Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Modeling Real-World Networks Using Intrinsic Vertex Fitness ERGMs

Alfonso Zack Robles Saldaña

alfonso.robles7829@alumnos.udg.mx

University of Guadalajara

Alexander Ilyich Nesterov

University of Guadalajara

Claudia Moreno Gonzalez

University of Guadalajara

Research Article

Keywords: Complex networks, Statistical mechanics, Real-world networks, Graph ensembles, Vertex fitness models, Hidden variables, Network temperature, ERGMs

Posted Date: March 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8823459/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Modeling Real-World Networks Using Intrinsic Vertex Fitness ERGMs

Alfonso Zack Robles[1*], Alexander I. Nesterov[1†] and Claudia Moreno[1†]

[1*]Department of Physics, CUCEI, University of Guadalajara, Av. Revolución 1500, Guadalajara, 44420, Jalisco, México.

*Corresponding author(s). E-mail(s): alfonso.robles7829@alumnos.udg.mx;
Contributing authors: alexander.nesterov@academicos.udg.mx; claudia.moreno@academico.udg.mx;
†These authors contributed equally to this work.

Abstract

In this work, we present a Fitness Exponential Random Graph Model (FERGMs) that brings together concepts from statistical mechanics and network science to describe complex networks beyond the traditional scale-free framework. By assigning each node an intrinsic "fitness" and viewing network formation as a thermodynamic process, we develop two complementary variants: a Configuration Model (CM) with bounded fitness and a Generalized Configuration Model (GCM) with unbounded fitness. These model are regulated by a temperature-like parameter $\boldsymbol{\tau}$ and a chemical potential $\boldsymbol{\mu}$, respectively. Together, these parameters control the balance between randomness and structural organization, as well as the overall link density in the network. We obtain closed-form expressions for the degree distributions across distinct thermodynamic regimes, showing that they become Poisson-like at high temperatures and transition to structured, weakly scale-free forms at low temperatures. We then examine 12 heterogeneous real-world networks—including social (YouTube, Orkut), technological (Skitter, web-Google), collaborative (DBLP), and infrastructural (road networks) systems—and find that FERGM surpasses conventional power-law models, lowering the relative error in average-degree prediction from as high as 768% to nearly zero, while also capturing clustering and higher-order statistics. This framework offers a principled, interpretable, and empirically supported alternative to scale-free models, opening up new opportunities for network design, inference, and analysis in domains ranging from epidemiology to computational social science.

Keywords: Complex networks, Statistical mechanics, Real-world networks, Graph ensembles, Vertex fitness models, Hidden variables, Network temperature, ERGMs

1 Introduction

Complex networks are used as models to represent numerous types of systems across various disciplines, including biology, sociology, computer science, and physics [1]. Their research has enabled a deeper understanding of phenomena such as how diseases propagate, how information circulates in social networks, and how ecological systems behave. Complex networks gained prominence during the 1990s and 2000s, when researchers applied graph theory [1–3] to explore interactions within various systems. Today, complex networks play a vital role in designing transportation infrastructures, improving communication systems, and examining financial markets, underscoring their value in both academic research and practical problem-solving [2, 4, 5].

Because complex networks are by nature large-scale systems, an entirely deterministic approach is insufficient to capture their emergent collective dynamics. Consequently, a statistical viewpoint provides a more suitable framework for their comprehensive analysis. Techniques derived from statistical mechanics enable the modeling and analysis of real-world networks by relating their dynamics to those of physical systems. The conceptual framework and mathematical methods developed in statistical physics are essential for clarifying the behavior and properties of complex networks [3, 6–8]. In particular, Exponential Random Graph Models (ERGMs) have been introduced as a framework for analyzing and interpreting the properties of real-world networks [9, 10]. ERGMs are applied across many disciplines; for instance, in the study of online communities, they are used to detect key influencers and to monitor how information propagates through the network [11, 12]. In the social sciences, they shed light on how people collaborate and how groups with shared opinions emerge [13–15]. For companies, they offer an understanding of customer preferences and how brand communities develop over time [16, 17]. Even areas such as epidemiology can take advantage of FERGMs, since they allow researchers to model how diseases propagate through contact networks [18]. ERGMs are based on the idea that specific network statistics can influence the probability of a connection forming between two actors (nodes). In the context of social networks, these statistics might include the number of existing connections, the number of mutual acquaintances, and any shared attributes. ERGMs have proven invaluable in studying their formation, particularly those dense, interconnected clusters characterized by a strong sense of homophily. This homophily, where actors within a cluster tend to share similar attributes, is a crucial aspect of social network dynamics [12, 19].

Various extensions of ERGMs have been proposed. For example, Temporal ERGMs (TERGMs) capture the dynamics of networks that evolve over time. Additional variants include Weighted ERGMs (WERGMs) and Generalized ERGMs (GERGMs), which are the two main approaches for incorporating edge weights into the model,

thereby facilitating the study of weighted graph. Lastly, there are Multiplex or Multilayer ERGMs (MP-ERGMs, ML-ERGMs, often collectively referred to as MERGMs). These networks feature nodes that are linked based on more than one attribute, representing an extension of MERGMs that can model such configurations in multilevel networks where interlayer or interlevel connections are present [9]. In contrast to conventional models that view ties as random events, FERGMs incorporate the notion of fitness —a measurable indicator of a node's appeal— that critically influences its likelihood of attracting new links [20]. In real-world networks, fitness is associated with the intrinsic attributes of the vertices, including rank, attractiveness, or wealth [21–24].

The fitness concept described above arose in parallel with another class of models known as Intrinsic Vertex Fitness Models (IVFMs). These models were proposed as an alternative mechanism to replicate the typical topology observed in scale-free networks. In such models, each vertex in the network is assigned a non-negative real variable, ε, drawn from a probability distribution $\rho(\varepsilon)$. This variable measures the "fitness" of the node, meaning how effectively it can compete to form new edges (connections).

Many real-world networks are found to have a degree distribution that follows a power-law, indicating that they are scale-free networks. This means there are a few nodes with very high connectivity (hubs) and many with low connectivity. Recent studies have shown that the universality of scale-free networks remains controversial and that strongly scale-free structures are actually uncommon [25–27].

An analysis of close to 1000 networks across social, biological, technological, transportation, and information domains showed that networks with a strongly scale-free topology are empirically uncommon [25]. Instead, log-normal and other distributions generally provide a better fit to the degree data than power laws for most networks. In addition, solid evidence indicates that social networks are, at most, weakly scale-free, whereas only a small subset of technological and biological networks displays strongly scale-free characteristics. Here, "weakly" denotes that the network does not strictly satisfy the formal definition of a scale-free network, which demands an exact power-law degree distribution, whereas the term "strong" is used when this condition holds [25].

Networks with Poisson-like degree distributions are typically those in which the connections between nodes are formed by a random process. In such networks, most nodes have a number of connections close to the average, but there are fewer nodes with a very high or very low number of connections. This distribution is characterized by a peak at the average value and a fast decay for higher values. Certain types of networks, such as Erdös–Rényi Random Networks, Collaboration Networks, Infrastructure Networks, Social Networks (like Twitter and Instagram), Telephone Call Networks, Biological Networks, Financial Networks, and Internet Networks, can exhibit the Poisson-like degree distribution [6, 28–33].

