper, the goal is to support collaborative efforts, promote transparency, and accelerate innovation in TDL. This culture of openness allows researchers from various fields to contribute and benefit from the advances in TDL.

Acknowledgements

This work is partially supported by DOE grant DE-SC0023157 (BW), NIH grants R01AI164266 (GWW) and R35GM148196 (GWW), and NSF grants 2134241 (NM, MH), DMS-2134223 (BW), IIS-2205418 (BW), CCF-2112665 (YW), CCF-2310411 (YW), DMS2052983 (GWW) and IIS-1900473 (GWW). This work was partially funded by the Italian NRRP (PE00000001 - program 'RESTART'), and by the 6G-GOALS project under the 6G SNS-JU Horizon program, n.101139232 (PDL). VM is partially funded by NSF DMS No 2012609, US Army Research Lab Contract No W911NF-21-2-0186, US Army Research Lab Contract No W911NF-22-2-0143, US Army Research Office No W911NF-21-1-0094, and The University of Tennessee Materials Research Science & Engineering Center – The Center for Advanced Materials and Manufacturing – NSF DMR No. 2309083. BR is supported by the Bavarian state government with funds from the High-tech Agenda Bavaria. SS is partly funded by the Sapienza grant RM1221816BD028D6 (DESMOS). MTS acknowledges funding by the European Union (ERC, HIGH-HOPeS, 101039827). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

The authors thank Cristian Bodnar for his input on the theoretical foundations of TDL, and the ICML reviewers, Thomas Kipf and Charles Blundell for their reviews and feedback.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of this work, none of which the authors feel must be specifically highlighted here.

References

Acosta, F., Sanborn, S., Duc, K. D., Madhav, M., and Miolane, N. Quantifying extrinsic curvature in neural manifolds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 610–619, 2023.

Ahmed, E., Saint, A., Shabayek, A. E. R., Cherenkova, K., Das, R., Gusev, G., Aouada, D., and Ottersten, B. A survey on deep learning advances on different 3D data representations. *arXiv preprint arXiv:1808.01462*, 2018.

Aktas, M. E. and Akbas, E. Graph classification via heat diffusion on simplicial complexes. *IEEE Access*, 9:12291–12300, 2021.

Alain, M., Takao, S., Paige, B., and Deisenroth, M. P. Gaussian processes on cellular complexes. *arXiv preprint arXiv:2311.01198*, 2023.

Ameneyro, B., Maroulas, V., and Siopsis, G. Quantum persistent homology. arXiv:2202.12965, 2022a.

Ameneyro, B., Siopsis, G., and Maroulas, V. Quantum persistent homology for time series. In *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*, pp. 387–392, 2022b.

Andreeva, R., Limbeck, K., Rieck, B., and Sarkar, R. Metric space magnitude and generalisation in neural networks. In *Proceedings of the 2nd Annual Workshop on Topology, Algebra, and Geometry in Machine Learning (TAG-ML)*, number 221 in Proceedings of Machine Learning Research, pp. 242–253. PMLR, 2023.

Arya, D. and Worring, M. Exploiting relational information in social networks using geometric deep learning on hypergraphs. In *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, pp. 117–125, 2018.

Attene, M., Biasotti, S., and Spagnuolo, M. Shape understanding by contour-driven retiling. *The Visual Computer*, 19(2):127–138, 2003.

Bai, J., Gong, B., Zhao, Y., Lei, F., Yan, C., and Gao, Y. Multi-scale representation learning on hypergraph for 3D shape retrieval and recognition. *IEEE Transactions on Image Processing*, 30:5327–5338, 2021a.

Bai, S., Zhang, F., and Torr, P. H. S. Hypergraph convolution and hypergraph attention. *Pattern Recognition*, 110: 107637, 2021b.

Bajaj, C. L., Pascucci, V., and Schikore, D. R. The contour spectrum. In *Proceedings of the 8th Conference on Visualization '97*, pp. 167–ff. IEEE Computer Society Press, 1997.

Banerjee, S., Magee, L., Wang, D., Li, X., Huo, B., Jayakumar, J., Matho, K., Lin, M., Ram, K., Sivaprakasam, M., Huang, J., Wang, Y., and Mitra, P. Semantic segmentation of microscopic neuroanatomical data by combining topological priors with encoder-decoder deep networks. *Nature Machine Intelligence*, 2:585–594, 2020.

Barannikov, S., Trofimov, I., Balabin, N., and Burnaev, E. Representation topology divergence: a method for comparing neural network representations. In *ICML*, 2022.

Barbarossa, S. and Sardellitti, S. Topological signal processing over simplicial complexes. *IEEE Transactions on Signal Processing*, 68:2992–3007, 2020a.

Barbarossa, S. and Sardellitti, S. Topological signal processing: making sense of data building on multiway relations. *IEEE Signal Processing Magazine*, 37(6):174–183, 2020b.

Barbarossa, S., Sardellitti, S., and Ceci, E. Learning from signals defined over simplicial complexes. In *2018 IEEE Data Science Workshop (DSW)*, pp. 51–55. IEEE, 2018.

Barbarossa, S., Comminiello, D., Grassucci, E., Pezone, F., Sardellitti, S., and Di Lorenzo, P. Semantic communications based on adaptive generative models and information bottleneck. *IEEE Communications Magazine*, 61 (11):36–41, 2023.

Batatia, I., Kovacs, D. P., Simm, G., Ortner, C., and Csanyi, G. MACE: higher order equivariant message passing neural networks for fast and accurate force fields. In *NeurIPS*, volume 35, pp. 11423–11436, 2022.

Battiloro, C., Di Lorenzo, P., and Ribeiro, A. Parametric dictionary learning for topological signal representation. In *2023 31st European Signal Processing Conference (EUSIPCO)*, pp. 1958–1962. IEEE, 2023a.

Battiloro, C., Spinelli, I., Telyatnikov, L., Bronstein, M., Scardapane, S., and Di Lorenzo, P. From latent graph to latent topology inference: differentiable cell complex module. *arXiv preprint arXiv:2305.16174*, 2023b.

Battiloro, C., Testa, L., Giusti, L., Sardellitti, S., Di Lorenzo, P., and Barbarossa, S. Generalized simplicial attention neural networks. *arXiv preprint arXiv:2309.02138*, 2023c.

Battiloro, C., Wang, Z., Riess, H., Di Lorenzo, P., and Ribeiro, A. Tangent bundle convolutional learning: from manifolds to cellular sheaves and back. *arXiv preprint arXiv:2303.11323*, 2023d.

Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.-G., and Petri, G. Networks beyond pairwise interactions: structure and dynamics. *Physics Reports*, 874:1–92, 2020.

Bauer, U. Ripser: efficient computation of Vietoris–Rips persistence barcodes. *Journal of Applied and Computational Topology*, 5(3):391–423, 2021.

Bazi, Y., Bashmal, L., Rahhal, M. M. A., Dayil, R. A., and Ajlan, N. A. Vision transformers for remote sensing image classification. *Remote Sensing*, 13(3), 2021.

Behrouz, A., Hashemi, F., Sadeghian, S., and Seltzer, M. CAT-Walk: inductive hypergraph learning via set walks. In *NeurIPS*, 2023.

Bengio, Y., Courville, A., and Vincent, P. Representation learning: a review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8): 1798–1828, 2013.

Benson, A. R., Gleich, D. F., and Higham, D. J. Higher-order network analysis takes off, fueled by classical ideas and new data. *arXiv preprint arXiv:2103.05031*, 2021.

Bernárdez, G., Telyatnikov, L., Alarcón, E., Cabellos-Aparicio, A., Barlet-Ros, P., and Liò, P. Topological network traffic compression. In *Proceedings of the 2nd on Graph Neural Networking Workshop 2023*, CoNEXT 2023. ACM, 2023.

Bianchini, M. and Scarselli, F. On the complexity of neural network classifiers: a comparison between shallow and deep architectures. *IEEE Trans. Neural Netw. Learn. Syst.*, 25(8):1533–1565, 2014.

Biasotti, S., De Floriani, L., Falcidieno, B., Frosini, P., Giorgi, D., Landi, C., Papaleo, L., and Spagnuolo, M. Describing shapes by geometrical-topological properties of real functions. *ACM Computing Surveys (CSUR)*, 40 (4):12, 2008.

Bick, C., Gross, E., Harrington, H. A., and Schaub, M. T. What are higher-order networks? *SIAM Review*, 65(3): 686–731, 2023.

Billings, J. C. W., Hu, M., Lerda, G., Medvedev, A. N., Mottes, F., Onicas, A., Santoro, A., and Petri, G. Simplex2Vec embeddings for community detection in simplicial complexes. *arXiv preprint arXiv:1906.09068*, 2019.

Birdal, T., Lou, A., Guibas, L. J., and Simsekli, U. Intrinsic dimension, persistent homology and generalization in neural networks. In *Advances in Neural Information Processing Systems*, volume 34, pp. 6776–6789. Curran Associates, Inc., 2021.

Biswas, A., Ranjan, D., Zubair, M., Zeil, S., Al Nasr, K., and He, J. An effective computational method incorporating multiple secondary structure predictions in topology determination for cryo-EM images. *IEEE/ACM transactions on computational biology and bioinformatics*, 14 (3):578–586, 2016.

Bodnar, C. *Topological deep learning: graphs, complexes, sheaves*. PhD thesis, Apollo - University of Cambridge Repository, 2022.

Bodnar, C., Frasca, F., Otter, N., Wang, Y., Lio, P., Montufar, G. F., and Bronstein, M. Weisfeiler and Lehman go cellular: CW networks. *Advances in Neural Information Processing Systems*, 34:2625–2640, 2021a.

Bodnar, C., Frasca, F., Wang, Y., Otter, N., Montufar, G. F., Lio, P., and Bronstein, M. Weisfeiler and Lehman go topological: message passing simplicial networks. In *Int. Conf. Mach. Learn.*, pp. 1026–1037. PMLR, 2021b.

