A Comprehensive Literature Review with Self-Reflection

Literature Review

October 8, 2025

Abstract

This literature review provides a comprehensive analysis of recent research in the field. The review synthesizes findings from 328 research papers, identifying key themes, methodological approaches, and future research directions.

Contents

1 Introduction

1.1 Evolution of Graph Representation Learning

The analysis of graph-structured data has historically presented a formidable challenge in machine learning, necessitating methods capable of discerning intricate relational patterns that underpin complex systems. Initially, approaches to graph analysis were rooted in traditional graph theory algorithms, such as shortest path computations, connectivity analysis, and various centrality measures, which provided foundational insights into network structure. Complementing these were statistical methods like spectral clustering ?, which leveraged the eigenvalues and eigenvectors of graph matrices to partition nodes, and graph kernels ??, which measured graph similarity by comparing substructures or paths. More recently, the concept of integrating graph kernels into GNNs has emerged to enhance interpretability and representation ability ?. Manifold learning methods, such as Laplacian Eigenmaps ?, further aimed to embed graph structures into lower-dimensional Euclidean spaces while preserving topological properties.

However, these pre-neural approaches, while valuable for specific tasks on smaller, static graphs, suffered from significant limitations. They often relied on extensive manual feature engineering, requiring domain expertise to define relevant graph properties. A major bottleneck was their high computational cost, frequently exhibiting polynomial complexity, such as $O(N^3)$ for many kernel methods or spectral decompositions on a graph with N nodes ?, rendering them impractical for large-scale graphs. Furthermore, these methods struggled inherently with generalization to unseen nodes or dynamic graph structures due to their transductive nature, and their capacity to capture complex, multi-hop structural information or integrate rich node features was constrained, leading to a bottleneck in extracting deeper insights from increasingly complex real-world graphs ?.

A crucial step towards automated feature extraction came with early graph embedding techniques, which sought to learn low-dimensional vector representations for nodes or entire graphs. Methods like matrix factorization (e.g., Singular Value Decomposition on adjacency matrices) provided a way to capture latent structural information. More

notably, random-walk based algorithms such as DeepWalk **?** and Node2Vec **?** offered a more scalable approach to learn node embeddings by simulating random walks and applying word2vec-like models. These were joined by other influential methods like LINE **?**, which aimed to preserve both first-order and second-order proximities. These techniques effectively captured proximity information within the graph, allowing for downstream tasks like node classification or link prediction by learning fixed embeddings for each node. However, these early embedding methods frequently operated in an unsupervised or semi-supervised manner, often overlooking explicit node attributes and struggling to capture the full spectrum of complex, multi-hop structural dependencies that are vital for many predictive tasks **?**. Their inherent limitations—particularly their transductive nature, meaning they cannot directly generalize to unseen nodes or dynamically evolving graphs without complete retraining **??**, insufficient capacity to leverage rich node features effectively, and limited ability for complex relational reasoning—highlighted a pressing need for a more powerful paradigm **?**.

This confluence of limitations necessitated a fundamental paradigm shift towards learning-based representations, where neural network models specifically designed for graphs emerged as a promising solution **?**. The conceptual genesis of Graph Neural Networks (GNNs) marked this pivotal moment, driven by the ambition to extend the success of deep learning from Euclidean data (like images and text) to the irregular, non-Euclidean structures inherent in graphs **?**. Pioneering works by Gori et al. **?** and Scarselli et al. **?** laid the theoretical foundations for this new class of models, proposing a framework to learn node representations by iteratively propagating and aggregating information across graph connections until a stable fixed-point was reached. This approach offered a powerful mechanism to automatically learn features and patterns directly from the graph topology and node attributes, moving beyond hand-crafted features or simple proximity embeddings. By enabling neural networks to process and aggregate information across graph connections, GNNs offered a powerful mechanism to overcome the constraints of prior methods, particularly in handling large-scale, complex, and dynamic graphs, and in capturing intricate, multi-hop structural dependencies. This foundational shift set the

essential context for the development and widespread adoption of GNNs, which form the core subject of this review.

1.2 Scope and Organization of the Review

This literature review offers a comprehensive and systematically organized exploration of Graph Neural Networks (GNNs), designed to provide readers with a coherent and pedagogically structured understanding of this rapidly evolving field. GNNs have emerged as a transformative paradigm in machine learning, adept at processing complex non-Euclidean data structures, leading to an explosion of research and applications across diverse domains ??. Given the remarkable breadth and depth of recent advancements, coupled with persistent challenges, this review aims to synthesize foundational concepts, delineate core architectural innovations, analyze theoretical underpinnings, address practical challenges, and highlight the diverse real-world applications of GNNs, culminating in a critical discussion of future directions and ethical considerations. The structured organization serves as a critical roadmap, guiding readers from the fundamental principles to the cutting-edge developments, thereby connecting disparate research threads into a cohesive narrative.

The review is meticulously structured to guide the reader through a logical progression, beginning with the fundamental principles and gradually advancing to more complex and specialized topics. This pedagogical approach ensures that both newcomers and experienced researchers can gain a holistic understanding of the field's trajectory and current state.

We commence in **Section 2: Foundational Concepts and Early GNN Models**, by establishing the essential graph theory basics and tracing the genesis of GNNs from their pioneering theoretical models to the development of the scalable message-passing paradigm. This initial section is crucial as it lays the groundwork, explaining *how* neural networks are fundamentally adapted to operate on graph structures, thereby setting the stage for all subsequent architectural advancements.

Building upon these foundations, **Section 3: Enhancing GNN Expressive Power and Theoretical Foundations** delves into the inherent capabilities and limitations of

GNNs. Here, we critically examine efforts to overcome the Weisfeiler-Leman test barrier, which restricts the discriminative power of many standard GNNs. We further explore the principles of geometric and equivariant GNNs, which are crucial for modeling physical systems by respecting inherent symmetries, and discuss the role of spectral graph theory in designing more powerful architectures. This section provides a rigorous understanding of *what* GNNs can model and *how* their discriminative power is theoretically enhanced, moving beyond empirical observations to principled design.

Recognizing that theoretical advancements must be validated through robust assessment, **Section 4: Evaluation and Benchmarking of Graph Neural Networks** is dedicated to the methodologies and frameworks for rigorously evaluating GNN models. We discuss the evolution of standardized evaluation protocols, metrics, and comprehensive benchmarking frameworks, such as those that have revealed the necessity of robust measures beyond simple accuracy **?**. This section critically underscores the community's drive towards reproducible research and fair comparisons, which are paramount for identifying truly impactful innovations and preventing misleading claims.

The review then transitions to the practical realities of deploying GNNs in **Section 5: Addressing Practical Challenges: Depth, Scalability, and Robustness**. This section tackles critical engineering and algorithmic hurdles, including strategies for building deeper GNNs that circumvent issues like oversmoothing and over-squashing **?**, techniques for scaling GNNs to massive graphs, and methods for enhancing robustness against structural heterogeneity and imperfect data. Furthermore, it explores advanced knowledge transfer paradigms like pre-training and prompt-based adaptation, which are vital for achieving data efficiency and generalization in real-world, often resource-constrained, scenarios.

A significant portion of this review is dedicated to the imperative of responsible AI, reflecting a critical shift in the field. **Section 6: Trustworthy GNNs: Explainability, Fairness, and Security** addresses the critical aspects of GNN trustworthiness. We explore methodologies for interpreting GNN decisions, enhancing their transparency, and discuss techniques to improve robustness against adversarial attacks, where many existing

defenses have been shown to be ineffective against adaptive adversaries **?**. Additionally, this section covers strategies to mitigate biases and protect sensitive information, ensuring fairness and privacy in GNN predictions, particularly relevant in distributed learning settings like Federated GNNs **?**. This dedicated focus highlights the growing importance of ethical considerations alongside performance.

The practical utility and transformative potential of GNNs are showcased in **Section 7: Applications of Graph Neural Networks**, which highlights their diverse and impactful real-world deployments. From revolutionizing recommender systems and accelerating scientific discovery in chemistry and materials science to analyzing time series and integrating with multi-modal data for semantic understanding, this section demonstrates *how* the theoretical and practical advancements discussed previously translate into tangible benefits across various industries and research fields.

Finally, **Section 8: Future Directions and Societal Impact** looks ahead, identifying pressing open challenges and promising research avenues that will shape the next generation of GNNs. This includes discussions on theoretical gaps, emerging architectural paradigms, and the integration of GNNs with other advanced AI techniques. We also critically examine the broader societal implications, advocating for responsible deployment and ethical considerations to ensure GNNs contribute positively to technological advancement. The review concludes in **Section 9** by synthesizing the major advancements and offering a forward-looking outlook on the field's trajectory.

By adopting this structured and progressive approach, this review aims to provide a comprehensive roadmap for researchers and practitioners. It not only details the "what" of GNN developments but critically examines the "why" behind their evolution and the "how" of their practical implementation, fostering a deeper understanding of their capabilities, limitations, and immense future potential.

2 Foundational Concepts and Early GNN Models

2.1 Graph Theory and Neural Network Basics

The analysis of structured data, particularly in the form of graphs, presents inherent challenges for conventional neural networks (NNs) which are primarily designed for processing Euclidean data with fixed-size, ordered inputs. Graph Neural Networks (GNNs) overcome this limitation by fundamentally integrating concepts from graph theory with the core principles of neural networks, thereby enabling models to learn rich representations directly from non-Euclidean graph structures **?**. This subsection establishes the foundational understanding of both graph theory and neural networks, elucidating how their conceptual fusion forms the bedrock of GNN architectures.

Graph theory provides the mathematical language for describing relationships and interactions within complex systems. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is formally defined by a set of nodes (or vertices) $\mathcal{V}$ and a set of edges $\mathcal{E}$ connecting pairs of nodes. Edges can be either undirected, signifying a symmetric relationship between two nodes, or directed, indicating an asymmetric flow of information or influence. Furthermore, graphs can be weighted, where each edge is associated with a numerical value representing the strength, cost, or capacity of the connection. The structural connectivity of a graph is typically represented by an adjacency matrix $\mathbf{A}$, where $\mathbf{A}\_ij = 1$ if an edge exists between node i and node j, and 0 otherwise (or the edge weight for weighted graphs). Beyond these basic types, real-world graphs often exhibit greater complexity, including heterogeneous graphs, where nodes and edges can belong to different types (e.g., users and items in a recommender system), and attributed graphs, where nodes and/or edges are endowed with feature vectors that provide additional descriptive information. A comprehensive understanding of these diverse structural properties is indispensable, as GNNs are explicitly designed to leverage this connectivity and associated features to derive meaningful insights.

Neural networks, conversely, are powerful computational models renowned for their ability to approximate complex functions and learn hierarchical representations from data. At their core, NNs consist of multiple layers, each performing a linear transformation

(a weighted sum of inputs, governed by learnable parameters or weights) followed by a non-linear activation function (e.g., ReLU, sigmoid, tanh). The network learns by iteratively adjusting these weights through an optimization process, typically stochastic gradient descent, guided by backpropagation to minimize a defined loss function. This layered architecture allows NNs to progressively extract abstract features from raw input data. However, the fundamental assumption of standard NNs—that inputs are fixed-size vectors with a predefined order—renders them incompatible with graph data. Graphs are inherently permutation-invariant (the order in which nodes are listed does not change the graph's structure) and variable-sized, making direct application of traditional NNs challenging.

The innovation of GNNs lies in adapting these fundamental neural network principles to operate directly on graph structures, thereby addressing the unique characteristics of non-Euclidean data ?. The concept of a neural network "layer" is reimagined in the graph domain. Instead of processing a flat vector, a GNN layer operates on node features within the context of their local graph neighborhood. The pioneering theoretical models, such as those proposed by ? and further formalized by ?, conceptualized GNNs as systems that iteratively propagate and aggregate information across the graph until node representations stabilize. This iterative process, often referred to as "message passing" or "neighborhood aggregation," forms the central mechanism of a generic GNN layer.

Abstractly, a GNN layer involves two primary steps for each node:

1. **Aggregation**: Each node gathers information (or "messages") from its immediate neighbors. This aggregation function must be permutation-invariant, meaning the order in which neighbors' features are collected does not affect the outcome. Common abstract aggregation operations might include summation, averaging, or taking the maximum of neighbor features.

2. **Update**: After aggregating information from its neighbors, a node combines this aggregated message with its own current feature representation. This combination typically involves a learnable transformation and a non-linear activation function, mirroring the operations within a standard neural network neuron. The output is

a new, enriched representation for the node.

This iterative application of aggregation and update steps across multiple layers allows GNNs to learn increasingly sophisticated and context-aware representations for each node, effectively encoding information from its multi-hop neighborhood into its feature vector. The parameters governing these aggregation and update functions within each graph layer are learnable and optimized end-to-end via backpropagation, just like in traditional neural networks. This adaptation enables GNNs to capture the intricate relational patterns and structural dependencies inherent in graph data, addressing the unique challenges posed by non-Euclidean structures and laying the groundwork for more advanced architectures discussed in subsequent sections.

2.2   The Genesis of Graph Neural Networks

The challenge of applying neural networks to non-Euclidean, graph-structured data necessitated a fundamental rethinking of traditional deep learning paradigms. The genesis of Graph Neural Networks (GNNs) can be traced back to pioneering efforts that sought to extend neural computation to arbitrary graph topologies, laying the theoretical groundwork for learning node and graph representations through iterative information propagation. These early conceptualizations established the core principle of aggregating local neighborhood information to derive stable, context-aware embeddings.

The earliest formal proposal for a "neural network for graphs" emerged with the work of ?. This seminal paper introduced a model based on a state-transition system, where each node in a graph maintains a state that is iteratively updated by aggregating information from its neighbors and its own previous state. The core idea was to propagate information across the graph until the node states reached a stable fixed-point, effectively encoding the structural and feature information of the entire neighborhood into each node's representation. While theoretically elegant, the computational demands of explicitly solving for a fixed-point limited its immediate widespread adoption.

Building upon this foundational concept, ? provided a comprehensive formalization of the Graph Neural Network (GNN) model. This work rigorously defined GNNs as

an extension of recursive neural networks, establishing a mathematical framework for learning functions on graphs. It formalized the iterative update process for node states, demonstrating that under certain conditions (e.g., using a contraction mapping), a unique fixed-point solution for node embeddings exists. The model leveraged a recurrent neural network-like structure to compute the node states, emphasizing the iterative exchange of messages between connected nodes. This formalization provided a robust theoretical basis, proving the GNN's universal approximation capabilities for functions on graphs, yet the practical challenges of training and scaling these fixed-point iterations remained a significant hurdle.

While the fixed-point iteration models provided the theoretical bedrock, early practical attempts to apply "convolutional" ideas to graphs also began to emerge, embodying the spirit of local information aggregation. ? introduced one of the earliest practical "convolutional" approaches for graphs, specifically tailored for learning molecular fingerprints. This model adapted the concept of summing features from a node's neighbors, followed by a non-linear transformation, to generate node embeddings. Although it departed from the strict fixed-point iteration of the earlier GNN models, it crucially demonstrated the effectiveness of learning representations by aggregating local information in a layer-wise fashion. This work served as a significant bridge, translating the theoretical underpinnings of GNNs into a more computationally tractable, albeit simplified, message-passing paradigm, paving the way for later, more sophisticated graph convolutional architectures.

In conclusion, the genesis of Graph Neural Networks was marked by a profound theoretical leap, establishing the principle of learning representations through iterative, local information propagation. The foundational models by ? and ? provided the mathematical rigor and conceptual framework, defining GNNs as fixed-point iterations of recursive neural networks. While computationally intensive, these efforts laid down the essential blueprint. Early practical instantiations, such as the molecular fingerprinting approach by ?, further demonstrated the viability of local aggregation. These pioneering works, despite their initial computational limitations, firmly established the core idea of learning on graphs through message passing, setting the stage for the subsequent explosion of

11

more scalable and expressive GNN architectures. The challenge of efficiently computing stable fixed-points and scaling these models to large graphs remained a key area for future development.

2.3   Message Passing Paradigm and Inductive Learning

The evolution of Graph Neural Networks (GNNs) marked a pivotal shift towards scalable and effective architectures rooted in the message-passing paradigm, fundamentally transforming how neural networks process graph-structured data. This paradigm, where nodes iteratively aggregate information from their neighbors, enabled GNNs to move beyond early theoretical models towards practical applications by defining a localized and computationally efficient mechanism for learning node representations.

A cornerstone of this transformation was the introduction of Graph Convolutional Networks (GCNs) by ?. This seminal work provided a highly efficient, localized, first-order approximation of spectral graph convolutions. Prior spectral methods, such as ChebyNet ?, relied on computationally intensive eigen-decomposition of the graph Laplacian or polynomial filters, limiting their applicability to large graphs. GCNs ingeniously simplified this by restricting the convolutional filter to a first-order neighborhood and introducing a self-loop, effectively propagating and transforming node features through a layer-wise aggregation of neighbor information. This simplification made graph convolutions practical and scalable, establishing GCNs as a foundational baseline for semi-supervised node classification. From a unified optimization perspective, GCNs' propagation mechanism can be interpreted as an optimal solution to an objective function that balances feature fitting with graph Laplacian regularization, essentially performing a low-pass filtering operation on node features ?. However, GCNs were primarily designed for transductive settings, requiring the full adjacency matrix during training, which limited their ability to generalize to unseen nodes or scale to very large, dynamic graphs.

Addressing these critical limitations, ? introduced GraphSAGE (SAmple and aggreGatE), a framework specifically designed for inductive representation learning on large graphs. GraphSAGE innovated by proposing an efficient neighbor sampling strategy, al-

lowing models to learn aggregation functions that generalize to unseen nodes and entire graphs without requiring the full graph structure at training time. By learning how to aggregate information from a sampled, fixed-size set of neighbors, GraphSAGE enabled the generation of node embeddings for new nodes by simply running the learned aggregation functions. This represented a significant step towards practical scalability and applicability in dynamic graph environments, where new nodes are constantly added. Subsequent work, such as GNNAutoScale ?, has continued to build on these principles, developing frameworks to scale arbitrary message-passing GNNs to even larger graphs while provably maintaining expressive power by leveraging historical embeddings.

Further enhancing the flexibility and expressive power of the message-passing paradigm, ? developed Graph Attention Networks (GATs). GATs introduced learned attention weights, allowing each node to assign varying importance to its neighbors during the aggregation process, rather than relying on fixed or uniformly weighted averages as in earlier GCN variants. This self-attention mechanism enables GATs to capture more nuanced relationships between nodes, providing greater flexibility in information aggregation and improving the model's ability to generalize to unseen nodes and graph structures by adaptively focusing on relevant neighbors. While GATs provided enhanced flexibility and often improved performance, this came at the cost of increased computational complexity and memory requirements during training and inference compared to simpler models like GCNs. The versatility of attention-based aggregation has been widely adopted, extending beyond traditional graph tasks to areas like robust feature matching in computer vision, as demonstrated by SuperGlue ?.

While these models significantly advanced the practical utility of GNNs, their theoretical expressive power became a subject of critical investigation. ? provided a fundamental analysis, demonstrating that many popular message-passing GNN variants, including GCNs, GraphSAGE, and GATs, are at most as powerful as the 1-Weisfeiler-Lehman (1-WL) graph isomorphism test. This work highlighted inherent limitations in their ability to distinguish certain complex non-isomorphic graph structures. To overcome this, the same paper introduced Graph Isomorphism Networks (GINs), which were provably shown to

be maximally powerful within the message-passing framework by employing a learnable, injective aggregation function (e.g., a multi-layer perceptron). Despite GINs matching the 1-WL test's discriminative power, even these advanced GNNs still face representational limits; for instance, they cannot compute fundamental graph properties such as girth, diameter, or the number of k-cliques ?.

In conclusion, the message-passing paradigm, spearheaded by GCNs, GraphSAGE, GATs, and GINs, marked a transformative period in GNN research, enabling scalable, inductive, and expressive learning on graph-structured data. The discovery of the Weisfeiler-Lehman test's upper bound on expressivity, and the subsequent development of GINs to meet this bound, catalyzed a new line of inquiry focused on designing architectures that could systematically transcend the 1-WL test barrier. This ongoing quest for higher discriminative power, alongside efforts to address practical challenges like scalability and robustness, forms the core of the advancements explored in the subsequent sections.

3   Enhancing GNN Expressive Power and Theoretical Foundations

3.1   Overcoming the Weisfeiler-Leman Barrier

Standard message-passing Graph Neural Networks (GNNs) have demonstrated remarkable empirical success across various domains, yet their theoretical discriminative power is fundamentally limited. A cornerstone of GNN theory establishes that many popular GNN architectures, such as Graph Convolutional Networks (GCNs) and GraphSAGE, are at most as powerful as the 1-Weisfeiler-Leman (1-WL) graph isomorphism test ??. This equivalence implies that these GNNs cannot distinguish between certain non-isomorphic graphs or capture fine-grained structural properties like triangle counts, girth, or diameter ???. This limitation restricts their ability to perform tasks requiring a deeper understanding of graph topology, such as substructure counting ?. Consequently, a significant research thrust has focused on developing novel GNN architectures and techniques to sur-

pass this 1-WL barrier, enhancing their expressive power for more complex graph learning tasks.

One direct approach to overcome the 1-WL limitation involves extending GNNs to operate on higher-order graph structures, mirroring the k-dimensional Weisfeiler-Leman (k-WL) tests. **?** pioneered this direction by introducing **k-GNNs**, which perform message passing on k-element subsets (k-tuples) of nodes rather than individual nodes. These k-GNNs are theoretically proven to be strictly more powerful than 1-GNNs, enabling them to distinguish graphs that confound the 1-WL test and capture higher-order motifs. While powerful, k-GNNs typically incur significantly higher computational and memory costs, growing exponentially with k, which limits their practical applicability for larger graphs. An alternative strategy to implicitly capture higher-order information efficiently is seen in GNNML3, which designs graph convolution supports in the spectral domain using non-linear functions of eigenvalues, allowing it to experimentally match the expressive power of 3-WL equivalent models while maintaining linear computational complexity after initial preprocessing **?**.

Another category of methods enhances GNN discriminative power by explicitly leveraging path or substructure information. Early work, such as the SEAL framework for link prediction, demonstrated the utility of learning high-order features from local enclosing subgraphs, supported by a β-`decaying heuristic theory` that justifies approximating complex heuristics from local contexts **?**. Building on the need for substructure awareness, **?** introduced Local Relational Pooling (LRP) to enable GNNs to effectively count substructures, a task proven to be beyond the capabilities of standard Message Passing Neural Networks (MPNNs) for many connected patterns. More recently, **Path Neural Networks (PathNNs)** were proposed, which update node representations by aggregating information from various paths, including single shortest, all shortest, and all simple paths **?**. By operating on "annotated sets of paths," PathNNs are shown to be strictly more powerful than 1-WL and can even distinguish graphs that are indistinguishable by the 3-WL algorithm. Similarly, **Substructure Aware Graph Neural Networks (SAGNNs)** inject subgraph-level structural information, derived from novel "Cut sub-

graphs" and random walk return probability encodings, into node features, proving to be strictly more powerful than 1-WL and achieving state-of-the-art results on tasks requiring fine-grained structural understanding ?. For efficiency, ELPH and BUDDY introduce "subgraph sketching," using compact, node-wise representations of neighborhoods to approximate structural features like distance-based node labels, thereby achieving the expressivity of subgraph GNNs (e.g., for triangle counting and distinguishing automorphic nodes) within a more efficient full-graph GNN framework ?.

