What Do LLMs Need to Understand Graphs:
A Survey of Parametric Representation of Graphs

Dongqi Fu* , **Liri Fang*** , **Zihao Li*** , **Hanghang Tong** , **Vetle I. Torvik** , **Jingrui He**

University of Illinois Urbana-Champaign

{dongqif2, lirif2, zihaoli5, htong, vtorvik, jingrui}@illinois.edu

Abstract

Graphs, as a relational data structure, have been widely used for various application scenarios, like molecule design and recommender systems. Recently, large language models (LLMs) are reorganizing in the AI community for their expected reasoning and inference abilities. Making LLMs understand graph-based relational data has great potential, including but not limited to (1) distillate external knowledge base for eliminating hallucination and breaking the context window limit for LLMs' inference during the retrieval augmentation generation process; (2) taking graph data as the input and directly solve the graph-based research tasks like protein design and drug discovery. However, **inputting the entire graph data to LLMs is not practical** due to its complex topological structure, data size, and the lack of effective and efficient semantic graph representations. A natural question arises: *Is there a kind of graph representation that can be described by natural language for LLM's understanding and is also easy to require to serve as the raw input for LLMs?* Based on statistical computation, **graph laws** pre-defines a set of parameters (e.g., degree, time, diameter) and identifies their relationships and values by observing the topological distribution of plenty of real-world graph data. We believe this kind of parametric representation of graphs, graph laws, can be a solution for making LLMs understand graph data as the input. In this survey, we first review the previous study of graph laws from multiple perspectives, i.e., *macroscope and microscope* of graphs, *low-order and high-order* graphs, *static and dynamic* graphs, *different observation spaces*, and *newly proposed graph parameters*. After we review various real-world applications benefiting from the guidance of graph laws, we conclude the paper with current challenges and future research directions.

---

*Equal Contribution

1 Introduction

1.1 Motivation of this Paper

Graphs serve as a fundamental relational data structure and are extensively utilized in a wide range of application scenarios, including molecule design, social network analysis, and recommender systems [Zhou *et al.*, 2020]. Their ability to represent complex interconnections among entities makes them indispensable in modeling real-world relationships. However, despite their widespread use, integrating graph-based data input with large language models (LLMs) remains a challenging problem, as shown in Figure 2 with a ChatGPT 4o [Hurst *et al.*, 2024] case study.

Recently, LLMs have revolutionized the AI community with their remarkable reasoning and inference capabilities [Touvron *et al.*, 2023; Guo *et al.*, 2025], and those models have demonstrated significant potential in various tasks, such as natural language understanding, machine translation, and knowledge extraction. Given the growing importance of LLMs, enabling them to comprehend and process graph-based relational data could open new frontiers in artificial intelligence research and applications. This integration holds immense potential for enhancing LLMs in multiple ways, including but not limited to:

- **Knowledge Distillation for LLMs**: Graph-based external knowledge bases can provide crucial insights, mitigating issues such as hallucinations in LLM-generated responses and overcoming the limitations imposed by fixed context windows. By incorporating structured graph data, LLMs can improve retrieval-augmented generation (RAG) techniques and enhance inference accuracy [Edge *et al.*, 2024; He *et al.*, 2024b].

- **Direct Graph-Based Problem Solving**: Many research domains, such as protein design and drug discovery, inherently rely on graph-based data representations [Liang *et al.*, 2023; Wang *et al.*,]. Equipping LLMs with the capability to understand and manipulate graph structures could significantly advance research in these fields by enabling direct problem-solving approaches.

Despite the clear advantages of incorporating graph data into LLMs, several challenges hinder this integration. The primary obstacles include (1) the complexity of graph topologies, (2) the size of graph datasets, and (3) the absence of ef-

fective semantic representations of graphs that LLMs can process efficiently. Unlike textual data, which LLMs are inherently designed to understand, graphs lack a straightforward natural language representation. This leads to a fundamental research question: *Is there a form of graph representation that is both interpretable in natural language for LLMs and compact enough to serve as a viable input format?*

A promising solution lies in the concept of **graph laws**. Graph laws refer to statistical principles that define relationships between key structural parameters of graphs, such as degree, clustering coefficients, diameter, and time. **Hence, a graph can be represented by a few parameters to reflect its properties well**. Correspondingly, the parameters' formal mathematical relations and values are captured by graph law researchers by analyzing real-world, large-scale graph data distributions. By encoding graph properties through predefined sets of parameters, graph laws offer a way to translate complex graph topologies into a form that LLMs can potentially comprehend. For example, according to the previous research [Leskovec *et al.*, 2005; Leskovec *et al.*, 2008], the relation between the possibility of a newly-arrived node connecting to an old node (parameter #1) and the degree of that old node (parameter #2) is studied by maximum likelihood estimation (MLE) based on the observed real-world graph data [Leskovec *et al.*, 2008].

1.2 Background of Graph Representations

To model real-world tasks within graphs, graph representations are indispensable middleware that provides the basis for specific and complex task-oriented computations. To be specific, graph representations can be decomposed into three aspects as shown in Figure 1, (1) graph embedding (i.e., vector representation), (2) graph law (i.e., parametric representation), and (3) graph visualization (i.e., visual representation).
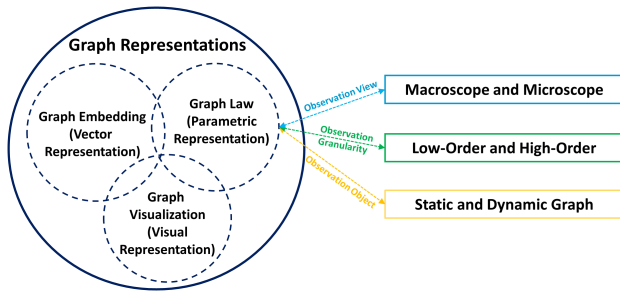

Figure 1: Position of Graph Law in Graph Representations.