Bodnar, C., Di Giovanni, F., Chamberlain, B., Lió, P., and Bronstein, M. Neural sheaf diffusion: a topological perspective on heterophily and oversmoothing in GNNs. In *Advances in Neural Information Processing Systems*, volume 35, pp. 18527–18541. Curran Associates, Inc., 2022.

Boyell, R. L. and Ruston, H. Hybrid techniques for real-time radar simulation. In *Proceedings of the November 12-14, 1963, Fall Joint Computer Conference*, pp. 445–458. ACM, 1963.

Bronstein, M. M., Bruna, J., Cohen, T., and Veličković, P. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

Bunch, E., You, Q., Fung, G., and Singh, V. Simplicial 2-complex convolutional neural nets. *NeurIPS Workshop on Topological Data Analysis and Beyond*, 2020.

Calmon, L., Schaub, M. T., and Bianconi, G. Higher-order signal processing with the Dirac operator. In *56th Asilomar Conference on Signals, Systems, and Computers*, pp. 925–929. IEEE, 2022.

Cang, Z. and Wei, G.-W. TopologyNet: Topology based deep convolutional and multi-task neural networks for biomolecular property predictions. *PLoS computational biology*, 13(7):e1005690, 2017.

Cang, Z., Mu, L., and Wei, G.-W. Representability of algebraic topology for biomolecules in machine learning based scoring and virtual screening. *PLoS computational biology*, 14(1):e1005929, 2018.

Carlsson, G. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

Carlsson, G., Zomorodian, A., Collins, A., and Guibas, L. J. Persistence barcodes for shapes. *International Journal of Shape Modeling*, 11(02):149–187, 2005.

Carr, H., Snoeyink, J., and van de Panne, M. Simplifying flexible isosurfaces using local geometric measures. In *IEEE Visualization*, pp. 497–504. IEEE, 2004.

Caselles, V. and Monasse, P. *Geometric description of images as topographic maps*. Springer, 2009.

Catanzaro, M. J., Curry, J. M., Fasy, B. T., Lazovskis, J., Malen, G., Riess, H., Wang, B., and Zabka, M. Moduli spaces of morse functions for persistence. *Journal of Applied and Computational Topology*, 4(3):353–385, 2020.

Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A., and Fiete, I. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature neuroscience*, 22(9):1512–1520, 2019.

Chen, C., Ni, X., Bai, Q., and Wang, Y. A topological regularizer for classifiers via persistent homology. In *AISTATS*, 2019.

Chen, J. and Wei, G.-W. Omicron BA. 2 (B. 1.1. 529.2): high potential for becoming the next dominant variant. *The journal of physical chemistry letters*, 13(17):3840–3849, 2022.

Chen, J., Wang, R., Wang, M., and Wei, G.-W. Mutations strengthened SARS-CoV-2 infectivity. *Journal of molecular biology*, 432(19):5212–5226, 2020.

Chen, J., Qiu, Y., Wang, R., and Wei, G.-W. Persistent Laplacian projected Omicron BA. 4 and BA. 5 to become new dominating variants. *Computers in Biology and Medicine*, 151:106262, 2022.

Chen, J., Woldring, D. R., Huang, F., Huang, X., and Wei, G.-W. Topological deep learning based deep mutational scanning. *Computers in Biology and Medicine*, 164: 107258, 2023.

Clough, J. R., Byrne, N., Oksuz, I., Zimmer, V. A., Schnabel, J. A., and King, A. P. A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 44(12):8766–8778, 2022.

Curry, J., Mio, W., Needham, T., Okutan, O. B., and Russold, F. Topologically attributed graphs for shape discrimination. In *Topological, Algebraic and Geometric Learning Workshops 2023*, pp. 87–101. PMLR, 2023.

Curto, C. What can topology tell us about the neural code? *Bulletin of the American Mathematical Society*, 54(1): 63–78, 2017.

Dabaghian, Y., Mémoli, F., Frank, L., and Carlsson, G. A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Computational Biology*, 8(8):e1002581, 2012.

Dey, T., Wang, J., and Wang, Y. Road network reconstruction from satellite images with machine learning supported by topological methods. In *Proc. 27th ACM SIGSPATIAL Intl. Conf. Adv. Geographic Information Systems (GIS)*, pp. 520–523, 2019.

Dey, T. K. and Wang, Y. *Computational topology for data analysis*. Cambridge University Press, 2022.

Dupuis, B., Deligiannidis, G., and Şimşekli, U. Generalization bounds with data-dependent fractal dimensions. In *ICML*, 2023.

Dwivedi, V. P. and Bresson, X. A generalization of transformer networks to graphs. *arXiv preprint arXiv:2012.09699*, 2020.

Ebli, S., Defferrard, M., and Spreemann, G. Simplicial neural networks. *NeurIPS Workshop on Topological Data Analysis and Beyond*, 2020.

Edelsbrunner, H. and Harer, J. *Computational topology: an introduction*. American Mathematical Society, Providence, RI, USA, 2010.

Feng, Y., You, H., Zhang, Z., Ji, R., and Gao, Y. Hypergraph neural networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):3558–3565, 2019.

Fey, M. and Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. In *ICLR Workshop on Representation Learning on Graphs and Manifolds*, 2019.

Fong, B., Spivak, D., and Tuyéras, R. Backprop as functor: a compositional perspective on supervised learning. In *Proceedings of the 34th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '19. IEEE Press, 2021.

Gabella, M., Afambo, N., Ebli, S., and Spreemann, G. Topology of learning in artificial neural networks. arXiv:1902.08160, 2019.

Gabrielsson, R. B., Nelson, B. J., Dwaraknath, A., and Skraba, P. A topology layer for machine learning. In *AISTATS*, 2020.

Gao, Y., Zhang, Z., Lin, H., Zhao, X., Du, S., and Zou, C. Hypergraph learning: methods and practices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

Gao, Y., Feng, Y., Ji, S., and Ji, R. HGNN+: general hypergraph neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

Gardner, R. J., Hermansen, E., Pachitariu, M., Burak, Y., Baas, N. A., Dunn, B. A., Moser, M.-B., and Moser, E. I. Toroidal topology of population activity in grid cells. *Nature*, 602(7895):123–128, 2022.

Georgiev, D., Brockschmidt, M., and Allamanis, M. HEAT: hyperedge attention networks. *Transactions on Machine Learning Research*, 2022. ISSN 2835-8856.

Ghose, A., Zhang, V., Zhang, Y., Li, D., Liu, W., and Coates, M. Generalizable cross-graph embedding for GNN-based congestion prediction. In *IEEE/ACM Int. Conf. Computer Aided Design, ICCAD*, pp. 1–9, 2021.

Ghrist, R. W. *Elementary applied topology*, volume 1. Createspace Seattle, 2014.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *Int. Conf. Mach. Learn.*, pp. 1263–1272. PMLR, 2017.

Girrbach, L., Christensen, A., Winther, O., Akata, Z., and Koepke, A. S. Addressing caveats of neural persistence with deep graph persistence. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856.

Giusti, C., Ghrist, R., and Bassett, D. S. Two's company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data. *Journal of Computational Neuroscience*, 41:1, 2016.

Giusti, L., Battiloro, C., Di Lorenzo, P., Sardellitti, S., and Barbarossa, S. Simplicial attention neural networks. *arXiv preprint arXiv:2203.07485*, 2022.

Giusti, L., Battiloro, C., Testa, L., Di Lorenzo, P., Sardellitti, S., and Barbarossa, S. Cell attention networks. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2023.

Goh, C. W. J., Bodnar, C., and Lio, P. Simplicial attention networks. In *ICLR 2022 Workshop on Geometrical and Topological Representation Learning*, 2022.

Gong, X., Higham, D. J., and Zygalakis, K. Generative hypergraph models and spectral embedding. *Scientific Reports*, 13(1):540, 2023.

Guo, X. and Zhao, L. A systematic survey on deep generative models for graph generation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5370–5390, 2022.

Gupta, S., Hu, X., Kaan, J., Jin, M., Mpoy, M., Chung, K., Singh, G., Saltz, M., Kurc, T., Saltz, J., Tassiopoulos, A., Prasanna, P., and Chen, C. Learning topological interactions for multi-class medical image segmentation. In *Computer Vision – ECCV 2022*, pp. 701–718, Cham, 2022. Springer Nature Switzerland.

Gupta, S., Zhang, Y., Hu, X., Prasanna, P., and Chen, C. Topology-aware uncertainty for image segmentation. In *NeurIPS*, 2023.

Haefeli, K. K., Martinkus, K., Perraudin, N., and Wattenhofer, R. Diffusion models for graphs benefit from discrete state spaces, 2022.

Hagberg, A., Swart, P., and S Chult, D. Exploring network structure, dynamics, and function using NetworkX. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

Hajij, M., Istvan, K., and Zamzmi, G. Cell complex neural networks. *NeurIPS 2020 Workshop TDA and Beyond*, 2020.

Hajij, M., Ramamurthy, K. N., Saenz, A. G., and Istvan, K. Topological deep learning. *NeurIPS Data-Centric AI Workshop*, 2021a.

Hajij, M., Zamzmi, G., Ramamurthy, K. N., and Saenz, A. G. Data-centric AI requires rethinking data notion. *NeurIPS Data-Centric AI Workshop*, 2021b.

Hajij, M., Ramamurthy, K. N., Guzmán-Sáenz, A., and Za, G. High skip networks: a higher order generalization of skip connections. In *ICLR 2022 Workshop on Geometrical and Topological Representation Learning*, 2022a.

Hajij, M., Zamzmi, G., Papamarkou, T., Maroulas, V., and Cai, X. Simplicial complex representation learning. In *Machine Learning on Graphs (MLoG) Workshop at 15th ACM International WSDM (2022) Conference, WSDM2022-MLoG*, January 2022b.