A third set of techniques injects unique identifiers or richer structural context to boost GNN discriminative power. A surprising finding by ? demonstrated that standard MPNNs, when augmented with **Random Node Initialization (RNI)**, become universal approximators for functions on graphs of a fixed order, effectively breaking the 1-WL barrier without the prohibitive cost of higher-order GNNs. This randomization individualizes nodes, allowing the GNN to distinguish otherwise isomorphic local structures. To provide more stable and learnable positional information, **Learnable Structural and Positional Encodings (LSPE)** decouple and simultaneously learn both structural and positional representations throughout the GNN layers ?. LSPE, often initialized with Random Walk Positional Encoding (RWPE), overcomes the sign ambiguity issues of Laplacian eigenvectors and significantly enhances expressivity. Complementing this, **Positional Encoding GNN (PEG)** uses separate channels for original node features and positional features (e.g., Laplacian Eigenmaps), ensuring O(p) equivariance for positional features and provable stability against graph perturbations, which is crucial for reliable discrimination in complex graphs ?. Beyond abstract graph structures, **E(n) Equivariant Graph Neural Networks (EGNNs)** directly incorporate geometric structural context by updating node coordinates within the message-passing framework, preserving E(n) equivariance efficiently without relying on computationally expensive higher-order representations ?. This is critical for tasks in physics and chemistry, as exemplified by NequIP, which uses E(3)-equivariant convolutions for interatomic potentials, achieving high accuracy with remarkable data efficiency ?. More dynamically, **Cooperative Graph Neural Networks (CO-GNNs)** allow nodes to adaptively choose their communication actions

(listen, broadcast, isolate), leading to dynamic and asynchronous message passing that is theoretically shown to be more expressive than 1-WL due to the variance introduced by action sampling **?**. Finally, **Spatio-Spectral Graph Neural Networks (S2GNNs)** combine local spatial message passing with global spectral filtering, inherently mitigating over-squashing and providing "free" positional encodings from partial eigendecomposition, making them strictly more expressive than 1-WL and highly effective for long-range interactions **?**.

In conclusion, the journey to overcome the Weisfeiler-Leman barrier has led to a diverse array of innovative GNN architectures. While k-GNNs offer a direct theoretical extension, their computational cost remains a challenge. More efficient solutions often involve injecting rich structural context through explicit path or substructure information, or by providing unique node identifiers and learnable positional encodings. Recent advancements explore dynamic message-passing paradigms and hybrid spatial-spectral approaches, which offer novel ways to enhance discriminative power and capture long-range dependencies. However, the trade-off between increased expressivity, computational efficiency, and generalizability to unseen graph structures remains an active area of research, pushing the boundaries of GNN design for fine-grained structural understanding.

3.2   Geometric and Equivariant GNNs

Graph Neural Networks (GNNs) are increasingly applied to tasks involving 3D structures like molecules, proteins, and physical systems, where respecting geometric symmetries and transformations is paramount. This section explores the development of GNNs specifically designed to incorporate these symmetries, leading to more data-efficient, robust, and physically consistent models.

Early efforts to enforce geometric symmetries often relied on complex higher-order representations or were limited to 3D spaces, posing challenges for scalability and efficiency. A significant advancement in this area is the introduction of E(n) Equivariant Graph Neural Networks (EGNNs) by **?**. The core innovation of EGNNs lies in their Equivariant Graph Convolutional Layer (EGCL), which directly updates node coordinates using

relative differences and aggregates messages based on squared relative distances, thereby preserving E(n) equivariance for both scalar features and vector positions without resorting to computationally expensive spherical harmonics. This simpler and more efficient architecture has demonstrated superior performance in tasks such as N-body system simulations, significantly outperforming prior methods like SE(3) Transformer and Tensor Field Networks. Building upon this foundation, **?** introduced Neural Equivariant Interatomic Potentials (NequIP), an E(3)-equivariant neural network that learns interatomic potentials for molecular dynamics simulations. NequIP leverages E(3)-equivariant convolutions for interactions of geometric tensors, achieving state-of-the-art accuracy and remarkable data efficiency, often requiring orders of magnitude less training data than other models, which is crucial for scientific applications using high-fidelity quantum chemical data.

While these architectural innovations proved highly effective, a comprehensive theoretical framework to characterize the expressive power of geometric GNNs was lacking, similar to the Weisfeiler-Leman (WL) test for non-geometric graphs **??**. To address this, **?** developed the Geometric Weisfeiler-Leman (GWL) test, a symmetry-aware generalization of the WL test for geometric graphs. This framework provides a theoretical upper bound for the expressivity of geometric GNNs and offers insights into the distinct roles and power of invariant versus equivariant layers and higher-order tensors, thereby guiding the principled design of more powerful architectures.

Further enhancing the robustness and expressivity of geometric GNNs, the stability of positional encodings (PEs) became a critical area of focus. Existing PE methods often suffered from instability, particularly when dealing with graph automorphisms or small eigengaps, and lacked guaranteed permutation equivariance. **?** tackled this by proposing the PEG architecture, which employs separate channels for node and positional features and imposes O(p) equivariance for positional features. This approach achieves provable stability by depending on a larger eigengap, ensuring that the positional information used by geometric GNNs is both consistent and reliable.

Despite these advancements, a fundamental limitation of spatial Message Passing

GNNs (MPGNNs) in geometric contexts is the "over-squashing" phenomenon, which restricts long-range information exchange and limits their receptive field and expressive power. To overcome this, ? introduced Spatio-Spectral Graph Neural Networks (S2GNNs), a novel hybrid architecture that synergistically combines local spatial message passing with global spectral filtering. S2GNNs are provably capable of vanquishing over-squashing, offer superior approximation bounds, and provide stable positional encodings "for free," representing a significant step towards more powerful and robust geometric GNNs for complex 3D structures.

In conclusion, the field of geometric and equivariant GNNs has rapidly evolved from practical architectural designs that enforce E(n) equivariance ? and demonstrate their data efficiency in scientific domains ?, to establishing rigorous theoretical foundations for their expressive power ?. Subsequent work has addressed crucial aspects like stable positional encoding ? and fundamental architectural limitations such as over-squashing ?. While significant progress has been made in building physically consistent and data-efficient models, ongoing challenges include further scaling these computationally intensive models to extremely large and dense geometric graphs, and developing more sophisticated hybrid architectures that can seamlessly integrate diverse geometric priors and multi-scale information.

3.3   Spectral Graph Neural Networks

Spectral Graph Neural Networks (GNNs) represent a foundational class of architectures that draw inspiration from spectral graph theory, analyzing graph signals in the frequency domain to capture global graph properties and long-range dependencies. This approach leverages the eigen-decomposition of the graph Laplacian, where eigenvectors form an orthonormal basis for graph signals, and eigenvalues correspond to frequencies, enabling principled filtering operations.

Early spectral convolutions, such as those pioneered by Bruna et al. ?, directly applied filters in the spectral domain. While theoretically sound, these methods faced significant practical limitations: the prohibitive computational complexity of eigen-decomposition

(typically $O(N^3)$ for dense graphs or $O(N^2)$ for sparse graphs) for large graphs, and the non-transferability of the learned spectral basis to graphs with different structures. To address these challenges, simplified polynomial approximations of spectral filters emerged. The most influential of these is the Graph Convolutional Network (GCN) by Kipf and Welling **?**. GCNs approximate spectral filters with low-order polynomials, achieving remarkable computational efficiency and localization in the node space. Crucially, this spectral simplification translates directly into the widely adopted spatial message-passing update rule, establishing a profound duality between spectral theory and practical spatial GNNs.

While effective, these simplified polynomial filters inherently act as low-pass filters, leading to a critical challenge: over-smoothing. As GNNs deepen, node representations become increasingly indistinguishable, converging to a non-informative constant, which severely limits their expressive power for complex tasks **?**. The theoretical underpinnings of this phenomenon were rigorously explored by Cai et al. **?**, who demonstrated that the Dirichlet energy of node embeddings, a measure of their discriminative power in the frequency domain, exponentially converges to zero with increasing layers. Beyond over-smoothing, another related but distinct challenge is *over-squashing*, where the hierarchical aggregation of information through local neighborhoods leads to an exponential shrinkage of the effective receptive field, causing information from distant nodes to become diluted or bottlenecked due to the tree-like structure of message passing **?**. Spectral methods, by their nature, can offer a principled way to mitigate over-squashing by providing a global perspective on information flow, effectively creating "shortcuts" in the frequency domain that bypass local topological bottlenecks. Architectural solutions to over-smoothing and over-squashing, such as decoupling propagation from transformation or using skip connections, are further discussed in Section 5.1.

To overcome the limitations of simple polynomial filters and enhance the frequency response, research has branched into several directions. Bianchi et al. **?** introduced Graph Neural Networks with Convolutional ARMA Filters. This work proposed a novel graph convolutional layer based on Auto-Regressive Moving Average (ARMA) filters, which offer

a more versatile class of rational filters capable of modeling a wider variety of frequency responses and capturing longer dynamics with fewer parameters than traditional polynomial filters. By approximating the ARMA filter through a recursive update rule with skip connections, they mitigated over-smoothing and maintained computational efficiency.

Further theoretical investigations have delved into the expressive power of spectral GNNs. Wang et al. ? rigorously analyzed linear spectral GNNs, demonstrating that even without non-linearities, they can achieve universal approximation under specific conditions (e.g., no multiple eigenvalues, no missing frequency components). Their work connects these universality conditions to the discriminative power of the 1-Weisfeiler-Leman test, bridging spectral expressivity with graph isomorphism testing. They also provide an optimization perspective on why different polynomial bases, despite having the same expressive power, lead to varying empirical performance, advocating for orthogonal bases like Jacobi polynomials (used in their proposed JacobiConv) for faster convergence. This highlights that the choice of filter basis is not merely an implementation detail but impacts both theoretical guarantees and practical performance. From a generalization perspective, Wang et al. ? provided a manifold perspective, proving that generalization bounds of GNNs in the spectral domain decrease with graph size and increase with the spectral continuity constants of filter functions, offering insights into practical GNN design.

More recent advancements have sought to synergistically combine the strengths of local spatial message passing with global spectral filtering, or to design more adaptive and powerful spectral filters. Spatio-Spectral Graph Neural Networks (S2GNNs) ? exemplify this trend, proposing architectures that operate directly in the spectral domain while incorporating local spatial information. These methods leverage eigen-decomposition to provide a principled way to capture global graph properties and enhance expressivity for long-range interactions, directly mitigating the over-squashing problem by enabling global information flow.

Beyond fixed filter designs, the field is moving towards more flexible and adaptive spectral approaches. Bo et al. ? introduced Specformer, which encodes the set of all eigenvalues and performs self-attention in the spectral domain, leading to a learn-

able set-to-set spectral filter. This approach moves beyond scalar-to-scalar functions and fixed-order polynomials, offering greater expressiveness and flexibility for non-local graph convolution. Similarly, for specific tasks like anomaly detection, Tang et al. **?** proposed the Beta Wavelet Graph Neural Network (BWGNN) with spectral and spatial localized band-pass filters. This design specifically addresses the "right-shift" phenomenon observed in anomalous graphs, where spectral energy concentrates more on high frequencies, demonstrating the utility of tailored spectral filters. Han et al. **?** further pushed the adaptivity, proposing Node-wise Filtering via Mixture of Experts (NODE-MOE) to dynamically apply distinct filters to individual nodes based on their specific structural patterns (homophilic or heterophilic). While NODE-MOE is a general adaptive filtering framework, its motivation to overcome the limitations of uniform global filters for mixed graph patterns is highly relevant to designing more versatile spectral filters that can adjust to local graph characteristics.

Novel mathematical foundations are also being explored to redefine signal propagation in the frequency domain. Kang et al. **?** introduced the FRactional-Order graph Neural Dynamical network (FROND) framework, which generalizes continuous GNNs by incorporating Caputo fractional derivatives. This allows for modeling non-local, memory-dependent dynamics in graph signals, leading to an algebraic rate of convergence to stationarity, which inherently mitigates over-smoothing more effectively than traditional integer-order diffusion processes.

Addressing the critical scalability challenge of eigen-decomposition for large graphs, Li et al. **?** proposed GloGNN. While not strictly a spectral GNN in the traditional sense, GloGNN achieves spectral-like effects by learning a signed coefficient matrix for global aggregation. This allows it to implicitly combine low-pass and high-pass filtering effects to handle heterophily, and crucially, it achieves linear time complexity for global aggregation, circumventing the computational bottleneck of explicit eigen-decomposition for large graphs. This represents a pragmatic solution to achieve global information capture with spectral-like properties at scale.

In conclusion, spectral GNNs have evolved from direct but computationally intensive

spectral filtering to efficient polynomial approximations, and further to more expressive rational filters, learnable set-to-set filters, and sophisticated hybrid architectures. These advancements, including S2GNNs, fractional calculus-based models, and adaptive filtering mechanisms, leverage the frequency domain to enhance expressivity for long-range interactions and provide a principled way to capture global graph properties, directly addressing challenges like over-smoothing and over-squashing. Despite these advancements, persistent challenges remain in developing truly adaptive spectral filters that can dynamically adjust to varying local graph structures and heterophily levels, and in addressing the fundamental scalability and transferability issues of explicit eigen-decomposition for very large and dynamic graphs. Future research will likely focus on more efficient approximations, hybrid spatial-spectral designs, and novel mathematical frameworks that offer the benefits of spectral analysis without its computational drawbacks.

4 Evaluation and Benchmarking of Graph Neural Networks

4.1 Standardized Evaluation Protocols and Metrics

The burgeoning field of Graph Neural Networks (GNNs) critically relies on rigorous and standardized evaluation protocols to ensure reliable comparisons, foster reproducible research, and accurately assess model generalization capabilities. The unique characteristics of graph-structured data, such as varying topologies, non-i.i.d. (independent and identically distributed) node features, and the inherent complexity of defining ground truth for structural patterns, pose significant challenges to consistent evaluation, often leading to inconsistent experimental setups and potentially misleading performance claims ?.

Early research underscored the necessity for statistical rigor in GNN evaluation. ? notably emphasized a robust experimental protocol for semi-supervised node classification. Their methodology involved conducting 100 runs across multiple random data splits and initializations, employing a fixed visible/test set split, optimizing hyperparameters

on a dedicated validation set, utilizing early stopping, and crucially, performing statistical significance testing (e.g., bootstrapping for confidence intervals and paired t-tests for p-values). This meticulous approach revealed that many reported performance gains could "vanish" under careful scrutiny, highlighting the importance of robust statistical validation over single-run accuracy reports. Similarly, **?**, while primarily investigating GNN expressive power, implicitly demonstrated the need for consistent empirical validation on diverse graph datasets to substantiate theoretical claims, advocating for a principled approach to experimental design.

A cornerstone of standardized evaluation lies in the experimental protocols for data splitting and validation. For graph data, a primary distinction exists between **transductive** and **inductive** settings. In a transductive setup, typically used for semi-supervised node classification, the entire graph structure is known during training, but only a subset of node labels is available. Evaluation then occurs on the unlabeled nodes within this *known* graph. Conversely, inductive evaluation, common in graph classification or link prediction involving unseen nodes/graphs, assesses a model's ability to generalize to *entirely new* graph structures or nodes not observed during training. The choice of splitting strategy is paramount; for instance, a simple random split of nodes might be appropriate for transductive node classification, but for graph classification, a split at the graph level is required. Challenges arise in ensuring that splits do not inadvertently leak information or create unrealistic evaluation scenarios. For example, in link prediction, ensuring that test edges are truly unseen and that negative samples are challenging and realistic is critical **?**.

Cross-validation, a standard practice in machine learning, also presents unique challenges for graph data. Traditional k-fold cross-validation can be problematic if graph connectivity is not carefully considered, potentially leading to information leakage between folds or breaking structural integrity. Therefore, repeated random splits, often with a fixed ratio for training, validation, and testing, are commonly employed, especially in transductive node-level tasks, to mitigate variance and provide more stable performance estimates. For graph-level tasks, k-fold cross-validation can be applied to the set

of graphs, ensuring each graph appears in the test set.

Beyond protocols, a comprehensive understanding of **evaluation metrics** is essential for interpreting GNN performance across different tasks:

* **Node Classification:** Common metrics include **Accuracy**, **Precision**, **Recall**, and **F1-score**. For datasets with significant class imbalance, raw accuracy can be misleading. In such cases, **Macro-F1** (averaging F1-scores per class) or **Balanced Accuracy** (average of recall for each class) are preferred over **Micro-F1** (calculating F1 globally) or overall accuracy, as they provide a more equitable assessment of performance across all classes **??**. **Area Under the Receiver Operating Characteristic curve (AUC-ROC)** and **Area Under the Precision-Recall Curve (AUC-PRC)** are also widely used, particularly for binary classification or highly imbalanced datasets, as they are less sensitive to class distribution. * **Link Prediction:** The primary goal is to predict the existence of edges between nodes. Metrics like **AUC-ROC**, **AUC-PRC**, and **Average Precision (AP)** are standard. A critical aspect here is the strategy for **negative sampling**. **?** highlighted that unrealistic or overly simple negative sampling can inflate reported performance. They proposed a standardized methodology and a novel Heuristic Related Sampling Technique (HeaRT) to generate more challenging and realistic negative samples, ensuring more reliable evaluations for link prediction models. * **Graph Classification/Regression:** For classification tasks, metrics similar to node classification (Accuracy, F1-score) are used, but applied at the graph level. For graph-level regression tasks, such as predicting molecular properties or power flow, **Mean Squared Error (MSE)**, **Mean Absolute Error (MAE)**, and **R2 score** are standard measures of prediction error **?**.

The drive for standardized evaluation extends to broader trustworthiness aspects of GNNs, necessitating specialized metrics and protocols:

Explainability: *The field of GNN explainability has historically suffered from a "lack of standardization," including inconsistent evaluation metrics and testbeds* **?**. *Efforts are underway to define metrics that assess the* faithfulness *(how accurately an explanation reflects the model's internal reasoning) and* plausibility* (how understandable and intuitive

an explanation is to a human). ***Robustness:*** *Evaluating GNN robustness against adversarial attacks requires a systematic methodology.* **?** *critically revealed that many GNN defenses are evaluated against weak, non-adaptive attacks, leading to overly optimistic robustness estimates. They proposed a systematic methodology for designing strong* adaptive attacks* (e.g., Meta-PGD) to truly test defenses, emphasizing that the choice of attack methodology is itself a crucial part of standardized evaluation for robustness. Metrics include attack success rate, perturbation budget, and the degradation of predictive performance under attack. * **Fairness:** Addressing bias in GNNs necessitates specific fairness metrics. **?** noted that "Accountability" in trustworthy GNNs requires "different evaluation standards and graph-based metrics." Metrics like **Statistical Parity Difference** (ΔSP) and **Equal Opportunity Difference** (ΔEO) are used to quantify disparities in outcomes across sensitive attributes (e.g., gender, race) in graph data, particularly in the presence of "distribution disparities" and "attribute imbalance" **?**.

Despite significant progress in establishing rigorous protocols and defining relevant metrics, challenges persist. The continuous evolution of GNN architectures and the increasing complexity of real-world graph data (e.g., dynamic graphs, extremely large-scale graphs) demand adaptable and extensible evaluation methodologies. Furthermore, the development of universally accepted metrics for complex tasks like graph generation, interpretability, and causality remains an active area of research, requiring a careful balance between theoretical soundness and practical applicability. The field continues to move towards more holistic and adaptive evaluation methodologies that encompass not just predictive performance but also crucial aspects of trustworthiness, fairness, and robustness, ensuring GNNs are not only powerful but also reliable and responsible for real-world deployment.

4.2   Comprehensive Benchmarking Frameworks and Datasets

The burgeoning field of Graph Neural Networks (GNNs) has witnessed an explosion of novel architectures, yet the absence of standardized evaluation protocols and robust, discriminative datasets historically hampered fair comparisons and the identification of truly

impactful advancements. Early GNN research often relied on a handful of small, homogeneous benchmark datasets like Cora, CiteSeer, and MUTAG. While foundational, these datasets frequently led to saturated performance, making it difficult to discern the genuine strengths of complex GNN models over simpler baselines or even graph-agnostic methods ?. This critical need for common ground and consistent experimental validation spurred the development of comprehensive benchmarking frameworks.

A pivotal moment in standardizing GNN evaluation was the introduction of the **Open Graph Benchmark (OGB) ?**. OGB marked a significant shift by providing a diverse collection of large-scale, real-world graph datasets spanning various domains, including molecular graphs, academic networks, and social networks. It standardized tasks for node classification, link prediction, and graph classification, offering unified data splits, evaluation metrics, and leaderboards. OGB's primary motivation was to move beyond the limitations of small datasets, enabling researchers to assess GNN scalability, generalization capabilities, and performance on more challenging, realistic graph structures. Its widespread adoption has been instrumental in fostering reproducible research and accelerating progress by providing a consistent and robust platform for evaluating new GNN models.

Complementing OGB's focus on large-scale, real-world challenges, other frameworks have emerged to address specific aspects of GNN evaluation. For instance, ? introduced a modular, open-source benchmarking framework built on PyTorch and DGL. This framework specifically targeted the need for controlled architectural comparisons by proposing standardized experimental settings, including fixed parameter budgets (e.g., 100k and 500k parameters). This crucial constraint ensures that observed performance differences are primarily attributable to architectural design rather than varying model capacity. The framework incorporates a diverse collection of 12 medium-scale datasets, encompassing both real-world graphs (e.g., ZINC, AQSOL, OGB-COLLAB, WikiCS) and synthetic mathematical graphs (e.g., PATTERN, CLUSTER, CSL, CYCLES). The inclusion of synthetic graphs is particularly valuable as they are designed to test specific theoretical graph properties and reveal the discriminative power of GNNs, thereby facilitating the

discovery of fundamental components like Graph Positional Encoding (PE) using Laplacian eigenvectors. This framework thus provides a complementary perspective to OGB, focusing on rigorous, controlled experiments to probe the theoretical underpinnings and design principles of GNNs.

Beyond general benchmarking, specific GNN tasks often necessitate tailored evaluation methodologies to reflect real-world complexities accurately. For link prediction, **?** critically examined existing evaluation practices, highlighting how issues like underreported baseline performance, inconsistent data splits, and unrealistic negative sampling strategies led to unreliable comparisons. To address these, they proposed a standardized and reproducible benchmarking methodology, introducing the **Heuristic Related Sampling Technique (HeaRT)**. HeaRT generates more challenging and realistic negative samples by restricting them to "corruptions" that share one node with the positive sample, making the link prediction task less trivial and more reflective of real-world scenarios. This work underscores that robust evaluation for specific tasks requires careful consideration of the inherent challenges and the design of evaluation protocols that genuinely test model capabilities. Similarly, for the critical problem of GNNs under label noise, **?** introduced **NoisyGL**, the first comprehensive benchmark specifically designed for this challenging setting. NoisyGL provides unified experimental settings and interfaces, enabling fair comparisons and deeper analysis of various Graph Neural Networks under Label Noise (GLN) methods, thereby fostering advancements in this specialized area.

Furthermore, the development of domain-specific datasets with unique characteristics is crucial for robust model assessment and for revealing the true strengths and weaknesses of GNNs in specialized applications. In network neuroscience, **?** introduced **BrainGB**, a benchmark tailored for brain network analysis with GNNs. This benchmark addresses challenges unique to brain networks, such as the lack of useful initial node features, the presence of real-valued and signed connection weights, and fixed ROI identities, providing standardized preprocessing pipelines and a modular GNN design space. Likewise, **?** introduced **PowerGraph**, a GNN-tailored benchmark dataset for electrical power grid applications. PowerGraph is comprehensive, covering node-level tasks (e.g., power flow,

optimal power flow) and graph-level tasks (e.g., cascading failure analysis) across real-world-based power grids. A key innovation of PowerGraph is its provision of empirical ground-truth explanations (e.g., cascading edges for failures), which is critical for benchmarking GNN explainability methods and gaining deeper insights into model behavior, moving beyond mere predictive accuracy.