In this work, we develop a statistical-mechanical framework for modeling real-world networks using FERGMs, where nodes possess intrinsic fitness values that govern connection probabilities. We derive analytical expressions for degree distributions across different temperature regimes and introduce two complementary formulations: a Configuration Model (CM) with bounded fitness and a Generalized Configuration Model

(GCM) with unbounded fitness. Through systematic comparison with twelve empirical networks spanning technological, social, and infrastructure clasess, we demonstrate that FERGMs outperform traditional power-law models in reproducing observed degree distributions.

The paper is organized as follows. Sec. 2 discusses the statistical properties of complex undirected, unweighted, simple uni-edge (also called fermionic) networks with a fixed number of vertices and a varying number of links, described by the grand canonical ensemble. In Sec. 3, we study a class of ERGMs fitness models in different temperature regimes and consider applications to scale-free networks by comparing different distributions to the model. In Sec. 4, we explore modeling real-world networks using databases such as KONECT [34] or SNAP [35] to study the network's features. Finally, in Sec. 5: In conclusion, we summarize our results and discuss possible generalizations of our approach.

2 Statistical description of complex networks

A network is a set of N nodes (or vertices) connected by L links (or edges). One can describe the network by an adjacency matrix, a_{ij}, where each existing or non-existing link between pairs of nodes (ij) [5, 6, 36, 37][1]. Individual nodes possess local properties such as node degree k_i and clustering coefficient c_i, respectively defined as:

$$k_i = \sum_j a_{ij}, \tag{1}$$

$$c_i = \sum_{jk} \frac{a_{ij} a_{jk} a_{ki}}{k_i (k_i - 1)}. \tag{2}$$

In contrast to the traditional methodology employed in statistical mechanics, wherein the Gibbs distribution is deduced through the examination of a system in weak interaction with its surrounding environment, the statistical characterization of complex networks is based on the informational Shannon-Gibbs entropy, constrained by specific parameters [7, 38, 39].

For context, a graph and a network refer to the same underlying concept. The distinction here is that a graph is an abstract graphical representation of a network, whereas a network is the realization of that graph in the physical world.

For a graph $G \in \mathcal{G}$, belonging to an ensemble of graphs $\mathcal{G}$, we denote by $P(G)$ the probability of obtaining this graph such that the Shannon-Gibbs entropy reads

$$S = -\sum_{G \in \mathcal{G}} P(G) \ln P(G). \tag{3}$$

[1] In graph theory, a node is typically denoted by n_i, while an edge represents a connection between a pair of nodes and is denoted by $e_{ij} = (i, j)$. The presence or absence of an edge between node i, n_i, and node j, n_j, is represented in the adjacency matrix by the entry in position (i, j), which takes the value 1 if the edge e_{ij} exists and 0 otherwise.

The definition of the expected value of any function, $X(G)$, can be calculated as follows:

$$\langle X \rangle = \sum_{G \in \mathcal{G}} X(G) P(G). \tag{4}$$

Assume that the following constraints are applied: the total probability of the distribution is normalized, and the expected value of the energy is determined from the Hamiltonian function as in Eq. (4):

$$\sum_{G} P(G) = 1, \tag{5}$$

$$\sum_{G} H(G) P(G) = E, \tag{6}$$

where $H(G)$ is the graph Hamiltonian and E is the expectation value of the energy of the system. From the principle of maximum entropy with constraints, we obtain the convenient Gibbs distribution:

$$P(G) = \frac{1}{Z} e^{-\beta H(G)}, \tag{7}$$

where the canonical partition function, expressed in terms of the graph ensemble, is given by

$$Z = \sum_{G \in \mathcal{G}} e^{-\beta H(G)}. \tag{8}$$

Employing the relation $\partial S / \partial E = 1/T$, we find that the Lagrange multiplier $\beta = 1/T$ is the inverse "temperature" of the network. It is worth noting to notice that these equations define the ERGM model [7, 9, 19, 40–42].

To fit the model to the real empirical network[2], G^*, one can apply the maximum-likelihood principle, which specifies the parameter choice for the given set of constraints [38]. The log-likelihood

$$\mathcal{L} = \ln P(G^*) = -\beta H(G^*) - \ln Z(\beta), \tag{9}$$

is maximized by a particular choice of β^* yielding $E^* = -\partial \ln Z / \partial \beta^*$, where E^* is the empirical value measured on the real network:

$$E(G^*) = E^* = \sum_{G \in \mathcal{G}} H(G) P(G|\beta^*), \tag{10}$$

this implies that the network's temperature can be determined directly from the empirical data.

[2]To facilitate comparison of our results, we introduce G^*, the empirical network, which serves as a reference for evaluating the model network G constructed within the adopted statistical framework.

For an undirected network with a fixed number of vertices N and a variable number of links L, the probability of realizing a particular graph G is given by [7, 8, 43, 44]

$$P(G) = \frac{1}{\mathcal{Z}} e^{\beta(\mu L(G) - H(G))}, \tag{11}$$

with the associated grand-canonical partition function as

$$\mathcal{Z} = \sum_{G \in \mathcal{G}} e^{\beta(\mu L(G) - H(G))}. \tag{12}$$

where μ denotes the chemical potential, and

$$L(G) = \sum_{i<j} a_{ij}, \tag{13}$$

represents the number of links in the graph G.

We associate to every edge e_{ij} an "energy" ε_{ij}. With this assignment, the graph Hamiltonian can be expressed as

$$H(G) = \sum_{i<j} \varepsilon_{ij} a_{ij}. \tag{14}$$

By inserting Eq. (13) and Eq. (14) into Eq. (12), and performing some algebraic manipulations, the partition function can be expressed as [44]

$$\mathcal{Z} = \sum_{G \in \mathcal{G}} \prod_{i<j} e^{\beta(\mu - \varepsilon_{ij}) a_{ij}}. \tag{15}$$

From Eq. (12), the connection probability of the existing link between nodes i and j are defined as

$$p_{ij} = a_{ij} = -\frac{\partial \ln \mathcal{Z}}{\partial(\beta \varepsilon_{ij})}, \tag{16}$$

this probability is equivalent to the expected number of edges between vertices i and j[3]. In what follows, we consider undirected ($a_{ij} = a_{ji}$), unweighted graphs without selfloops ($a_{ii} = 0$), simple (uni-edge) (*fermionic*) graphs ($a_{ij} \leq 1$); in other words, we permit only one edge between any pair of nodes (excluding self-edges) [7]. Considering these aspects, we restrict ourselves to the terms in the summation that are equal to one, so that the evaluation of the partition function in Eq. (15) yields

$$\mathcal{Z} = \prod_{i<j} \left(1 + e^{\beta(\mu - \varepsilon_{ij})}\right). \tag{17}$$

[3]The expected number of edges $\overline{a}_{ij} = \langle a_{ij} \rangle$ is equivalent to the adjacency matrix a_{ij} in this case.