Hajij, M., Zamzmi, G., Papamarkou, T., Guzmán-Sáenz, A., Birdal, T., and Schaub, M. T. Combinatorial complexes: bridging the gap between cell complexes and hypergraphs. *arXiv preprint arXiv:2312.09504*, 2023a.

Hajij, M., Zamzmi, G., Papamarkou, T., Miolane, N., Guzmán-Sáenz, A., Ramamurthy, K. N., Birdal, T., Dey, T. K., Mukherjee, S., Samaga, S. N., Livesay, N., Walters, R., Rosen, P., and Schaub, M. T. Topological deep learning: going beyond graph data. *arXiv*, 2023b.

Hajij, M., Papillon, M., Frantzen, F., Agerberg, J., AlJabea, I., Ballester, R., Battiloro, C., Bernárdez, G., Birdal, T., Brent, A., Chin, P., Escalera, S., Fiorellino, S., Gardaa, O. H., Gopalakrishnan, G., Govil, D., Hoppe, J., Karri, M. R., Khouja, J., Lecha, M., Livesay, N., Meißner, J., Mukherjee, S., Nikitin, A., Papamarkou, T., Prílepok, J., Ramamurthy, K. N., Rosen, P., Guzmán-Sáenz, A., Salatiello, A., Samaga, S. N., Scardapane, S., Schaub, M. T., Scofano, L., Spinelli, I., Telyatnikov, L., Truong, Q., Walters, R., Yang, M., Zaghen, O., Zamzmi, G., Zia, A., and Miolane, N. TopoX: a suite of Python packages for machine learning on topological domains. *arXiv preprint arXiv:2402.02441*, 2024.

Hamilton, W. L., Ying, R., and Leskovec, J. Representation learning on graphs: methods and applications. *IEEE Data Eng. Bull.*, 40(3):52–74, 2017.

Hanocka, R., Hertz, A., Fish, N., Giryes, R., Fleishman, S., and Cohen-Or, D. MeshCNN: a network with an edge. *ACM Trans. Graph.*, 38(4):1–12, 2019.

Hansen, J. and Ghrist, R. Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology*, 3(4):315–358, 2019.

Hatcher, A. *Algebraic topology*. Cambridge University Press, 2005.

Hayhoe, M., Riess, H., Preciado, V. M., and Ribeiro, A. Stable and transferable hyper-graph neural networks. *arXiv preprint arXiv:2211.06513*, 2022.

He, S., Li, S., Jiang, S., and Jiang, W. HMSM-Net: hierarchical multi-scale matching network for disparity estimation of high-resolution satellite stereo images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188: 314–330, 2022.

Hensel, F., Moor, M., and Rieck, B. A survey of topological machine learning methods. *Frontiers in Artificial Intelligence*, 4:681108, 2021.

Hoef, L. V., Adams, H., King, E. J., and Ebert-Uphoff, I. A primer on topological data analysis to support image analysis tasks in environmental science. *Artificial Intelligence for the Earth Systems*, 2(1):e220039, 2023.

Hofer, C., Kwitt, R., Niethammer, M., and Uhl, A. Deep learning with topological signatures. In *Adv. Neural Inform. Process. Syst.*, pp. 1634–1644, 2017.

Hofer, C. D., Kwitt, R., Niethammer, M., and Dixit, M. Connectivity-optimized representation learning via persistent homology. In *ICML*, 2019.

Hofer, C. D., Graf, F., Niethammer, M., and Kwitt, R. Topologically densified distributions. In *ICML*, 2020a.

Hofer, C. D., Graf, F., Rieck, B., Niethammer, M., and Kwitt, R. Graph filtration learning. In *Proceedings of the 37th International Conference on Machine Learning*, number 119 in Proceedings of Machine Learning Research, pp. 4314–4323. PMLR, 2020b.

Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887. PMLR, 2022.

Horn, M., De Brouwer, E., Moor, M., Moreau, Y., Rieck, B., and Borgwardt, K. Topological graph neural networks. In *International Conference on Learning Representations*, 2022.

Horoi, S., Huang, J., Rieck, B., Lajoie, G., Wolf, G., and Krishnaswamy, S. Exploring the geometry and topology of neural network loss landscapes. In *Advances in Intelligent Data Analysis XX*, pp. 171–184, Cham, Switzerland, 2022. Springer.

Hu, W., Fey, M., Zitnik, M., Dong, Y., Ren, H., Liu, B., Catasta, M., and Leskovec, J. Open graph benchmark: datasets for machine learning on graphs. In *Advances in*

Neural Information Processing Systems, volume 33, pp. 22118–22133. Curran Associates, Inc., 2020.

Hu, W., Fey, M., Ren, H., Nakata, M., Dong, Y., and Leskovec, J. OGB-LSC: a large-scale challenge for machine learning on graphs. *arXiv preprint arXiv:2103.09430*, 2021a.

Hu, X., Li, F., Samaras, D., and Chen, C. Topology-preserving deep image segmentation. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Hu, X., Wang, Y., Fuxin, L., Samaras, D., and Chen, C. Topology-aware segmentation using discrete Morse theory. In *ICLR*, 2021b.

Hu, X., Lin, X., Cogswell, M., Yao, Y., Jha, S., and Chen, C. Trigger hunting with a topological prior for Trojan detection. In *International Conference on Learning Representations*, 2022.

Huang, J. and Yang, J. UniGNN: a unified framework for graph and hypergraph neural networks. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI*, 2021.

Ingraham, J., Garg, V., Barzilay, R., and Jaakkola, T. Generative models for graph-based protein design. *Advances in neural information processing systems*, 32, 2019.

Jiang, C., Lansigan, D., Marcus, P., Nießner, M., et al. Ddsl: Deep differentiable simplex layer for learning geometric signals. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8769–8778, 2019a.

Jiang, J., Wei, Y., Feng, Y., Cao, J., and Gao, Y. Dynamic hypergraph neural networks. In *IJCAI*, pp. 2635–2641, 2019b.

Jiang, X., Lim, L.-H., Yao, Y., and Ye, Y. Statistical ranking and combinatorial hodge theory. *Mathematical Programming*, 127(1):203–244, 2011.

Jiang, Y., Chen, D., Chen, X., Li, T., Wei, G.-W., and Pan, F. Topological representations of crystalline compounds for the machine-learning prediction of materials properties. *npj computational materials*, 7(1):28, 2021.

Jing, B., Corso, G., Chang, J., Barzilay, R., and Jaakkola, T. Torsional diffusion for molecular conformer generation, February 2023.

Jones, G. M., Story, B., Maroulas, V., and Vogiatzis, K. D. *Molecular representations for machine learning*. American Chemical Society, Washington, DC, USA, 2023.

Kahng, A. B. Machine learning for CAD/EDA: the road ahead. *IEEE Design and Test*, 40(1):8–16, 2023.

Kim, E.-S., Kang, W. Y., On, K.-W., Heo, Y.-J., and Zhang, B.-T. Hypergraph attention networks for multimodal learning. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 14581–14590, 2020.

Kweon, I. S. and Kanade, T. Extracting topographic terrain features from elevation maps. *CVGIP: image understanding*, 59(2):171–182, 1994.

Laclau, C., Redko, I., Choudhary, M., and Largeron, C. All of the fairness for edge prediction with optimal transport. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 1774–1782. PMLR, 2021.

Landry, N. W., Lucas, M., Iacopini, I., Petri, G., Schwarze, A., Patania, A., and Torres, L. XGI: a Python package for higher-order interaction networks. *Journal of Open Source Software*, 8(85):5162, 2023.

Lee, H., Chung, M. K., Kang, H., Kim, B.-N., and Lee, D. S. Computing the shape of brain networks using graph filtration and Gromov-Hausdorff metric. *International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 302–309, 2011.

Lim, L.-H. Hodge Laplacians on graphs. *SIAM Review*, 62 (3):685–715, 2020.

Lin, Z.-H., Huang, S.-Y., and Wang, Y.-C. F. Convolution in the cloud: Learning deformable kernels in 3D graph convolution networks for point cloud analysis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1800–1809, 2020.

Liu, J., Yang, C., Lu, Z., Chen, J., Li, Y., Zhang, M., Bai, T., Fang, Y., Sun, L., Yu, P. S., et al. Towards graph foundation models: A survey and beyond. *arXiv preprint arXiv:2310.11829*, 2023a.

Liu, X. T., Firoz, J., Lumsdaine, A., Joslyn, C., Aksoy, S., Praggastis, B., and Gebremedhin, A. H. Parallel algorithms for efficient computation of high-order line graphs of hypergraphs. In *2021 IEEE 28th International Conference on High Performance Computing, Data, and Analytics (HiPC)*, pp. 312–321. IEEE, 2021.

Liu, Z., Feng, Y., Black, M. J., Nowrouzezahrai, D., Paull, L., and Liu, W. MeshDiffusion: score-based generative 3D mesh modeling. In *International Conference on Learning Representations*, 2023b.

Love, E. R., Filippenko, B., Maroulas, V., and Carlsson, G. Topological convolutional layers for deep learning. *Journal of Machine Learning Research*, 24(59):1–35, 2023.

Lucic, A., Ter Hoeve, M. A., Tolomei, G., De Rijke, M., and Silvestri, F. CF-GNNExplainer: counterfactual explanations for graph neural networks. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 4499–4511. PMLR, 2022.

Lum, P. Y., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., Carlsson, J., and Carlsson, G. Extracting insights from the shape of complex data using topology. *Scientific reports*, 3:1236, 2013.

Luo, D., Cheng, W., Xu, D., Yu, W., Zong, B., Chen, H., and Zhang, X. Parameterized explainer for graph neural network. *Advances in Neural Information Processing Systems*, 33, 2020.