The increasing emphasis on trustworthiness in AI has also driven the need for benchmarks that evaluate aspects beyond predictive performance. Evaluating GNN explainability, for instance, has been challenging due to the lack of reliable ground-truth explanations in existing datasets. To address this, **?** introduced **GRAPHXAI**, a general-purpose framework that includes **SHAPEGG EN**, a novel synthetic graph generator capable of creating diverse datasets with guaranteed reliable ground-truth explanations. GRAPHXAI provides a comprehensive suite of metrics for evaluating Graph Explanation Accuracy, Faithfulness, Stability, and crucially, Fairness, revealing significant limitations of existing explainers on heterophilic graphs, large explanations, and fairness properties. This highlights the need for dedicated benchmarks to assess the quality and ethical implications of GNN explanations. Similarly, in the realm of robustness, **?** critically evaluated GNN defenses, demonstrating that most existing defenses offer only marginal improvement against *adaptive attacks*. While not a benchmark framework itself, this work underscores the necessity of rigorous, adversarial benchmarking methodologies to truly assess the robustness of GNNs against sophisticated threats. These efforts align with the broader call for trustworthy GNNs, encompassing privacy, robustness, fairness, and explainability, as comprehensively surveyed by **?**, which emphasizes the interconnectedness of these dimensions.

Collectively, these comprehensive benchmarking frameworks and specialized datasets are instrumental in fostering a more scientific and reproducible research environment for GNNs. They provide the common ground and consistent protocols necessary for fair comparisons, allowing researchers to identify true architectural advancements and understand the nuanced strengths and weaknesses of different GNN models across diverse graph structures and tasks. However, the continuous evolution of GNNs presents ongoing

29

challenges for benchmarking. Future efforts must focus on expanding these benchmarks to encompass even larger-scale, dynamic/temporal graphs, multi-modal data integration, and out-of-distribution generalization capabilities. Crucially, as GNNs are deployed in high-stakes applications, benchmarks must also evolve to comprehensively capture the full spectrum of model trustworthiness, including privacy, fairness, and robustness, beyond traditional predictive accuracy metrics.

5 Addressing Practical Challenges: Depth, Scalability, and Robustness

5.1 Building Deeper GNN Architectures

The development of deep Graph Neural Networks (GNNs) is crucial for capturing complex, long-range dependencies in graph-structured data, yet it is significantly hampered by fundamental challenges such as 'oversmoothing' and 'over-squashing'. Oversmoothing causes node representations to become indistinguishable with increasing layers, akin to Laplacian smoothing, while over-squashing limits information flow over long distances by compressing exponentially growing receptive fields into fixed-size vectors. Addressing these issues is paramount for enhancing the representational capacity of GNNs.

Early efforts to enable deeper GNNs primarily focused on decoupling the feature transformation from the propagation step. ? pioneered this with Personalized Propagation of Neural Predictions (PPNP) and its scalable approximation APPNP, which leveraged Personalized PageRank to allow arbitrary propagation depth without increasing the neural network's parameter count per layer, effectively mitigating oversmoothing. Building on this, ? further analyzed the performance degradation in deep GNNs, arguing that the entanglement of transformation and propagation was the primary culprit at moderate depths, and demonstrated that decoupling allowed for significantly deeper models before severe oversmoothing occurred.

To provide a more rigorous understanding of oversmoothing, ? introduced a novel

analytical technique based on Dirichlet energy, proving its exponential convergence to zero with increasing layers across a broader range of non-linearities. This theoretical foundation was further expanded by **?**, who provided an axiomatic, unified definition of oversmoothing, emphasizing exponential convergence of node-similarity measures like Dirichlet energy, and critically evaluating various mitigation strategies. Leveraging these insights, **?** proposed Energetic Graph Neural Networks (EGNNs), which employ a Dirichlet energy-constrained learning principle to guide deep GNN training through orthogonal weight control, lower-bounded residual connections, and Shifted ReLU (SReLU) activation, enabling models with up to 64 layers to achieve state-of-the-art performance.

Beyond direct oversmoothing mitigation, other architectural innovations have sought to enhance GNN depth and expressivity. **?** introduced Graph Neural Networks with Convolutional ARMA Filters, offering a more flexible frequency response than polynomial filters and capturing longer-range dynamics with fewer parameters through a recursive Graph Convolutional Skip (GCS) layer, implicitly aiding deeper architectures. To address the memory bottleneck that still limits *very* deep GNNs, **?** introduced Grouped Reversible GNNs, which reduce memory complexity for activations to be independent of network depth, enabling the training of GNNs with over 1000 layers.

The problem of over-squashing, distinct from oversmoothing, was formally diagnosed by **?**, who demonstrated that the exponential growth of a node's receptive field leads to information compression and loss in fixed-size message passing vectors, particularly for tasks requiring long-range interactions. As a simple mitigation, they proposed adding a "fully-adjacent layer" to alleviate this bottleneck. Complementing this, **?** introduced a novel principle to decouple the depth and scope of GNNs, proposing SHADOW-GNNs that apply deep GNNs on shallow, localized subgraphs. This approach theoretically prevents oversmoothing and enhances expressivity beyond the 1-Weisfeiler-Lehman (1-WL) test while significantly reducing computational costs.

More recently, fundamentally different approaches have emerged. **?** proposed Random Walk with Unifying Memory (RUM), a non-convolutional GNN architecture that processes graph information via random walks and RNNs. RUM jointly tackles limited

expressiveness, over-smoothing, and over-squashing, demonstrating superior theoretical properties and competitive empirical performance. Building on continuous GNNs, **?** introduced FROND, which leverages fractional calculus to generalize the integer-order differential equations, allowing GNNs to model non-local, memory-dependent dynamics and algebraically mitigate oversmoothing, thereby enhancing expressivity and robustness. Furthermore, **?** introduced Cooperative Graph Neural Networks (CO-GNNs), where nodes dynamically choose communication actions, leading to flexible, asynchronous information flow that addresses information bottlenecks and enhances expressive power beyond 1-WL. Finally, **?** proposed Spatio-Spectral Graph Neural Networks (S2GNNs), which synergistically combine local spatial message passing with global spectral filtering, provably vanquishing over-squashing and achieving superior approximation bounds for long-range interactions.

Despite these advancements, building truly robust and universally applicable deep GNNs remains an active area of research. Challenges persist in balancing the computational cost of advanced architectures with the need for high expressivity, especially for extremely large and dynamic graphs. Future directions include developing more adaptive and interpretable mechanisms for controlling information flow, exploring novel mathematical frameworks beyond traditional calculus, and designing architectures that can dynamically adjust their depth and receptive field based on the specific task and graph characteristics.

5.2 Scaling GNNs for Large-Scale Graphs

The application of Graph Neural Networks (GNNs) to real-world scenarios, particularly in industrial settings, is often hampered by significant computational and memory bottlenecks. Graphs with billions of nodes and edges, common in domains like social networks or recommender systems, necessitate advanced techniques to enable efficient processing, training, and inference. Addressing these scalability challenges is paramount for the practical deployment of GNNs.

One of the most widely adopted strategies for scaling GNNs involves efficient neigh-

borhood sampling. Early inductive methods like GraphSAGE **?** laid the groundwork by sampling a fixed number of neighbors for each node, thereby bounding the computational cost per node. Building upon this, PinSage **?** emerged as a pioneering industrial solution for web-scale recommender systems at Pinterest. PinSage introduced several innovations: "on-the-fly" convolutions that dynamically sample computation graphs for each minibatch, eliminating the need for the full graph Laplacian; a random walk-based neighborhood sampling that generates "importance scores" for neighbors, leading to a novel "importance pooling" aggregation strategy; and a producer-consumer architecture coupled with a MapReduce pipeline for scalable inference. This framework enabled the generation of billions of node embeddings, demonstrating practical solutions for achieving real-world scalability and efficiency. Beyond PinSage, other sampling-based methods include Cluster-GCN **?**, which partitions the graph into subgraphs and performs GNN computations on these clusters to reduce memory footprint and improve training efficiency, and GraphSAINT **?**, which samples entire subgraphs or edges instead of individual nodes, providing a more unbiased gradient estimation by preserving local graph structure. While sampling methods significantly reduce computational complexity, they often introduce variance into the training process and can lead to information loss, necessitating careful design of sampling strategies and aggregation functions.

Another critical approach to combat scalability issues is through graph reduction and sparsification. Graph Condensation (GCOND) **?** exemplifies this by learning a small, synthetic graph that preserves the essential information of a large original graph for GNN training. GCOND achieves remarkable size reduction (e.g., >99.9

For truly massive graphs that exceed the capacity of a single machine, system-level optimizations and distributed training frameworks are indispensable. These approaches typically involve partitioning the graph across multiple GPUs or CPU clusters. DistGNN **?** provides a notable example, optimizing the Deep Graph Library (DGL) for full-batch training on CPU clusters. It leverages efficient shared memory implementations, communication reduction via minimum vertex-cut graph partitioning, and communication

avoidance through delayed-update algorithms. This allows DistGNN to achieve significant speed-ups (up to 97x on 128 CPU sockets), demonstrating the power of distributed computing in making full-batch GNN training feasible for large graphs. The complexity of managing graph partitions, ensuring data consistency, and minimizing communication overhead across distributed nodes remains a significant engineering challenge, but it is fundamental for scaling GNNs to web-scale datasets.

Beyond sampling and graph reduction, some architectural designs inherently prioritize scalability. SIGN (Scalable Inception Graph Neural Networks) ? offers an alternative to sampling-based methods by leveraging an Inception-like module. Its key innovation lies in precomputing graph diffusion operators (e.g., powers of adjacency matrices or Personalized PageRank) on node features offline. This decouples the graph structure from the online training and inference phases, effectively reducing their complexity to that of a Multi-Layer Perceptron (MLP). SIGN avoids sampling biases and achieves significant speed-ups (even an order of magnitude) compared to sampling methods, particularly on extremely large graphs like ogbn-papers100M, by making the forward and backward passes independent of the graph structure. This approach challenges the notion that deep GNNs are always necessary for irregular graphs, suggesting that shallow architectures with expressive precomputed operators can be highly effective and scalable.

The practical impact of these scaling techniques is profoundly evident in industrial applications. PinSage ? remains a prime example of deploying GNNs for web-scale recommender systems, significantly enhancing user engagement. The principles of scalable GNN design are also crucial in other domains, such as radio resource management in wireless networks ?, where GNNs are designed with permutation equivariance and high computational efficiency to solve large-scale optimization problems within milliseconds on a single GPU. Despite significant progress, the tension between maximizing expressive power and maintaining efficiency for truly massive, dynamic, and noisy real-world graphs persists. Future research will likely focus on developing more adaptive, hybrid approaches that dynamically balance local and global information, further optimize memory footprint, and integrate hardware-aware designs. Furthermore, emerging paradigms like

Graph Neural Networks on Quantum Computers **?** offer a long-term, speculative avenue, proposing quantum algorithms for GNNs that could potentially achieve exponential reductions in space complexity, pushing the boundaries of scalability to unprecedented levels.

5.3   Handling Structural Heterogeneity and Imperfect Data

Real-world graphs rarely conform to idealized structures, often exhibiting a complex interplay of homophilic and heterophilic patterns, alongside pervasive issues of missing information, noisy connections, and weak supervision. Developing Graph Neural Networks (GNNs) that are robust to these complexities is paramount for reliable performance and broad applicability. This subsection explores advanced GNN architectures and techniques designed to adapt to structural disparities and learn effectively from imperfect graph data.

A significant challenge for conventional GNNs stems from their implicit assumption of homophily, where connected nodes share similar attributes or labels. **?** critically challenged this strict assumption, demonstrating that GCNs can perform well on certain heterophilous graphs by distinguishing between "good" and "bad" heterophily based on neighborhood patterns. This nuanced understanding paved the way for more adaptive designs, which were systematically categorized in a comprehensive survey by **?**, highlighting key design principles for GNNs operating in heterophilic environments. To explicitly address mixed structural patterns, early adaptive filtering mechanisms emerged. **?** introduced the Adaptive Channel Mixing (ACM) framework, which augments baseline GNNs by adaptively combining low-pass, high-pass, and identity filters in a node-wise and local manner, allowing GNNs to extract richer, localized information. Building on this, **?** rigorously demystified "structural disparity" within single graphs, demonstrating that a "one-size-fits-all" filtering approach leads to significant performance disparities on minority structural patterns. Directly addressing this limitation, **?** proposed NODE-MOE, a novel Mixture of Experts approach for node-wise filtering. NODE-MOE employs a sophisticated gating model to dynamically select and apply different "expert" GNN filters (e.g., low-pass, high-pass) to individual nodes, significantly enhancing GNN robustness to

mixed homophilic and heterophilic patterns. Complementing these node-wise approaches, ? developed GloGNN and GloGNN++, which find global homophily in heterophilous graphs by learning a signed coefficient matrix for all nodes, effectively combining low-pass and high-pass filtering with linear time complexity. Furthermore, ? introduced GNNs with Convolutional ARMA Filters, offering more flexible frequency responses than traditional polynomial filters, which can implicitly better handle diverse structural patterns. More recently, ? proposed Graph Neural Networks with Soft Association between Topology and Attribute (GNN-SATA), which utilizes separate embeddings for attributes and structures and establishes interconnections through soft association. GNN-SATA further employs a Graph Pruning Module (GPM) and Graph Augmentation Module (GAM) to dynamically remove or add edges, making the model better fit for graphs with varying degrees of homophily or heterophily, demonstrating improved accuracy, especially on highly heterophilic datasets.

Beyond adapting to existing but complex structural patterns, a separate and equally critical challenge arises when the graph structure itself is unreliable, incomplete, or altogether absent, or when node features and labels are scarce. Strategies for learning with such imperfect data primarily include structure learning, data augmentation, and self-supervised learning. When the graph structure is noisy, incomplete, or entirely absent, ? introduced Iterative Deep Graph Learning (IDGL), an end-to-end framework that jointly and iteratively learns optimal graph structures and GNN parameters, even from raw features. This approach is robust to initial graph imperfections and offers a scalable anchor-based version (IDGL-ANCH). Similarly, ? proposed SLAPS, which leverages self-supervision through a denoising autoencoder to guide structure learning, effectively addressing the "supervision starvation" problem where many edges receive insufficient supervision from the primary task. For enhancing GNN resilience through data augmentation, ? developed the GAUG framework. GAUG uses a learned edge predictor (e.g., a Graph Auto-Encoder) to strategically add missing intra-class edges and remove noisy inter-class edges, thereby "denoising" and "completing" the graph structure.

The broader paradigm of self-supervised learning (SSL) has emerged as a powerful tool

for learning with weak or incomplete graph information, as comprehensively reviewed by
?. SSL methods for GNNs aim to leverage abundant unlabeled graph data by generating
self-supervision signals, often through contrastive learning or predictive tasks, to learn
robust representations. Addressing the challenging scenario of "extreme weak informa-
tion," where structure, features, and labels are simultaneously deficient, **?** introduced
D2PT (Dual-channel Diffused Propagation then Transformation). D2PT employs a dual-
channel architecture, combining information from the input graph with a learned global
graph (to connect stray nodes), and uses prototype contrastive alignment to ensure mu-
tual benefit between channels. This aligns with the principles of SSL by creating auxiliary
tasks to learn from limited explicit supervision. Similarly, **?** proposed a self-supervised
GNN model specifically for heterogeneous information networks, aiming to flexibly com-
bine different types of additional information and mine deep features, thereby improving
adaptability to graph diversity and complexity, especially when initial structure and at-
tribute information might be insufficient or redundant. Furthermore, to enhance GNN
resilience against "structure shift" (where test graph topology differs from training), **?**
proposed the Cluster Information Transfer (CIT) mechanism. CIT learns invariant repre-
sentations by applying spectral clustering and transferring node representations between
cluster statistics in the embedding space, without explicit graph modification, implicitly
addressing potential imperfections in the structural alignment between training and test
data. In a more specific application of handling missing data, **?** introduced GRIN, a
novel GNN architecture for multivariate time series imputation. GRIN reconstructs miss-
ing temporal data by learning spatio-temporal representations through message passing,
demonstrating the utility of GNNs in filling gaps in real-world datasets where relational
information is crucial.

In conclusion, significant strides have been made in making GNNs robust to the com-
plexities of real-world graphs. Adaptive filtering mechanisms, from node-wise channel
mixing **?** to sophisticated mixture-of-experts approaches **?** and dynamic edge modifi-
cations **?**, are increasingly capable of handling structural heterogeneity and mixed ho-
mophilic/heterophilic patterns. Concurrently, advancements in structure learning **??**,

data augmentation **?**, and particularly self-supervised learning paradigms **???** are enhancing GNN resilience to missing information, noisy connections, and weak supervision, even in challenging scenarios like extreme weak information **?** or structural shifts **?**. Despite these advancements, challenges remain, particularly regarding the computational cost and theoretical guarantees for dynamically learned structures, and ensuring these adaptive and learning mechanisms generalize reliably across highly diverse and unseen graph topologies. Future research will likely focus on developing more efficient and theoretically grounded adaptive mechanisms, potentially integrating meta-learning for rapid adaptation to novel graph patterns, and exploring hybrid approaches that seamlessly combine explicit structure learning with robust adaptive filtering for truly resilient GNNs.

5.4   Pre-training and Prompt-based Adaptation

The efficacy of Graph Neural Networks (GNNs) in real-world applications is often hampered by the scarcity of labeled graph data. To overcome this, advanced techniques for efficient knowledge transfer and adaptation have emerged, primarily through foundational self-supervised pre-training and the innovative paradigm of prompt tuning. These strategies aim to learn generalizable representations from abundant unlabeled graph data, minimize the 'objective gap' between pre-training and downstream tasks, and enable data-efficient adaptation.

Early efforts laid the groundwork for pre-training GNNs to learn robust, transferable representations. **?** pioneered a systematic investigation into GNN pre-training strategies, proposing a combined node- and graph-level self-supervised approach. Their work introduced novel self-supervised tasks like Context Prediction and Attribute Masking, demonstrating significant generalization improvements and effectively mitigating "negative transfer" often encountered in naive pre-training. Building upon this, **?** advanced the field by proposing GPT-GNN, a generative pre-training framework that explicitly models the intricate dependency between node attributes and graph structure during an attributed graph generation task. This approach provided a more holistic and semantically rich pre-training objective, crucial for complex real-world graphs. Further refining

the pre-training paradigm, **?** introduced L2P-GNN, which leverages meta-learning to explicitly optimize for rapid adaptation during the pre-training process. By framing pre-training as an optimization-based meta-learning problem, L2P-GNN directly addresses the objective divergence between pre-training and fine-tuning, ensuring the learned initialization is inherently optimized for quick transferability to new tasks.

Despite these advancements in learning robust initializations, a persistent "objective gap" often remained between pre-training pretext tasks and diverse downstream tasks, necessitating costly and inefficient fine-tuning. This challenge spurred the development of prompt-based adaptation for GNNs, a paradigm shift inspired by its success in natural language processing. **?** introduced GPPT, the first framework to apply prompt tuning to GNNs. It ingeniously reformulates downstream node classification tasks to mimic the pre-training objective (e.g., masked edge prediction) by using a novel "graph prompting function" and "token pairs," thereby bridging the objective gap and enabling efficient knowledge transfer without tedious fine-tuning. However, GPPT's prompt design was somewhat specialized. Addressing this limitation, **?** proposed Universal Prompt Tuning for GNNs, introducing Graph Prompt Feature (GPF) which modifies the *input graph's feature space* rather than altering its structure or reformulating tasks. This innovation provided a universal and theoretically grounded approach, compatible with *any* pre-trained GNN and *any* pre-training strategy, demonstrating superior performance over fine-tuning with significantly fewer tunable parameters.

Further extending the versatility of prompt tuning, **?** introduced GraphPrompt, a framework that unifies pre-training and diverse downstream tasks (node and graph classification) into a common "subgraph similarity" template. GraphPrompt employs novel learnable prompts that guide the `ReadOut` aggregation operation, allowing a single pre-trained GNN to adapt to various tasks in few-shot settings by adaptively fusing node representations. Most recently, the paradigm has expanded to multi-modal learning. **?** proposed Morpher, a multi-modal prompt learning framework that aligns independently pre-trained GNNs with Large Language Models (LLMs) under *extremely weak text supervision*. This groundbreaking work enables GNNs to gain real-world semantic under-

standing and achieve CLIP-style zero-shot generalization to unseen classes, significantly enhancing data efficiency and model versatility by bridging the conceptual gap between graph structure and natural language semantics.

In conclusion, the evolution from foundational self-supervised pre-training to sophisticated prompt-based adaptation marks a critical intellectual trajectory in GNN research. While initial pre-training strategies focused on learning generalizable representations and mitigating negative transfer, the advent of prompt tuning has revolutionized adaptation by minimizing the objective gap and enabling highly efficient, few-shot, and even zero-shot generalization. The latest advancements, particularly in multi-modal prompt learning, are pushing GNNs towards greater semantic awareness and versatility. However, an ongoing challenge lies in developing truly universal pre-training objectives that inherently support diverse downstream tasks without requiring complex prompt engineering, paving the way for more intrinsically adaptable and robust GNN architectures.

6 Trustworthy GNNs: Explainability, Fairness, and Security

6.1 Interpreting GNN Decisions

The inherent complexity of Graph Neural Networks (GNNs), arising from their non-linear message-passing mechanisms and intricate interactions within graph structures, often renders them "black-box" models. This opacity hinders trust, limits adoption in critical domains, and impedes debugging, necessitating robust methods for interpreting their decisions. The field of GNN interpretability aims to unveil *why* a GNN makes a particular prediction, encompassing both instance-level explanations for specific predictions and model-level insights into general patterns learned by the network.

Initial efforts in GNN interpretability were systematically organized by ?, which provided a foundational taxonomic survey, categorizing methods and highlighting the unique challenges posed by graph data. This survey underscored the critical need for both

instance-level explanations, which pinpoint influential subgraphs or features for individual predictions, and model-level explanations, which uncover general patterns learned by the GNN.

Pioneering instance-level explanations, ? introduced GNNExplainer, a general and model-agnostic framework. GNNExplainer identifies a compact subgraph structure and a small subset of node features that are most influential for a given prediction by maximizing the mutual information between the GNN's output and potential explanations. While effective for individual instances, understanding the overall model behavior required examining numerous examples, which is often impractical. Addressing this, ? proposed XGNN, the first method for model-level explanations. XGNN trains a graph generator, guided by reinforcement learning, to produce graph patterns that maximally activate a specific prediction score of the target GNN, thereby revealing general patterns associated with a particular class.

However, many early explanation methods, including GNNExplainer, often relied on additive feature attribution, implicitly assuming feature independence. ? challenged this by introducing PGM-Explainer, which leverages Probabilistic Graphical Models (specifically Bayesian Networks) to explicitly model the *dependencies* and *conditional probabilities* among features contributing to a GNN's prediction. This provides a richer, more nuanced understanding of feature interactions beyond simple importance scores. Further, the pursuit of truly faithful explanations led to a focus on causal patterns. ? introduced Discovering Invariant Rationale (DIR), an invariant learning strategy that identifies *causal rationales* for GNNs by generating interventional distributions. This approach ensures explanations are based on stable, causal patterns rather than spurious correlations, leading to improved out-of-distribution generalization and more trustworthy interpretations.