First of all, graph representations can be in the form of embedding matrices, i.e., the graph topological information and attributes are encoded into matrices, which has been widely discussed and studied in the research community and usually can be referred to as *graph representation learning*. [Hamilton, 2020]. Then, graphs can also be represented by plotting directly for a better human-understandable illustration. For example, one interesting research topic is how to plot the graph topological structures into the 2D space with less structure distortion. More interesting works can be referred to [Fu and He, 2022]. Last but not least, graphs can also be represented by a few key parameters such as Erdős-Rényi random

Figure 2: A Case Study of ChatGPT 4o's Explaining about Why it Needs Graph Laws

graph $G(n, p)$ [Drobyshevskiy and Turdakov, 2020], where n stands for the number of nodes in the graph, and p stands for the independent edge connection probability in the graph. As shown in Figure 1, these three representations can have

overlapping to mutually contribute to each other [Fu and He, 2022].

1.3 Organization of this Paper

Graph law is the study of investigating the statistical properties of graphs. In this survey, we introduce the graph laws studies in the macroscopic view and microscopic view, plus multiple angles like low-order and high-order connections, static and dynamic graphs, as shown in Table 1.

- **Macroscopic and Microscopic Views**. The macroscopic graph laws describe the graph properties in a global view, like how the total degree (or eigenvalues) distribution of the entire graph looks like [Leskovec *et al.*, 2005]; while the microscopic laws try to focus on the individual behavior and investigate their behaviors as part in the entire graph [Leskovec *et al.*, 2008].

- **Low-Order and High-Order Connections**. Most graph laws are based on the node-level connections (i.e., low-order connections), while some graph law investigations are based on the group activities (high-order connections), i.e., motif in [Paranjape *et al.*, 2017; Zeno *et al.*, 2020], hyperedge in [Do *et al.*, 2020; Kook *et al.*, 2020], and simplex in [Benson *et al.*, 2018; Comrie and Kleinberg, 2021].

- **Static and Dynamic Graphs**. Compared with static graphs, dynamic graphs allow the graph components like topology structure and node attributes to evolve over time. Correspondingly, some graph laws study how the graph parameters change over time and their temporal relations. Note that, in some graph research, the dynamics are created by the algorithms, like adding virtual nodes to preserve graph representations [Abraham *et al.*, 2007; Liu *et al.*, 2022]; this kind of research is beyond the discussion of this paper, and we focus on the graph law with natural time.

Furthermore, in Section 4, we survey different real-world applications that would benefit from the guidance of graph laws. In Section 5, we introduce different observation spaces and newly proposed parameters, whose corresponding laws are not fully discovered yet. Also, in Section 6, we review the related work of graph law surveys and compare the differences. Finally, in Section 7, we conclude the survey with current challenges and future directions.

2 Macroscopic Graph Laws

In this section, we introduce the graph laws from the macroscope and microscope. In detail, we will introduce what is the intuition of researchers proposing or using graph statistical properties as parameters and how they fit the value of parameters against real-world observations.

Several classical theories model the growth of graphs, for example, Barabasi-Albert model [Barabási and Albert, 1999; Barabâsi *et al.*, 2002] assumes that the graphs follow the uniform growth pattern in terms of the number of nodes, and the Bass model [Mahajan *et al.*, 1990] and the Susceptible-Infected model [Anderson and May, 1992] follow the Sigmoid growth (more random graph models can be founded in [Drobyshevskiy and Turdakov, 2020]). However, these pre-defined graph growths have been tested that they could not handle the complex real-world network growth patterns very well [Kwak *et al.*, 2010; Zang *et al.*, 2018]. To this end, researchers begin to fit the graph growth on real-world networks directly, to discover graph laws.

2.1 Low-Order Macroscopic Laws

Based on fitting nine real-world temporal graphs from four different domains, the authors in [Leskovec *et al.*, 2005] found two temporal graph laws, called (1) *Densification Laws* and (2) *Shrinking Diameters*, respectively. First, the densification law states as follows.

$$e(t) \propto n(t)^{\alpha} \tag{1}$$

where $e(t)$ denotes the number of edges at time t, $n(t)$ denotes the number of nodes at time t, $\alpha \in [1, 2]$ is an exponent representing the density degree. The second law, shrinking diameters, states that *the effective diameter is decreasing as the network grows, in most cases*. Here, the diameter means the node-pair shortest distance, and the effective diameter of the graph means the minimum distance d such that approximately 90% of all connected pairs are reachable by a path of length at most d. Later, in [Zang *et al.*, 2018], the densification law gets in-depth confirmed on four different real social networks, the research shows that the number of nodes and number of edges both grown exponentially with time, i.e., following the power-law distribution.

2.2 High-Order Macroscopic Laws

Above discoveries are based on the node-level connections (i.e., low-order connections), then several researchers start the investigation based on the group activities, for example, motifs [Paranjape *et al.*, 2017], simplices [Benson *et al.*, 2018] and hyperedges [Do *et al.*, 2020; Kook *et al.*, 2020]. Motif is defined as a subgraph induced by a sequence of selected temporal edges in [Paranjape *et al.*, 2017], where the authors discovered that *different domain networks have significantly different numbers of similar motifs, and different motifs usually occur at different time*. Similar laws are also discovered in [Benson *et al.*, 2018] that the authors study 19 graph data sets from domains like biology, medicine, social networks, and the web, to characterize how high-order structure emerges and differs in different domains. They discovered that the higher-order Egonet features can discriminate the domain of the graph, and the probability of simplicial closure events typically increases with additional edges or tie strength.