Luo, Z., Tabaghi, P., Koh, D., Rezaei, E., Defferrard, M., Carey, R., Davis, R., Jain, R., and Wang, Y. DE-HNN: an effective neural model for circuit netlist representation. In *27th Intl. Conf. Artificial Intelligence and Statistics*, 2024. To appear.

Maggs, K., Hacker, C., and Rieck, B. Simplicial representation learning with neural k-forms. arXiv:2312.08515, 2023.

Maria, C., Boissonnat, J.-D., Glisse, M., and Yvinec, M. The Gudhi library: simplicial complexes and persistent homology. In Hong, H. and Yap, C. (eds.), *Mathematical Software – ICMS 2014*, pp. 167–174, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

Maroulas, V., Mike, J. L., and Oballe, C. Nonparametric estimation of probability density functions of random persistence diagrams. *Journal of Machine Learning Research*, 20(151):1–49, 2019.

Maroulas, V., Nasrin, F., and Oballe, C. A Bayesian framework for persistent homology. *SIAM Journal on Mathematics of Data Science*, 2(1):48–74, 2020.

Meirom, S. and Bobrowski, O. Unsupervised geometric and topological approaches for cross-lingual sentence representation and comparison. In *Workshop on Representation Learning for NLP*, 2022.

Mendel, J. M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79(3): 278–305, 1991.

Min, E., Chen, R., Bian, Y., Xu, T., Zhao, K., Huang, W., Zhao, P., Huang, J., Ananiadou, S., and Rong, Y. Transformer for graphs: an overview from architecture perspective. *arXiv preprint arXiv:2202.08455*, 2022.

Mirhoseini, A., Goldie, A., Yazgan, M., Jiang, J. W., Songhori, E., Wang, S., Lee, Y.-J., Johnson, E., Pathak, O., Nazi, A., Pak, J., Tong, A., Srinivasa, K., Hang, W., Tuncer, E., Le, Q. V., Laudon, J., Ho, R., Carpenter, R., and Dean, J. A graph placement methodology for fast chip design. *Nature*, 594(7862):207–212, 2021.

Mitchell, E. C., Story, B., Boothe, D., Franaszczuk, P. J., and Maroulas, V. A topological deep learning framework for neural spike decoding. *Biophysical Journal*, 2024. doi: https://doi.org/10.1016/j.bpj.2024.01.025.

Moor, M., Horn, M., Rieck, B., and Borgwardt, K. Topological autoencoders. In *Proceedings of the 37th International Conference on Machine Learning*, number 119 in Proceedings of Machine Learning Research, pp. 7045–7054. PMLR, 2020.

Morris, C., Ritzert, M., Fey, M., Hamilton, W. L., Lenssen, J. E., Rattan, G., and Grohe, M. Weisfeiler and Leman go neural: higher-order graph neural networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01): 4602–4609, 2019.

Musaelian, A., Batzner, S., Johansson, A., Sun, L., Owen, C. J., Kornbluth, M., and Kozinsky, B. Learning local equivariant representations for large-scale atomistic dynamics. *Nature Communications*, 14(1):579, 2023.

Na, G., Farzana, N., Yong, W., Huibin, W., David J, K., Vasileios, M., and Orlando, R. Persistent homology on electron backscatter diffraction data in nano/ultrafine-grained 18Cr–8Ni stainless steel. *Materials Science and Engineering: A*, 829:142172, 2022.

Nadimpalli, K. V., Chattopadhyay, A., and Rieck, B. Euler characteristic transform based topological loss for reconstructing 3D images from single 2D slices. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 571–579, 2023.

Naitzat, G., Zhitnikov, A., and Lim, L.-H. Topology of deep neural networks. *Journal of Machine Learning Research*, 21(1), 2020.

Nasrin, F., Oballe, C., Boothe, D., and Maroulas, V. Bayesian topological learning for brain state classification. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pp. 1247–1252, 2019.

Newman, M. E. J. Mixing patterns in networks. *Phys. Rev. E*, 67:026126, Feb 2003.

Nguyen, D. D. and Wei, G.-W. Dg-gl: Differential geometry-based geometric learning of molecular datasets. *International journal for numerical methods in biomedical engineering*, 35(3):e3179, 2019.

Nguyen, D. D., Cang, Z., Wu, K., Wang, M., Cao, Y., and Wei, G.-W. Mathematical deep learning for pose and binding affinity prediction and ranking in D3R grand challenges. *Journal of computer-aided molecular design*, 33:71–82, 2019.

Nguyen, D. D., Gao, K., Wang, M., and Wei, G.-W. MathDL: mathematical deep learning for D3R grand challenge 4. *Journal of computer-aided molecular design*, 34: 131–147, 2020.

Nguyen, K., Hieu, N. M., Nguyen, V. D., Ho, N., Osher, S., and Nguyen, T. M. Revisiting over-smoothing and over-squashing using Ollivier-Ricci curvature. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 25956–25979. PMLR, 2023.

Nicolau, M., Levine, A. J., and Carlsson, G. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proceedings of the National Academy of Sciences*, 108 (17):7265–7270, 2011.

Oballe, C., Boothe, D., Franaszczuk, P. J., and Maroulas, V. Tofu: Topology functional units for deep learning. *Foundations of Data Science*, 4(4):641–665, 2022a.

Oballe, C., Cherne, A., Boothe, D., Kerick, S., Franaszczuk, P. J., and Maroulas, V. Bayesian topological signal processing. *Discrete and Continuous Dynamical Systems - S*, 15(4):797–817, 2022b.

Olah, C. Neural networks, manifolds, and topology. *Blog post*, 2014.

Papamarkou, T., Nasrin, F., Lawson, A., Gong, N., Rios, O., and Maroulas, V. A random persistence diagram generator. *Statistics and Computing*, 32(5):88, 2022.

Papillon, M., Hajij, M., Myers, A., Frantzen, F., Zamzmi, G., Jenne, H., Mathe, J., Hoppe, J., Schaub, M., Papamarkou, T., Guzmán-Sáenz, A., Rieck, B., Livesay, N., Dey, T., Rabinowitz, A., Brent, A., Salatiello, A., Nikitin, A., Zia, A., Battiloro, C., Gavrilev, D., Bökman, G., Magai, G., Bazhenov, G., Bernardez, G., Spinelli, I., Agerberg, J., Nadimpalli, K., Telyatninkov, L., Scofano, L., Testa, L., Lecha, M., Yang, M., Hassanin, M., Gardaa, O. H., Zaghen, O., Hausner, P., Snopoff, P., Melnyk, P., Ballester, R., Barikbin, S., Escalera, S., Fiorellino, S., Kvinge, H., Meissner, J., Ramamurthy, K. N., Scholkemper, M., Rosen, P., Walters, R., Samaga, S. N., Mukherjee, S., Sanborn, S., Emerson, T., Doster, T., Birdal, T., Grande, V., Khamis, A., Scardapane, S., Singh, S., Malygina, T., Yue, Y., and Miolane, N. ICML 2023 topological deep learning challenge: design and results. In *Proceedings of 2nd Annual Workshop on Topology, Algebra, and Geometry in Machine Learning (TAG-ML)*, volume 221, pp. 3–8. PMLR, 2023a.

Papillon, M., Sanborn, S., Hajij, M., and Miolane, N. Architectures of topological deep learning: a survey on topological neural networks. *arXiv preprint arXiv:2304.10031*, 2023b.

Perez, I. and Reinauer, R. The topological BERT: transforming attention into topology for natural language processing. arXiv preprint arXiv:2206.15195, 2022.

Pérez, J. B., Hauke, S., Lupo, U., Caorsi, M., and Dassatti, A. giotto-ph: a Python library for high-performance computation of persistent homology of Vietoris-Rips filtrations. *arXiv preprint arXiv:2107.05412*, 2021.

Petri, G. and Leitão, A. On the topological expressive power of neural networks. In *TDA & Beyond NeurIPS Workshop*, 2020.

Poole, B., Jain, A., Barron, J. T., and Mildenhall, B. Dreamfusion: text-to-3D using 2D diffusion. *arXiv preprint arXiv:2209.14988*, 2022.

Prokopchik, K., Benson, A. R., and Tudisco, F. Nonlinear feature diffusion on hypergraphs. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 17945–17958. PMLR, 2022.

Purvine, E., Brown, D., Jefferson, B., Joslyn, C., Praggastis, B., Rathore, A., Shapiro, M., Wang, B., and Zhou, Y. Experimental observations of the topology of convolutional neural network activations. *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI)*, 2023.

Qi, C. R., Su, H., Mo, K., and Guibas, L. J. PointNet: deep learning on point sets for 3D classification and segmentation. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 652–660, 2017.

Qiu, Y. and Wei, G.-W. Persistent spectral theory-guided protein engineering. *Nature Computational Science*, 3(2): 149–163, 2023.

Rahman, T., Surma, B., Backes, M., and Zhang, Y. Fairwalk: towards fair graph embedding. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, IJCAI'19, pp. 3289–3295. AAAI Press, 2019.

Ramamurthy, K. N., Varshney, K., and Mody, K. Topological data analysis of decision boundaries with application to model selection. In *International Conference on Machine Learning*, pp. 5351–5360. PMLR, 2019.

Ramamurthy, K. N., Guzmán-Sáenz, A., and Hajij, M. Topomlp: A simplicial network without message passing. In

ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5. IEEE, 2023.

Rathore, A., Zhou, Y., Srikumar, V., and Wang., B. TopoBERT: Exploring the topology of fine-tuned word representations. *Information Visualization*, 2023.

Reiser, P., Neubert, M., Eberhard, A., Torresi, L., Zhou, C., Shao, C., Metni, H., van Hoesel, C., Schopmans, H., Sommer, T., and Friederich, P. Graph neural networks for materials science and chemistry. *Communications Materials*, 3(1):93, 2022.