Expanding on model-level interpretability, ? moved towards unveiling "global interactive patterns" across entire graphs, particularly for graph-level tasks. Their method employs graph coarsening and learnable graph prototypes to identify inter-cluster interactions, providing insights into long-range dependencies and global structural influences on predictions.

41

More recently, the field has moved towards establishing rigorous theoretical foundations to ensure the faithfulness and structure-awareness of explanations, addressing limitations in prior approaches. ? critically observed that traditional Shapley-based methods for GNNs often fail to account for graph structure and high-order interactions, leading to biased attributions. They introduced the **Myerson-Taylor interaction index**, an axiomatically grounded measure that inherently incorporates graph connectivity and captures high-order interactions. Their proposed MAGE explainer leverages this index to identify both positively and *negatively* influential motifs, offering a comprehensive and theoretically sound approach to instance-level explanations. Concurrently, ? provided a profound theoretical critique of existing attention-based interpretable GNNs. They introduced the **Subgraph Multilinear Extension (SubMT)** framework and rigorously proved that many attention-based XGNNs suffer from "approximation failures" for non-linear GNNs, meaning their explanations may not faithfully represent the model's true decision-making process. To overcome this, they proposed **Graph Multilinear ne T (GMT)**, an architecture that uses random subgraph sampling to more accurately approximate SubMT, leading to provably more powerful and generalizable interpretations.

Despite these significant advancements, challenges remain. The computational cost of generating high-fidelity, structure-aware explanations, especially for very large graphs or complex high-order interactions, is still a bottleneck. Furthermore, while theoretical frameworks like SubMT and the Myerson-Taylor index provide rigorous foundations, bridging the gap between mathematical guarantees and intuitive, actionable insights for human users remains an active area of research. Future work will likely focus on developing more scalable and efficient algorithms for these theoretically sound frameworks, exploring multi-modal explanations that combine structural and feature-based insights, and developing standardized human-centric evaluation protocols to ensure explanations truly build user trust and provide actionable understanding.

6.2 Robustness to Adversarial Attacks and Data Bias

The increasing deployment of Graph Neural Networks (GNNs) in critical applications necessitates a thorough understanding of their trustworthiness, particularly concerning robustness to adversarial attacks and mitigation of data bias ?. GNNs are inherently vulnerable to subtle, often imperceptible, perturbations in their graph structure or node features, which can drastically alter predictions and undermine their reliability. This section explores various attack strategies, corresponding defense mechanisms, and methods for learning invariant rationales to enhance GNN resilience.

Initial investigations into GNN vulnerabilities highlighted the susceptibility to adversarial manipulations. ? pioneered the study of global poisoning attacks, formulating them as a bilevel optimization problem solved via meta-gradients to degrade the overall performance of GNNs on discrete graph structures. Concurrently, ? addressed topology attacks by proposing an optimization-based framework that leverages convex relaxation to enable gradient-based perturbations on discrete graph structures, demonstrating effective attack generation and laying groundwork for adversarial training. However, these early attack methods often faced scalability challenges, limiting their applicability to large real-world graphs. To overcome this, ? introduced sparsity-aware first-order optimization attacks (PR-BCD and GR-BCD) and novel surrogate losses (Masked Cross Entropy, tanh margin loss), enabling global GNN attacks on graphs orders of magnitude larger than previously possible and achieving significantly stronger attacks. Beyond general performance degradation, specific attack vectors target sensitive information or introduce stealthy backdoors. ? demonstrated "link stealing attacks," revealing that GNN outputs inherently encode significant structural information about their training graphs, even under black-box access, posing a critical privacy threat. Furthermore, ? introduced subgraph-based backdoor attacks for graph classification, where a predefined trigger pattern injected into training graphs causes the GNN to misclassify any triggered test graph to a target label, highlighting a stealthy and persistent form of manipulation.

In response to these vulnerabilities, various defense mechanisms have been proposed. ? introduced GNNGuard, a general algorithm that defends against training-time poisoning

attacks by dynamically estimating neighbor importance and employing layer-wise graph memory to prune suspicious edges and revise message passing. ? proposed Elastic Graph Neural Networks, which enhance robustness by integrating L1-based graph smoothing into their message passing scheme. This approach allows for adaptive local smoothness, preserving discontinuities while providing robustness against adversarial attacks, a limitation of traditional L2-based smoothing methods. Adversarial training, as proposed by ?, also emerged as a defense, where GNNs are trained to minimize loss under worst-case topology perturbations. However, a critical re-evaluation by ? exposed a significant limitation in the evaluation of many existing GNN defenses. Their comprehensive analysis, using custom-designed adaptive attacks, revealed that most defenses offer "no or only marginal improvement" against adversaries aware of the defense mechanism, challenging earlier optimistic claims and emphasizing the need for more rigorous evaluation against adaptive threats.

Beyond adversarial robustness, GNNs must also contend with inherent data biases that can lead to unfair or unreliable predictions. ? addressed this by proposing EDITS, a novel, model-agnostic framework that mitigates bias directly in the input attributed network by defining and optimizing against "attribute bias" and "structural bias." This approach shifts the focus from debiasing GNN models or outcomes to debiasing the input data itself, overcoming the limitation of model-specific debiasing. Addressing fairness under practical constraints, ? introduced FairGNN, a framework for fair node classification that effectively handles limited sensitive attribute information by estimating missing values and employing an adversarial network to ensure predictions are independent of sensitive attributes. Building on this, ? further improved fairness by mitigating "sensitive attribute leakage" during feature propagation, a phenomenon where non-sensitive features become correlated with sensitive ones after GNN message passing. More recently, ? rethought fair GNNs from a re-balancing perspective, proposing FairGB, which uses Counterfactual Node Mixup (CNM) to generate unbiased ego-networks and Contribution Alignment Loss (CAL) to balance group contributions, demonstrating superior fairness and utility with minimal architectural changes.

To enhance overall reliability and prevent malicious manipulation, research is also focusing on discovering and learning invariant rationales, making GNNs more resilient to spurious correlations and out-of-distribution (OOD) shifts. **?** proposed Discovering Invariant Rationale (DIR), a novel invariant learning strategy for intrinsically interpretable GNNs. DIR identifies causal patterns that are stable across different data distributions by generating "interventional distributions" through causal interventions on non-causal graph components. This approach helps GNNs learn from true causal features rather than shortcut features or data biases, thereby improving OOD generalization and making explanations more faithful.

In conclusion, while significant progress has been made in understanding and mitigating adversarial attacks and data bias in GNNs, the field faces ongoing challenges. The critical finding that many existing defenses are vulnerable to adaptive attacks underscores the need for more robust and theoretically grounded defense mechanisms. Future directions must focus on developing inherently robust GNN architectures, certified robustness guarantees, and advanced causal learning frameworks that can consistently identify invariant rationales to ensure GNNs are not only performant but also truly trustworthy, fair, and resilient to malicious manipulation in high-stakes real-world applications.

6.3 Fairness and Privacy in GNNs

The increasing deployment of Graph Neural Networks (GNNs) in critical domains necessitates a rigorous examination of their ethical implications, particularly concerning fairness and privacy. GNNs, by their inherent nature of aggregating and propagating information across interconnected entities, can inadvertently perpetuate or even amplify societal biases and expose sensitive information, demanding robust mitigation strategies **??**. This section delves into methodologies for ensuring fair outcomes and protecting sensitive data, exploring techniques that identify and mitigate biases, alongside mechanisms designed to safeguard private information.

6.3.1 Ensuring Fairness in GNNs

Ensuring fair outcomes in GNNs is a significant challenge, as biases can originate from imbalanced graph structures, node features, or the message-passing mechanism itself, often exacerbated by the homophilous nature of many real-world graphs. The field distinguishes between group fairness (fairness across predefined demographic groups) and individual fairness (similar individuals receiving similar predictions), both of which present unique challenges in graph contexts ??.

Early efforts, such as FairGNN by ?, addressed GNN bias, particularly when sensitive attribute information (e.g., gender, race) is limited. FairGNN employs a GCN-based sensitive attribute estimator to infer missing protected attributes, enabling an adversarial debiasing framework to learn fair node representations. This in-processing approach aims to make the learned representations independent of sensitive attributes. Building on the understanding of bias sources, ? shifted the focus to a data-centric, pre-processing approach with EDITS. This method models and mitigates both "attribute bias" (from node features) and "structural bias" (from graph topology) directly in the input attributed network using Wasserstein-1 distance, offering a model-agnostic solution that can be applied before any GNN model.

While group fairness is crucial, individual fairness presents distinct complexities. ? introduced REDRESS, a ranking-based framework to enhance individual fairness in GNNs. This approach circumvents the practical difficulties associated with traditional Lipschitz-condition-based definitions of individual fairness by reframing it from a ranking perspective, making it a plug-and-play solution for various GNN architectures. Further refining the understanding of bias propagation, ? identified "sensitive attribute leakage" as a critical issue where GNN feature propagation dynamically alters feature correlations, exacerbating discrimination. Their FairVGNN framework mitigates this by using a generative adversarial debiasing module to mask sensitive-correlated features and an adaptive weight clamping module for the GNN encoder.

More recently, researchers have explored causal and invariant learning perspectives for debiasing. ? re-evaluated fair GNNs from a re-balancing perspective, proposing FairGB.

This method uses Counterfactual Node Mixup (CNM) to generate unbiased ego-networks and a Contribution Alignment Loss (CAL) to balance group contributions, offering a robust alternative to simple re-balancing and adversarial methods by providing theoretical insights into causal paths. Complementing these efforts, **?** introduced Discovering Invariant Rationale (DIR), an invariant learning strategy for intrinsically interpretable GNNs. DIR aims to identify causal patterns that are stable across different data distributions, thus ensuring fairness and robustness by filtering out spurious correlations that might lead to biased decisions. Similarly, **?** proposed DisC (Debiasing Graph Neural Networks via Learning Disentangled Causal Substructure) to tackle severe bias by explicitly disentangling causal and bias substructures using an edge mask generator and dual GNNs, further decorrelating them in the embedding space. This approach is particularly effective in scenarios where bias is severe and dominates GNN learning.

The choice between these methods often involves trade-offs between model utility, the strength of fairness guarantees, and computational overhead. Pre-processing methods like EDITS offer model agnosticism but might lose information, while in-processing adversarial methods like FairGNN and FairVGNN can be more tightly integrated but are model-specific. Causal and invariant learning approaches like FairGB, DIR, and DisC aim for more fundamental debiasing by targeting the underlying causal mechanisms, promising better generalization and interpretability.

6.3.2 Protecting Privacy in GNNs

Privacy is another paramount concern in GNNs, especially given the sensitive nature of graph connections and node attributes. GNNs can inadvertently leak private information through their outputs or learned representations, leading to various privacy attacks **??**.

A major vulnerability is the potential for **link stealing attacks**, where an adversary can infer private connections from the black-box outputs of a GNN. **?** pioneered this area by demonstrating that GNN outputs inherently encode significant structural information about their training graphs, even with limited adversary knowledge. Their work introduced a comprehensive threat model and various attack methodologies (unsu-

pervised, supervised, and transfer attacks), showing high success rates in inferring links. This highlights a critical privacy leakage that demands robust countermeasures. Beyond link stealing, GNNs are also susceptible to other inference attacks:

- **Membership Inference Attacks**: These attacks aim to determine whether a specific node or subgraph was part of the GNN's training dataset **??**. This can reveal sensitive information about individuals or entities whose data was used in training.

- **Attribute Inference Attacks**: Adversaries can infer sensitive node features (e.g., income, health status) from publicly available information or GNN outputs, leveraging the message-passing mechanism that propagates feature information across the graph **??**.

To counter these threats, various privacy-preserving mechanisms have been explored. **Differential Privacy (DP)** stands out as a strong, mathematically provable privacy guarantee. DP mechanisms add carefully calibrated noise to data or computations to obscure individual contributions while preserving overall statistical properties **??**. In GNNs, DP can be applied at different levels:

- **Node Feature Perturbation**: Adding noise directly to node features before training, ensuring that the presence or absence of a single node's features does not significantly alter the model's output.

- **Gradient Perturbation**: Applying DP-SGD (Differential Private Stochastic Gradient Descent) during GNN training, adding noise to gradients to protect individual contributions to the model updates.

- **Graph Structure Perturbation**: This is particularly challenging due to the discrete and interconnected nature of graphs. Techniques involve adding or deleting edges with certain probabilities, or perturbing the adjacency matrix, to protect the privacy of connections **?**. However, perturbing graph structure can significantly impact graph utility and model performance.

The main challenge with DP in GNNs lies in balancing the privacy budget with model utility, especially for graph structures where small perturbations can have far-reaching effects due to message passing.

Another promising approach is **Federated Learning (FL)**, which enables collaborative training of GNNs across multiple data owners without centralizing raw data **?**. In FL for GNNs, local GNN models are trained on private graph partitions, and only model updates (e.g., gradients or aggregated parameters) are shared with a central server. This inherently protects raw node features and graph structures from being exposed, making it suitable for scenarios with distributed and sensitive graph data, such as medical networks or financial transaction graphs. However, FL itself can still be vulnerable to inference attacks if not combined with additional privacy mechanisms like DP.

The comprehensive surveys by **?** and **?** synthesize these concerns, providing taxonomies for privacy attacks and defense mechanisms, and highlighting the unique challenges of graph data for trustworthiness. Despite significant progress in identifying and mitigating fairness and privacy issues, challenges remain in developing universally applicable, scalable, and provably robust solutions that do not incur significant trade-offs in model utility. Future research must continue to explore advanced differential privacy mechanisms tailored for graph structures, more sophisticated adversarial training methods that withstand adaptive attacks, and integrated frameworks that jointly optimize for fairness, privacy, and utility, ensuring GNNs operate ethically, securely, and without perpetuating or amplifying societal inequalities.

7 Applications of Graph Neural Networks

7.1 Recommender Systems

Recommender systems, essential for navigating vast digital content and alleviating information overload, have been profoundly transformed by Graph Neural Networks (GNNs) **???**. The inherent graph structure of user-item interactions, social connections, and item knowledge graphs makes GNNs particularly adept at learning rich latent factors and cap-

turing complex relational patterns. This capability leads to more personalized, accurate, and context-aware recommendations that consistently outperform traditional methods by explicitly modeling collaborative signals and multi-hop relationships **?**.

A pioneering challenge for GNNs in recommendation was scaling to web-scale platforms with billions of users and items. **PinSage ?** addressed this by developing a novel Graph Convolutional Network (GCN) algorithm specifically for Pinterest's massive graph. PinSage overcame the limitations of prior GCNs by introducing efficient, localized convolutions that operate "on-the-fly" through random walk-based neighborhood sampling and an "Importance Pooling" aggregation. This, combined with a producer-consumer minibatch architecture and a MapReduce inference pipeline, enabled the largest-ever application of deep graph embeddings at the time, demonstrating substantial improvements in offline metrics and user engagement in A/B tests. The technical details of these scaling mechanisms, which are critical for industrial deployment, are further elaborated in Section 5.2.

Beyond sheer scale, enhancing recommendation quality often necessitates integrating heterogeneous information. Social recommender systems, for instance, leverage both user-item interactions and user-user social relations to better understand user tastes, capitalizing on phenomena like homophily and social influence **?**. **GraphRec ?** is a notable GNN framework that adeptly integrates user-item interactions, associated opinions (e.g., ratings), and heterogeneous social relations. It learns user latent factors from both "item-space" and "social-space" perspectives, employing opinion embedding vectors and attention mechanisms (item, social, and user attention) to dynamically weigh the contributions of diverse information sources. This approach captures nuanced user preferences and social influence, leading to more comprehensive user representations. The comprehensive survey by **?** further categorizes and reviews GNN-based social recommender systems, highlighting their architectural variations and input types for leveraging social graphs as side information.

GNNs also excel at learning richer latent factors and capturing complex relational patterns from interaction graphs, which is vital for precise recommendations. For instance,

in link prediction tasks—a core component of many recommender systems aimed at predicting future interactions—**SEAL ?** leverages GNNs to learn general graph structure features from local enclosing subgraphs. This approach, underpinned by a novel "β-decaying heuristic theory," theoretically justifies learning high-order features from small local contexts. SEAL significantly outperforms traditional heuristic and latent feature methods by adaptively learning patterns specific to link formation, thereby improving the accuracy of predicting new user-item connections.

The dynamic nature of user preferences and item catalogs necessitates GNNs capable of handling **sequential and session-based recommendation tasks**. Traditional sequential models often struggle with complex, non-linear item transitions within a user session. **SR-GNN ?** pioneered the application of GNNs to session-based recommendation by modeling item sequences as graph-structured data. It uses GNNs to capture complex item transitions within a session, representing each session as a composition of global preference and current interest via an attention network. Building on this, **GCE-GNN ?** further enhanced session-based recommendation by learning two levels of item embeddings: session-level from the current session graph and global-level from an aggregated global graph of all sessions. A session-aware attention mechanism recursively incorporates neighbors' embeddings on the global graph, allowing for a more subtle exploitation of item transitions across all sessions. More recently, **SURGE ?** transforms loose item sequences into dynamic item-item interest graphs, employing metric learning with ϵ-sparseness for robust graph construction and utilizing cluster- and query-aware graph attentive convolutional layers. Its dynamic graph pooling layer adaptively extracts activated core preferences, effectively modeling evolving user interests from long, noisy historical sequences and outperforming state-of-the-art sequential recommenders.

To further enhance the robustness and data efficiency of GNNs in recommendation, especially in data-scarce or cold-start scenarios, **pre-training strategies** are increasingly important. Self-supervised pre-training, as investigated by **?**, can learn generalizable representations from abundant unlabeled graph data (e.g., through context prediction or attribute masking). Such pre-trained GNNs are more robust and adaptable for downstream

recommendation tasks, even with limited labeled user feedback. This topic, including foundational self-supervised strategies, is further discussed in Section 5.4.

Despite these significant advancements, several challenges persist in the application of GNNs to recommender systems ???. Scaling GNNs to truly massive, dynamic graphs with real-time updates while maintaining high expressivity and interpretability remains an active research area. Developing GNN architectures that can seamlessly handle extreme data sparsity and cold-start problems, particularly for new users or items, continues to be a critical hurdle. Furthermore, understanding and mitigating biases inherent in graph data and GNN models to ensure fair and diverse recommendations is crucial for responsible deployment ?. Research into individual fairness for GNNs, such as the ranking-based approach of REDRESS ? or frameworks like FairGNN ? that address fairness with limited sensitive attribute information, directly contributes to this goal, as detailed in Section 6.3. The comprehensive surveys by ??? provide systematic taxonomies and highlight critical challenges in graph construction, network design, optimization, and computational efficiency, thereby guiding future research in this rapidly evolving domain.

7.2   Scientific Domains (Chemistry, Materials, Physics)

Graph Neural Networks (GNNs) have emerged as a transformative paradigm in scientific research, particularly within chemistry, materials science, and physics, owing to the inherently graph-like nature of molecular and crystal structures. Their capacity to directly learn from structural data, coupled with the ability to capture physical symmetries and interactions, makes them invaluable for accelerating discovery, drug design, and material innovation. This section details the application of GNNs across these domains, focusing on their practical impact in molecular property prediction, interatomic potential modeling, and materials discovery.

One of the most prominent applications of GNNs is in **molecular property prediction**, a cornerstone of drug discovery and chemical engineering. Molecules are naturally represented as graphs, with atoms as nodes and chemical bonds as edges, allowing GNNs to directly process their topology and features. For instance, ? demonstrated an

attention-based GNN approach for predicting Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties directly from Simplified Molecular Input Line Entry System (SMILES) notation. This "bottom-up" method processes information from substructures to the whole molecule, bypassing the computationally expensive retrieval of molecular descriptors and showing effectiveness across various regression and classification tasks crucial for early-stage drug development. However, the landscape of molecular property prediction is complex, and GNNs do not always universally outperform traditional methods. A comprehensive comparison by ? across 11 diverse datasets revealed that, on average, descriptor-based models (e.g., SVM, XGBoost, RF, DNN) could still outperform graph-based models (e.g., GCN, GAT, MPNN, Attentive FP) in terms of prediction accuracy and computational efficiency, especially for smaller datasets. This highlights the critical need for careful benchmarking and understanding the specific strengths and weaknesses of GNNs in different chemical contexts, suggesting that while GNNs offer a powerful framework, their superiority is not guaranteed across all tasks and data regimes.

Beyond static property prediction, GNNs are revolutionizing **the modeling of interatomic potentials (IAPs)**, which are essential for high-fidelity molecular dynamics (MD) simulations. Traditional *ab initio* methods like Density Functional Theory (DFT) are accurate but computationally prohibitive for large systems or long timescales. GNNs offer a data-driven alternative to learn IAPs from *ab initio* calculations, enabling accelerated simulations. A significant advancement in this area is Neural Equivariant Interatomic Potentials (NequIP) by ?. NequIP employs E(3)-equivariant convolutions, which ensure that the learned potential respects fundamental physical symmetries (rotations, translations, reflections) without relying on computationally expensive higher-order representations. This approach achieved state-of-the-art accuracy on diverse molecules and materials, demonstrating remarkable data efficiency by outperforming existing models with up to three orders of magnitude fewer training data. Such data efficiency challenges the notion that deep neural networks always require massive datasets and enables the construction of accurate potentials using high-level quantum chemical theories. Complementing NequIP, ? introduced GemNet, a universal directional GNN for molecules.

GemNet's innovation lies in its explicit incorporation of full geometric information, including interatomic distances, angles, and crucially, dihedral angles, through a novel two-hop message passing scheme based on directed edge embeddings. This architectural design, coupled with a theoretical proof of universality for spherical representations, led to significant performance improvements (e.g., 20-41

In **materials science**, GNNs are instrumental in designing new materials and analyzing crystal structures. Crystal structures, with their periodic arrangements of atoms, can be effectively represented as graphs with periodic boundary conditions. GNNs are used to predict properties of novel materials, guide their synthesis, and identify stable structures. The integration of GNNs with other advanced AI paradigms is also emerging, as demonstrated by ? with Hybrid-LLM-GNNs for enhanced materials property prediction. This approach combines the structural learning capabilities of GNNs with the semantic reasoning of Large Language Models (LLMs), showcasing a promising direction for leveraging multi-modal data and contextual information to improve predictive accuracy in complex materials informatics tasks.

In summary, GNNs have profoundly impacted scientific domains by offering a powerful framework to learn directly from the intrinsic graph-like nature of molecules and materials. The development of geometrically equivariant architectures, such as NequIP and GemNet, has been crucial for capturing physical symmetries and interactions efficiently, leading to data-efficient and accurate predictions. While challenges remain, including scaling to extremely large and heterogeneous systems, integrating diverse multi-modal data, and enhancing interpretability to provide mechanistic insights, the continuous evolution of GNNs promises further acceleration of scientific discovery and innovation across chemistry, materials science, and physics.

7.3   Time Series and Dynamic Graphs

The analysis of time series data, particularly in multivariate settings, presents a significant challenge due to the complex spatial-temporal dependencies that often exist be-

tween variables. Traditional time series models frequently struggle to explicitly capture these non-Euclidean relationships, necessitating the emergence of Graph Neural Networks (GNNs) as a powerful paradigm for modeling such interconnected systems ??. This subsection examines the application of GNNs to two distinct yet related problems: first, Spatio-Temporal Graph Neural Networks (STGNNs) for multivariate time series analysis on largely static graph structures, and second, GNNs for truly dynamic graphs where the underlying topology itself evolves over time.