In hypergraphs, each hyperedge could connect an arbitrary number of nodes, rather than two [Do *et al.*, 2020], where the authors found that real-world static hypergraphs obey the following properties: (1) *Giant Connected Components*, that there is a connected component comprising a large proportion of nodes, and this proportion is significantly larger than that of the second-largest connected component. (2) *Heavy-Tailed Degree Distributions*, that high-degree nodes are more likely to form new links. (3) *Small Effective Diameters*, that most connected pairs can be reached by a small distance (4) *High Clustering Coefficients*, that the global average of local clustering coefficient is high. (5) *Skewed Singularvalue Distributions*, that the singular-value distribution is usually

Table 1: A summary of parameteric representations of graphs. Some laws have multiple aspects and are indexed by numbers in parentheses.

Input	Law	Parameter	Scope	Order	Temporality	Description
Graphs	Densification Law [Leskovec et al., 2005]	Density degree α	Macro	Low	Dynamic	$e(t) \propto n(t)^\alpha, \alpha \in [1,2]$, $e(t)$ is # edges at t
	Shrinking Law [Leskovec et al., 2005]	Effective diameter d	Macro	Low	Dynamic	$d_{t+1} < d_t$, d decreases as network grows
	Motif Differing Law(1) [Paranjape et al., 2017]	Numbers of similar motifs n	Macro	High	Dynamic	$n_1 \neq n_2$ for different domains
	Motif Differing Law(2) [Paranjape et al., 2017]	Motif occurring timestamp t	Macro	High	Dynamic	$t_1 \neq t_2$ for different motifs
	Egonet Differing Law [Benson et al., 2018]	Features of Egonets X	Macro	High	Static	$X_1 \neq X_2$ for different domains
	Simplicial Closure Law [Benson et al., 2018]	Simplicial closure probability p	Macro	High	Static	p increases with additional edges or tie strength
	Spectral Power Law(1) [Eikmeier and Gleich, 2017]	Degree, SVD, eigen distributions	Macro	High	Static	These distributions usually follow power-law
	Spectral Power Law(2) [Eikmeier and Gleich, 2017]	Degree, SVD, eigen distributions	Macro	High	Static	If one follow power-law, usually others follow
	Edge Attachment Law(1) [Leskovec et al., 2008]	Node degree d, edge create $p_e(d)$	Micro	Low	Dynamic	$p_e(d) \propto d$ for node with degree d
	Edge Attachment Law(2) [Leskovec et al., 2008]	Node age $a(u)$, edge create $p_e(d)$	Micro	Low	Dynamic	$p_e(d)$ seems to be non-decreasing with $a(u)$
	Triangle Closure Law(1) [Huang et al., 2018]	Triangular connections e_1, e_2, e_3	Micro	Low	Dynamic	Strong $e_3 \Rightarrow$ unlikely e_1/e_2 will be weakened
	Triangle Closure Law(2) [Huang et al., 2018]	Triangular connections e_1, e_2, e_3	Micro	Low	Dynamic	Strong $e_1/e_2 \Rightarrow$ unlikely they will be weakened
	Local Closure Law [Yin et al., 2019]	Local closure coefficient $H(u)$	Micro	Low	Static	Please refer to Section 5 for details
	Spectral Density Law [Dong et al., 2019]	Density of states $\mu(\lambda)$	Macro	High	Static	Please refer to Section 5 for details
	Motif Activity Law(1) [Zeno et al., 2020]	Motif type	Micro	High	Dynamic	Motifs do not transit from one type to another
	Motif Activity Law(2) [Zeno et al., 2020]	Motif re-appear rate	Micro	High	Dynamic	Motifs re-appear with configured rates
Hypergraphs	Degree Distribution Law [Do et al., 2020]	Node degree, edge link probability	Macro	High	Dynamic	High-degree nodes are likely to form new links
	SVD Distribution Law [Do et al., 2020]	Singular value distribution	Macro	High	Static	Singular value distribution usually heavy-tailed
	Diminishing Overlaps [Kook et al., 2020]	density of interactions $DoI(\mathcal{H}(t))$	Macro	High	Dynamic	Overall hyperedge overlaps decrease over time
	Densification Law [Kook et al., 2020]	Density degree α	Macro	High	Dynamic	$e(t) \propto n(t)^\alpha, \alpha \geq 1$, $e(t)$ is # hyperedges at t
	Shrinking Law [Kook et al., 2020]	Hypergraph effective diameter d	Macro	High	Dynamic	$d_{t+1} < d_t$, d decreases as network grows
	Edge Interacting Law [Comrie and Kleinberg, 2021]	Edge interacting rate	Micro	High	Dynamic	Temporally adjacent interactions highly similar
Heterographs	Densification Law [Wang et al., 2019]	Density degree α, # meta-path	Macro	Low	Dynamic	$e(t) \propto n(t)^\alpha, \alpha \geq 1$ for some meta-path
	Non-densification Law [Wang et al., 2019]	Density degree α, # meta-path	Macro	Low	Dynamic	Maybe, for some meta-path, $e(t) \not\propto n(t)^\alpha$

heavy-tailed. Later, the evolution of real-word hypergraphs is investigated in [Kook et al., 2020], and the following laws are discovered.

- *Diminishing Overlaps*: The overall overlaps of hyperedges decrease over time.

- *Densification*: The average degrees increase over time.

- *Shrinking Diameter*: The effective diameters decrease over time.