For the connection probability given in Eq. (16), we obtain

$$p_{ij} = \frac{1}{e^{\beta(\varepsilon_{ij}-\mu)}+1}. \tag{18}$$

The expected number of links, as defined in Eq. (13), is obtained by applying the result from Eq. (18) as follows:

$$L = \sum_{i<j} \frac{1}{e^{\beta(\varepsilon_{ij}-\mu)}+1}. \tag{19}$$

The expected node degree[4] $\bar{k}_i$, of a vertex i, and the average node degree $\langle k \rangle$ in the whole network are denoted as

$$\bar{k}_i = \sum_j p_{ij}, \tag{20}$$

$$\langle k \rangle = \frac{1}{N} \sum_i \bar{k}_i. \tag{21}$$

By substituting Eq. (18) into Eq. (20) and subsequently inserting the result into Eq. (21), we obtain

$$\begin{aligned} \langle k \rangle &= \frac{1}{N} \sum_{i,j} \frac{1}{e^{\beta(\varepsilon_{ij}-\mu)}+1}, \\ &= \frac{2}{N} \sum_{i<j} \frac{1}{e^{\beta(\varepsilon_{ij}-\mu)}+1}. \end{aligned} \tag{22}$$

By comparing the last expression with Eq. (19), we derive the following relationship between the expected number of links and the average node degree: $L = N\langle k \rangle/2$. Suppose that $\varepsilon_{ij} = \varepsilon_i + \varepsilon_j$, where ε_i is an "energy" assigned to each vertex i, such that this energy ε_{ij} is the sum of both independent energy contributions of each vertex i and j. For $N \gg 1$ one can replace the sum by an integral, $\frac{2}{N(N-1)} \sum_{i<j} \to \iint$, and recast Eq. (19) as

$$L = \frac{N(N-1)}{2} \iint p(\varepsilon, \varepsilon')\rho(\varepsilon)d\varepsilon\rho(\varepsilon')d\varepsilon', \tag{23}$$

where $\rho(\varepsilon)$ is the density of states and $\varepsilon'' = \varepsilon + \varepsilon'$ is now the "energy" assigned to each vertex in the continuous limit. The connection probability Eq. (18) for the continuous limit is recast as

$$p(\varepsilon, \varepsilon') = \frac{1}{e^{\beta(\varepsilon+\varepsilon'-\mu)}+1}. \tag{24}$$

[4]It is necessary to notice for the reader the distinction between the node degree k_i and the expected node degree $\bar{k}_i$, which, in this work, in this specific case of study, are the same.

In the same limit, the expected degree of a vertex with energy ε in Eq. (20) is given by

$$\bar{k}(\varepsilon) = (N-1)\int p(\varepsilon,\varepsilon')\rho(\varepsilon')d\varepsilon'. \tag{25}$$

For the average node degree in the whole network, Eq. (21) yields

$$\langle k\rangle = \int \bar{k}(\varepsilon)\rho(\varepsilon)d\varepsilon = (N-1)\iint p(\varepsilon,\varepsilon')\rho(\varepsilon)d\varepsilon\rho(\varepsilon')d\varepsilon', \tag{26}$$

thus, the average node degree per node κ is given by the following implicit function

$$\kappa = \iint p(\varepsilon,\varepsilon')\rho(\varepsilon)\rho(\varepsilon')d\varepsilon d\varepsilon', \tag{27}$$

where we can define $\kappa = \langle k\rangle/(N-1)$.

2.1 Clustering coefficient

In network science, a measure called the clustering coefficient shows how many nodes in a network tend to group. It provides insight into the local cohesiveness of connections in a network. A high clustering coefficient indicates a network with a community structure where nodes form tightly interconnected groups. In contrast, a low clustering coefficient implies a more random or decentralized network structure [6, 28, 36, 45–61].

A local clustering coefficient (LCC) of a node i is the ratio of the number of connections between the neighbors of the node to the total possible connections among them. Graph theory is defined as:

$$c_i = \frac{2\times \text{number of triangles centered on node } i}{\text{degree of } i\cdot(\text{ degree of } i-1)}. \tag{28}$$

The global clustering coefficient (GCC) measures the tendency for nodes in a graph to cluster together. The GCC ranges from 0 to 1, with 0 indicating that the graph is entirely unclustered, while 1 suggests that every node is part of a closed triangle [5, 6, 36, 37]. There are two possible definitions of the GCC [7, 62]. The first one is the definition of the GCC as the average of the local clustering coefficients of all nodes in the network:

$$C_1 = \frac{1}{N}\sum_i c_i \quad \rightarrow \quad C_1 = \int c(\varepsilon)\,\rho(\varepsilon)\,d\varepsilon, \tag{29}$$

where N denotes the total number of nodes and c_i represents the LCC of node i (or $c(\varepsilon)$ in the continuous limit). The second definition is given by

$$C_2 = \frac{3\times \text{ number of triangles}}{\text{number of connected triples}}. \tag{30}$$

The GCC definitions introduced above are not mutually equivalent; in particular, there are cases where one finds $C_1 = 1$ while $C_2 = 0$ (see, for example, the discussion in Refs. [63, 64]).

For a specific node i with energy ε_i, the local clustering coefficient $c(\varepsilon_i)$ is given by [65],

$$c_i = c(\varepsilon_i) = \frac{\sum_{j,k} p(\varepsilon_i, \varepsilon_j) p(\varepsilon_j, \varepsilon_k) p(\varepsilon_k, \varepsilon_i)}{\left(\sum_j p(\varepsilon_i, \varepsilon_j)\right)^2}. \tag{31}$$

From this result, the global clustering coefficient (GCC) can be expressed as [63]

$$C_2 = \frac{\sum_i \bar{k}(\varepsilon_i)(\bar{k}(\varepsilon_i) - 1) c(\varepsilon_i)}{\sum_i \bar{k}(\varepsilon_i)(\bar{k}(\varepsilon_i) - 1)}. \tag{32}$$

In the continuum limit, Eqs. (31) and (32) are expressed by

$$c(\varepsilon) = \frac{\iint p(\varepsilon, \varepsilon') p(\varepsilon', \varepsilon'') p(\varepsilon, \varepsilon'') \rho(\varepsilon') \rho(\varepsilon'') d\varepsilon' d\varepsilon''}{\left(\int p(\varepsilon, \varepsilon') \rho(\varepsilon') d\varepsilon'\right)^2}, \tag{33}$$

$$C_2 = \frac{\iiint p(\varepsilon, \varepsilon') p(\varepsilon', \varepsilon'') p(\varepsilon, \varepsilon'') \rho(\varepsilon) \rho(\varepsilon') \rho(\varepsilon'') d\varepsilon d\varepsilon' d\varepsilon''}{\int \left(\int p(\varepsilon, \varepsilon') \rho(\varepsilon') d\varepsilon'\right)^2 \rho(\varepsilon) d\varepsilon}. \tag{34}$$