7.3.1 Spatio-Temporal Graph Neural Networks for Multivariate Time Series Analysis

In many real-world scenarios, multivariate time series data is inherently structured as a graph, such as sensor networks, road networks, climate monitoring stations, or even brain activity recordings. Here, the spatial relationships between variables are often static or slowly changing, while the features (e.g., traffic speed, temperature, brain signals) evolve dynamically. Spatio-Temporal Graph Neural Networks (STGNNs) are designed to capture both the spatial dependencies (via graph convolutions) and temporal patterns (via recurrent, convolutional, or attention mechanisms) simultaneously ??. The versatility of STGNNs extends beyond mere forecasting to critical tasks like classification, imputation, and anomaly detection, addressing a broader spectrum of time series challenges ?.

Early and foundational STGNN architectures often integrated Graph Convolutional Networks (GCNs) with Recurrent Neural Networks (RNNs) or Convolutional Neural Networks (CNNs). Models like Diffusion Convolutional Recurrent Neural Network (DCRNN) ? and Spatio-Temporal Graph Convolutional Network (STGCN) ? exemplify this approach. DCRNN combines graph convolutions based on diffusion processes with a sequence-to-sequence RNN framework, allowing information to propagate across the graph and evolve over time. STGCN, on the other hand, interweaves graph convolutions with 1D convolutional layers for temporal feature extraction, offering a more parallelizable alternative to RNNs for capturing local temporal patterns. These methods demonstrated significant improvements in tasks like traffic flow forecasting by explicitly modeling the

road network structure and its dynamic conditions **?**. For instance, the GNN estimator deployed in Google Maps for ETA prediction leverages the road network's graph structure, employing sophisticated featurization and robust training regimes to predict travel times across "supersegments," achieving substantial reductions in ETA errors **?**.

Critically, the choice of temporal backbone (RNN, CNN, or attention) presents distinct trade-offs. RNN-based models excel at capturing long-range temporal dependencies but suffer from sequential processing, limiting parallelization and increasing computational cost for long sequences. CNN-based models offer better parallelizability and capture local temporal patterns efficiently but may struggle with very long-range dependencies without deep stacking or dilated convolutions. Attention mechanisms, particularly those inspired by Transformers, have gained prominence for their ability to capture global, long-range temporal dependencies and dynamic spatial relationships more flexibly **?**. Attention-based STGNNs can adaptively weigh the importance of different temporal steps or spatial neighbors, providing highly adaptive spatio-temporal modeling. For example, **?** proposed a framework that automatically extracts uni-directed relations among variables through a graph learning module and employs a mix-hop propagation layer alongside a dilated inception layer to effectively capture both spatial and temporal dependencies.

Beyond forecasting, STGNNs are increasingly applied to other vital time series tasks:

- **Imputation:** Addressing missing values in multivariate time series is crucial. GNNs can reconstruct missing data by exploiting both temporal patterns and spatial correlations within the graph. For instance, GRIN **?** is a novel GNN architecture specifically designed for multivariate time series imputation, learning spatio-temporal representations through message passing to reconstruct missing data. More recently, Casper **?** introduced a Causality-Aware STGNN that leverages causal inference to block confounders and discover sparse causal relationships, enhancing imputation accuracy by avoiding non-causal correlations.

- **Anomaly Detection:** GNNs are powerful for identifying anomalous patterns in time series, which can manifest as unusual node features or structural deviations. This is particularly relevant in domains like Industrial Internet of Things (IIoT)

for smart transportation, energy, and factories **?**. GNN-based methods leverage the expressive power of message passing to learn normal patterns and detect deviations **?**. For example, the Beta Wavelet Graph Neural Network (BWGNN) **?** analyzes anomalies through the graph spectrum, observing a "right-shift" phenomenon in spectral energy distribution, and uses spectral and spatial localized band-pass filters to better handle this.

- **Classification:** While less explicitly covered in the provided papers, STGNNs are also employed for time series classification, such as classifying different types of human activities from sensor data or medical conditions from physiological signals, by learning discriminative spatio-temporal features **?**.

A crucial development across these applications is the ability of STGNNs to infer and adapt graph structures from sequential data itself, moving beyond reliance on pre-defined static graphs **??**. This addresses scenarios where explicit graph connectivity is unknown or noisy.

7.3.2 GNNs for Truly Dynamic Graphs

While STGNNs often assume a static or slowly changing underlying graph structure, truly dynamic graphs involve scenarios where the graph topology (nodes, edges, or their attributes) evolves rapidly and significantly over time. This includes interaction networks, social graphs, or evolving knowledge graphs. This area requires specialized architectures and formalizations to capture the temporal dynamics of the graph structure itself **?**. A comprehensive survey by **?** formalizes various learning settings and tasks for Temporal GNNs (TGNNs), categorizing them based on temporal representation (snapshot-based vs. event-based) and processing mechanisms.

One critical aspect in dynamic graphs is the learning and adaptation of graph structures. When initial graph structures are noisy, incomplete, or entirely absent, GNNs must infer them from data. This challenge is particularly acute in dynamic settings where relationships are constantly changing. Iterative Deep Graph Learning (IDGL) **?** is a framework that jointly and iteratively learns optimal graph structures and GNN param-

eters, making models more robust to imperfect topologies. Complementing this, SLAPS ? leverages self-supervision to improve structure learning for GNNs, addressing the "supervision starvation" problem inherent in learning graph topologies from limited labels (see also Section 5.3 for a broader discussion on handling imperfect data). For sequential recommendation, where user-item interactions form evolving graphs, ? transformed loose item sequences into dynamic item-item interest graphs, utilizing metric learning and attention for robust interest fusion (a topic further explored in Section 7.1). These methods underscore the shift from relying on fixed graph structures to adaptively learning them from evolving data.

Capturing complex temporal dynamics and memory effects beyond instantaneous changes is another significant challenge. Traditional continuous GNNs, relying on integer-order differential equations, often model Markovian updates, which might not suffice for systems with long-range dependencies or historical influence. To overcome this, FROND ? introduced a framework that generalizes continuous GNNs by incorporating Caputo fractional derivatives. This allows FROND to inherently capture non-local and memory-dependent dynamics, mitigating over-smoothing algebraically and offering a more nuanced understanding of real-world graph evolution. While fractional calculus has also been explored to mitigate oversmoothing in deep GNNs (see Section 5.1), its application here is distinct, focusing on modeling long-term memory in temporal dynamics. Similarly, ? proposed GNNs with Convolutional ARMA Filters, which offer a more flexible frequency response than polynomial filters, enabling the capture of longer-range dynamics with fewer parameters.

The formalization of dynamic graphs also distinguishes between discrete-time dynamic graphs (DTDGs), represented as a sequence of static snapshots, and continuous-time dynamic graphs (CTDGs), where changes occur as a stream of timestamped events ?. CTDGs offer richer temporal information but pose greater computational challenges. Frameworks like TGLite ? are emerging to provide lightweight programming abstractions and optimizations specifically for CTDG-based Temporal GNNs, facilitating the exploration of new designs and accelerating training and inference for these complex models. Further-

more, ensuring robust predictions in evolving systems requires GNNs to generalize well under "structure shift," where the underlying graph distribution changes. **?** addressed this by proposing the Cluster Information Transfer (CIT) mechanism, which learns invariant representations by manipulating node embeddings in the latent space, thereby enhancing GNN generalization to unseen graph structures.

In conclusion, the application of GNNs to time series and dynamic graphs has evolved significantly. STGNNs have become a cornerstone for multivariate time series forecasting, classification, imputation, and anomaly detection on structured data, leveraging various temporal modeling techniques combined with graph convolutions. Meanwhile, the field of truly dynamic graphs focuses on the more complex problem of evolving graph topologies, with advancements in dynamic graph structure learning, modeling memory-dependent dynamics through novel mathematical formalisms, and enhancing generalization under structural shifts. Despite these strides, challenges remain in developing unified frameworks that seamlessly handle both feature and topology evolution, scaling adaptive mechanisms to extremely large and rapidly changing graphs, and ensuring the interpretability of complex dynamic graph learning processes, especially in real-time, high-stakes applications.

7.4   Multi-modal and Semantic Understanding Tasks

While Graph Neural Networks (GNNs) have traditionally excelled in tasks driven by structural patterns and local attribute propagation, a pivotal frontier in their evolution lies in achieving deeper semantic understanding and effectively processing multi-modal data. This involves moving beyond purely topological insights to integrate GNNs with other advanced AI paradigms, particularly Large Language Models (LLMs) and computer vision models, to imbue them with real-world semantic knowledge, enable cross-domain reasoning, and facilitate robust generalization, including zero-shot capabilities. This section highlights cutting-edge applications where such integrations push GNNs into new domains of intelligence and versatility.

The foundation for leveraging GNNs in semantically rich and multi-modal contexts often builds upon advanced knowledge transfer paradigms, such as the pre-training and

prompt-based adaptation strategies discussed in Section 5.4. These methods enable GNNs to learn generalizable representations from abundant unlabeled data and efficiently adapt to diverse downstream tasks with limited supervision. With this established capability, the focus shifts to how GNNs can interpret and interact with external semantic information.

One significant avenue for semantic understanding involves the application of GNNs to text-rich graphs and Knowledge Graphs (KGs). In text classification, GNNs are increasingly employed to model complex relationships between words, documents, and entire corpora, moving beyond sequential models to capture global, contextual-aware word relations ?. By constructing graphs where nodes represent words or documents and edges denote relationships (e.g., co-occurrence, semantic similarity), GNNs can learn rich representations that encode the semantic context of text. Initial node features are often derived from powerful language models (e.g., BERT, ELMo), which GNNs then refine by propagating information across the graph structure, thereby enhancing the semantic understanding for classification tasks ?.

Knowledge Graphs, which explicitly represent factual information and relationships between entities, are inherently multi-relational graphs that demand deep semantic understanding. GNNs have become indispensable for various KG tasks, including link prediction, knowledge graph alignment, and knowledge graph reasoning ?. By learning embeddings for entities and relations, GNNs enable models to infer missing links or answer complex queries, effectively performing symbolic reasoning over structured semantic knowledge. For instance, GNNs can identify complex interaction patterns in drug-drug interaction graphs or enhance recommender systems by leveraging rich semantic relationships between items and users within a KG ?. Beyond explicit KGs, methods like Knowledge-Enhanced GNNs (KE-GNN) integrate propositional logic rules into the learning process, mapping knowledge into a semantic space and using a teacher-student scheme to guide GNN predictions, thereby enhancing semantic understanding and robustness for tasks like fault scenario identification in communication networks ?. These approaches demonstrate how GNNs, when augmented with structured knowledge, can achieve more accurate and interpretable semantic reasoning.

60

The synergy between GNNs and Large Language Models (LLMs) represents a cutting-edge direction for deeper semantic understanding. LLMs, with their vast world knowledge and powerful inference capabilities, can significantly enhance GNNs. One approach is to leverage LLMs to enrich node features or improve graph structure. For example, LLMs can generate more informative textual embeddings for graph nodes, providing a richer semantic context than traditional attribute vectors. Furthermore, LLMs can contribute to the robustness of GNNs against adversarial attacks by identifying malicious edges or inferring missing important connections to recover a more robust graph structure ?. The LLM4RGNN framework, for instance, distills GPT-4's inference capabilities into a local LLM to identify perturbed edges, showcasing how LLMs can semantically "reason" about graph integrity and improve GNN resilience, even in highly perturbed scenarios ?.

A particularly groundbreaking paradigm for integrating GNNs and LLMs is multi-modal prompt learning, which aims to align independently pre-trained GNNs with the rich semantic embedding spaces of LLMs, even under extremely weak text supervision. This approach enables GNNs to perform CLIP-style zero-shot generalization, classifying unseen graph categories by leveraging the semantic understanding encoded in LLMs. A notable example is **Morpher (Multi-modal Prompt Learning for Graph Neural Networks) ?**. Morpher addresses critical challenges in graph-text alignment, such as data scarcity and the weak nature of text supervision (e.g., single-word labels), which hinder joint pre-training. Its innovations include an improved, stable graph prompt design that carefully balances cross-connections between prompt tokens and input graph nodes with the original graph's inner-connections, preventing prompt features from overwhelming the graph structure. Coupled with tunable text prompts and a cross-modal projector, Morpher aligns graph and text representations in a shared semantic space using contrastive learning, keeping the GNN and LLM backbones frozen. This allows GNNs to dynamically adapt to semantic cues from LLMs, achieving significant accuracy improvements in few-shot and zero-shot graph classification tasks across diverse datasets (e.g., molecular, bioinformatic, citation networks) ?. Morpher's ability to imbue GNNs with real-world semantic knowledge without extensive paired data or fine-tuning represents a significant

61

leap towards more intelligent and versatile graph-based AI systems.

Beyond text, GNNs are also crucial for multi-modal understanding in computer vision, particularly in tasks involving vision and language. In these applications, GNNs model the relationships between visual entities (e.g., objects in an image, points in a 3D scene) and their linguistic descriptions. For instance, GNNs are used to construct scene graphs from images, where nodes represent objects and edges represent their interactions or attributes, which can then be used for image captioning, visual question answering, or complex scene understanding ?. By capturing the relational structure inherent in visual data, GNNs provide a powerful mechanism to integrate visual and linguistic information, enabling models to reason about complex visual scenes and generate semantically coherent textual descriptions. The survey by ? highlights how GNNs and Graph Transformers are applied across various computer vision tasks, including those that explicitly bridge vision and language modalities.

In conclusion, the integration of GNNs with other advanced AI paradigms, particularly LLMs and computer vision models, is rapidly expanding their capabilities beyond purely structural analysis. From leveraging GNNs for semantic understanding in text classification and Knowledge Graphs to employing LLMs for enhancing GNN robustness and enabling multi-modal prompt learning for zero-shot generalization, these applications are transforming GNNs into more semantically aware and versatile systems. Future research will likely explore more complex multi-modal reasoning tasks, develop more robust and interpretable multi-modal prompt designs, and extend zero-shot generalization capabilities to a wider array of graph learning problems, ultimately fostering GNNs that can truly "understand" and interact with the world's knowledge across diverse modalities.

8 Future Directions and Societal Impact

8.1 Open Challenges and Emerging Trends

Despite the remarkable progress in Graph Neural Networks (GNNs), the field continues to grapple with fundamental unresolved tensions and theoretical gaps that drive ongoing

research, while new paradigms and interdisciplinary integrations are rapidly emerging. These challenges span from the intrinsic limitations of GNN architectures to their robustness in complex real-world scenarios and their capacity for advanced reasoning.

A persistent and central challenge lies in the quest for maximum expressive power while maintaining scalability and depth. While the 1-Weisfeiler-Leman (1-WL) test provides a foundational benchmark for GNN discriminative power, the quest for truly expressive yet tractable architectures remains a central unresolved tension **?**. Seminal work established that many message-passing GNNs are fundamentally limited to 1-WL power, restricting their ability to distinguish complex non-isomorphic graphs **??**. Even advanced GNNs, including those incorporating spatial cues like port numbering, have been theoretically shown to be unable to compute fundamental graph properties like girth, diameter, or k-clique counts, revealing intrinsic representational limits **?**. This highlights a critical gap: how can GNNs be designed to capture such global or higher-order structural information without resorting to computationally expensive k-GNNs or sacrificing scalability **??** Furthermore, the challenge of building deeper, more expressive GNNs is compounded by oversmoothing, where node representations converge, limiting the network's discriminative capacity **??**. This phenomenon, even observed in attention-based GNNs **?**, necessitates architectural innovations. While solutions like Personalized Propagation of Neural Predictions (PPNP) **?** and Energetic GNNs (EGNN) **?** mitigate oversmoothing, a deeper theoretical understanding of information diffusion in deep GNNs, perhaps through the lens of Partial Differential Equations (PDEs) **?**, is crucial for principled design. The overarching challenge is to develop a unified theoretical framework that bridges expressivity and scalability, ensuring that GNNs can both discern intricate graph structures and operate efficiently on massive, real-world datasets, a relationship further explored in recent theoretical work **?**.

Beyond architectural expressivity, a paramount challenge for GNNs is achieving robust generalization to out-of-distribution (OOD) data, where underlying data generation processes or graph structures shift between training and test environments. While pre-training strategies **??** and adaptive filtering mechanisms for heterophily **??** offer practical

improvements, the core problem lies in learning *invariant causal mechanisms* rather than spurious correlations. GNNs are susceptible to learning from "shortcut features" or biases in training data, leading to poor OOD performance **??**. For instance, models can struggle to disentangle causal substructures from biased ones, especially under severe bias conditions **?**. This necessitates methods like Discovering Invariant Rationale (DIR), which employs causal interventions to identify stable patterns across synthetic environments **?**. Similarly, the Cluster Information Transfer (CIT) mechanism aims to learn representations robust to "structure shift" **?**. A critical open question is how to develop GNNs that inherently perform causal discovery and reasoning on graphs, moving beyond correlation to robustly infer underlying generative processes, as exemplified by efforts in fraud detection using causal temporal GNNs **?**.

The integration of GNNs with other advanced AI paradigms represents a rapidly emerging frontier. The synergy with Large Language Models (LLMs) is particularly transformative, enabling GNNs to leverage vast external semantic knowledge for tasks like zero-shot generalization through multi-modal prompt learning **?**. However, this raises new challenges: how to achieve true compositional reasoning by seamlessly combining structural and semantic knowledge, prevent "semantic hijacking" where LLM priors overshadow valid structural insights, and establish appropriate evaluation protocols for such hybrid models. Mathematically, the exploration of novel frameworks like fractional calculus, as seen in the FROND framework, offers a promising path to model non-local, memory-dependent graph dynamics and algebraically mitigate oversmoothing **?**. The broader applicability and theoretical guarantees of such continuous GNNs remain an active area. Furthermore, hybrid architectural paradigms, such as Spatio-Spectral GNNs **?** and non-convolutional approaches like Random Walk with Unifying Memory (RUM) networks **?**, seek to combine the strengths of different filtering mechanisms or depart from message-passing entirely. The principled design of these hybrid models, particularly for tasks like combinatorial optimization where physics-inspired GNNs show promise **??**, requires a deeper understanding of how to optimally combine local and global information.

The dynamic and evolving nature of real-world graphs presents another significant

challenge. While Temporal GNNs (TGNNs) are being developed to capture evolving connections and features **?**, a key unresolved tension is the development of *lifetime-learning GNNs*. These models must continuously adapt to non-stationary graph dynamics without catastrophic forgetting, efficiently incorporating new nodes, edges, and attribute changes in an online fashion, which is crucial for applications like evolving social networks or real-time recommendation systems **?**.

Ultimately, the field faces the grand challenge of bridging the theoretical understanding of GNN capabilities with practical architectural design. While theoretical studies have begun to characterize the intricate relationship between expressivity and generalization, suggesting a potential trade-off **?**, empirical evidence often presents a more complex picture. A unified theory that balances expressive power, computational efficiency, and transferability across diverse graph distributions remains elusive. Moreover, moving beyond post-hoc explanations to *interpretable-by-design* GNNs, and developing robust evaluation frameworks that can assess these multifaceted capabilities, are crucial for the responsible deployment and continued advancement of GNNs in high-stakes applications. The decidability of GNNs via logical characterizations also represents a nascent but critical theoretical frontier for understanding their fundamental computational limits **?**.

In conclusion, the GNN landscape is defined by a vibrant interplay between persistent theoretical challenges and exciting emerging trends. The ongoing quest to reconcile maximal expressive power with scalability and depth, coupled with the critical need for robust OOD generalization through causal reasoning, continues to drive fundamental research. Simultaneously, the integration of GNNs with powerful AI paradigms like LLMs, the exploration of novel mathematical frameworks, and the demand for adaptive models for dynamic graphs are opening new frontiers. Addressing these multifaceted challenges requires interdisciplinary efforts, pushing the boundaries towards more intelligent, reliable, and ethically sound graph learning solutions that can truly unlock the potential of interconnected data in an increasingly complex world.

8.2 Ethical Considerations and Responsible AI

As Graph Neural Networks (GNNs) continue their pervasive integration into high-stakes societal domains such as social networks, healthcare, finance, and critical infrastructure, the imperative for a proactive and holistic focus on ethical considerations becomes paramount. This subsection, positioned within the future outlook, synthesizes the critical open challenges and interdependencies in ensuring GNNs are not only performant but also transparent, accountable, fair, private, and robust against misuse. While Section 6 provides a detailed exposition of current methodologies for trustworthy GNNs, this section projects forward, emphasizing the need for inherently responsible AI systems and addressing the complex trade-offs that define the frontier of ethical graph learning.

A central future direction involves navigating the intricate **interdependencies and potential trade-offs** among these ethical pillars. Achieving a robust GNN might inadvertently reduce its explainability or introduce biases, while stringent privacy measures could impact model utility or fairness. The comprehensive survey by **?** underscores these complex connections, advocating for a unified framework that considers these dimensions synergistically. Future research must move beyond isolated solutions, striving to design GNNs that are *inherently* fair, private, robust, and explainable by design, rather than relying on post-hoc patches. This requires novel architectural inductive biases and training paradigms that embed ethical principles from the ground up.

Fairness remains a critical, evolving challenge, particularly given the inherent biases amplified by graph structures and message-passing mechanisms. While Section 6.3 details foundational fairness concepts, future work must address more nuanced aspects. For instance, ensuring **individual fairness**, where similar individuals receive similar predictions regardless of sensitive attributes, is more complex than group fairness on interconnected data. **?** introduced REDRESS, a ranking-based framework to promote individual fairness by refining its definition and integrating it into an end-to-end training process, demonstrating a promising direction. However, scaling such approaches to diverse graph topologies and defining "similarity" in complex, multi-relational graphs remains an open problem. Furthermore, practical deployment often faces the challenge of **limited sen-**

sitive attribute information. ? proposed FairGNN, which tackles this by estimating missing sensitive attributes via a GCN-based estimator, enabling adversarial debiasing. This highlights a crucial future avenue: developing robust methods for fairness when sensitive data is scarce or incomplete. Beyond model-specific interventions, **data-centric debiasing** presents a powerful future frontier. ? introduced EDITS, a model-agnostic framework that mitigates bias directly in the input attributed network by defining and optimizing against attribute and structural bias metrics. This pre-processing approach offers a universal solution, but its effectiveness relies on accurate bias modeling and the ability to maintain utility. A more profound understanding of **causal inference** in GNNs is also critical for fairness. ? addressed the problem of severe bias by learning disentangled causal substructures, enabling GNNs to generalize better to out-of-distribution data. Future research needs to further integrate causal reasoning into GNN architectures to distinguish genuine causal patterns from spurious correlations, thereby building inherently fairer models.

In terms of **privacy**, while federated GNNs (FedGNNs) offer a promising paradigm for collaborative learning without raw data sharing ?, significant challenges persist. As discussed in Section 6.3, FedGNNs are crucial for sensitive domains like healthcare, where frameworks like that by ? enable disease classification on PPI networks while preserving patient privacy. However, future research must address the heterogeneity of graph data across clients, develop more robust and communication-efficient aggregation mechanisms, and provide stronger theoretical privacy guarantees against inference attacks on the aggregated models. Advanced **differential privacy (DP)** techniques tailored for graph structures are also needed, considering the unique challenges of perturbing nodes, edges, and features while preserving graph utility and avoiding utility collapse.