To be specific, given a hypergraph $G(t) = (V(t), E(t))$, the density of interactions is stated as follows.

$$DoI(G(t)) = \frac{| \{\{e_i, e_j\} \mid e_i \cap e_j \neq \emptyset \text{ for } e_i, e_j \in E(t)\} |}{\{\{e_i, e_j\} | e_i, e_j \in E(t)\}} \tag{2}$$

and the densification is stated as follows.

$$|E(t)| \propto |V(t)|^s \tag{3}$$

where $s > 1$ stands for the density term.

In heterogeneous information networks (where nodes and edges can have multiple types), the power law distribution is also discovered [Wang et al., 2019]. For example, for the triplet "author-paper-venue" (i.e., A-P-V), the number of authors is power law distributed w.r.t the number of A-P-V instances composed by an author.

3 Microscopic Graph Laws

In contrast to representing the whole distribution of the entire graph, many researchers try to model individual behavior and investigate how they interact with each other to see the evolution pattern microscopically.

3.1 Low-Order Microscopic Laws

In [Leskovec et al., 2008], the authors view temporal graphs in a three-fold process, i.e., node arrival (determining how many nodes will be added), edge initiation (how many edges will be added), and edge destination (where are each edge will be added). They ignore the deletion of nodes and edges, and they assign variables (models) to parameterize this process.

- *Edge Attachment with Locality* (an inserted edge closing an open triangle): It is responsible for the edge destination.

- *Node Lifetime* and *Time Gap between Emitting Edges*: It is responsible for edge initiation.

- *Node Arrival Rate*: It is responsible for the node arrival.

To model the individual behaviors, there are many candidate models for selection. For example, in edge attachment, the probability of a newcomer u to connect the node v can be proportional to v's current degree or v's current age or the combination. Based on fitting each model to the real-world observation under the supervision of MLE principle, the authors empirically choose the *random-random* model for edge

attachment with locality, i.e., first, let node u choose a neighbor v uniformly and let v uniform randomly choose u's neighbor w to close a triangle. And node lifetime and time gap between emitting edges are defined as follows.

$$a(u) = t_{d(u)}(u) - t_1(u) \quad (4)$$

where $a(u)$ stands for the age of node u, $t_k(u)$ is the time when node u links its kth edge, $d_t(u)$ denote the degree of node u at time t, and $d(u) = d_T(u)$. T is the final timestamp of the data.

$$\delta_u(d) = t_{d+1}(u) - t_d(u) \quad (5)$$

where $\delta_u(d)$ records the time gap between the current time and the time when that node emits its last edge. Finding the node arrival is a regression process in [Leskovec *et al.*, 2008], for example, in Flickr graph $N(t) = exp(0.25t)$, and $N(t) = 3900t^2 + 76000t - 130000$ in LinkedIn graph.

In [Yang *et al.*, 2013; Park and Kim, 2018], the selection of edge attachment gets flourished where the authors propose several variants of edge attachment models for preserving the graph properties. With respect to the triangle closure phenomenon, several in-depth researches follow up. For example, in [Huang *et al.*, 2018], researchers found that (1) *the stronger the third tie (the interaction frequency of the closed edge) is, the less likely the first two ties are weakened*; (2) *when the stronger the first two ties are, the more likely they are weakened*.

3.2 High-Order Microscopic Laws

Hypergraph ego-network [Comrie and Kleinberg, 2021] is a structure defined to model the high-order interactions involving an individual node. The star ego-network $T(u)$ is defined as follows.

$$T(u) = \{s : (u \in s)\}, \forall s \in S \quad (6)$$

where S is the set of all hyperedges (or simplices). Also, in [Comrie and Kleinberg, 2021], there are other hypergraph ego-networks, like radial ego-network $R(u)$ and contracted ego-network $C(u)$. The relationship between them is as follows.

$$T(u) \subseteq R(u) \subseteq C(u) \quad (7)$$

In [Comrie and Kleinberg, 2021], authors observe that contiguous hyperedges (simplices) in an ego-network tend to have relatively large interactions with each other, which suggests that *temporally adjacent high-order interactions have high similarity, i.e., the same nodes tend to appear in neighboring simplices*.

In [Zeno *et al.*, 2020], authors try to model the temporal graph growth in terms of motif evolution activities. In brief, this paper investigates how many motifs change and what are the exact motif types in each time interval and fits the arrival rate parameter of each type of motif against the whole observed temporal graph.

4 Law-Guided Research Tasks

The discovered graph laws describe the graph property, which provides guidance to many downstreaming tasks. Some examples are discussed below.

4.1 Graph Generation

If not all, in most of graph law studies [Leskovec *et al.*, 2005; Zang *et al.*, 2018; Do *et al.*, 2020; Kook *et al.*, 2020; Leskovec *et al.*, 2008; Park and Kim, 2018; Zeno *et al.*, 2020], after the law (i.e., evolution pattern) is discovered, a follow-up action is to propose the corresponding graph generative model to test whether there is a realizable graph generator could generate graphs while preserving the discovered law in terms of graph properties. Also, graph generation tasks have impactful application scenarios like drug design and protein discovery [Zhang and et al., 2023].

For example, in [Leskovec *et al.*, 2005], the Forest Fire model is proposed to preserve the macroscopic graph law while larges preserve the discovered evolution pattern.

- First, node v first chooses an ambassador (i.e., node w) uniformly random, and establish a link to w;

- Second, node v generates a random value x, and selects x links of node w, where selecting in-links r times less than out-links;

- Third, node v forms links to w's neighbors; this step executes recursively (neighbors of neighbors) until v dies out.

This proposed Forest Fire model holds the following graph properties most of time.

- *Heavy-tailed In-degrees*: The highly linked nodes can easily get reached, i.e., "rich get richer".