Rieck, B. and Leitte, H. Persistent homology for the evaluation of dimensionality reduction schemes. *Computer Graphics Forum*, 34(3):431–440, 2015.

Rieck, B., Bock, C., and Borgwardt, K. A persistent Weisfeiler–Lehman procedure for graph classification. In Chaudhuri, K. and Salakhutdinov, R. (eds.), *Proceedings of the 36th International Conference on Machine Learning (ICML)*, number 97 in Proceedings of Machine Learning Research, pp. 5448–5458. PMLR, 2019a.

Rieck, B., Togninalli, M., Bock, C., Moor, M., Horn, M., Gumbsch, T., and Borgwardt, K. Neural persistence: a complexity measure for deep neural networks using algebraic topology. In *International Conference on Learning Representations (ICLR)*, 2019b.

Rieck, B., Yates, T., Bock, C., Borgwardt, K., Wolf, G., Turk-Browne, N., and Krishnaswamy, S. Uncovering the topology of time-varying fMRI data using cubical persistence. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, pp. 6900–6912. Curran Associates, Inc., 2020.

Roddenberry, T. M. and Segarra, S. HodgeNet: graph neural networks for edge data. In *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 220–224. IEEE, 2019.

Roddenberry, T. M., Glaze, N., and Segarra, S. Principled simplicial neural networks for trajectory prediction. In *Int. Conf. Mach. Learn.*, pp. 9020–9029. PMLR, 2021.

Roddenberry, T. M., Schaub, M. T., and Hajij, M. Signal processing on cell complexes. *ICASSP*, 2022.

Roell, E. and Rieck, B. Differentiable Euler characteristic transforms for shape classification. arXiv:2310.07630, 2023.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. High-resolution image synthesis with latent diffusion models, 2022.

Rosen, P., Wang, B., Seth, A., Mills, B., Ginsburg, A., Kamenetzky, J., Kern, J., and Johnson, C. R. Using contour trees in the analysis and visualization of radio astronomy data cubes. *arXiv preprint arXiv:1704.04561*, pp. 1–7, 2017.

Rozemberczki, B., Kiss, O., and Sarkar, R. Karate Club: an API oriented open-source Python framework for unsupervised learning on graphs. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM '20)*, pp. 3125–3132. ACM, 2020.

Sanford, C., Hsu, D., and Telgarsky, M. Representational strengths and limitations of transformers. In *NeurIPS*, 2023.

Sardellitti, S., Barbarossa, S., and Testa, L. Topological signal processing over cell complexes. *Proceeding IEEE Asilomar Conference. Signals, Systems and Computers*, 2021.

Scardapane, S., Spinelli, I., and Lorenzo, P. D. Distributed training of graph convolutional networks. *IEEE Transactions on Signal and Information Processing over Networks*, 7:87–100, 2021.

Schaub, M. T. and Segarra, S. Flow smoothing and denoising: graph signal processing in the edge-space. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 735–739, 2018.

Schaub, M. T., Benson, A. R., Horn, P., Lippner, G., and Jadbabaie, A. Random walks on simplicial complexes and the normalized hodge 1-Laplacian. *SIAM Review*, 62 (2):353–391, 2020.

Schaub, M. T., Zhu, Y., Seby, J.-B., Roddenberry, T. M., and Segarra, S. Signal processing on higher-order networks: livin'on the edge... and beyond. *Signal Processing*, 187: 108149, 2021.

Schaub, M. T., Seby, J.-B., Frantzen, F., Roddenberry, T. M., Zhu, Y., and Segarra, S. Signal processing on simplicial complexes. In *Higher-Order Systems*, pp. 301–328. Springer, 2022.

Shi, Z., Peng, S., Xu, Y., Geiger, A., Liao, Y., and Shen, Y. Deep generative models on 3D representations: a survey. *arXiv preprint arXiv:2210.15663*, 2022.

Southern, J., Wayland, J., Bronstein, M., and Rieck, B. Curvature filtrations for graph generative model evaluation. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, 2023.

Spannaus, A., Law, K. J., Luszczek, P., Nasrin, F., Micucci, C. P., Liaw, P. K., Santodonato, L. J., Keffer, D. J., and Maroulas, V. Materials fingerprinting classification. *Computer Physics Communications*, 266:108019, 2021.

Spinelli, I., Scardapane, S., Hussain, A., and Uncini, A. FairDrop: biased edge dropout for enhancing fairness in graph representation learning. *IEEE Transactions on Artificial Intelligence*, 3(3):344–354, 2022.

Sun, S., Chen, W., Wang, L., Liu, X., and Liu, T.-Y. On the depth of deep neural networks: a theoretical view. In *AAAI*, 2016.

Sunada, T. *Topological crystallography with a view towards discrete geometric analysis*. Springer, 2013.

Sverrisson, F., Feydy, J., Correia, B. E., and Bronstein, M. M. Fast end-to-end learning on protein surfaces. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15267–15276, 2021.

Tang, B., Chen, S., and Dong, X. Hypergraph-MLP: learning on hypergraphs without message passing. *arXiv preprint arXiv:2312.09778*, 2023a.

Tang, B., Chen, S., and Dong, X. Hypergraph structure inference from data under smoothness prior. *arXiv preprint arXiv:2308.14172*, 2023b.

Tauzin, G., Lupo, U., Tunstall, L., PÃ©rez, J. B., Caorsi, M., Medina-Mardones, A. M., Dassatti, A., and Hess, K. giotto-tda: a topological data analysis toolkit for machine learning and data exploration. *Journal of Machine Learning Research*, 22(39):1–6, 2021.

Telyatnikov, L., Bucarelli, M. S., Bernardez, G., Zaghen, O., Scardapane, S., and Lio, P. Hypergraph neural networks through the lens of message passing: a common perspective to homophily and architecture design. *arXiv preprint arXiv:2310.07684*, 2023.

Topaz, C. M., Ziegelmeier, L., and Halverson, T. Topological data analysis of biological aggregation models. *PloS one*, 10(5):e0126383, 2015.

Topping, J., Giovanni, F. D., Chamberlain, B. P., Dong, X., and Bronstein, M. M. Understanding over-squashing and bottlenecks on graphs via curvature. In *International Conference on Learning Representations*, 2022.

Townsend, J., Micucci, C. P., Hymel, J. H., Maroulas, V., and Vogiatzis, K. D. Representation of molecular structures with persistent homology for machine learning applications in chemistry. *Nature communications*, 11(1): 3230, 2020.

Trofimov, I., Cherniavskii, D., Tulchinskii, E., Balabin, N., Burnaev, E., and Barannikov, S. Learning topology-preserving data representations. In *ICLR*, 2023.

Tulchinskii, E., Kuznetsov, K., Kushnareva, L., Cherniavskii, D., Nikolenko, S., Burnaev, E., Barannikov, S., and Piontkovskaya, I. Intrinsic dimension estimation for robust detection of AI-generated texts. In *37th Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

Veličković, P. Message passing all the way up. In *ICLR 2022 Workshop on Geometrical and Topological Representation Learning*, 2022.

Veličković, P. Everything is connected: graph neural networks. *Current Opinion in Structural Biology*, 79:102538, 2023.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. Graph attention networks. In *International Conference on Learning Representations*, 2018.

Verdon, G., McCourt, T., Luzhnica, E., Singh, V., Leichenauer, S., and Hidary, J. Quantum graph neural networks. *arXiv preprint arXiv:1909.12264*, 2019.

Vignac, C., Krawczuk, I., Siraudin, A., Wang, B., Cevher, V., and Frossard, P. DiGress: discrete denoising diffusion for graph generation. *arXiv preprint arXiv:2209.14734*, 2023.

von Rohrscheidt, J. and Rieck, B. Topological singularity detection at multiple scales. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, number 202 in Proceedings of Machine Learning Research, pp. 35175–35197. PMLR, 2023.

Wagner, A., Solomon, E., and Bendich, P. Improving metric dimensionality reduction with distributed topology, 2021. arXiv:2106.07613.

Waibel, D. J. E., Atwell, S., Meier, M., Marr, C., and Rieck, B. Capturing shape information with multi-scale topological loss terms for 3D reconstruction. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 150–159, Cham, Switzerland, 2022. Springer.

Wang, H., Yin, H., Zhang, M., and Li, P. Equivariant and stable positional encoding for more powerful graph neural networks. In *International Conference on Learning Representations*, 2022.

Wang, M., Zheng, D., Ye, Z., Gan, Q., Li, M., Song, X., Zhou, J., Ma, C., Yu, L., Gai, Y., Xiao, T., He, T., Karypis, G., Li, J., and Zhang, Z. Deep Graph Library: a graph-centric, highly-performant package for graph neural networks. *arXiv preprint arXiv:1909.01315*, 2019.

Wang, P., Yang, S., Liu, Y., Wang, Z., and Li, P. Equivariant hypergraph diffusion neural operators. In *International Conference on Learning Representations (ICLR)*, 2023.

Wang, R., Nguyen, D. D., and Wei, G.-W. Persistent spectral graph. *International journal for numerical methods in biomedical engineering*, 36(9):e3376, 2020.

Wang, R., Chen, J., and Wei, G.-W. Mechanisms of SARS-CoV-2 evolution revealing vaccine-resistant mutations in Europe and America. *The journal of physical chemistry letters*, 12(49):11850–11857, 2021.

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976): 1089–1100, 2023.

Wei, X. and Wei, G.-W. Persistent sheaf Laplacians. *arXiv preprint arXiv:2112.10906*, 2021.

Wheeler, M., Bouza, J., and Bubenik, P. Activation landscapes as a topological summary of neural network performance. In *2021 IEEE International Conference on Big Data (Big Data)*, pp. 3865–3870, 2021.

Wu, H. and Ng, M. K. Hypergraph convolution on nodes-hyperedges network for semi-supervised node classification. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 16(4):1–19, 2022.