The landscape of **robustness** against adversarial attacks is continuously evolving, demanding adaptive and proactive defense mechanisms. While Section 6.2 covers existing defenses, the critical assessment by ? highlights that many are brittle against adaptive adversaries. Future work requires more rigorous evaluation protocols and the development of truly adaptive and certifiable robust GNNs. New attack vectors, such as **Graph**

67

Injection Attacks (GIA), where adversaries inject new nodes without modifying existing structures, pose significant real-world threats **?**. Addressing such evolving threats necessitates GNNs designed with inherent resilience, perhaps through novel architectural inductive biases or by leveraging invariant learning strategies **?** that can distinguish causal patterns from spurious correlations. Furthermore, the theoretical understanding of GNN robustness is still nascent. **?** provided a theoretical definition of expected robustness for attributed graphs and derived upper bounds for GCNs and GINs against node feature attacks, proposing GCORNs as a more robust variant. This highlights the need for more principled, theoretically grounded approaches to robustness. The emergence of **backdoor attacks** on GNNs, where a trigger subgraph can be injected into training data to manipulate predictions during inference **?**, further complicates the security landscape, demanding novel detection and defense mechanisms beyond traditional adversarial training.

Finally, **explainability and transparency** remain critical for fostering accountability and trust. As detailed in Section 6.1, methods like GNNExplainer **?** and XGNN **?** offer insights, but theoretical limitations persist, with **?** demonstrating approximation failures in prevalent attention-based interpretable GNNs. Future research must focus on developing more faithful, structure-aware explanation methods, such as MAGE **?**, which accounts for graph connectivity and high-order interactions. Crucially, explanations must not only be accurate but also private and robust against manipulation, ensuring they genuinely contribute to understanding and debugging GNN behavior rather than providing misleading insights.

In conclusion, the journey towards truly responsible AI in GNNs is multifaceted, requiring a paradigm shift from addressing individual problems to developing integrated, holistic solutions. Future research must prioritize understanding and managing the inherent trade-offs between fairness, privacy, robustness, and explainability. This includes advancing federated learning for privacy-preserving graph analysis, developing inherently robust architectures against evolving adversarial threats, and creating more faithful and robust explanation methods. Ultimately, achieving equitable and secure outcomes for all users will necessitate not only technological advancements but also interdisciplinary

approaches that integrate legal, ethical, and sociological perspectives into the design and deployment of GNNs.

9 Conclusion

9.1 Summary of Key Advancements

The past decade has been a period of transformative growth for Graph Neural Networks (GNNs), evolving from nascent theoretical constructs into a cornerstone of modern artificial intelligence for structured data. This rapid evolution, comprehensively documented by recent surveys ???, reflects a continuous intellectual journey marked by foundational breakthroughs, critical self-assessment of limitations, and the development of sophisticated solutions across multiple fronts.

The initial conceptualizations of GNNs, pioneered by works like ? and ?, established the mathematical framework for learning on graphs through iterative information propagation. While theoretically sound, these early models faced significant computational hurdles, limiting their practical applicability. A pivotal shift occurred with the advent of the message-passing paradigm, which simplified and localized graph convolutions. Architectures like Graph Convolutional Networks (GCNs) and GraphSAGE ? democratized GNNs by enabling scalable and inductive learning, allowing models to generalize to unseen nodes and graphs. This architectural innovation quickly translated into practical success, exemplified by the deployment of GNNs in web-scale recommender systems, demonstrating their capacity for massive real-world applications ?.

As GNNs gained prominence, a deeper theoretical understanding of their expressive power became imperative. This led to a critical intellectual reckoning, notably with the realization that many message-passing GNNs are fundamentally limited by their equivalence to the 1-Weisfeiler-Leman (1-WL) graph isomorphism test ??. This insight exposed their inability to distinguish certain non-isomorphic graphs, directly catalyzing a new wave of research aimed at surpassing this "WL barrier." Solutions emerged through higher-order GNNs, such as k-GNNs ?, which operate on k-tuples of nodes to capture richer struc-

tural information, and methods that explicitly leverage path or substructure patterns **?**. Concurrently, for applications in scientific domains like chemistry and physics, the development of geometric and E(n) equivariant GNNs (e.g., EGNNs **?** and NequIP **?**) marked a significant advancement, ensuring that models inherently respect physical symmetries, leading to more data-efficient and physically consistent predictions. This period also saw a growing emphasis on rigorous theoretical guarantees for GNNs, with new PAC-Bayesian bounds offering tighter and more realistic quantifications of generalization performance, moving beyond empirical observations to principled understanding **?**.

Addressing practical challenges has been another central theme in GNN research. The "oversmoothing" problem, where node representations become indistinguishable in deep GNNs, was initially identified as a major impediment to building deeper architectures capable of capturing long-range dependencies. Early solutions like PPNP **?** decoupled propagation from prediction, while theoretical analyses formalized oversmoothing using Dirichlet energy **??**, guiding the design of deeper models **?**. More recently, a critical re-evaluation has suggested that the trainability challenges of the underlying MLPs, rather than oversmoothing alone, might be the dominant factor limiting deep GNN performance **?**. This deeper understanding has spurred the development of fundamentally non-convolutional architectures, such as RUM **?** and FROND **?**, which aim to simultaneously address expressivity, oversmoothing, and over-squashing by moving beyond traditional message-passing paradigms. Scalability for massive graphs, a persistent industrial concern, has been tackled through innovations like graph condensation **?** and efficient sampling techniques, enabling GNNs to operate on billions of nodes and edges. Furthermore, the field has developed robust mechanisms to handle the complexities of real-world graphs, which often exhibit heterophily (dissimilar connected nodes), missing information, or noisy connections **??**. Adaptive filtering mechanisms **??**, joint structure learning **??**, and data augmentation techniques **?** have significantly enhanced GNN resilience to diverse and imperfect graph topologies.

The maturation of GNNs is also evident in the increasing focus on trustworthiness and efficient knowledge transfer. Moving beyond mere predictive accuracy, research has

increasingly emphasized interpretability, with comprehensive taxonomies and standardized testbeds emerging to explain GNN decisions and build user trust **?**. Efforts to enhance robustness against adversarial attacks and mitigate inherent biases, alongside strategies to protect privacy, are crucial for responsible deployment in high-stakes applications. Moreover, the paradigm of knowledge transfer has seen significant advancements, with self-supervised pre-training strategies **??** learning generalizable representations from abundant unlabeled graph data. The emerging field of prompt-based adaptation further enables pre-trained GNNs to generalize to diverse downstream tasks with minimal labeled data, often by integrating with large language models for semantic understanding **?**.

Finally, the field has embraced a more principled approach to evaluation and design. The establishment of comprehensive, open-source benchmarking frameworks, such as OGB **?**, has set rigorous standards for comparing GNN architectures on discriminative datasets, fostering reproducibility and accelerating progress. Unifying optimization frameworks **?** have provided deeper insights into the propagation mechanisms of various GNNs, offering a principled path for future model design. Despite these significant advancements, unresolved tensions persist, particularly in balancing maximum expressive power with computational efficiency and achieving robust out-of-distribution generalization. The ongoing integration of GNNs into complex, multi-modal AI systems and the continuous pursuit of more adaptive and interpretable mechanisms will undoubtedly shape the next generation of graph learning solutions.

9.2 Future Outlook

The trajectory of Graph Neural Network (GNN) research is rapidly evolving, moving towards a future defined by increasingly intelligent, adaptive, and responsible graph-based AI systems. This concluding outlook synthesizes the most promising avenues and persistent challenges, emphasizing a continuous drive towards GNNs that are not only powerful but also robust, interpretable, and capable of deep reasoning across complex, real-world scenarios.

A central theme for the next phase of GNN development is the pursuit of architectures

that strike an optimal balance between maximal expressive power and computational efficiency. While significant strides have been made in overcoming the limitations of the 1-Weisfeiler-Leman (1-WL) test **??** through higher-order GNNs **?** and geometric principles **?**, the quest for truly universal and scalable expressivity continues. Future work will delve into novel mathematical frameworks, such as fractional calculus, which offers a principled approach to model non-local, memory-dependent graph dynamics and algebraically mitigate issues like oversmoothing, thereby enabling deeper and more expressive GNNs **?**. Concurrently, the theoretical understanding of GNN generalization, including their VC dimension and statistical properties on manifolds, will become paramount for designing robust models that generalize to unseen data over diverse graph structures **??**. Scalability to web-scale graphs remains a critical practical challenge, necessitating further advancements in efficient approximation techniques, graph condensation methods **?**, and the exploration of novel graph structures like superhypergraphs to model increasingly complex relationships **?**. The ultimate goal is to develop adaptive GNNs that can dynamically adjust their complexity and receptive field based on the inherent structure and task requirements.

Another defining frontier lies in the development of truly foundational GNN models capable of universal pre-training and intrinsically adaptable generalization. Current pre-training strategies, while effective, often require substantial fine-tuning or intricate prompt engineering **??**. The future envisions GNNs that can generalize across a wide spectrum of graph types and tasks with minimal adaptation, effectively minimizing the "objective gap" between pre-training and downstream applications. This will be significantly propelled by multi-modal learning, particularly the deeper integration of GNNs with large language models (LLMs) to imbue graph representations with rich semantic understanding and enable zero-shot or few-shot generalization across novel graph distributions **??**. The advancement of universal prompt tuning methods **??** will continue to enhance parameter-efficient adaptation, pushing towards GNNs that require less human intervention in task reformulation.

Crucially, GNNs must evolve to reason and generalize robustly across diverse, complex

real-world scenarios, moving beyond learning spurious correlations to identifying invariant causal mechanisms. This involves not only addressing structural complexities like heterophily and imperfect data through adaptive filtering and structure learning **???** but also a fundamental shift towards causal reasoning. Future GNNs will aim to discover invariant causal substructures, making them inherently more robust to out-of-distribution shifts and biases **???**. This causal perspective is indispensable for high-stakes applications where understanding the true underlying mechanisms, rather than mere statistical associations, is vital for reliable decision-making and preventing unintended consequences **?**.

Finally, the increasing interdisciplinary nature of GNN research will unlock new capabilities in AI, emphasizing a continuous drive towards more intelligent and responsible graph-based systems. This involves deeper integration with other AI paradigms, such as federated learning to ensure privacy and distributed intelligence **?**, and neuro-symbolic AI to augment GNNs' pattern recognition with formal reasoning capabilities. The maturation of comprehensive benchmarking frameworks **?** will continue to accelerate principled research, while a heightened focus on trustworthiness will ensure that GNNs are not only powerful but also explainable **?**, fair, and secure against adversarial attacks **?**. This holistic approach, prioritizing ethical deployment alongside technological advancement, will solidify GNNs as indispensable tools for scientific discovery, personalized systems, and complex societal problem-solving, shaping the future of artificial intelligence with increasing sophistication and responsibility.

References

Yuanqing Wang, and Kyunghyun Cho (2024). *Non-convolutional Graph Neural Networks.* Neural Information Processing Systems.

Xiang Li, Renyu Zhu, Yao Cheng, et al. (2022). *Finding Global Homophily in Graph Neural Networks When Meeting Heterophily.* International Conference on Machine Learning.

Qiyu Kang, Kai Zhao, Qinxu Ding, et al. (2024). *Unleashing the Potential of Fractional Calculus in Graph Neural Networks with FROND.* International Conference on Learning Representations.

Chen Gao, Xiang Wang, Xiangnan He, et al. (2022). *Graph Neural Networks for Recommender System.* Web Search and Data Mining.

Juanhui Li, Harry Shomer, Haitao Mao, et al. (2023). *Evaluating Graph Neural Networks for Link Prediction: Current Pitfalls and New Benchmarking.* Neural Information Processing Systems.

Gaspard Michel, Giannis Nikolentzos, J. Lutzeyer, et al. (2023). *Path Neural Networks: Expressive and Accurate Graph Neural Networks.* International Conference on Machine Learning.

Chaoqi Chen, Yushuang Wu, Qiyuan Dai, et al. (2022). *A Survey on Graph Neural Networks and Graph Transformers in Computer Vision: A Task-Oriented Perspective.* IEEE Transactions on Pattern Analysis and Machine Intelligence.

Hao Yuan, Haiyang Yu, Jie Wang, et al. (2021). *On Explainability of Graph Neural Networks via Subgraph Explorations.* International Conference on Machine Learning.

Yushun Dong, Jian Kang, H. Tong, et al. (2021). *Individual Fairness for Graph Neural Networks: A Ranking based Approach.* Knowledge Discovery and Data Mining.

Quentin Cappart, D. Chételat, Elias Boutros Khalil, et al. (2021). *Combinatorial optimization and reasoning with graph neural networks*. International Joint Conference on Artificial Intelligence.

Yushun Dong, Ninghao Liu, B. Jalaeian, et al. (2021). *EDITS: Modeling and Mitigating Data Bias for Graph Neural Networks*. The Web Conference.

Zhixun Li, Yushun Dong, Qiang Liu, et al. (2024). *Rethinking Fair Graph Neural Networks from Re-balancing*. Knowledge Discovery and Data Mining.

Tong Zhao, Yozen Liu, Leonardo Neves, et al. (2020). *Data Augmentation for Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

Chaitanya K. Joshi, and Simon V. Mathis (2023). *On the Expressive Power of Geometric Graph Neural Networks*. International Conference on Machine Learning.

Mingchen Sun, Kaixiong Zhou, Xingbo He, et al. (2022). *GPPT: Graph Pre-training and Prompt Tuning to Generalize Graph Neural Networks*. Knowledge Discovery and Data Mining.

Austin Derrow-Pinion, Jennifer She, David Wong, et al. (2021). *ETA Prediction with Graph Neural Networks in Google Maps*. International Conference on Information and Knowledge Management.

Yu Chen, Lingfei Wu, and Mohammed J. Zaki (2020). *Iterative Deep Graph Learning for Graph Neural Networks: Better and Robust Node Embeddings*. Neural Information Processing Systems.

Hanqing Zeng, Muhan Zhang, Yinglong Xia, et al. (2022). *Decoupling the Depth and Scope of Graph Neural Networks*. Neural Information Processing Systems.

Haonan Yuan, Jiliang Tang, Xia Hu, et al. (2020). *XGNN: Towards Model-Level Explanations of Graph Neural Networks*. Knowledge Discovery and Data Mining.

Yaochen Xie, Zhao Xu, Zhengyang Wang, et al. (2021). *Self-Supervised Learning of Graph Neural Networks: A Unified Review*. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Shaswata Mitra, Trisha Chakraborty, Subash Neupane, et al. (2024). *Use of Graph Neural Networks in Aiding Defensive Cyber Operations*. arXiv.org.

Muhan Zhang, and Pan Li (2021). *Nested Graph Neural Networks*. Neural Information Processing Systems.

Hongya Wang, Haoteng Yin, Muhan Zhang, et al. (2022). *Equivariant and Stable Positional Encoding for More Powerful Graph Neural Networks*. International Conference on Learning Representations.

Yuanfu Lu, Xunqiang Jiang, Yuan Fang, et al. (2021). *Learning to Pre-train Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

Shaohua Fan, Xiao Wang, Yanhu Mo, et al. (2022). *Debiasing Graph Neural Networks via Learning Disentangled Causal Substructure*. Neural Information Processing Systems.

Zaixi Zhang, Jinyuan Jia, Binghui Wang, et al. (2020). *Backdoor Attacks to Graph Neural Networks*. ACM Symposium on Access Control Models and Technologies.

Hejie Cui, Wei Dai, Yanqiao Zhu, et al. (2022). *BrainGB: A Benchmark for Brain Network Analysis With Graph Neural Networks*. IEEE Transactions on Medical Imaging.

Ngoc Bui, Hieu Trung Nguyen, Viet Anh Nguyen, et al. (2024). *Explaining Graph Neural Networks via Structure-aware Interaction Index*. International Conference on Machine Learning.

Chuang Liu, Yibing Zhan, Chang Li, et al. (2022). *Graph Pooling for Graph Neural Networks: Progress, Challenges, and Opportunities*. International Joint Conference on Artificial Intelligence.

Ming Jin, Huan Yee Koh, Qingsong Wen, et al. (2023). *A Survey on Graph Neural Networks for Time Series: Forecasting, Classification, Imputation, and Anomaly Detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence.

Rex Ying, Dylan Bourgeois, Jiaxuan You, et al. (2019). *GNNExplainer: Generating Explanations for Graph Neural Networks.* Neural Information Processing Systems.

Meng Liu, Hongyang Gao, and Shuiwang Ji (2020). *Towards Deeper Graph Neural Networks.* Knowledge Discovery and Data Mining.

Antonio Longa, Veronica Lachi, G. Santin, et al. (2023). *Graph Neural Networks for temporal graphs: State of the art, open challenges, and opportunities.* Trans. Mach. Learn. Res..

P. Papp, Karolis Martinkus, Lukas Faber, et al. (2021). *DropGNN: Random Dropouts Increase the Expressiveness of Graph Neural Networks.* Neural Information Processing Systems.

Jianxin Chang, Chen Gao, Y. Zheng, et al. (2021). *Sequential Recommendation with Graph Neural Networks.* Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.

Felix Mujkanovic, Simon Geisler, Stephan Gunnemann, et al. (2023). *Are Defenses for Graph Neural Networks Robust?.* Neural Information Processing Systems.

Jiaxuan You, Jonathan M. Gomes-Selman, Rex Ying, et al. (2021). *Identity-aware Graph Neural Networks.* AAAI Conference on Artificial Intelligence.

Dongsheng Luo, Tianxiang Zhao, Wei Cheng, et al. (2024). *Towards Inductive and Efficient Explanations for Graph Neural Networks.* IEEE Transactions on Pattern Analysis and Machine Intelligence.

Hejie Cui, Wei Dai, Yanqiao Zhu, et al. (2022). *Interpretable Graph Neural Networks for Connectome-Based Brain Disorder Analysis.* International Conference on Medical Image Computing and Computer-Assisted Intervention.

Enyan Dai, Wei-dong Jin, Hui Liu, et al. (2022). *Towards Robust Graph Neural Networks for Noisy Graphs with Sparse Labels*. Web Search and Data Mining.

Kunze Wang, Yihao Ding, and S. Han (2023). *Graph Neural Networks for Text Classification: A Survey*. Artificial Intelligence Review.

Bharti Khemani, S. Patil, K. Kotecha, et al. (2024). *A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions*. Journal of Big Data.

Chirag Agarwal, Owen Queen, Himabindu Lakkaraju, et al. (2022). *Evaluating explainability for graph neural networks*. Scientific Data.

Vijay Prakash Dwivedi, Chaitanya K. Joshi, T. Laurent, et al. (2023). *Benchmarking Graph Neural Networks*. Journal of machine learning research.

Ralph Abboud, .Ismail .Ilkan Ceylan, Martin Grohe, et al. (2020). *The Surprising Power of Graph Neural Networks with Random Node Initialization*. International Joint Conference on Artificial Intelligence.

Yixin Liu, Kaize Ding, Jianling Wang, et al. (2023). *Learning Strong Graph Neural Networks with Weak Information*. Knowledge Discovery and Data Mining.

Xiaorui Liu, W. Jin, Yao Ma, et al. (2021). *Elastic Graph Neural Networks*. International Conference on Machine Learning.

M. Balcilar, P. Héroux, Benoit Gaüzère, et al. (2021). *Breaking the Limits of Message Passing Graph Neural Networks*. International Conference on Machine Learning.

Weihua Hu, Bowen Liu, Joseph Gomes, et al. (2019). *Strategies for Pre-training Graph Neural Networks*. International Conference on Learning Representations.

B. Chamberlain, S. Shirobokov, Emanuele Rossi, et al. (2022). *Graph Neural Networks for Link Prediction with Subgraph Sketching*. International Conference on Learning Representations.

Patrick Reiser, Marlen Neubert, Andr'e Eberhard, et al. (2022). *Graph neural networks for materials science and chemistry.* Communications Materials.

Guohao Li, Matthias Müller, Bernard Ghanem, et al. (2021). *Training Graph Neural Networks with 1000 Layers.* International Conference on Machine Learning.

Xiyuan Wang, and Muhan Zhang (2022). *How Powerful are Spectral Graph Neural Networks.* International Conference on Machine Learning.

Zaixin Zhang, Qi Liu, Hao Wang, et al. (2021). *ProtGNN: Towards Self-Explaining Graph Neural Networks.* AAAI Conference on Artificial Intelligence.

Vikas K. Garg, S. Jegelka, and T. Jaakkola (2020). *Generalization and Representational Limits of Graph Neural Networks.* International Conference on Machine Learning.

Bahare Fatemi, Layla El Asri, and Seyed Mehran Kazemi (2021). *SLAPS: Self-Supervision Improves Structure Learning for Graph Neural Networks.* Neural Information Processing Systems.

Yuyu Zhang, Xinshi Chen, Yuan Yang, et al. (2021). *Graph Neural Networks.* Deep Learning on Graphs.

Anna Varbella, Kenza Amara, B. Gjorgiev, et al. (2024). *PowerGraph: A power grid benchmark dataset for graph neural networks.* Neural Information Processing Systems.

T. Konstantin Rusch, Michael M. Bronstein, and Siddhartha Mishra (2023). *A Survey on Oversmoothing in Graph Neural Networks.* arXiv.org.

Zhengdao Chen, Lei Chen, Soledad Villar, et al. (2020). *Can graph neural networks count substructures?.* Neural Information Processing Systems.

He Zhang, Bang Wu, Xingliang Yuan, et al. (2022). *Trustworthy Graph Neural Networks: Aspects, Methods, and Trends.* Proceedings of the IEEE.

Haoyu Han, Juanhui Li, Wei Huang, et al. (2024). *Node-wise Filtering in Graph Neural Networks: A Mixture of Experts Approach.* arXiv.org.

Emanuele Rossi, Fabrizio Frasca, B. Chamberlain, et al. (2020). *SIGN: Scalable Inception Graph Neural Networks*. arXiv.org.

Yingmin Wu, Xiang Wang, An Zhang, et al. (2022). *Discovering Invariant Rationales for Graph Neural Networks*. International Conference on Learning Representations.

Christopher Morris, Martin Ritzert, Matthias Fey, et al. (2018). *Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

Enyan Dai, Tianxiang Zhao, Huaisheng Zhu, et al. (2022). *A Comprehensive Survey on Trustworthy Graph Neural Networks: Privacy, Robustness, Fairness, and Explainability*. Machine Intelligence Research.

Yuwen Wang, Shunyu Liu, Tongya Zheng, et al. (2024). *Unveiling Global Interactive Patterns across Graphs: Towards Interpretable Graph Neural Networks*. Knowledge Discovery and Data Mining.

Haotian Ju, Dongyue Li, Aneesh Sharma, et al. (2023). *Generalization in Graph Neural Networks: Improved PAC-Bayesian Bounds on Graph Diffusion*. International Conference on Artificial Intelligence and Statistics.

Zewen Liu, Guancheng Wan, B. A. Prakash, et al. (2024). *A Review of Graph Neural Networks in Epidemic Modeling*. Knowledge Discovery and Data Mining.

Muhan Zhang, and Yixin Chen (2018). *Link Prediction Based on Graph Neural Networks*. Neural Information Processing Systems.

F. Bianchi, Daniele Grattarola, L. Livi, et al. (2019). *Graph Neural Networks With Convolutional ARMA Filters*. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Yao Ma, Xiaorui Liu, Neil Shah, et al. (2021). *Is Homophily a Necessity for Graph Neural Networks?*. International Conference on Learning Representations.

Li, Lecheng Zheng, Bowen Jin, et al. (2024). *Can Graph Neural Networks Learn Language with Extremely Weak Text Supervision?*. Annual Meeting of the Association for Computational Linguistics.

Xinlei He, Jinyuan Jia, M. Backes, et al. (2020). *Stealing Links from Graph Neural Networks*. USENIX Security Symposium.

Taoran Fang, Yunchao Zhang, Yang Yang, et al. (2022). *Universal Prompt Tuning for Graph Neural Networks*. Neural Information Processing Systems.

Yongqiang Chen, Yatao Bian, Bo Han, et al. (2024). *How Interpretable Are Interpretable Graph Neural Networks?*. International Conference on Machine Learning.