2.2 Degree distribution

The degree distribution $P(k)$ describes how the connections in a network are distributed among nodes. Specifically, it provides the probability distribution of the degrees (the number of connections or edges) across all nodes in the network. We employ the generating function method to derive the degree distribution of configuration models with hidden variables [66]. For such networks, the generating function takes the form (see Appendix A for details):

$$G_0(z) = \int_0^{\varepsilon_0} e^{(z-1)\,\bar{k}(\varepsilon)}\, \rho(\varepsilon)\, d\varepsilon, \tag{35}$$

where z is the function variable and $\bar{k}(\varepsilon)$ represents the expected degree of a node specified by the hidden variable ε. Using Eq. (35), the degree distribution can be written as

$$P(k) = \frac{1}{k!} \frac{d^k}{dz^k} G_0(z)\Big|_{z=0} = \frac{1}{k!} \int_0^{\varepsilon_0} e^{-\bar{k}(\varepsilon)} \left(\bar{k}(\varepsilon)\right)^k \rho(\varepsilon)\, d\varepsilon, \tag{36}$$

where k is the degree. It is straightforward to verify that the degree distribution $P(k)$ is properly normalized:

$$\sum_{k=0}^{\infty} P(k) = 1. \tag{37}$$

Using the definition of degree distribution $P(k)$, the calculation of the average node degree $\langle k \rangle$ can be defined as:

$$\langle k \rangle = \sum_{k=0}^{\infty} kP(k), \tag{38}$$

where $\langle k \rangle$ is given by Eq. (26). If a network contains no isolated nodes, the degree k starts from one, and $P(k)$ should be replaced by the following degree distribution $\mathcal{P}(k)$:

$$\mathcal{P}(k) = \frac{1}{\mathfrak{Z}} P(k), \tag{39}$$

where the partition function $\mathfrak{Z}$ is given by

$$\mathfrak{Z} = \sum_{k=1}^{\infty} \mathcal{P}(k) = 1 - \int_0^{\varepsilon_0} e^{-\bar{k}(\varepsilon)} \rho(\varepsilon) d\varepsilon. \tag{40}$$

The calculation of the average node degree, $\mathcal{D}$, yields

$$\mathcal{D} = \sum_{k=1}^{\infty} k\mathcal{P}(k) = \frac{\langle k \rangle}{\mathfrak{Z}}. \tag{41}$$

2.3 Generalized Degree distribution

The distinction between the configuration model (Section 2.2) and the generalized configuration model introduced in this work corresponds to fundamentally different physical scenarios within the framework of statistical ensembles. The bounded model with $0 \leq \varepsilon \leq \varepsilon_0$ (hard configuration model) corresponds to a microcanonical-like ensemble in which node fitness is subject to a strict upper bound. This hard cutoff reflects systems characterized by finite resources or intrinsic capacity constraints (i.e., systems with a maximum number of allowable connections), thereby providing a more realistic representation of many empirical networks in which time, attention, or physical resources are limited. In contrast, the generalized model with $0 \leq \varepsilon < \infty$ (soft configuration model) corresponds to a canonical ensemble in which, in principle, arbitrarily large fitness values are admissible. However, such high values are statistically disfavored, as they are exponentially suppressed by the Boltzmann factor encoded in the density of states $\rho(\varepsilon)$. This exponential decay guaranties the convergence of the underlying mathematical formalism while simultaneously enabling the model to represent infrequent, high-connectivity phenomena (i.e., extreme hub nodes) that can emerge in large-scale networks. By applying the generating function method introduced in the previous section to derive the degree distribution of the configuration model (see Appendix C for details), we obtain:

$$G_0(z) = \int_0^{\infty} e^{(z-1)\bar{k}(\varepsilon)} \rho(\varepsilon) d\varepsilon. \tag{42}$$

Employing Eq. (C33), we obtain that

$$P(k) = \frac{1}{k!}\int_0^\infty e^{-\bar{k}(\varepsilon)}\left(\bar{k}(\varepsilon)\right)^k \rho(\varepsilon)d\varepsilon. \tag{43}$$

It is straightforward to verify that the degree distribution $P(k)$ is normalized, as given in Eq. (37). The computation of the mean node degree is defined by Eq. (38).

If a network does not have isolated nodes, then the degree k ranges from one, and one should replace $P(k)$ with the distribution described in Eq. (39), where the partition function $\mathfrak{Z}$ is now given by

$$\mathfrak{Z} = 1 - \int_0^\infty e^{-\bar{k}(\varepsilon)}\rho(\varepsilon)d\varepsilon. \tag{44}$$

following the same way of thinking as in previous section the computation of the average node degree, $\mathcal{D}$, can be recovered by the same expression in Eq. (41).

The distinction between these two formulations is warranted for several important reasons. In particular, the bounded formulation ensures mathematical rigor and mitigates potential normalization issues, whereas the unbounded formulation exhibits a more direct conceptual connection to previously proposed vertex fitness models in the literature [21–24] and enables a more tractable analytical treatment across different temperature regimes. Both formulations are thermodynamically equivalent in the large-N limit for well-behaved distributions, though they may exhibit different behaviors near critical points or in finite-size systems.

3 Fitness Model

3.1 Configuration Model

In previous work [67], motivated by the importance of scale-free networks, they explored configuration models consistent with scale-free networks. In this context, they examine a restricted scenario by considering the model in which the energy is limited by the chemical potential: $0 \leq \varepsilon \leq \mu_0$.

Now we will discuss a similar configuration model with an exponential distribution of the density of states $\rho(\varepsilon) \sim e^{\alpha\beta\varepsilon}$, assuming that limited energy is restricted by the chemical potential but is limited by a maximum value of ε_0 so that $0 \leq \varepsilon \leq \varepsilon_0$. By imposing the standard normalization condition,

$$\int_0^{\varepsilon_0} \rho(\varepsilon)\, d\varepsilon = 1, \tag{45}$$

where

$$\rho(\varepsilon) = \frac{\alpha\,\beta\, e^{\alpha\beta\,(\varepsilon-\varepsilon_0/2)}}{2\sinh(\alpha\,\beta\,\varepsilon_0/2)}, \tag{46}$$

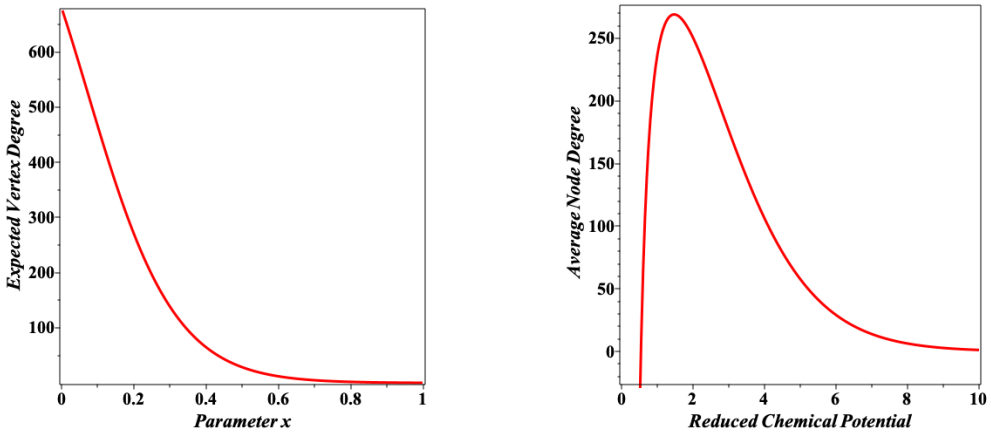
and α is a parameter related to the unspecified temperature, and β is, once more, the inverse temperature. To be able to compare our results we are using $\varepsilon_0 = \mu$.