Wu, K. and Wei, G.-W. Quantitative toxicity prediction using topology based multitask deep neural networks. *Journal of chemical information and modeling*, 58(2): 520–531, 2018.

Wu, Q., Yang, C., Zeng, K., Nie, F., Bronstein, M., and Yan, J. Advective diffusion transformers for topological generalization in graph learning. *arXiv preprint arXiv:2310.06417*, 2023.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24, 2020.

Xia, K. and Wei, G.-W. Persistent homology analysis of protein structure, flexibility, and folding. *International journal for numerical methods in biomedical engineering*, 30(8):814–844, 2014.

Xia, K. and Wei, G.-W. Persistent topology for cryo-EM data analysis. *International Journal for Numerical Methods in Biomedical Engineering*, 31(8), 2015.

Xie, Z., Liang, R., Xu, X., Hu, J., Duan, Y., and Chen, Y. Net2: a graph attention network method customized for pre-placement net length estimation. In *Asia and South Pacific Design Automation Conference*, ASPDAC, pp. 671–677, 2021.

Xu, M., Yu, L., Song, Y., Shi, C., Ermon, S., and Tang, J. GeoDiff: a geometric diffusion model for molecular conformation generation, 2022.

Yang, M. and Isufi, E. Convolutional learning on simplicial complexes. *arXiv preprint arXiv:2301.11163*, 2023.

Yang, S., Yang, Z., Li, D., Zhang, Y., Zhang, Z., Song, G., and Hao, J. Versatile multi-stage graph neural network for circuit representation. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Yim, J., Trippe, B. L., De Bortoli, V., Mathieu, E., Doucet, A., Barzilay, R., and Jaakkola, T. SE(3) diffusion model with application to protein backbone generation. *arXiv preprint arXiv:2302.02277*, 2023.

Yun, S., Jeong, M., Kim, R., Kang, J., and Kim, H. J. Graph transformer networks. In *NeurIPS*, 2019.

Zeng, S., Graf, F., Hofer, C., and Kwitt, R. Topological attention for time series forecasting. In *Advances in Neural Information Processing Systems*, volume 34, pp. 24871–24882. Curran Associates, Inc., 2021.

Zeng, X., Vahdat, A., Williams, F., Gojcic, Z., Litany, O., Fidler, S., and Kreis, K. Lion: Latent point diffusion models for 3D shape generation. *arXiv preprint arXiv:2210.06978*, 2022.

Zhan, Y., Chen, Z., Yu, J., Yu, B., Tao, D., and Luo, Y. Hyper-relationship learning network for scene graph generation. *arXiv preprint arXiv:2202.07271*, 2022.

Zhang, M., Qamar, M., Kang, T., Jung, Y., Zhang, C., Bae, S.-H., and Zhang, C. A survey on graph diffusion models: generative AI in science for molecule, protein and material. *arXiv preprint arXiv:2304.01565*, 2023.

Zhao, R., Wang, M., Chen, J., Tong, Y., and Wei, G.-W. The de Rham-Hodge analysis and modeling of biomolecules. *Bulletin of mathematical biology*, 82:1–38, 2020.

Zhao, Y., Birdal, T., Deng, H., and Tombari, F. 3D point capsule networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1009–1018, 2019.

Zheng, S., Zhang, Y., Wagner, H., Goswami, M., and Chen, C. Topological detection of Trojaned neural networks. In *NeurIPS*, 2021.

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. Graph neural networks: a review of methods and applications. *AI Open*, 1:57–81, 2020.

Zhou, Y., Zhou, Y., Ding, J., and Wang, B. Visualizing and analyzing the topology of neuron activations in deep adversarial training. *Topology, Algebra, and Geometry*

in Machine Learning (TAGML) Workshop at ICML, pp. 134–145, 2023.

Zia, A., Khamis, A., Nichols, J., Tayab, U. B., Hayder, Z., Rolland, V., Stone, E., and Petersson, L. Topological deep learning: a review of an emerging paradigm. *Artificial Intelligence Review*, 57(4):77, 2024.

A. Literature Review

This appendix provides a brief literature review focused on TDL, while also broadly discussing the role of topology in machine learning. Three main themes are considered, namely higher-order deep learning models, research at the intersection of the topology of data and neural networks, and research that connects TDA with neural networks.

A.1. Higher-Order Deep Learning Models

Recent years have witnessed a growing interest in higher-order networks (Mendel, 1991; Battiston et al., 2020; Bick et al., 2023) due to their effective ability to capture higher-order interactions. In signal processing and deep learning, various approaches, such as Hodge-theoretic methods, message-passing schemes, and skip connections, have been developed for higher-order deep learning models.

The use of Hodge Laplacians for data analysis has been investigated by Jiang et al. (2011); Lim (2020) and has been extended to a signal processing context, for example, in Schaub et al. (2021); Sardellitti et al. (2021); Roddenberry et al. (2022) for simplicial and cellular complexes. Also, edge-centric convolutional neural networks (CNNs) leveraging the 1-Hodge Laplacian operator for linear filtering have been defined (Roddenberry & Segarra, 2019; Barbarossa et al., 2018; Schaub & Segarra, 2018; Barbarossa & Sardellitti, 2020a;b; Schaub et al., 2021).

Convolutional operators and message-passing algorithms for higher-order neural networks have been developed. For hypergraphs, a convolutional operator has been proposed (Jiang et al., 2019b; Feng et al., 2019; Arya & Worring, 2018), and has been further investigated in Wu & Ng (2022); Bai et al. (2021a); Jiang et al. (2019b); Bai et al. (2021b); Gao et al. (2020); Giusti et al. (2023); Gong et al. (2023). In addition, a unified framework for learning on graphs and hypergraphs has been developed (Huang & Yang, 2021), and general hypergraph neural networks have been introduced (Gao et al., 2022).

In the context of message passing on complexes, a general message-passing scheme on topological spaces has been developed by Hajij et al. (2020), encompassing previous proposals (Gilmer et al., 2017; Ebli et al., 2020; Bunch et al., 2020; Hayhoe et al., 2022) and employing various topological neighborhood aggregation schemes. The expressive power of simplicial and cellular message passing neural networks has been studied by Bodnar (2022) and Bodnar et al. (2022), respectively. The work of Hajij et al. (2023b) provides the mathematical blueprint of topological deep learning, unifying existing deep learning architectures with a common mathematical language. Papillon et al. (2023b) has conducted a review of message-passing topological neural network architectures, employing foundational concepts from Hajij et al. (2023b). This survey highlights key challenges and promising directions for future research in the field. Zia et al. (2024) have carried out a review of TDL, focusing primarily on how the use of TDA techniques has evolved over time to support deep learning frameworks. Recurrent simplicial neural networks for trajectory prediction have been introduced (Mitchell et al., 2024), and a multi-signal approach on higher-order networks utilizing the Dirac operator has been developed (Calmon et al., 2022; Hajij et al., 2023a).

While the dominant trend of topological neural networks represents a special case of higher-order message passing (Hajij et al., 2023b), recently, several notable topological neural network methods have been introduced that do not rely on this paradigm. For example, Maggs et al. (2023) have introduced neural k-forms, providing a novel approach to operate on geometric simplicial complexes and graphs without relying on traditional message passing. In addition, Ramamurthy et al. (2023) have introduced an MLP-based simplicial neural network algorithm to learn the representation of elements in a simplicial complex without explicitly relying on message passing, ensuring fast inference time and high robustness against the lack of connectivity information during inference.

To facilitate the training of higher-order DNNs, a generalization of skip connections to simplicial complexes has been proposed (Hajij et al., 2022a). Additionally, a higher-order GNN considering higher-order graph structures at multiple scales has been presented (Morris et al., 2019).

Most attention-based models are designed primarily for graphs, with some recent exceptions that have been introduced in higher-order domains (Bai et al., 2021b; Kim et al., 2020; Georgiev et al., 2022; Giusti et al., 2022; Goh et al., 2022; Giusti et al., 2023). For instance, Goh et al. (2022) have proposed a generalization of the graph attention model of Veličković et al. (2018), while Giusti et al. (2022) have introduced an attention model for simplicial complexes based on Hodge decomposition, similar to Roddenberry et al. (2021). Furthermore, Giusti et al. (2022) have introduced a simplicial network architecture leveraging masked self-attention layers. Expanding this work, Battiloro et al. (2023c) have introduced a generalized attention architecture that integrates the interplay among simplices of different orders through the Dirac operator and its Dirac decomposition, adding a layer of complexity and depth to the understanding of attention mechanisms. Attention

on cell complexes has been introduced by Giusti et al. (2023) exploiting higher-order topological information through feature lifting and attention mechanisms over lower and upper neighborhoods. In a related work, a generalized attention mechanism on combinatorial complexes has been introduced by Hajij et al. (2023b). Hypergraph attention models in Bai et al. (2021b); Kim et al. (2020) provide alternative generalizations of the graph attention model by Veličković et al. (2018). These models collectively represent a generalization of the traditional graph-based attention framework to more complex and higher-dimensional structures.

Finally, tangent neural networks that operate over the tangent bundle of Riemann manifolds have recently been introduced by Battiloro et al. (2023d). Battiloro et al. (2023b) have extracted meaningful latent topology information tailored to specific tasks, providing a versatile approach for topology-aware neural networks.

A.2. Topology of Data and Neural Networks

Interpreting Euclidean data as a sample set from a topological space has a long history (Carlsson, 2009). Some early points on the interplay between topology and neural networks appeared in a blog of Olah (2014). In particular, Olah (2014) has performed topological experiments that illustrate the importance of considering the topology of the underlying data when choosing a neural network. Approximately at the same time, Bianchini & Scarselli (2014) have published their seminal work related to the expressivity of DNNs, later extended by Sun et al. (2016). In Naitzat et al. (2020), the activations of a binary classification neural network have been considered as point clouds on which the layer functions of the neural network act. The topologies of these activations have been studied using homological tools such as persistent homology (Edelsbrunner & Harer, 2010). Rieck et al. (2019b) have studied the topology of individual layers in MLPs, developing neural persistence as a way to describe deep learning phenomena based on topological concepts. Some caveats of this approach have recently been observed and rectified by Girrbach et al. (2023), leaving the open question of how to effectively use topology to describe neural network training. Hajij et al. (2021a) have used topological notions to provide insights into the supervised classification problem in the context of neural networks. Similarly, Oballe et al. (2022a); Love et al. (2023) have used topological properties of data to infer deep learning architectures.