Zemin Liu, Xingtong Yu, Yuan Fang, et al. (2023). *GraphPrompt: Unifying Pre-Training and Downstream Tasks for Graph Neural Networks*. The Web Conference.

Guimin Dong, Mingyue Tang, Zhiyuan Wang, et al. (2022). *Graph Neural Networks in IoT: A Survey*. ACM Trans. Sens. Networks.

Wenqi Fan, Yao Ma, Qing Li, et al. (2019). *Graph Neural Networks for Social Recommendation*. The Web Conference.

Jiaxuan You, Rex Ying, and J. Leskovec (2020). *Design Space for Graph Neural Networks*. Neural Information Processing Systems.

Chen Cai, and Yusu Wang (2020). *A Note on Over-Smoothing for Graph Neural Networks*. arXiv.org.

Lukas Gosch, Simon Geisler, Daniel Sturm, et al. (2023). *Adversarial Training for Graph Neural Networks*. Neural Information Processing Systems.

Xiang Zhang, and M. Zitnik (2020). *GNNGuard: Defending Graph Neural Networks against Adversarial Attacks*. Neural Information Processing Systems.

Uri Alon, and Eran Yahav (2020). *On the Bottleneck of Graph Neural Networks and its Practical Implications*. International Conference on Learning Representations.

Meiqi Zhu, Xiao Wang, C. Shi, et al. (2021). *Interpreting and Unifying Graph Neural Networks with An Optimization Framework*. The Web Conference.

Xu Zou, Qinkai Zheng, Yuxiao Dong, et al. (2021). *TDGIA: Effective Injection Attacks on Graph Neural Networks*. Knowledge Discovery and Data Mining.

Kaidi Xu, Hongge Chen, Sijia Liu, et al. (2019). *Topology Attack and Defense for Graph Neural Networks: An Optimization Perspective*. International Joint Conference on Artificial Intelligence.

Donglin Xia, Xiao Wang, Nian Liu, et al. (2024). *Learning Invariant Representations of Graph Neural Networks via Cluster Generalization*. Neural Information Processing Systems.

Shiwen Wu, Fei Sun, Fei Sun, et al. (2020). *Graph Neural Networks in Recommender Systems: A Survey*. ACM Computing Surveys.

F. Bianchi, and Veronica Lachi (2023). *The expressive power of pooling in Graph Neural Networks*. Neural Information Processing Systems.

Minh N. Vu, and M. Thai (2020). *PGM-Explainer: Probabilistic Graphical Model Explanations for Graph Neural Networks*. Neural Information Processing Systems.

Chen Gao, Yu Zheng, Nian Li, et al. (2021). *A Survey of Graph Neural Networks for Recommender Systems: Challenges, Methods, and Directions*. Trans. Recomm. Syst..

Alaa Bessadok, M. Mahjoub, and I. Rekik (2021). *Graph Neural Networks in Network Neuroscience*. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Xiao Wang, Hongrui Liu, Chuan Shi, et al. (2021). *Be Confident! Towards Trustworthy Graph Neural Networks via Confidence Calibration*. Neural Information Processing Systems.

Simon Geisler, Arthur Kosmala, Daniel Herbst, et al. (2024). *Spatio-Spectral Graph Neural Networks*. Neural Information Processing Systems.

DingYi Zeng, Wanlong Liu, Wenyu Chen, et al. (2023). *Substructure Aware Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

Wei Jin, Yao Ma, Xiaorui Liu, et al. (2020). *Graph Structure Learning for Robust Graph Neural Networks*. Knowledge Discovery and Data Mining.

Enyan Dai, and Suhang Wang (2020). *Say No to the Discrimination: Learning Fair Graph Neural Networks with Limited Sensitive Attribute Information*. Web Search and Data Mining.

Johannes Klicpera, Florian Becker, and Stephan Gunnemann (2021). *GemNet: Universal Directional Graph Neural Networks for Molecules*. Neural Information Processing Systems.

Vijay Prakash Dwivedi, A. Luu, T. Laurent, et al. (2021). *Graph Neural Networks with Learnable Structural and Positional Representations*. International Conference on Learning Representations.

Jiarui Feng, Yixin Chen, Fuhai Li, et al. (2022). *How Powerful are K-hop Message Passing Graph Neural Networks*. Neural Information Processing Systems.

Victor Garcia Satorras, Emiel Hoogeboom, and M. Welling (2021). *E(n) Equivariant Graph Neural Networks*. International Conference on Machine Learning.

Haitao Mao, Zhikai Chen, Wei Jin, et al. (2023). *Demystifying Structural Disparity in Graph Neural Networks: Can One Size Fit All?*. Neural Information Processing Systems.

Daniel Zügner, and Stephan Günnemann (2019). *Adversarial Attacks on Graph Neural Networks via Meta Learning*. International Conference on Learning Representations.

Hao Yuan, Haiyang Yu, Shurui Gui, et al. (2020). *Explainability in Graph Neural Networks: A Taxonomic Survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Ben Finkelshtein, Xingyue Huang, Michael M. Bronstein, et al. (2023). *Cooperative Graph Neural Networks*. International Conference on Machine Learning.

Ana Lucic, Maartje ter Hoeve, Gabriele Tolomei, et al. (2021). *CF-GNNExplainer: Counterfactual Explanations for Graph Neural Networks*. International Conference on Artificial Intelligence and Statistics.

Xin Zheng, Yixin Liu, Shirui Pan, et al. (2022). *Graph Neural Networks for Graphs with Heterophily: A Survey*. arXiv.org.

Enyan Dai, M. Lin, Xiang Zhang, et al. (2023). *Unnoticeable Backdoor Attacks on Graph Neural Networks*. The Web Conference.

G. Jin, Yuxuan Liang, Yuchen Fang, et al. (2023). *Spatio-Temporal Graph Neural Networks for Predictive Learning in Urban Computing: A Survey*. IEEE Transactions on Knowledge and Data Engineering.

Rex Ying, Ruining He, Kaifeng Chen, et al. (2018). *Graph Convolutional Neural Networks for Web-Scale Recommender Systems*. Knowledge Discovery and Data Mining.

Ziniu Hu, Yuxiao Dong, Kuansan Wang, et al. (2020). *GPT-GNN: Generative Pre-Training of Graph Neural Networks*. Knowledge Discovery and Data Mining.

Sitao Luan, Chenqing Hua, Qincheng Lu, et al. (2022). *Revisiting Heterophily For Graph Neural Networks*. Neural Information Processing Systems.

Johannes Klicpera, Aleksandar Bojchevski, and Stephan Günnemann (2018). *Predict then Propagate: Graph Neural Networks meet Personalized PageRank*. International Conference on Learning Representations.

Deli Chen, Yankai Lin, Wei Li, et al. (2019). *Measuring and Relieving the Over-smoothing Problem for Graph Neural Networks from the Topological View*. AAAI Conference on Artificial Intelligence.

Yu Wang, Yuying Zhao, Yushun Dong, et al. (2022). *Improving Fairness in Graph Neural Networks via Mitigating Sensitive Attribute Leakage.* Knowledge Discovery and Data Mining.

Kaixiong Zhou, Xiao Huang, D. Zha, et al. (2021). *Dirichlet Energy Constrained Learning for Deep Graph Neural Networks.* Neural Information Processing Systems.

S. Jegelka (2022). *Theory of Graph Neural Networks: Representation and Learning.* arXiv.org.

Wei Jin, Lingxiao Zhao, Shichang Zhang, et al. (2021). *Graph Condensation for Graph Neural Networks.* International Conference on Learning Representations.

Simon Geisler, Tobias Schmidt, Hakan cSirin, et al. (2021). *Robustness of Graph Neural Networks at Scale.* Neural Information Processing Systems.

Zonghan Wu, Shirui Pan, Fengwen Chen, et al. (2019). *A Comprehensive Survey on Graph Neural Networks.* IEEE Transactions on Neural Networks and Learning Systems.

Keyulu Xu, Weihua Hu, J. Leskovec, et al. (2018). *How Powerful are Graph Neural Networks?.* International Conference on Learning Representations.

Jie Zhou, Ganqu Cui, Zhengyan Zhang, et al. (2018). *Graph Neural Networks: A Review of Methods and Applications.* AI Open.

Simon L. Batzner, Albert Musaelian, Lixin Sun, et al. (2021). *E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials.* Nature Communications.

Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, et al. (2019). *SuperGlue: Learning Feature Matching With Graph Neural Networks.* Computer Vision and Pattern Recognition.

Zonghan Wu, Shirui Pan, Guodong Long, et al. (2020). *Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks.* Knowledge Discovery and Data Mining.

Shu Wu, Yuyuan Tang, Yanqiao Zhu, et al. (2018). *Session-based Recommendation with Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

Jiong Zhu, Yujun Yan, Lingxiao Zhao, et al. (2020). *Beyond Homophily in Graph Neural Networks: Current Limitations and Effective Designs*. Neural Information Processing Systems.

Minjie Wang, Da Zheng, Zihao Ye, et al. (2019). *Deep Graph Library: A Graph-Centric, Highly-Performant Package for Graph Neural Networks.*. Unpublished manuscript.

Mengzhang Li, and Zhanxing Zhu (2020). *Spatial-Temporal Fusion Graph Neural Networks for Traffic Flow Forecasting*. AAAI Conference on Artificial Intelligence.

Victor Garcia Satorras, and Joan Bruna (2017). *Few-Shot Learning with Graph Neural Networks*. International Conference on Learning Representations.

Yaqin Zhou, Shangqing Liu, J. Siow, et al. (2019). *Devign: Effective Vulnerability Identification by Learning Comprehensive Program Semantics via Graph Neural Networks*. Neural Information Processing Systems.

Kenta Oono, and Taiji Suzuki (2019). *Graph Neural Networks Exponentially Lose Expressive Power for Node Classification*. International Conference on Learning Representations.

Lei Shi, Yifan Zhang, Jian Cheng, et al. (2019). *Skeleton-Based Action Recognition With Directed Graph Neural Networks*. Computer Vision and Pattern Recognition.

Zhanghao Wu, Paras Jain, Matthew A. Wright, et al. (2022). *Representing Long-Range Context for Graph Neural Networks with Global Attention*. Neural Information Processing Systems.

Ziyang Wang, Wei Wei, G. Cong, et al. (2020). *Global Context Enhanced Graph Neural Networks for Session-based Recommendation*. Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.

Hongwei Wang, Fuzheng Zhang, Mengdi Zhang, et al. (2019). *Knowledge-aware Graph Neural Networks with Label Smoothness Regularization for Recommender Systems*. Knowledge Discovery and Data Mining.

Peixiang Zhong, Di Wang, and C. Miao (2019). *EEG-Based Emotion Recognition Using Regularized Graph Neural Networks*. IEEE Transactions on Affective Computing.

Tianxiang Zhao, Xiang Zhang, and Suhang Wang (2021). *GraphSMOTE: Imbalanced Node Classification on Graphs with Graph Neural Networks*. Web Search and Data Mining.

Qingsong Lv, Ming Ding, Qiang Liu, et al. (2021). *Are we really making much progress?: Revisiting, benchmarking and refining heterogeneous graph neural networks*. Knowledge Discovery and Data Mining.

Yue Yu, Jie Chen, Tian Gao, et al. (2019). *DAG-GNN: DAG Structure Learning with Graph Neural Networks*. International Conference on Machine Learning.

Jianheng Tang, Jiajin Li, Zi-Chao Gao, et al. (2022). *Rethinking Graph Neural Networks for Anomaly Detection*. International Conference on Machine Learning.

Jianan Zhao, Xiao Wang, C. Shi, et al. (2021). *Heterogeneous Graph Structure Learning for Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

R. Keisler (2022). *Forecasting Global Weather with Graph Neural Networks*. arXiv.org.

Maosen Li, Siheng Chen, Yangheng Zhao, et al. (2020). *Dynamic Multiscale Graph Neural Networks for 3D Skeleton Based Human Motion Prediction*. Computer Vision and Pattern Recognition.

Lingfei Wu, P. Cui, Jian Pei, et al. (2022). *Graph Neural Networks: Foundation, Frontiers and Applications*. Knowledge Discovery and Data Mining.

Zhenguang Liu, Peng Qian, Xiaoyang Wang, et al. (2021). *Combining Graph Neural Networks With Expert Knowledge for Smart Contract Vulnerability Detection*. IEEE Transactions on Knowledge and Data Engineering.

Hengrui Zhang, Qitian Wu, Junchi Yan, et al. (2021). *From Canonical Correlation Analysis to Self-supervised Graph Neural Networks*. Neural Information Processing Systems.

Yifei Shen, Yuanming Shi, Jun Zhang, et al. (2020). *Graph Neural Networks for Scalable Radio Resource Management: Architecture Design and Theoretical Analysis*. IEEE Journal on Selected Areas in Communications.

Yufeng Zhang, Xueli Yu, Zeyu Cui, et al. (2020). *Every Document Owns Its Structure: Inductive Text Classification via Graph Neural Networks*. Annual Meeting of the Association for Computational Linguistics.

Qian Huang, Horace He, Abhay Singh, et al. (2020). *Combining Label Propagation and Simple Models Out-performs Graph Neural Networks*. International Conference on Learning Representations.

S. Schaefer, Daniel Gehrig, and D. Scaramuzza (2022). *AEGNN: Asynchronous Event-based Graph Neural Networks*. Computer Vision and Pattern Recognition.

Tianwen Chen, and R. C. Wong (2020). *Handling Information Loss of Graph Neural Networks for Session-based Recommendation*. Knowledge Discovery and Data Mining.

Yifei Shen, Jun Zhang, Shenghui Song, et al. (2022). *Graph Neural Networks for Wireless Communications: From Theory to Practice*. IEEE Transactions on Wireless Communications.

Kartik Sharma, Yeon-Chang Lee, S. Nambi, et al. (2022). *A Survey of Graph Neural Networks for Social Recommender Systems*. ACM Computing Surveys.

Tianlong Chen, Yongduo Sui, Xuxi Chen, et al. (2021). *A Unified Lottery Ticket Hypothesis for Graph Neural Networks*. International Conference on Machine Learning.

Cen Chen, Kenli Li, Wei Wei, et al. (2022). *Hierarchical Graph Neural Networks for Few-Shot Learning*. IEEE transactions on circuits and systems for video technology (Print).

Jiachen Li, Siheng Chen, Xiaoyong Pan, et al. (2022). *Cell clustering for spatial transcriptomics data with graph neural networks*. Nature Computational Science.

Seongjun Yun, Seoyoon Kim, Junhyun Lee, et al. (2022). *Neo-GNNs: Neighborhood Overlap-aware Graph Neural Networks for Link Prediction*. Neural Information Processing Systems.

Asiri Wijesinghe, and Qing Wang (2022). *A New Perspective on "How Graph Neural Networks Go Beyond Weisfeiler-Lehman?"*. International Conference on Learning Representations.

Andrea Cini, Ivan Marisca, and C. Alippi (2021). *Filling the $G_a p_s$: $MultivariateTimeSeriesImputationbyGraphNeuralNetworks.InternationalConferenceonLearni$*

Lingfei Wu, Yu Chen, Kai Shen, et al. (2023). *Graph Neural Networks for Natural Language Processing: A Survey*. Found. Trends Mach. Learn..

Tianfu Li, Zheng Zhou, Sinan Li, et al. (2022). *The emerging graph neural networks for intelligent fault diagnostics and prognostics: A guideline and a benchmark study*. Mechanical systems and signal processing.

Petar Velickovic (2023). *Everything is Connected: Graph Neural Networks*. Current Opinion in Structural Biology.

Dejun Jiang, Zhenxing Wu, Chang-Yu Hsieh, et al. (2020). *Could graph neural networks learn better molecular representation for drug discovery? A comparison study of descriptor-based and graph-based models*. Journal of Cheminformatics.

Xiangguo Sun, Hongtao Cheng, Jia Li, et al. (2023). *All in One: Multi-Task Prompting for Graph Neural Networks*. Knowledge Discovery and Data Mining.

Xiao-Meng Zhang, Li Liang, Lin Liu, et al. (2021). *Graph Neural Networks and Their Current Applications in Bioinformatics*. Frontiers in Genetics.

Aleksandar Bojchevski, Johannes Klicpera, Bryan Perozzi, et al. (2020). *Scaling Graph Neural Networks with Approximate PageRank*. Knowledge Discovery and Data Mining.

Jun Xia, Chengshuai Zhao, Bozhen Hu, et al. (2023). *Mole-BERT: Rethinking Pre-training Graph Neural Networks for Molecules*. International Conference on Learning Representations.

Saeed Rahmani, Asiye Baghbani, N. Bouguila, et al. (2023). *Graph Neural Networks for Intelligent Transportation Systems: A Survey*. IEEE transactions on intelligent transportation systems (Print).

Hao Chen, Yuan-Qi Bei, Qijie Shen, et al. (2024). *Macro Graph Neural Networks for Online Billion-Scale Recommender Systems*. The Web Conference.

Wenlong Liao, B. Bak-Jensen, J. Pillai, et al. (2021). *A Review of Graph Neural Networks and Their Applications in Power Systems*. Journal of Modern Power Systems and Clean Energy.

David Hin, Andrey Kan, Huaming Chen, et al. (2022). *LineVD: Statement-level Vulnerability Detection using Graph Neural Networks*. IEEE Working Conference on Mining Software Repositories.

Anton Tsitsulin, John Palowitch, Bryan Perozzi, et al. (2020). *Graph Clustering with Graph Neural Networks*. Journal of machine learning research.

Victor Fung, Jiaxin Zhang, Eric Juarez, et al. (2021). *Benchmarking graph neural networks for materials chemistry*. npj Computational Materials.

Yongxin Wang, Kris Kitani, and Xinshuo Weng (2021). *Joint Object Detection and Multi-Object Tracking with Graph Neural Networks*. IEEE International Conference on Robotics and Automation.

Danqing Wang, Pengfei Liu, Y. Zheng, et al. (2020). *Heterogeneous Graph Neural Networks for Extractive Document Summarization*. Annual Meeting of the Association for Computational Linguistics.

Kanchan Jha, S. Saha, and Hiteshi Singh (2022). *Prediction of protein–protein interaction using graph neural networks*. Scientific Reports.

M. Schuetz, J. K. Brubaker, and H. Katzgraber (2021). *Combinatorial optimization with physics-inspired graph neural networks*. Nature Machine Intelligence.

Meng Shen, Jinpeng Zhang, Liehuang Zhu, et al. (2021). *Accurate Decentralized Application Identification via Encrypted Traffic Analysis Using Graph Neural Networks*. IEEE Transactions on Information Forensics and Security.

Deyu Bo, Chuan Shi, Lele Wang, et al. (2023). *Specformer: Spectral Graph Neural Networks Meet Transformers*. International Conference on Learning Representations.

Mengqi Zhang, Shu Wu, Xueli Yu, et al. (2021). *Dynamic Graph Neural Networks for Sequential Recommendation*. IEEE Transactions on Knowledge and Data Engineering.

Jianjun Wei, Yue Liu, Xin Huang, et al. (2024). *Self-Supervised Graph Neural Networks for Enhanced Feature Extraction in Heterogeneous Information Networks*. 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA).

Feng Yu, Yanqiao Zhu, Q. Liu, et al. (2020). *TAGNN: Target Attentive Graph Neural Networks for Session-based Recommendation*. Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.

Chaoyang He, Keshav Balasubramanian, Emir Ceyani, et al. (2021). *FedGraphNN: A Federated Learning System and Benchmark for Graph Neural Networks*. arXiv.org.

Yulei Wu, Hongning Dai, and Haina Tang (2021). *Graph Neural Networks for Anomaly Detection in Industrial Internet of Things*. IEEE Internet of Things Journal.

Miltiadis Kofinas, Boris Knyazev, Yan Zhang, et al. (2024). *Graph Neural Networks for Learning Equivariant Representations of Neural Networks*. International Conference on Learning Representations.

Shuangli Li, Jingbo Zhou, Tong Xu, et al. (2021). *Structure-aware Interactive Graph Neural Networks for the Prediction of Protein-Ligand Binding Affinity*. Knowledge Discovery and Data Mining.

M. Balcilar, G. Renton, P. Héroux, et al. (2021). *Analyzing the Expressive Power of Graph Neural Networks in a Spectral Perspective.* International Conference on Learning Representations.

Muhan Zhang, Pan Li, Yinglong Xia, et al. (2020). *Labeling Trick: A Theory of Using Graph Neural Networks for Multi-Node Representation Learning.* Neural Information Processing Systems.

Tristan Bilot, Nour El Madhoun, K. A. Agha, et al. (2023). *Graph Neural Networks for Intrusion Detection: A Survey.* IEEE Access.

Qitian Wu, Yiting Chen, Chenxiao Yang, et al. (2023). *Energy-based Out-of-Distribution Detection for Graph Neural Networks.* International Conference on Learning Representations.

P. Huang, Han-Hung Lee, Hwann-Tzong Chen, et al. (2021). *Text-Guided Graph Neural Networks for Referring 3D Instance Segmentation.* AAAI Conference on Artificial Intelligence.

Susheel Suresh, Vinith Budde, Jennifer Neville, et al. (2021). *Breaking the Limit of Graph Neural Networks by Improving the Assortativity of Graphs with Local Mixing Patterns.* Knowledge Discovery and Data Mining.

R. Liu, and Han Yu (2022). *Federated Graph Neural Networks: Overview, Techniques, and Challenges.* IEEE Transactions on Neural Networks and Learning Systems.

Lijing Wang, A. Adiga, Jiangzhuo Chen, et al. (2022). *CausalGNN: Causal-Based Graph Neural Networks for Spatio-Temporal Epidemic Forecasting.* AAAI Conference on Artificial Intelligence.

Fan Zhou, and Chengtai Cao (2021). *Overcoming Catastrophic Forgetting in Graph Neural Networks with Experience Replay.* AAAI Conference on Artificial Intelligence.

Vasimuddin, Sanchit Misra, Guixiang Ma, et al. (2021). *DistGNN: Scalable Distributed*

Training for Large-Scale Graph Neural Networks. International Conference for High Performance Computing, Networking, Storage and Analysis.

Moshe Eliasof, E. Haber, and Eran Treister (2021). *PDE-GCN: Novel Architectures for Graph Neural Networks Motivated by Partial Differential Equations.* Neural Information Processing Systems.

Kexin Huang, Ying Jin, E. Candès, et al. (2023). *Uncertainty Quantification over Graph with Conformalized Graph Neural Networks.* Neural Information Processing Systems.

Matthias Fey, J. E. Lenssen, F. Weichert, et al. (2021). *GNNAutoScale: Scalable and Expressive Graph Neural Networks via Historical Embeddings.* International Conference on Machine Learning.

Van-Anh Nguyen, D. Q. Nguyen, Van Nguyen, et al. (2021). *ReGVD: Revisiting Graph Neural Networks for Vulnerability Detection.* 2022 IEEE/ACM 44th International Conference on Software Engineering: Companion Proceedings (ICSE-Companion).

Nouhaila Innan, Abhishek Sawaika, Ashim Dhor, et al. (2023). *Financial Fraud Detection using Quantum Graph Neural Networks.* Quantum Machine Intelligence.

Jia Guo, and Chenyang Yang (2022). *Learning Power Allocation for Multi-Cell-Multi-User Systems With Heterogeneous Graph Neural Networks.* IEEE Transactions on Wireless Communications.

M. Maurizi, Chao Gao, and F. Berto (2022). *Predicting stress, strain and deformation fields in materials and structures with graph neural networks.* Scientific Reports.

Zi Ye, Y. J. Kumar, G. O. Sing, et al. (2022). *A Comprehensive Survey of Graph Neural Networks for Knowledge Graphs.* IEEE Access.