We can now compute the expected vertex degree $\overline{k}\,(\varepsilon)$, that is, the average node degree $\langle k\rangle$. After carrying out some algebraic manipulations, we obtain

$$\overline{k}\,(\varepsilon)=\frac{N-1}{2\sinh\left(\alpha\beta\mu/2\right)}\left[e^{\alpha\beta\mu/2}\;{}_2F_1\left(1,\alpha;1+\alpha;-e^{\beta\varepsilon}\right)-e^{-\alpha\beta\mu/2}\;{}_2F_1\left(1,\alpha;1+\alpha;-e^{\beta(\varepsilon-\mu)}\right)\right], \tag{47}$$

$$\langle k\rangle=\frac{(N-1)}{4\sinh^2\left(\alpha\beta\mu/2\right)}\Bigg[e^{\alpha\beta\mu}\;{}_3F_2\left(1,\alpha,\alpha;1+\alpha,1+\alpha;-e^{\beta\mu}\right)-2\;{}_3F_2\left(1,\alpha,\alpha;1+\alpha,1+\alpha;-1\right)$$
$$+e^{-\alpha\beta\mu}\;{}_3F_2\left(1,\alpha,\alpha;1+\alpha,1+\alpha;-e^{-\beta\mu}\right)\Bigg], \tag{48}$$

where ${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;z)$ is the generalized hypergeometric function [68, 69], and $\alpha\sim T/T_c$ can be associated as the reduced temperature. Note that the critical temperature T_c and chemical potential μ are model-dependent parameters.



(a) (b)

Fig. 1: In this images, we present (a) the expected vertex degreee $\overline{k}\,(x)$ versus the parameter x; as the energy increases, the expected vertex degree decreases monotonically. In case (b), the behavior of the average node degree $\langle k\rangle$ as a function of μ^* shows a peak at a specific value of μ^* and then declines as the chemical potential is increased further.

In Fig. 1, we present in the first panel, the expected vertex degree $\overline{k}\,(x)$, as defined in Eq. (47), as a function of the parameter x, and in the second panel we show the average node degree $\langle k\rangle$, as defined in Eq. (48), as a function of μ^*. Here, $\beta\varepsilon=x$ represents the parameter x, and $\beta\mu=\mu^*$ denotes the reduced chemical potential. The parameter values are $N=10^3$, $\alpha=1.1$ (with $\mu^*=10$ used only in the case shown in Fig. 1b). For this particular set of conditions, the behavior of the expected vertex

degree, shown in Fig. 1a, indicates that its maximum value is attained at low energy; as the energy increases, the expected vertex degree decreases monotonically. Similarly, for the average node degree in Fig. 1b, the curve exhibits a maximum at a particular value of the reduced chemical potential and subsequently decreases as the chemical potential continues to increase.

3.1.1 Degree Distribution

To obtain an analytical expression for the degree distribution $P(k)$, in this case, one must evaluate the integral appearing in Eq. (36), using the definition of the expected vertex degree $\overline{k}(\varepsilon)$ given in Eq. (47). It is straightforward to see that this is not a simple problem, and providing a fully analytical treatment seems to be quite challenging.

Since the general case for the degree distribution $P(k)$ is complicated to solve, we are looking for a limit case: $e^{\beta\mu} \gg 1$ and $e^{\beta\varepsilon} \gg 1$ that fit our case of study. Using Eqs. (47)-(48), we rewrite the generating function Eq. (35) and perform the integration for this limit case. In order to obtain the degree distribution Eq. (36), we evaluate the derivatives at the point $z = 0$ and obtain,

$$P(k) = \left(-\frac{e^{-\alpha\mu^*/2}}{2\sinh(\alpha\mu^*/2)}\right)\frac{\alpha\Lambda^\alpha}{k!}\left[\Gamma(k-\alpha,\Lambda) - \Gamma\left(k-\alpha,\Lambda e^{-\mu^*}\right)\right], \tag{49}$$

where

$$\Lambda = \frac{(N-1)}{2}\frac{e^{\alpha\mu^*/2}}{\sinh(\alpha\mu^*/2)}\frac{\alpha}{(\alpha-1)}, \tag{50}$$

and $\Gamma(\alpha, x)$ denotes the upper incomplete Gamma function, recalling that $\beta\mu = \mu^*$.

In Fig. 2, we present the degree distribution $P(k)$ for the configuration model in the limiting regime $\beta\mu \gg 1$, $\beta\varepsilon \gg 1$ in the main panel, and the corresponding log–log representation for the same regime in the inset defined by Eq. (36). The analytical solution for the degree distribution given by Eq. (49) exhibits identical functional forms. The analytical prediction and the exact numerical result coincide within graphical accuracy. The blue curve depicts $P(k)$ defined by the exact formula (36). The red dashed curve presents the degree distribution described by Eq. (49); the number of nodes is $N = 10^3$, $\alpha = 1.1$, and $\mu = 20$.

3.2 Generalized Configuration Model

As an application of the developed approach, we consider the FERGM proposed in Ref. [23]. In this model, a real non-negative hidden variable called the *fitness* is assigned to each vertex. The FERGM belongs to the class of soft configuration networks described above with exponential decay of the density of states, $\rho(\varepsilon) \sim e^{-\tau\beta\varepsilon}$. Imposing the standard normalization condition,