A.3. Topological Data Analysis and Neural Networks

Although an old field, algebraic topology (Hatcher, 2005) has only recently begun to solve problems in various domains. For example, persistent homology (Edelsbrunner & Harer, 2010) succeeded in solving complex data problems (Attene et al., 2003; Bajaj et al., 1997; Boyell & Ruston, 1963; Carr et al., 2004; Curto, 2017; Dabaghian et al., 2012; Giusti et al., 2016; Kweon & Kanade, 1994; Lee et al., 2011; Lum et al., 2013; Nicolau et al., 2011; Rosen et al., 2017). Deep learning models rooted in topology have left their imprint in various domains, including topological data signatures (Biasotti et al., 2008; Carlsson et al., 2005; Rieck & Leitte, 2015), neuroscience (Curto, 2017), bioscience (Topaz et al., 2015) time series forecasting (Zeng et al., 2021), Trojan detection (Hu et al., 2022), image segmentation (Hu et al., 2019), 3D reconstruction (Waibel et al., 2022), and time-varying setups (Rieck et al., 2020). TDA (Edelsbrunner & Harer, 2010; Carlsson, 2009; Ghrist, 2014; Dey & Wang, 2022; Love et al., 2023) has relied on topological tools to analyze data and generate machine learning algorithms.

In Rieck et al. (2019b), the analysis of neural networks and their generalization performance has been approached through persistent homology, providing valuable observational insights. Building on this work, Hofer et al. (2020b) have introduced a topology-based graph and simplicial complex readout function, paving the way for a deeper understanding of data representations. In pursuit of efficient representation learning, Moor et al. (2020) have focused on topology-based methods, particularly emphasizing the use of the 1-skeleton for performance reasons while actively working to improve the speed of simplicial complexes. Horoi et al. (2022) have adopted topological concepts to analyze neural network behaviors, providing observational insights into the learning process. Moving toward end-to-end learning, Horn et al. (2022) have introduced a novel perspective, incorporating topological information into graph and simplicial complex learning. The work of Horn et al. (2022) breaks free from the confines of traditional graph isomorphism tests, enhancing the understanding of complex learning structures. In the context of 2D-to-3D image reconstruction, Waibel et al. (2022) have demonstrated how topological information, specifically cubical complexes in 3D, leads to more efficient models and higher quality reconstruction results. Exploring has continued in Nadimpalli et al. (2023), focusing on quantized and approximated Euler characteristic transforms for image reconstruction, contributing to the arsenal of tools to enhance reconstruction processes. Taking a singular approach, von Rohrscheidt & Rieck (2023) have formulated a method for detecting multi-scale singularities in data using local persistent homology. This observational study has linked singularity information with generalization performance, adding depth to understanding the difficulty of classification tasks. In pursuit of high-performance representation learning,

Roell & Rieck (2023) have introduced a fully differentiable variant of the Euler characteristic transform, facilitating tasks on graphs and combinatorial complexes. The use of curvature-based topological information has been showcased by Southern et al. (2023), offering a perspective on the evaluation of generative graph models. The interplay between probabilistic machine learning and topology has recently been investigated in several works. Specifically, Maroulas et al. (2019; 2020); Oballe et al. (2022b); Papamarkou et al. (2022) have developed sampling techniques for distributions of topological summaries either in a non-parametric way or from a Bayesian perspective. Using variations of Cech and Vietoris-Rips, the topology of decision boundaries of neural networks has been studied by Ramamurthy et al. (2019) to quantify the complexity of DNNs and to enable the matching of datasets to pre-trained models.

B. Application Areas for TDL

This appendix highlights numerous domains of application in which TDL has shown or may show its potential. In particular, TDL is a promising tool for data compression, NLP, computer vision and computer graphics, quantum TDL, chemistry, biological imaging, virus evolution, drug design, neuroscience, protein engineering, chip design, semantic communications, satellite imagery, and materials science.

Data compression. TDL has the potential to compress relational data by exploiting multi-way correlations between elements to obtain effective lower-dimensional representations. As an example, Bernárdez et al. (2023) have proposed a TDL-based approach for compressing graph signals, demonstrating superior performance compared to both GNNs and feed-forward architectures when applied to two real-world internet datasets. Additionally, Battiloro et al. (2023a) have introduced a topological dictionary learning algorithm for sparse signal representation over regular cell complexes. This method has demonstrated advantages in edge flow compression tasks, further illustrating the potential of topological methods compared to graph-based methods. These studies suggest that TDL holds great promise for advancing data compression techniques by leveraging its ability to capture complex correlations in relational data.

Natural language processing. Deep learning has recently made impressive advances in the field of NLP, a rich and less explored domain to apply TDL. Specifically, contextualized word embeddings from large language models have revolutionized tasks such as text generation, sentiment analysis, and machine translation. Exploring the topological structures of these general-purpose learned embeddings can help NLP practitioners unlock the hidden structures of their models. TDL could be crucial to understanding how these models organize hierarchical class knowledge across neural network layers (Purvine et al., 2023). TDL can also probe the lexical, syntactic and semantic regularities of the embedding space (Hajij et al., 2021b; Rathore et al., 2023), and estimate its intrinsic dimensionality (Tulchinskii et al., 2023), thus providing insights into model effectiveness. Additionally, the scope of TDL extends beyond word embeddings. It can be used to compare sentence embeddings from language models (Meirom & Bobrowski, 2022), study the parameter space of neural networks (Gabella et al., 2019), and investigate deep adversarial training (Perez & Reinauer, 2022; Zhou et al., 2023). This can help NLP practitioners understand the complex structures of their models and further advance language processing technologies.

Computer vision and computer graphics. Several recent TDL works focus on learning directly on topological spaces. In the domains of vision and graphics, these spaces often manifest naturally as 3D point clouds, graphs, or meshes, either acquired in the real world through 3D sensors or designed in computer-aided design software. To date, the field has distinctly differentiated between these data representations, with the optimal choice often remaining elusive and highly dependent on the specific application (Ahmed et al., 2018; Shi et al., 2022). The lack of a unifying representation has led vision and graphics researchers to develop specialized architectures for different data types. Examples include PointNet for point clouds (Qi et al., 2017), MeshCNN for meshes (Hanocka et al., 2019) and 3D GCNs for spatial graphs (Lin et al., 2020). Limited efforts have been made to consider a unified domain approach. Jiang et al. (2019a) have attempted to bridge this gap via simplex mesh-based geometry representations, achieving success in shape optimization and segmentation on real-world datasets. Hajij et al. (2022b) have employed simplicial complexes in representation learning, although their evaluation has been confined to meshes. The concept of combinatorial complexes (Hajij et al., 2023b) has offered a unifying view and has been tested on mesh, point cloud and graph segmentation.

Quantum TDL. Quantum neural networks (QNNs) are promising architectures for handling various types of datasets, which have the potential for exponential speed-up compared to classical neural network algorithms. Drawing inspiration from the success of GDL, a QNN has been introduced by Verdon et al. (2019), allowing for both quantum inference and classical probabilistic inference on data characterized by a graph-geometric structure. The computation of topological features often presents a formidable challenge, necessitating the subsampling of underlying data. A breakthrough in this domain comes

from the work of Ameneyro et al. (2022a), which has introduced an efficient quantum computation method for persistent Betti numbers, using a persistent Dirac operator, whose square yields the persistent combinatorial Laplacian (Wang et al., 2020). The approach of Ameneyro et al. (2022a) unveils the underlying persistent Betti numbers that capture the enduring features of the data. To enhance its applicability to time series data, Ameneyro et al. (2022b) have introduced a quantum Takens delay embedding algorithm, which transforms time series into point clouds within a quantum framework. The work of Ameneyro et al. (2022b) broadens algorithmic utility and shows adaptability in various data structures. While existing QNNs do not explicitly cater to data defined over topological spaces, and overall quantum topological methods are in their infancy, the advent of TDL suggests a fertile research direction in this realm. Indeed, the emerging field of TDL may pave the way for tailored QNNs designed to process data within topological spaces, or develop QNN architectures that embody topological properties in their structures.

Chemistry. TDL can offer significant benefits to chemistry and molecular applications by providing a flexible framework that can effectively capture complex molecular structures, long-range dependencies, and higher-order interactions. As discussed in Jiang et al. (2021); Jones et al. (2023), capturing topological properties offers a promising way to understand the underlying structure of a molecule. For example, Jiang et al. (2021) have demonstrated the power of methods based on algebraic topology in constructing unique representations of crystalline compounds. These methods not only capture intricate pairwise and many-body interactions, but also reveal the topological structure within groups of atoms at various scales. This approach, which goes beyond a molecular graph representation, has been successfully applied in molecular discovery, as presented in Townsend et al. (2020), helping to uncover novel molecular structures and properties. Furthermore, TDL has been shown to be valuable for assessing and predicting the safety and potential risks associated with various chemical compounds (Wu & Wei, 2018). Another molecular application of TDL has been explored by Bodnar et al. (2021a), who have introduced a neural network based on a cell complex to predict molecular properties. The proposed network has provided enhanced expressivity, principled modeling of higher-order signals, and improved distance compression, leading to the achievement of state-of-the-art results in various synthetic and real-world molecular benchmarks. These studies suggest that topology plays a crucial role in supramolecular chemistry, particularly in the study of spatially organized multimolecular complexes characterized by diverse bonding patterns, such as mechanically interlocked and interpenetrated networks, thus highlighting the growing significance of TDL in the advancement of molecular and chemistry-related research.