Zemin Liu, Trung-Kien Nguyen, and Yuan Fang (2021). *Tail-GNN: Tail-Node Graph Neural Networks.* Knowledge Discovery and Data Mining.

Lun Du, Xiaozhou Shi, Qiang Fu, et al. (2021). *GBK-GNN: Gated Bi-Kernel Graph Neural Networks for Modeling Both Homophily and Heterophily.* The Web Conference.

Weizhi Xu, Jun Wu, Qiang Liu, et al. (2022). *Evidence-aware Fake News Detection with Graph Neural Networks*. The Web Conference.

Shaocong Wang, Yi Li, Dingchen Wang, et al. (2023). *Echo state graph neural networks with analogue random resistive memory arrays*. Nature Machine Intelligence.

Rui Bing, Guan Yuan, Mu Zhu, et al. (2022). *Heterogeneous graph neural networks analysis: a survey of techniques, evaluations and applications*. Artificial Intelligence Review.

Ziyu Lyu, Yue Wu, Junjie Lai, et al. (2023). *Knowledge Enhanced Graph Neural Networks for Explainable Recommendation*. IEEE Transactions on Knowledge and Data Engineering.

Hao Peng, Ruitong Zhang, Yingtong Dou, et al. (2021). *Reinforced Neighborhood Selection Guided Multi-Relational Graph Neural Networks*. ACM Trans. Inf. Syst..

Ying Xia, Chun-Qiu Xia, Xiaoyong Pan, et al. (2021). *GraphBind: protein structural context embedded rules learned by hierarchical graph neural networks for recognizing nucleic-acid-binding residues*. Nucleic Acids Research.

Aosong Feng, Chenyu You, Shiqiang Wang, et al. (2022). *KerGNNs: Interpretable Graph Neural Networks with Graph Kernels*. AAAI Conference on Artificial Intelligence.

Unknown Authors (2022). *Graph Neural Networks: Foundations, Frontiers, and Applications*. Unpublished manuscript.

Sitao Luan, Chenqing Hua, Qincheng Lu, et al. (2021). *Is Heterophily A Real Nightmare For Graph Neural Networks To Do Node Classification?*. arXiv.org.

Lilapati Waikhom, and Ripon Patgiri (2022). *A survey of graph neural networks in various learning paradigms: methods, applications, and challenges*. Artificial Intelligence Review.

Siyi Tang, Jared A. Dunnmon, Khaled Saab, et al. (2021). *Self-Supervised Graph Neural Networks for Improved Electroencephalographic Seizure Analysis*. International Conference on Learning Representations.

Veronika Thost, and Jie Chen (2021). *Directed Acyclic Graph Neural Networks*. International Conference on Learning Representations.

Ziwei Chai, Siqi You, Yang Yang, et al. (2022). *Can Abnormality be Detected by Graph Neural Networks?*. International Joint Conference on Artificial Intelligence.

Chengcheng Sun, Chenhao Li, Xiang Lin, et al. (2023). *Attention-based graph neural networks: a survey*. Artificial Intelligence Review.

Mengqi Zhang, Shu Wu, Meng Gao, et al. (2022). *Personalized Graph Neural Networks With Attention Mechanism for Session-Aware Recommendation*. IEEE Transactions on Knowledge and Data Engineering.

Sai Munikoti, D. Agarwal, L. Das, et al. (2022). *Challenges and Opportunities in Deep Reinforcement Learning With Graph Neural Networks: A Comprehensive Review of Algorithms and Applications*. IEEE Transactions on Neural Networks and Learning Systems.

Ting-Li Huoh, Yan Luo, Peilong Li, et al. (2023). *Flow-Based Encrypted Network Traffic Classification With Graph Neural Networks*. IEEE Transactions on Network and Service Management.

Jiaqi Han, Yu Rong, Tingyang Xu, et al. (2022). *Geometrically Equivariant Graph Neural Networks: A Survey*. arXiv.org.

Hwan Kim, Byung Suk Lee, Won-Yong Shin, et al. (2022). *Graph Anomaly Detection With Graph Neural Networks: Current Status and Challenges*. IEEE Access.

Zeyang Zhang, Xin Wang, Ziwei Zhang, et al. (2022). *Dynamic Graph Neural Networks Under Spatio-Temporal Distribution Shift*. Neural Information Processing Systems.

Yang Zhou, Jiuhong Xiao, Yuee Zhou, et al. (2022). *Multi-Robot Collaborative Perception With Graph Neural Networks*. IEEE Robotics and Automation Letters.

Xinyi Wu, A. Ajorlou, Zihui Wu, et al. (2023). *Demystifying Oversmoothing in Attention-Based Graph Neural Networks*. Neural Information Processing Systems.

Yahui Long, Min Wu, Yong Liu, et al. (2022). *Pre-training graph neural networks for link prediction in biomedical networks*. Bioinform..

Andrea Cini, Ivan Marisca, F. Bianchi, et al. (2022). *Scalable Spatiotemporal Graph Neural Networks*. AAAI Conference on Artificial Intelligence.

Zhen Zhang, Mohammed Haroon Dupty, Fan Wu, et al. (2023). *Factor Graph Neural Networks*. Journal of machine learning research.

Jianxin Chang, Chen Gao, Xiangnan He, et al. (2023). *Bundle Recommendation and Generation With Graph Neural Networks*. IEEE Transactions on Knowledge and Data Engineering.

J. Wang (2023). *A survey on graph neural networks*. EAI Endorsed Trans. e Learn..

Xusheng Zhao, Jia Wu, Hao Peng, et al. (2022). *Deep Reinforcement Learning Guided Graph Neural Networks for Brain Network Analysis*. Neural Networks.

Zahraa Al Sahili, and M. Awad (2023). *Spatio-Temporal Graph Neural Networks: A Survey*. arXiv.org.

R. Levie (2023). *A graphon-signal analysis of graph neural networks*. Neural Information Processing Systems.

Pengcheng Wang, Linping Tao, Mingwei Tang, et al. (2024). *Incorporating syntax and semantics with dual graph neural networks for aspect-level sentiment analysis*. Engineering applications of artificial intelligence.

Hu Dong, Longjie Li, Dongwen Tian, et al. (2024). *Dynamic link prediction by learning the representation of node-pair via graph neural networks*. Expert systems with applications.

Pengju Zhao, Wenjie Liao, Yuli Huang, et al. (2024). *Beam layout design of shear wall structures based on graph neural networks*. Automation in Construction.

Ming Chen, Yajian Jiang, Xiujuan Lei, et al. (2024). *Drug-Target Interactions Prediction Based on Signed Heterogeneous Graph Neural Networks*. Chinese journal of electronics.

P. Foroutan, and Salim Lahmiri (2024). *Deep Learning-Based Spatial-Temporal Graph Neural Networks for Price Movement Classification in Crude Oil and Precious Metal Markets.* Machine Learning with Applications.

Brook Wander, Muhammed Shuaibi, John R. Kitchin, et al. (2024). *CatTSunami: Accelerating Transition State Energy Calculations with Pretrained Graph Neural Networks.* ACS Catalysis.

Duantengchuan Li, Yuxuan Gao, Zhihao Wang, et al. (2024). *Homogeneous graph neural networks for third-party library recommendation.* Information Processing Management.

Yifan Duan, Guibin Zhang, Shilong Wang, et al. (2024). *CaT-GNN: Enhancing Credit Card Fraud Detection via Causal Temporal Graph Neural Networks.* arXiv.org.

R. Praveen, Aktalina Torogeldieva, B. Saravanan, et al. (2024). *Enhancing Intellectual Property Rights(IPR) Transparency with Blockchain and Dual Graph Neural Networks.* 2024 First International Conference on Software, Systems and Information Technology (SSITCON).

Huiwei Wang, Tianhua Liu, Ziyu Sheng, et al. (2024). *Explanatory subgraph attacks against Graph Neural Networks.* Neural Networks.

Baoyu Jing, Dawei Zhou, Kan Ren, et al. (2024). *Causality-Aware Spatiotemporal Graph Neural Networks for Spatiotemporal Time Series Imputation.* International Conference on Information and Knowledge Management.

Zhongjian Zhang, Xiao Wang, Huichi Zhou, et al. (2024). *Can Large Language Models Improve the Adversarial Robustness of Graph Neural Networks?.* Knowledge Discovery and Data Mining.

Charilaos I. Kanatsoulis, and Alejandro Ribeiro (2024). *Counting Graph Substructures with Graph Neural Networks.* International Conference on Learning Representations.

Rajat Mishra, and S. Shridevi (2024). *Knowledge graph driven medicine recommendation system using graph neural networks on longitudinal medical records.* Scientific Reports.

Zhenyao Fang, and Qimin Yan (2024). *Towards accurate prediction of configurational disorder properties in materials using graph neural networks.* npj Computational Materials.

Jintu Zhang, Luigi Bonati, Enrico Trizio, et al. (2024). *Descriptor-Free Collective Variables from Geometric Graph Neural Networks..* Journal of Chemical Theory and Computation.

Nan Yin, Mengzhu Wang, Li Shen, et al. (2024). *Continuous Spiking Graph Neural Networks.* arXiv.org.

Liuxi Yan, and Yaoqun Xu (2024). *XGBoost-Enhanced Graph Neural Networks: A New Architecture for Heterogeneous Tabular Data.* Applied Sciences.

Xu Shen, P. Lió, Lintao Yang, et al. (2024). *Graph Rewiring and Preprocessing for Graph Neural Networks Based on Effective Resistance.* IEEE Transactions on Knowledge and Data Engineering.

S. K. Manivannan, Venkatesh Kavididevi, D. Muthukumaran, et al. (2024). *Graph Neural Networks for Resource Allocation Optimization in Healthcare Industry.* 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS).

Xingyang He (2024). *Graph neural networks in recommender systems.* Applied and Computational Engineering.

Zhe Zhao, Pengkun Wang, Haibin Wen, et al. (2024). *A Twist for Graph Classification: Optimizing Causal Information Flow in Graph Neural Networks.* AAAI Conference on Artificial Intelligence.

Yafeng Yan, Shuyao He, Zhou Yu, et al. (2024). *Investigation of Customized Medical Decision Algorithms Utilizing Graph Neural Networks.* 2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE).

Zaishuo Xia, Han Yang, Binghui Wang, et al. (2024). *GNNCert: Deterministic Certification of Graph Neural Networks against Adversarial Perturbations.* International Conference on Learning Representations.

Yicheng Zhou, P. Wang, Hao Dong, et al. (2024). *Make Graph Neural Networks Great Again: A Generic Integration Paradigm of Topology-Free Patterns for Traffic Speed Prediction.* International Joint Conference on Artificial Intelligence.

Shengyao Lu, Keith G. Mills, Jiao He, et al. (2024). *GOAt: Explaining Graph Neural Networks via Graph Output Attribution.* International Conference on Learning Representations.

Zhiyang Wang, J. Cerviño, and Alejandro Ribeiro (2024). *A Manifold Perspective on the Statistical Generalization of Graph Neural Networks.* arXiv.org.

Dilong Li, Chenghui Lu, Zi-xing Chen, et al. (2024). *Graph Neural Networks in Point Clouds: A Survey.* Remote Sensing.

A. Castro-Ospina, M. Solarte-Sanchez, L. Vega-Escobar, et al. (2024). *Graph-Based Audio Classification Using Pre-Trained Models and Graph Neural Networks.* Italian National Conference on Sensors.

Tianyi Zhao, Jian Kang, and Lu Cheng (2024). *Conformalized Link Prediction on Graph Neural Networks.* Knowledge Discovery and Data Mining.

Wei Duan, Jie Lu, Yu Guang Wang, et al. (2024). *Layer-diverse Negative Sampling for Graph Neural Networks.* Trans. Mach. Learn. Res..

Xuexiong Luo, Jia Wu, Jian Yang, et al. (2024). *Graph Neural Networks for Brain Graph Learning: A Survey.* International Joint Conference on Artificial Intelligence.

Alessandro De Carlo, D. Ronchi, Marco Piastra, et al. (2024). *Predicting ADMET Properties from Molecule SMILE: A Bottom-Up Approach Using Attention-Based Graph Neural Networks.* Pharmaceutics.

Sahab Zandi, Kamesh Korangi, Mar'ia 'Oskarsd'ottir, et al. (2024). *Attention-based Dynamic Multilayer Graph Neural Networks for Loan Default Prediction.* European Journal of Operational Research.

Haihong Zhao, Bo Yang, Jiaxu Cui, et al. (2024). *Effective Fault Scenario Identification for Communication Networks via Knowledge-Enhanced Graph Neural Networks.* IEEE Transactions on Mobile Computing.

Rufan Yao, Zhenhua Shen, Xinyi Xu, et al. (2024). *Knowledge mapping of graph neural networks for drug discovery: a bibliometric and visualized analysis.* Frontiers in Pharmacology.

Tuan Vinh, Loc Nguyen, Quang H. Trinh, et al. (2024). *Predicting Cardiotoxicity of Molecules Using Attention-Based Graph Neural Networks.* Journal of Chemical Information and Modeling.

Inaam Ashraf, Janine Strotherm, L. Hermes, et al. (2024). *Physics-Informed Graph Neural Networks for Water Distribution Systems.* AAAI Conference on Artificial Intelligence.

Zachary Smith, Michael Strobel, Bodhi P. Vani, et al. (2024). *Graph Attention Site Prediction (GrASP): Identifying Druggable Binding Sites Using Graph Neural Networks with Attention.* Journal of Chemical Information and Modeling.

S. Abadal, Pablo Galván, Alberto Mármol, et al. (2024). *Graph neural networks for electroencephalogram analysis: Alzheimer's disease and epilepsy use cases.* Neural Networks.

Maximilian Pflueger, David J. Tena Cucala, and Egor V. Kostylev (2024). *Recurrent Graph Neural Networks and Their Connections to Bisimulation and Logic.* AAAI Conference on Artificial Intelligence.

H. Mohammadi, and Waldemar Karwowski (2024). *Graph Neural Networks in Brain Connectivity Studies: Methods, Challenges, and Future Directions.* Brain Science.

Yongduo Sui, Xiang Wang, Tianlong Chen, et al. (2024). *Inductive Lottery Ticket Learning for Graph Neural Networks*. Journal of Computational Science and Technology.

Jie Peng, Runlin Lei, and Zhewei Wei (2024). *Beyond Over-smoothing: Uncovering the Trainability Challenges in Deep Graph Neural Networks*. International Conference on Information and Knowledge Management.

Shan Zhao, Ioannis Prapas, Ilektra Karasante, et al. (2024). *Causal Graph Neural Networks for Wildfire Danger Prediction*. arXiv.org.

M. A. Nabian (2024). *X-MeshGraphNet: Scalable Multi-Scale Graph Neural Networks for Physics Simulation*. arXiv.org.

Jiacheng Cen, Anyi Li, Ning Lin, et al. (2024). *Are High-Degree Representations Really Unnecessary in Equivariant Graph Neural Networks?*. Neural Information Processing Systems.

Yachao Yang, Yanfeng Sun, Shaofan Wang, et al. (2024). *Graph Neural Networks with Soft Association between Topology and Attribute*. AAAI Conference on Artificial Intelligence.

Youjia Li, Vishu Gupta, Muhammed Nur Talha Kilic, et al. (2024). *Hybrid-LLM-GNN: Integrating Large Language Models and Graph Neural Networks for Enhanced Materials Property Prediction*. Digital Discovery.

Zhenbei Guo, Fuliang Li, Jiaxing Shen, et al. (2024). *ConfigReco: Network Configuration Recommendation With Graph Neural Networks*. IEEE Network.

A. Gnanabaskaran, K. Bharathi, S. P. Nandakumar, et al. (2024). *Enhanced Drug-Drug Interaction Prediction with Graph Neural Networks and SVM*. 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS).

Beibei Wang, Bo Jiang, and Chris H. Q. Ding (2024). *FL-GNNs: Robust Network Representation via Feature Learning Guided Graph Neural Networks*. IEEE Transactions on Network Science and Engineering.

Daniel Abode, Ramoni O. Adeogun, and Gilberto Berardinelli (2024). *Power Control for 6G In-Factory Subnetworks With Partial Channel Information Using Graph Neural Networks*. IEEE Open Journal of the Communications Society.

Tianxiang Zhao, Xiang Zhang, and Suhang Wang (2024). *Disambiguated Node Classification with Graph Neural Networks*. The Web Conference.

Christian Hausleitner, Heimo Mueller, Andreas Holzinger, et al. (2024). *Collaborative weighting in federated graph neural networks for disease classification with the human-in-the-loop*. Scientific Reports.

Shan Zhao, Zhaiyu Chen, Zhitong Xiong, et al. (2024). *Beyond Grid Data: Exploring graph neural networks for Earth observation*. IEEE Geoscience and Remote Sensing Magazine.

T. Konstantin Rusch, Nathan Kirk, M. Bronstein, et al. (2024). *Message-Passing Monte Carlo: Generating low-discrepancy point sets via graph neural networks*. Proceedings of the National Academy of Sciences of the United States of America.

Fali Wang, Tianxiang Zhao, and Suhang Wang (2024). *Distribution Consistency based Self-Training for Graph Neural Networks with Sparse Labels*. Web Search and Data Mining.

Bingyao Liu, Iris Li, Jianhua Yao, et al. (2024). *Unveiling the Potential of Graph Neural Networks in SME Credit Risk Assessment*. 2024 5th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI).

Zhenyao Fang, and Qimin Yan (2024). *Leveraging Persistent Homology Features for Accurate Defect Formation Energy Predictions via Graph Neural Networks*. Chemistry of Materials.

Michael Benedikt, Chia-Hsuan Lu, Boris Motik, et al. (2024). *Decidability of Graph Neural Networks via Logical Characterizations*. International Colloquium on Automata, Languages and Programming.

Yuelin Zhang, Jiacheng Cen, Jiaqi Han, et al. (2024). *Improving Equivariant Graph Neural Networks on Large Geometric Graphs via Virtual Nodes Learning.* International Conference on Machine Learning.

Caterina Graziani, Tamara Drucks, Fabian Jogl, et al. (2024). *The Expressive Power of Path-Based Graph Neural Networks.* International Conference on Machine Learning.

Dai Shi, Andi Han, Lequan Lin, et al. (2024). *Design Your Own Universe: A Physics-Informed Agnostic Method for Enhancing Graph Neural Networks.* International Journal of Machine Learning and Cybernetics.

Jiang Yuan, Shanxiong Chen, Bofeng Mo, et al. (2024). *R-GNN: recurrent graph neural networks for font classification of oracle bone inscriptions.* Heritage Science.

Haitao Wang, Zelin Liu, Mingjun Li, et al. (2024). *A Gearbox Fault Diagnosis Method Based on Graph Neural Networks and Markov Transform Fields.* IEEE Sensors Journal.

Ferial Abuhantash, Mohd Khalil Abu Hantash, and Aamna AlShehhi (2024). *Comorbidity-based framework for Alzheimer's disease classification using graph neural networks.* Scientific Reports.

Yassine Abbahaddou, Sofiane Ennadir, J. Lutzeyer, et al. (2024). *Bounding the Expected Robustness of Graph Neural Networks Subject to Node Feature Attacks.* International Conference on Learning Representations.

Renhong Huang, Jiarong Xu, Xin Jiang, et al. (2024). *Measuring Task Similarity and Its Implication in Fine-Tuning Graph Neural Networks.* AAAI Conference on Artificial Intelligence.

Yue Jiang, Changkong Zhou, Vikas Garg, et al. (2024). *Graph4GUI: Graph Neural Networks for Representing Graphical User Interfaces.* International Conference on Human Factors in Computing Systems.

Bin Wang, Yadong Xu, Manyi Wang, et al. (2024). *Gear Fault Diagnosis Method Based*

on the Optimized Graph Neural Networks. IEEE Transactions on Instrumentation and Measurement.

Thiago H. Silva, and Daniel Silver (2024). *Using graph neural networks to predict local culture*. Environment and Planning B Urban Analytics and City Science.

Xin Zhang, Zhen Xu, Yue Liu, et al. (2024). *Robust Graph Neural Networks for Stability Analysis in Dynamic Networks*. 2024 3rd International Conference on Cloud Computing, Big Data Application and Software Engineering (CBASE).

Mengfang Sun, Wenying Sun, Ying Sun, et al. (2024). *Applying Hybrid Graph Neural Networks to Strengthen Credit Risk Analysis*. 2024 3rd International Conference on Cloud Computing, Big Data Application and Software Engineering (CBASE).

Xin Zeng, Fan-Fang Meng, Meng-Liang Wen, et al. (2024). *GNNGL-PPI: multi-category prediction of protein-protein interactions using graph neural networks based on global graphs and local subgraphs*. BMC Genomics.

Ziang Chen, Xiaohan Chen, Jialin Liu, et al. (2024). *Expressive Power of Graph Neural Networks for (Mixed-Integer) Quadratic Programs*. arXiv.org.

Takaaki Fujita (2024). *Superhypergraph Neural Networks and Plithogenic Graph Neural Networks: Theoretical Foundations*. arXiv.org.

Mahdi Saleh, Michael Sommersperger, N. Navab, et al. (2024). *Physics-Encoded Graph Neural Networks for Deformation Prediction under Contact*. IEEE International Conference on Robotics and Automation.

Mohammed Aburidi, and Roummel F. Marcia (2024). *Topological Adversarial Attacks on Graph Neural Networks Via Projected Meta Learning*. IEEE Conference on Evolving and Adaptive Intelligent Systems.

Zhonghao Wang, Danyu Sun, Sheng Zhou, et al. (2024). *NoisyGL: A Comprehensive Benchmark for Graph Neural Networks under Label Noise*. Neural Information Processing Systems.

Rostislav Horcík, and Gustav Šír (2024). *Expressiveness of Graph Neural Networks in Planning Domains*. International Conference on Automated Planning and Scheduling.

Jianshan Sun, Suyuan Mei, Kun Yuan, et al. (2024). *Prerequisite-Enhanced Category-Aware Graph Neural Networks for Course Recommendation*. ACM Transactions on Knowledge Discovery from Data.

Langsha Li, Feng Qiang, and Li Ma (2024). *Advancing Cybersecurity: Graph Neural Networks in Threat Intelligence Knowledge Graphs*. International Conference on Algorithms, Software Engineering, and Network Security.

Renqiang Luo, Huafei Huang, Shuo Yu, et al. (2024). *FUGNN: Harmonizing Fairness and Utility in Graph Neural Networks*. Knowledge Discovery and Data Mining.

Shouheng Li, F. Geerts, Dongwoo Kim, et al. (2024). *Towards Bridging Generalization and Expressivity of Graph Neural Networks*. International Conference on Learning Representations.

Yidong Liao, Xiao-Ming Zhang, and Chris Ferrie (2024). *Graph Neural Networks on Quantum Computers*. arXiv.org.

Yufeng Wang, and Charith Mendis (2024). *TGLite: A Lightweight Programming Framework for Continuous-Time Temporal Graph Neural Networks*. International Conference on Architectural Support for Programming Languages and Operating Systems.

Ping Liu, Haichao Wei, Xiaochen Hou, et al. (2024). *LinkSAGE: Optimizing Job Matching Using Graph Neural Networks*. Knowledge Discovery and Data Mining.

Alan John Varghese, Zhen Zhang, and G. Karniadakis (2024). *SympGNNs: Symplectic Graph Neural Networks for identifiying high-dimensional Hamiltonian systems and node classification*. Neural Networks.

Giuseppe Alessio D'Inverno, M. Bianchini, and F. Scarselli (2024). *VC dimension of Graph Neural Networks with Pfaffian activation functions*. Neural Networks.