- *Communities*: A newcomer copies neighbors of its ambassador.

- *Heavy-tailed Out-degrees*: The recursive nature produces large out-degree.

- *Densification Law*: A newcomer will have a lot of links near the community of its ambassador.

- *Shrinking Diameter*: It may not always hold.

In [Leskovec *et al.*, 2008], the authors combine the microscopic edge destination model, edge initiation model, and node arrival rate together, to model the real-world temporal network's growth. The parameters of these three models are fitted against the partial observation. i.e., $G_{\frac{T}{2}}$, which is the half of the entire evolving graph. Then they three produce the residual part of G'_T. Finally, the generated G'_T is compared with the ground truth G_T, to see if the growth pattern is fully or near fully captured by these microscopic models. The procedures are stated as follows.

- First, nodes arrive using the node arrival function obtained from $G_{\frac{T}{2}}$;

- Second, node u arrives and samples its lifetime a from the age distribution of $G_{\frac{T}{2}}$;

- Third, node u adds the first edge to node v with probability proportional to node v's degree;

- Fourth, node u with degree d samples a time gap δ from the distribution of time gap in $G_{\frac{T}{2}}$;

- When a node wakes up, if its lifetime has not expired yet, it creates a two-hop edge using the "random-random" triangle closing model;

- If a node's lifetime has expired, then it stops adding edges; otherwise, it repeats from Step 4.

The generated graph G'_T is tested based on the comparison with the ground truth G_T, in terms of degree distribution, clustering coefficient, and diameter distribution. Taking the Flickr graph for example, the generated graph is very similar to the ground truth with aforementioned metrics [Leskovec *et al.*, 2008].

4.2 Link Prediction

To learn node representation vectors for predicting links between node pairs and contributing latent applications like recommender systems, CAW-N [Wang *et al.*, 2021] is proposed by inserting causal anonymous walks (CAWs) into the representation learning process. The CAW is a sequence of time -aware adjacent nodes, the authors claim that the extracted CAW sequence obeys the triadic closure law. To be specific, the temporal opening and closed triangles can be preserved in the extracted CAW sequence W. Further, to realize the inductive link prediction, CAW-N replaces the identification of each node in W with the relative position information, such that the CAW sequence W is transferred into anonymous $\hat{W}$. Then, the entire $\hat{W}$ is inserted into an RNN-like model and gets the embedding vector of each node, the loss function states as follows.

$$enc(\hat{W}) = \mathbf{RNN}(\{f_1(I_{CAW}(w_i)) \oplus f_2(t_{i-1}-t_i)\}_{i=0,1,...,|\hat{W}|})$$
(8)

where $I_{CAW}(w_i)$ is the anonymous identification of node i in $\hat{W}$, f_1 is the node embedding function realized by a multi-layer perceptron, f_2 is the time kernel function for representing a discrete time by a vector, and $\otimes$ denotes the concatenation operation. The training loss comes from predicting negative (disconnected) node pairs and positive (connected) node pairs.

Also, there are some related link prediction models based on the guidance of static graph laws during the representation learning process, for example, SEAL [Zhang and Chen, 2018] and HHNE [Wang *et al.*, 2019].

In the SEAL framework [Zhang and Chen, 2018], for each target link, SEAL extracts a local enclosing subgraph around it, and uses a GNN to learn general graph structure features for link prediction. The corresponding graph parameters include but are not limited to

- *Common Neighbors*: Number of common neighbors of two nodes.

- *Jaccard*: Jaccard similarity on the set of neighbors of two nodes.

- *Preferential Attachment*: The product of the cardinal of the sets of neighbors of two nodes.

- *Katz Index*: The summarization over the collection of paths of two nodes.

5 Some New Observation Space and Newly Discovered Graph Parameters

5.1 New Different Spaces

In [Eikmeier and Gleich, 2017], the power law is revisited based on the eigendecomposition and singular value decomposition to provide guidance on the presence of power laws in terms of the degree distribution, singular value (of adjacency matrix) distribution, and the eigenvalue (of Laplacian matrix) distribution. The authors [Eikmeier and Gleich, 2017] discovered that (1) degree distribution, singular value distribution, and eigenvalue distribution follow power law distribution in many real-world networks they collected; (2) and a significant power law distribution of degrees usually indicates power law distributed singular values and power law distributed eigenvalues with a high probability.

5.2 New Parameters

Currently, if not all, most graph law research focuses on the traditional graph properties, like the number of nodes, number of edges, degrees, diameters, eigenvalues, and singular values. Here, we provide some recently proposed graph properties, although they have not yet been tested on the scale for fitting the graph law on real-world networks.

The local closure coefficient [Yin *et al.*, 2019] is defined as the fraction of length-2 paths (wedges) emanating from the head node (of the wedge) that induce a triangle, i.e., starting from a seed node of a wedge, how many wedges are closed. According to [Yin *et al.*, 2019], features extracted within the constraints of the local closure coefficient can improve the link prediction accuracy. The local closure efficient of node u is defined as follows.

$$H(u) = \frac{2T(u)}{W^h(u)}$$

where $W^{(h)}(u)$ is the number of wedges where u stands for the head of the wedge, and $T(u)$ denotes the number of triangles that contain node u.

The density of states (or spectral density) [Dong *et al.*, 2019] is defined as follows.

$$\mu(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i), \quad \int f(\lambda)\mu(\lambda) = \text{trace}(f(H)) \quad (9)$$

where H denotes any symmetric graph matrix, $\lambda_1, \ldots, \lambda_N$ denote the eigenvalues of H in the ascending order, δ stands for the Dirac delta function and f is any analytic test function.