$$\int_0^\infty \rho(\varepsilon)d\varepsilon = 1, \tag{51}$$

we obtain that

$$\rho(\varepsilon) = \tau\beta e^{-\tau\beta\varepsilon}. \tag{52}$$

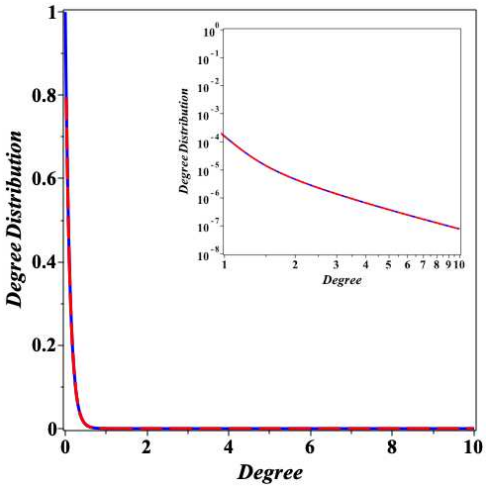


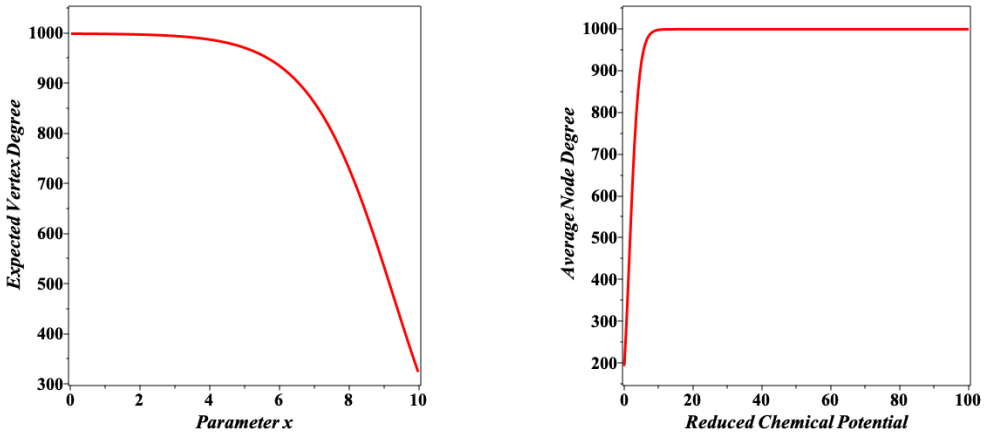
Fig. 2: In this figure, we present the degree distribution $P(k)$ versus degree k for a special limiting case as a function of k in the regime $\beta\mu \gg 1$ and $\beta\varepsilon \gg 1$. The inset is the log-log representation. The blue curve is obtained from an exact expression, while the red dashed curves indicate the limiting form of the degree distribution. The analytical prediction and the exact numerical result agree to within graphical resolution.

The expected vertex degree, $\bar{k}(\varepsilon)$, and the average node degree, $\langle k \rangle$, are defined by

$$\bar{k}(\varepsilon) = \frac{\tau(N-1)e^{\beta(\mu-\varepsilon)}}{1+\tau}\, {}_2F_1\big(1, 1+\tau; 2+\tau; -e^{\beta(\mu-\varepsilon)}\big), \tag{53}$$

$$\langle k \rangle = \frac{(N-1)\tau^2 e^{\beta\mu}}{(1+\tau)^2}\, {}_3F_2\big(1, 1+\tau, 1+\tau; 2+\tau, 2+\tau; -e^{\beta\mu}\big). \tag{54}$$

In Fig. 3, the behavior of the expected vertex degree, shown in Fig. 3a, indicates that its maximum value is maintained at low energy up to a certain energy threshold, beyond which it starts to decrease as the energy increases. Similarly, for the average node degree in Fig. 3b $\langle k \rangle$ defined by Eq. (54), the curve exhibits a maximum at a specific value of the reduced chemical potential and remains at this maximal value as the chemical potential continues to increase. Here, $\beta\varepsilon = x$ defines the parameter x, and $\beta\mu = \mu^*$ represents the reduced chemical potential. In this case, the parameters are set to $N = 10^3$, $\tau = 1.1$, and $\mu^* = 10$.



(a) (b)

Fig. 3: Plot (a) shows the expected vertex degree as a function of the parameter x, the maximum value remains constant at low energies up to a specific energy threshold, after which it begins to decrease as the energy continues to rise. In (b), the average node degree is plotted as a function of the reduced chemical potential. The curve reaches a maximum at a particular value of the reduced chemical potential and then stays at this maximal level as the chemical potential increases further.

3.2.1 Generalized Degree Distribution

As established in the previous section, our next step is to derive the degree distribution by applying the definition in Eq. (43) and inserting the expression for the expected vertex degree from Eq. (53).

We once more examine the limiting case characterized by an arbitrary reduced temperature τ and by $e^{\mu^*} \ll 1$. As is evident, the temperature dependence of the degree distribution is governed by the reduced temperature τ and the reduced chemical potential $\mu^*(\tau) = \beta\mu$. In Appendix B, we derive an approximate analytical form of the degree distribution Eq. (36) in this limiting regime we obtain

$$P(k) = \frac{\tau b^{-\tau} \gamma(k+\tau, b)}{k!}, \tag{55}$$

where

$$b = \frac{(1+\tau)}{\tau} \langle k \rangle, \tag{56}$$

and $\gamma(a, z)$ denotes the incomplete γ function [69].

In Fig. 4, we present the degree distribution $P(k)$ for the generalized configuration model, with the corresponding log–log representation for the same set of parameters shown in the inset. The blue curve depicts $P(k)$ computed from the exact formula in Eq. (36). The red dashed curves correspond to the asymptotic form of the degree distribution $P(k)$ for a generic reduced temperature τ and $e^{\mu^*} \ll 1$, as provided by

Eq. (55), where the analytical prediction coincides with the exact numerical result to graphical accuracy. The chosen parameters are $\mu = 20$, $\tau = 1.05$, and $N = 10^2$.

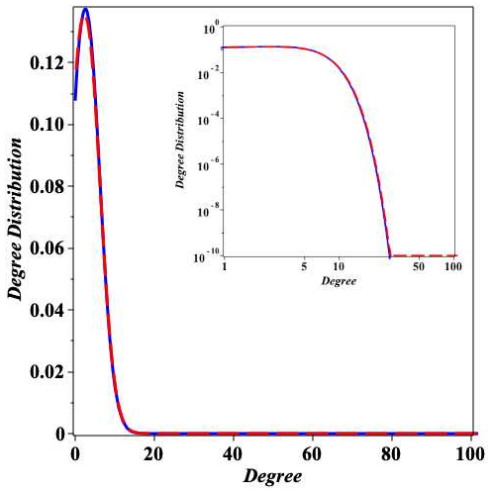


Fig. 4: We display the degree distribution $P(k)$ as a function of the degree for the generalized configuration model. The inset shows the corresponding log–log plot. The blue curve is obtained from an exact expression, while the red dashed curves indicate the limiting form of the degree distribution. The analytical prediction and the exact numerical result agree to within graphical resolution.

3.3 Applications to scale-free networks

Scale-free networks that exhibit a strongly scale-free topology are defined by a degree distribution that closely adheres to a power-law form, $P(k) \sim k^{-\gamma}$, over a broad interval of degree values k. For a scale-free network with power-law exponent γ and degree $k \in (1, \ldots, \infty)$, the degree distribution can be expressed as:

$$P_{sf}(k) = \frac{k^{-\gamma}}{\zeta(\gamma)}, \tag{57}$$

$$d_{sf} = \sum_{k=1}^{\infty} k P_{fs}(k) = \frac{\zeta(\gamma - 1)}{\zeta(\gamma)}, \tag{58}$$

where $\zeta(z)$ is the Riemann ζ-function [69].