Biological imaging. Capturing 3D configurations of minuscule biological entities, such as proteins, presents a significant imaging challenge. A pivotal advancement in this area has been made through cryo-electron microscopy (cryo-EM), which has notably transformed structural biology. This technique enables the imaging of molecules within a solution, achieving a level of detail that was previously unattainable. The primary purpose of using cryo-EM is to reconstruct the 3D shape of biomolecules from the 2D images it produces. However, this reconstruction task is complex, as the initial images obtained from cryo-EM are noisy 2D projections of the underlying 3D biomolecules. A promising approach to enhance the accuracy of this reconstruction is the application of robust topological or geometric principles, acting as guides for what the 3D structure of the biomolecule might plausibly be. For example, using a known 3D structure of the protein as a reference, conceptualized as a cellular complex with nodes symbolizing the residues, edges representing covalent bonds, and faces indicating the rings, can significantly aid in the reconstruction process. Any further analysis or processing of these reconstructed structures can benefit from TDL, leveraging its capabilities to refine and understand complex biomolecular shapes derived from cryo-EM (Xia & Wei, 2015; Biswas et al., 2016; Zhao et al., 2020). In addition, TDL has shown to be useful in scenarios involving high intra-subject variance, such as in magnetic resonance imaging (MRI). It facilitates the comparison and analysis of such complex datasets, thereby improving the effectiveness of learning tasks in these areas. The incorporation of topological structures as a means of regularizing various tasks has led to notable advances in areas such as image segmentation (Hu et al., 2019; 2021b; Clough et al., 2022; Gupta et al., 2022) and image-based shape reconstruction (Waibel et al., 2022). The application of topological concepts as a prior in these tasks has contributed to improved inference. These studies highlight the potential of TDL in biomedical imaging applications.

Virus evolution. During the COVID-19 pandemic, understanding the evolutionary trend of the SARS-CoV-2 virus was a great challenge. TDL has played a vital role in unraveling the evolutionary mechanisms of SARS-CoV-2, including the natural selection process that strengthens the infectivity, as elucidated in Chen et al. (2020), and the development of antibody-resistant spike mutations, as explained in Wang et al. (2021). This mechanistic understanding has been instrumental in forecasting the emergence of dominant variants such as BA.2 (Chen & Wei, 2022) and BA.4, BA.5 (Chen et al., 2022) approximately two months in advance. The contribution of TDL to understanding these mutations and their implications has been indispensable to addressing the challenges posed by the rapidly evolving SARS-CoV-2 virus. It exemplifies how TDL can be used effectively to analyze virus evolution and adapt to emerging threats.

Drug design. In life sciences, the development of effective drugs to combat diseases is of utmost importance. The central goals of molecular bioscience and biophysics are to decipher the molecular complexities of human diseases and to create drugs that can mitigate these diseases without causing harmful side effects. A crucial aspect of drug design and discovery is the ability to predict how a given molecule will interact with biomolecules such as proteins or DNA. This interaction is key to determining whether a drug can activate or inhibit the function of a biomolecule, leading to therapeutic benefits. The D3R Grand Challenges, an annual global competition in computer-aided drug design, served as a platform for testing and evaluating various methodologies from physical, chemical, statistical, and computer science perspectives in terms of scoring, ranking, and docking capabilities in drug design. In this challenge, TDL was a winner, outperforming other techniques in accurately predicting molecular interactions (Nguyen et al., 2019; 2020). The effectiveness of TDL is particularly notable in drug discovery and other applications involving molecular structures, given the intricate balance between geometry (the physical space of molecules) and topology (the inherent connectivity of molecules). Benchmark studies have demonstrated that both dynamic TDL methods (Horn et al., 2022) and static topological features (Rieck et al., 2019a) can result in enhanced predictive performance. This positions TDL as a promising area for future research and development in drug discovery.

Neuroscience. Exploring neuroscience through a topology-based lens proves instrumental in tackling the substantial intra-subject variance inherent in brain imaging modalities. The work of Rieck et al. (2020) has showcased the efficacy of static topological features in functional MRI (fMRI) data analysis and prediction tasks. The neural manifold hypothesis, as stated in Bengio et al. (2013), posits that neuronal activity forms a low-dimensional manifold reflecting the structure of encoded task variables. Notably, the topology of the neural manifold aligns intricately with the task's topology. Examples include the topological ring in neural circuits that encode the direction of an animal's head and the torus in the grid cell circuit for spatial navigation (Chaudhuri et al., 2019; Gardner et al., 2022). TDL, with its application to reveal the scientific meaning behind neural code topology, emerges as a crucial tool. Learning the topology serves as a fundamental first step for subsequent analysis, including geometric analysis (Acosta et al., 2023). An intriguing facet of spatial information encoding involves grid cells, layered neurons facilitating environment-based navigation. Mitchell et al. (2024) have captured the firing structure as simplicial complexes, decoding grid cell data by developing a recurrent neural network based on simplicial complexes. Furthermore, studies by Nasrin et al. (2019) and Maroulas et al. (2019) have addressed brain wave data stochasticity by detecting topological fingerprints in electroencephalograms and estimating the topological noise in such fingerprints. The comprehension, representation, and decoding of neural structures benefit from models with higher-order connectivity, making TDL integral to understanding intricate patterns in neuroscience.

Protein engineering. Protein engineering is at the forefront of biotechnology, presenting a transformative potential in diverse domains such as antibody design, drug discovery, food security, ecology and beyond. Taking advantage of extensive protein databases, machine learning models have significantly accelerated the pace of protein engineering and directed evolution. Recently, TDA approaches have been used to dramatically reduce the structural complexity of mutant and wide-type proteins (Qiu & Wei, 2023). TDL based on persistent homology has shown to be a top performer in an extensive study involving 34 datasets, but a persistent Laplacian approach outperforms persistent homology in protein engineering applications (Qiu & Wei, 2023). TDL has also facilitated deep mutational scanning (Chen et al., 2023).

Chip design. Chip design has been accelerated by the combination of machine learning with geometric and topological methods. In this field, the central element is the so-called netlist, which is a network comprising cells (logic gates) and nets (connections among cells). Netlists can be effectively modeled as (directed) hypergraphs, where nodes and hyperedges represent cells and nets, respectively. A key goal in chip design is to efficiently lay out netlists within a specified 2D area. The aim is to optimize several properties, such as minimizing total wire length and reducing congested areas known as 'hotspots'. However, optimizing netlists, including their placement and routing, is a complex and time-consuming task. It involves multiple steps (including placement, global routing, detailed routing) and repeated iterations. This complexity requires data-driven methods that can speed up these processes (Kahng, 2023; Mirhoseini et al., 2021). For example, there is growing interest in predicting the properties of a synthesized netlist without undergoing the lengthy placement and routing process (Ghose et al., 2021; Xie et al., 2021; Yang et al., 2022). However, netlists pose several challenges for existing graph learning models. Such models are often large, containing millions of nodes, and the long-range interactions and path information are crucial for determining netlist properties, such as timing and routability. Moreover, various symmetry structures in netlists must be considered. These complexities highlight the need to integrate topological and geometric concepts into the learning process. Luo et al. (2024) have made a step towards this direction by integrating persistent homology and spectral features within a carefully designed neural network based on directed hypergraphs to predict netlist properties.

Semantic communications. A key challenge in communication systems is balancing the complexity of data representation with the significance of symbols transmitted to convey the intended meaning or semantics within an allowable margin of error or distortion. Efficiently representing semantic knowledge involves mapping out the relations between the elements of a language, broadly defined, by creating a corresponding topological space (Barbarossa et al., 2023). TDL can play a crucial role in semantic communications, particularly in extracting relational semantic features from data. This is vital for transmitting information that is strictly necessary to convey the desired meaning or to execute a specific task, such as inference, control, or actuation, while meeting certain performance criteria, such as latency, energy efficiency, and accuracy. The integration of TDL in this context has the potential to revolutionize the way of understanding and optimizing the transmission of meaning in communication systems.

Satellite imagery. Satellite imagery, typically characterized by noise, sparsity, and a plethora of geophysical features on different spatial scales, presents unique challenges for data analysis. Currently, CNNs are the preferred choice for creating automated detection models that identify phenomena in satellite imagery. However, a CNN designed for a specific spatial scale might not be effective for images of different scales. While there are several methods that address multiscale processing, such as hierarchical CNNs (He et al., 2022) and vision transformers (Bazi et al., 2021), TDL offers an alternative approach to multiscale modeling. The use of topological invariants in this context can be particularly beneficial for several reasons. For example, topological invariants are not affected by geometric transformations such as translation, rotation, scaling, or shearing, making them highly robust and reliable for satellite imagery analysis (Hoef et al., 2023). This attribute of TDL can potentially enhance the accuracy and efficiency of analyzing complex satellite data. Dey et al. (2019) have used topological techniques to extract and analyze features from satellite images, showcasing the practical effectiveness of TDL in real-world scenarios.

Materials science. Using topological properties in experimental materials data analysis presents a useful approach to streamline the complex processing-structure-property relation in materials, as highlighted by Na et al. (2022); Papamarkou et al. (2022). Embedding material structure into topology-based supervised or unsupervised learning algorithms reveals crucial features in intricate processing-structure-property relations, for example, in datasets such as those found in high-entropy alloys (Spannaus et al., 2021). Taking advantage of topological crystallographic properties (Sunada, 2013), atom-specific machine learning models based on persistent homology have been developed, demonstrating improved accuracy in predicting the formation energy of crystalline compounds (Jiang et al., 2021). TDL may have the potential to uncover latent properties in quantum or other materials because it takes into account the underlying topological properties, such as symmetry, in the structure of the materials.