6 Related Work

To the best of our knowledge, there are only a few survey papers on graph laws, with none published after 2022, marking the beginning of the foundation model era. A 2006 survey [Chakrabarti and Faloutsos, 2006] primarily focused on graph laws for mining patterns, discussing the Densification Law and Shrinking Law. In 2016, another survey [Drakopoulosa *et al.*, 2016] shifted its focus towards the generation of large graphs using different graph modeling methods, including the Erdos-Renyi model, Watts-Strogatz

model, and Albert-Barabasi model. More recently, in 2019, the authors in [Drobyshevskiy and Turdakov, 2020] offered a broader perspective on random graph modeling, covering generative, feature-driven, and domain-specific approaches. In contrast to these earlier surveys, which were published before the advent of graph neural networks and prior to the discovery of several significant graph laws [Yin *et al.*, 2019; Dong *et al.*, 2019; Zeno *et al.*, 2020; Do *et al.*, 2020; Kook *et al.*, 2020; Comrie and Kleinberg, 2021; Wang *et al.*, 2019], our work represents the first survey to explore the potential of graph laws in the context of foundation models. We emphasize how graph laws can address domain inconsistencies across different graph data types and contribute to multimodal representation learning. Additionally, this survey is the first to offer an overview of high-order graph laws and heterogeneous graph laws, marking a novel contribution to the literature.

7 Future Directions

Here, we list several interesting research directions of graph parametric representation in modern graph research.

7.1 Graph Laws on Temporal Graphs

Discovering accurate temporal graph laws from real-world networks heavily relies on the number of networks and the size of networks (e.g., number of nodes, number of edges, and time duration). However, some of the temporal graph law studies mentioned above usually consider the number of graphs ranging from 10 to 20, when they discover the evolution pattern. The existence of time-dependent structure and feature information increases the difficulty of collecting real-world temporal graph data. To obtain robust and accurate (temporal) graph laws, we may need a considerably large amount of (temporal) network data available. Luckily, we have seen some pioneering work like TGB [Huang *et al.*, 2023b] and TUDataset [Morris *et al.*, 2020].

7.2 Graph Laws on Heterogeneous Networks

Though many graph laws have been proposed and verified on homogeneous graphs, real-world networks are usually heterogeneous [Shi *et al.*, 2017] and contain a large number of interacting, multi-typed components. While the existing work [Wang *et al.*, 2019] only studied 2 datasets to propose and verify the heterogeneous graph power law, the potential exists for a transition in graph laws from homogeneous networks to heterogeneous networks, suggesting the presence of additional parameters contributing to the comprehensive information within heterogeneous networks. For example, in an academic network, the paper citation subgraph and the author collaboration subgraph may have their own subgraph laws which affect other subgraphs' laws. Furthermore, Knowledge graphs, as a special group of heterogeneous networks, have not yet attracted much attention from the research community to study their laws.

7.3 Transferability of Graph Laws

As we can see in the front part of the paper, many nascent graph laws are described verbally without the exact mathematical expression, which hinders the transfer from the graph law to the numerical constraints for the representation learning process. One latent reason for this phenomenon is that selecting appropriate models and parameters and fitting the exact values of parameters on large evolving graphs are very computationally demanding.

7.4 Taxonomy of Graph Laws

After we discovered many graph laws, is there any taxonomy or hierarchy of those? For example, graph law A stands in the superclass of graph law B, and when we preserve graph law A during the representation, we actually have already preserved graph law B. For example, there is a hierarchy of different computer vision tasks, recently discovered [Zamir *et al.*, 2018]. And corresponding research on graph law development seems like a promising direction.

7.5 Domain-Specific Graph Laws

Since graphs serve as general data representations with extreme diversity, it is challenging to find universal graph laws that fit all graph domains because each domain may be internally different from another [Zhang *et al.*, 2023]. In fact, in many cases, we have prior knowledge about the domain of a graph, which can be a social network, a protein network, or a transportation network. Thus, it is possible to study the domain-specific graph laws that work well on only a portion of graphs and then apply the graph laws only on those graphs.

7.6 Graph Laws with LLMs

In the background of large language models (LLMs) developments, an interesting question attracts much research interest nowadays, i.e., **can LLMs replace GNNs as the backbone model for graphs?** To answer this question, many recent works show the great efforts [Huang *et al.*, 2023a; Ye *et al.*, 2024; He *et al.*, 2024a], where the key point is how to represent the structural information as the input for LLMs.

For example, Instruct-GLM [Ye *et al.*, 2024] follows the manner of instruction tuning and makes the template $\mathcal{T}$ of a 2-hop connection for a *central node* v as follows.

$$\mathcal{T}(v, \mathcal{A}) = \{v\} \text{ is connected with } \{|v_2|_{v_2 \in \mathcal{A}_2^v}\} \text{ within two hops.}$$
(10)

where $\mathcal{A}_k^v$ represents the list of node v's k-hop neighbors.

As discussed above, the topological information (e.g., 1-hop or 2-hop connections) can serve as external modality information to contribute to (e.g., through prompting) the reasoning ability of large language models (LLMs) [Huang *et al.*, 2023a] and achieve state-of-the-art on low-order tasks like node classification and link prediction.

8 Conclusion

Motivated by LLM's need to understand graphs, we first review the concepts and development progress of graph parametric representations (i.e., graph laws) from different perspectives like macroscope and microscope, low-order and high-order connections, and static and temporal graphs. We then discuss various real-world application tasks that can benefit the study of graph parametric representations. Finally, we envision the latent challenges and opportunities of graph parametric representations in modern graph research with several interesting and possible future directions.