Let the degree k vary from 1 up to k_m, where k_m denotes the maximum degree. Then, by applying the summation formula [70]

$$\sum_{k=1}^{n} \frac{1}{k^s} = \zeta(s) - \zeta(s, n+1), \tag{59}$$

we obtain

$$P_{sf}(k) = \frac{k^{-\gamma}}{\zeta(\gamma) - \zeta(\gamma, k_m + 1)}, \tag{60}$$

$$d_{sf} = \frac{\zeta(\gamma - 1) - \zeta(\gamma - 1, k_m + 1)}{\zeta(\gamma) - \zeta(\gamma, k_m + 1)}, \tag{61}$$

where $\zeta(s, a)$ is the Hurwitz zeta function [69]. Weakly scale-free networks share certain traits with scale-free networks but do not strictly follow a power-law degree distribution over the full degree range. In Ref. [25], scale-free models were compared with several alternative distributions—log-normal, exponential, Weibull, and power law with cutoff. The study demonstrated that, for most networks, log-normal distributions provide a better fit to empirical data than pure power laws.

We will contrast the exact solution defined by Eq. (36) with these alternative distributions in order to assess the robustness of the model. To demonstrate the correspondence showing that our model is also applicable to these networks, which more closely follow some of the next distributions.

- **Log-normal degree distribution:**

$$P(k) = \frac{e^{-\frac{(\ln(k) - \delta)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma}k}. \tag{62}$$

- **Exponential degree distribution:**

$$P(k) = Ce^{-\lambda k}. \tag{63}$$

- **Weibull degree distribution:**

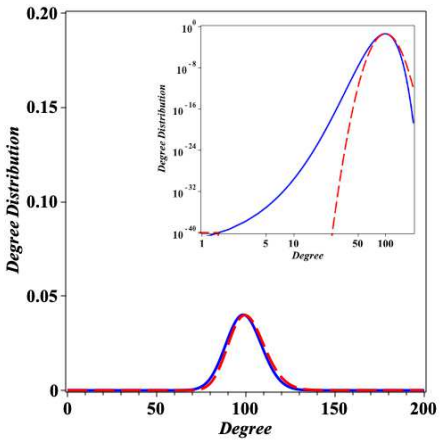
$$P(k) = Ce^{-(\lambda k)^a}. \tag{64}$$

- **Power law with cutoff degree distribution:**

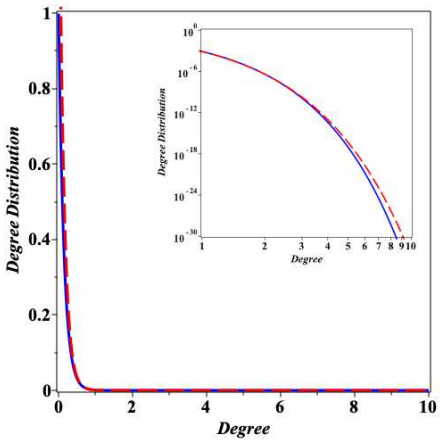
$$P(k) = \frac{k^{-\delta}e^{-\lambda k}}{\mathrm{Li}_\delta(\mathrm{e}^{-\lambda})}, \tag{65}$$

where $\mathrm{Li}_s(z)$ is the polylogarithm function [69] (also defined in Appendix D, Eq. (D50)).

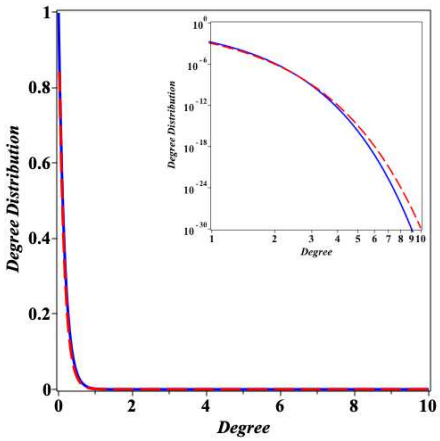
In Fig. 5, we display the degree distribution $P(k)$ as a function of the degree k. The blue curve represents $P(k)$ calculated from the exact expression in Eq. (43), while the red dashed curves depict the alternative degree distributions specified in Eqs. (62)-(65). Where we use the next parameters (a) $\mu = 20$, $\tau = 0.11$, $\delta = 4.61$, $\sigma = 0.1$, $N = 10^2$, (b) $\mu = 20$, $\tau = 1.375$, $\lambda = 7.63$, $C = 1.63$, $N = 10^3$, (c) $\mu = 20$, $\tau = 1.595$, $C = 1$, $\delta = 1$, $\lambda = 6.9$, $N = 10^6$, and (d) $\mu = 20$, $\tau = 1.32$, $C = 0.11$, $\delta = 1$, $\lambda = 5.5$, $N = 10^2$. We compare our FERGMs-based theoretical predictions with these results and find that they agree well with the alternative distributions.



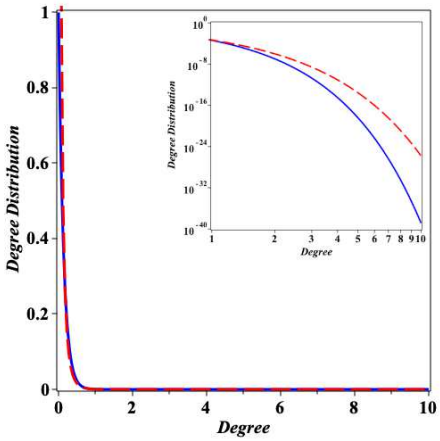
(a) Log-Normal Distribution.



(b) Exponential Distribution.



(c) Weibull Distribution.



(d) Power-Law with cutoff Distribution

Fig. 5: We present the degree distribution $P(k)$, blue curve, in function of degree k. The red dashed curves represent each of the alternative degree distributions. It's easy to see that the model and the alternative distributions' results agree to within graphical resolution; it's not until we see the inset part of this figures with the log-log representation that we can notice the main differences, and how accurate the model can be, some of the start differing by a factor of 10^{-12}. For example, the Exponential and Weibull distributions, and for the Log-normal, the difference between the distributions is visual. The main conclusion can be that the FERGMs reach fater zero, except for the log-normal distribution wich is the opposite.

4 Modeling Real-World Networks

Rigorous model validation requires comprehensive statistical metrics beyond visual comparison. The first four moments—mean, variance, skewness, and kurtosis—fully characterize distribution shape, capturing central tendency, spread, asymmetry, and tail behavior, which are critical for heterogeneous networks. Relative error in average degree measures the reproduction of network scale, while residual standard deviation quantifies the overall goodness-of-fit across all degrees. These metrics enable objective comparison between theoretical predictions and empirical data, revealing whether models capture full structural complexity or merely approximate certain features. This approach is essential when comparing FERGMs with power-law models, as networks may exhibit scale-free-like behavior in limited ranges while deviating in global properties.

We evaluate the predictions of the FERGMs against those of the scale-free model by comparing the standard deviation σ, the relative error ϵ, and the first four statistical moments—the mean μ_1, the variance σ^2, the skewness μ_3, and the kurtosis μ_4—and by visually contrasting both models with the empirical network data reported in Table 1. The datasets are sourced from the Stanford Network Analysis Project (SNAP) [35] and the Koblenz Network Collection (KONECT) [34].