References

[Abraham *et al.*, 2007] Ittai Abraham, Mahesh Balakrishnan, Fabian Kuhn, Dahlia Malkhi, Venugopalan Ramasubramanian, and Kunal Talwar. Reconstructing approximate tree metrics. In *Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing*, pages 43–52, 2007.

[Anderson and May, 1992] Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[Barabási and Albert, 1999] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[Barabâsi *et al.*, 2002] Albert-Laszlo Barabâsi, Hawoong Jeong, Zoltan Néda, Erzsebet Ravasz, Andras Schubert, and Tamas Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical mechanics and its applications*, 311(3-4):590–614, 2002.

[Benson *et al.*, 2018] Austin R. Benson, Rediet Abebe, Michael T. Schaub, Ali Jadbabaie, and Jon M. Kleinberg. Simplicial closure and higher-order link prediction. *Proc. Natl. Acad. Sci. USA*, 115(48):E11221–E11230, 2018.

[Chakrabarti and Faloutsos, 2006] Deepayan Chakrabarti and Christos Faloutsos. Graph mining: Laws, generators, and algorithms. *ACM Comput. Surv.*, 38(1):2, 2006.

[Comrie and Kleinberg, 2021] Cazamere Comrie and Jon Kleinberg. Hypergraph ego-networks and their temporal evolution. *arXiv preprint arXiv:2112.03498*, 2021.

[Do *et al.*, 2020] Manh Tuan Do, Se-eun Yoon, Bryan Hooi, and Kijung Shin. Structural patterns and generative models of real-world hypergraphs. In Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash, editors, *KDD 2020*, pages 176–186. ACM, 2020.

[Dong *et al.*, 2019] Kun Dong, Austin R. Benson, and David Bindel. Network density of states. In Ankur Teredesai, Vipin Kumar, Ying Li, Rómer Rosales, Evimaria Terzi, and George Karypis, editors, *KDD 2019*, pages 1152–1161. ACM, 2019.

[Drakopoulosa *et al.*, 2016] Georgios Drakopoulosa, Stavros Kontopoulosb, Christos Makrisb, and Vasileios Megalooikonomoua. Large graph models: A review. *arXiv preprint arXiv:1601.06444*, 2016.

[Drobyshevskiy and Turdakov, 2020] Mikhail Drobyshevskiy and Denis Turdakov. Random graph modeling: A survey of the concepts. *ACM Comput. Surv.*, 52(6):131:1–131:36, 2020.

[Edge *et al.*, 2024] Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, and Jonathan Larson. From local to global: A graph rag approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*, 2024.

[Eikmeier and Gleich, 2017] Nicole Eikmeier and David F. Gleich. Revisiting power-law distributions in spectra of real world networks. In *KDD 2017*, pages 817–826. ACM, 2017.

[Fu and He, 2022] Dongqi Fu and Jingrui He. Natural and artificial dynamics in graphs: Concept, progress, and future. *Frontiers in Big Data*, 5, 2022.

[Guo *et al.*, 2025] Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.

[Hamilton, 2020] William L Hamilton. *Graph representation learning*. Morgan & Claypool Publishers, 2020.

[He *et al.*, 2024a] Xiaoxin He, Xavier Bresson, Thomas Laurent, Adam Perold, Yann LeCun, and Bryan Hooi. Harnessing explanations: Llm-to-lm interpreter for enhanced text-attributed graph representation learning. In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net, 2024.

[He *et al.*, 2024b] Xiaoxin He, Yijun Tian, Yifei Sun, Nitesh V Chawla, Thomas Laurent, Yann LeCun, Xavier Bresson, and Bryan Hooi. G-retriever: Retrieval-augmented generation for textual graph understanding and question answering. *arXiv preprint arXiv:2402.07630*, 2024.

[Huang *et al.*, 2018] Hong Huang, Yuxiao Dong, Jie Tang, Hongxia Yang, Nitesh V. Chawla, and Xiaoming Fu. Will triadic closure strengthen ties in social networks? *ACM Trans. Knowl. Discov. Data*, 12(3):30:1–30:25, 2018.

[Huang *et al.*, 2023a] Jin Huang, Xingjian Zhang, Qiaozhu Mei, and Jiaqi Ma. Can llms effectively leverage graph structural information: When and why. *CoRR*, abs/2309.16595, 2023.

[Huang *et al.*, 2023b] Shenyang Huang, Farimah Poursafaei, Jacob Danovitch, Matthias Fey, Weihua Hu, Emanuele Rossi, Jure Leskovec, Michael M. Bronstein, Guillaume Rabusseau, and Reihaneh Rabbany. Temporal graph benchmark for machine learning on temporal graphs. *CoRR*, abs/2307.01026, 2023.

[Hurst *et al.*, 2024] Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*, 2024.

[Kook *et al.*, 2020] Yunbum Kook, Jihoon Ko, and Kijung Shin. Evolution of real-world hypergraphs: Patterns and models without oracles. In Claudia Plant, Haixun Wang, Alfredo Cuzzocrea, Carlo Zaniolo, and Xindong Wu, editors, *ICDM 2020*, pages 272–281. IEEE, 2020.

[Kwak *et al.*, 2010] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue B. Moon. What is twitter, a social network or a news media? In Michael Rappa, Paul Jones, Juliana Freire, and Soumen Chakrabarti, editors, *WWW 2010*, pages 591–600. ACM, 2010.

[Leskovec *et al.*, 2005] Jure Leskovec, Jon M. Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In

Robert Grossman, Roberto J. Bayardo, and Kristin P. Bennett, editors, *SIGKDD 2005*, pages 177–187. ACM, 2005.

[Leskovec *et al.*, 2008] Jure Leskovec, Lars Backstrom, Ravi Kumar, and Andrew Tomkins. Microscopic evolution of social networks. In Ying Li, Bing Liu, and Sunita Sarawagi, editors, *SIGKDD 2008*, pages 462–470. ACM, 2008.

[Liang *et al.*, 2023] Youwei Liang, Ruiyi Zhang, Li Zhang, and Pengtao Xie. Drugchat: towards enabling chatgpt-like capabilities on drug molecule graphs. *arXiv preprint arXiv:2309.03907*, 2023.

[Liu *et al.*, 2022] Xin Liu, Jiayang Cheng, Yangqiu Song, and Xin Jiang. Boosting graph structure learning with dummy nodes. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato, editors, *ICML 2022*, volume 162 of *Proceedings of Machine Learning Research*, pages 13704–13716. PMLR, 2022.

[Mahajan *et al.*, 1990] Vijay Mahajan, Eitan Muller, and Frank M Bass. New product diffusion models in marketing: A review and directions for research. *Journal of marketing*, 54(1):1–26, 1990.

[Morris *et al.*, 2020] Christopher Morris, Nils M. Kriege, Franka Bause, Kristian Kersting, Petra Mutzel, and Marion Neumann. Tudataset: A collection of benchmark datasets for learning with graphs. In *ICML 2020 Workshop on Graph Representation Learning and Beyond (GRL+ 2020)*, 2020.

[Paranjape *et al.*, 2017] Ashwin Paranjape, Austin R. Benson, and Jure Leskovec. Motifs in temporal networks. In Maarten de Rijke, Milad Shokouhi, Andrew Tomkins, and Min Zhang, editors, *WSDM 2017*, pages 601–610. ACM, 2017.

[Park and Kim, 2018] Himchan Park and Min-Soo Kim. Evograph: An effective and efficient graph upscaling method for preserving graph properties. In Yike Guo and Faisal Farooq, editors, *KDD 2018*, pages 2051–2059. ACM, 2018.

[Shi *et al.*, 2017] Chuan Shi, Yitong Li, Jiawei Zhang, Yizhou Sun, and Philip S. Yu. A survey of heterogeneous information network analysis. *IEEE Trans. Knowl. Data Eng.*, 29(1):17–37, 2017.

[Touvron *et al.*, 2023] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

[Wang *et al.*,] Runze Wang, Mingqi Yang, and Yanming Shen. Graph2token: Make llms understand molecule graphs. In *ICML 2024 Workshop on Efficient and Accessible Foundation Models for Biological Discovery*.

[Wang *et al.*, 2019] Xiao Wang, Yiding Zhang, and Chuan Shi. Hyperbolic heterogeneous information network embedding. In *AAAI 2019*, pages 5337–5344. AAAI Press, 2019.

[Wang *et al.*, 2021] Yanbang Wang, Yen-Yu Chang, Yunyu Liu, Jure Leskovec, and Pan Li. Inductive representation learning in temporal networks via causal anonymous walks. In *ICLR 2021*. OpenReview.net, 2021.

[Yang *et al.*, 2013] Yang Yang, Yuxiao Dong, and Nitesh V. Chawla. Microscopic evolution of social networks by triad position profile. *CoRR*, abs/1310.1525, 2013.

[Ye *et al.*, 2024] Ruosong Ye, Caiqi Zhang, Runhui Wang, Shuyuan Xu, and Yongfeng Zhang. Language is all a graph needs. In Yvette Graham and Matthew Purver, editors, *Findings of the Association for Computational Linguistics: EACL 2024, St. Julian's, Malta, March 17-22, 2024*, pages 1955–1973. Association for Computational Linguistics, 2024.

[Yin *et al.*, 2019] Hao Yin, Austin R. Benson, and Jure Leskovec. The local closure coefficient: A new perspective on network clustering. In J. Shane Culpepper, Alistair Moffat, Paul N. Bennett, and Kristina Lerman, editors, *WSDM 2019*, pages 303–311. ACM, 2019.

[Zamir *et al.*, 2018] Amir Roshan Zamir, Alexander Sax, William B. Shen, Leonidas J. Guibas, Jitendra Malik, and Silvio Savarese. Taskonomy: Disentangling task transfer learning. In *CVPR 2018*, pages 3712–3722. Computer Vision Foundation / IEEE Computer Society, 2018.

[Zang *et al.*, 2018] Chengxi Zang, Peng Cui, Christos Faloutsos, and Wenwu Zhu. On power law growth of social networks. *IEEE Trans. Knowl. Data Eng.*, 30(9):1727–1740, 2018.

[Zeno *et al.*, 2020] Giselle Zeno, Timothy La Fond, and Jennifer Neville. Dynamic network modeling from motif-activity. In Amal El Fallah Seghrouchni, Gita Sukthankar, Tie-Yan Liu, and Maarten van Steen, editors, *Companion of The 2020 Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, pages 390–397. ACM / IW3C2, 2020.

[Zhang and Chen, 2018] Muhan Zhang and Yixin Chen. Link prediction based on graph neural networks. In Samy Bengio, Hanna M. Wallach, Hugo Larochelle, Kristen Grauman, Nicolò Cesa-Bianchi, and Roman Garnett, editors, *NeurIPS 2018*, pages 5171–5181, 2018.

[Zhang and et al., 2023] Xuan Zhang and et al. Artificial intelligence for science in quantum, atomistic, and continuum systems. *CoRR*, abs/2307.08423, 2023.

[Zhang *et al.*, 2023] Ziwei Zhang, Haoyang Li, Zeyang Zhang, Yijian Qin, Xin Wang, and Wenwu Zhu. Large graph models: A perspective. *CoRR*, abs/2308.14522, 2023.

[Zhou *et al.*, 2020] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI open*, 1:57–81, 2020.