For the FERGMs, we adopt the CM whose degree distribution $\mathcal{P}(k)$ is specified in Eq. (39), with the corresponding partition function $\mathfrak{Z}$ defined in Eq. (40). In contrast, for the GCM the degree distribution $\mathcal{P}(k)$ follows Eq. (43), and its partition function $\mathfrak{Z}$ is given by Eq. (44). Finally, for the Scale-free model we employ the degree distribution $P_{sf}(k)$ as defined in Eq. (60).

The empirical networks analyzed here fall within the category of sparse networks, defined by a low average degree per node, $\kappa \ll 1$. Under this condition, one has $e^{\mu^*} \ll 1$, in this limit, Eq. (55) leads to a Poisson-like degree distribution for the GCM. In contrast, the opposite regime $\kappa \gg 1$, implying $e^{\mu^*} \gg 1$, corresponds to dense networks. For this second class, the CM is instead governed by Eq. (49).

To align the model with real-world networks, we fix the parameters τ or α for CM or GCM, respectively, along with the choice of μ^*, by imposing the following constraints:

$$d = \frac{\langle k \rangle}{\mathfrak{Z}}, \tag{66}$$

$$\mathcal{C}_2 = \frac{(\tau+1)^2}{(\tau+2)^2} e^{\mu^*}, \tag{67}$$

where d is the empirical average node degree, and $\mathcal{C}_2$ is the empirical global clustering coefficient defined in Table 1.

We then compare the normalized degree distributions generated by the FERGMs and by the power-law model with the empirical degree distribution for both configurations. For the Power-law Distribution (PLD), we quantify the relative error, ϵ, using the standard expression

$$\epsilon = \frac{|d_{sf} - d|}{d} \times 100\%, \tag{68}$$

where the mean node degree d_{sf} is obtained from Eq. (61). In the case of FERGMs, the relative error $\epsilon = 0$ arises because the average degree $\mathcal{D} = d$ is enforced through model constraints. In addition, we calculate the residual standard deviation, σ, for FERGM and PLD, along with the fourth-order statistical moments: the centered and standardized mean μ_1, the variance $\mu_2 = \sigma^2$, the skewness μ_3, and the kurtosis μ_4, incorporating the sampling adjustment defined as:

$$\mu_1 = \frac{1}{M}\sum_i x_i, \tag{69}$$

$$\sigma^2 = \frac{1}{(M-1)}\sum_i \left(x_i - \mu_1\right)^2, \tag{70}$$

$$\mu_3 = \frac{M}{(M-1)(M-2)}\sum_i \frac{\left(x_i - \mu_1\right)^3}{\sigma^3}, \tag{71}$$

$$\mu_4 = \frac{M(M+1)}{(M-1)(M-2)(M-3)}\sum_i \frac{\left(x_i - \mu_1\right)^4}{\sigma^4} - \frac{3(M-1)^2}{(M-2)(M-3)}, \tag{72}$$

where x_i is the normalized empirical data and M is the size of the sample defined by the distribution of each real-world network. Note that these four statistical moments are chosen because some of the samples of the networks are small. When network samples are huge $M \to \infty$, we recover centered standard first fourth moments without sampling adjustment.

We present in the main text the empirical data defined for the following networks shown in Table 1, collected by the databases KONECT and SNAP:

1. **Skitter** – This is the undirected network of autonomous systems on the Internet connected to each other from the Skitter project.
2. **Amazon** – This is the co-purchase network of Amazon based on the "customers who bought this also bought" feature. Nodes are products and an undirected edge between two nodes shows that the corresponding products have been frequently bought together.
3. **DBLP** – This is the co-authorship network of the DBLP (Digital Bibliography & Library Project) computer science bibliography. Nodes are authors and an undirected edge between two nodes exists if the corresponding authors have published at least one paper together.
4. **Orkut** – This is the social network of Orkut users and their connections. The network is undirected. The dataset was crawled from the orkut website and may thus be incomplete.
5. **Youtube** – This is the friendship network of the video-sharing site Youtube. Nodes are users, and an undirected edge between two nodes indicates a friendship.
6. **Brightkite** – This undirected network contains user–user friendship relations from Brightkite, a former location-based social network where users shared their locations. A node represents a user, and an edge indicates that a friendship exists between the user represented by the left node and the user represented by the right node.

Table 1: Empirical values of some networks from different categories, such as social networks, infrastructure networks, collaboration networks, friendship networks, hyperlinks networks, extracted from the KONECT and SNAP datasets.

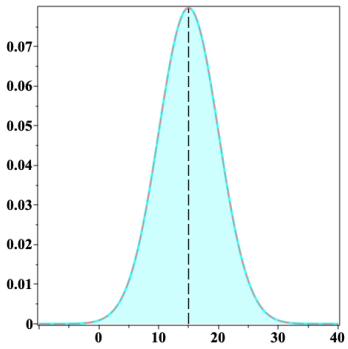
Network name	Network size, N (number of nodes)	Network density L (number of edges)	Average node degree, d	Maximum node degree, d_m	Power law exponent, γ	Tail power law exponent, γ_{tail}	Clustering Coefficient $\mathcal{C}_2$
skitter	1696415	11095298	13.08087703	35455	1.610	2.291	0.2581
com-amazon	334863	925872	5.529855493	549	1.691	3.581	0.3967
com-dblp	317080	1049866	6.622089063	343	1.720	3.261	0.6324
com-orkut	3072441	117185083	76.28142119	33313	1.272	1.272	0.1666
com-youtube	1134890	2987624	5.265045952	28754	2.429	2.141	0.0808
loc-brightkite	58228	214078	7.353094731	1134	1.905	2.481	0.1723
loc-gowalla	196591	950327	9.668062119	14730	1.730	2.651	0.2367
roadNet-CA	1965206	2766607	2.815589816	12	2.055	8.991	0.0464
roadNet-PA	1088092	1541898	2.834131673	9	2.052	8.991	0.0465
roadNet-TX	1379917	1921660	2.785182007	12	2.073	8.901	0.0470
soc-pokec †	1632803	30622564	37.50919615	14854	1.404	3.061	0.1094
web-Google †	875713	5105039	11.65916002	6332	1.617	2.731	0.5143

The symbol † denotes directed networks, such as soc-pokec, in the table. We also report the maximum node outdegree, indicated by $d_m \to d_m^+$.

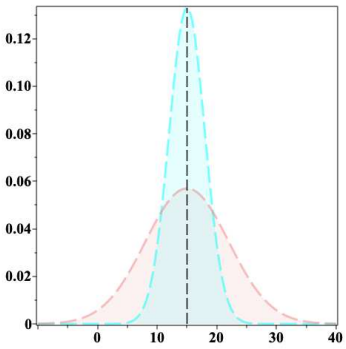
7. **Gowalla** – This undirected network contains user–user friendship relations from Gowalla, a former location-based social network where user shared their locations. A node represents a user, and an edge indicates that a friendship exists between the user represented by the left node and the user represented by the right node.
8. **California Roads** – This is the road network of the State of California in the United States of America. The nodes of the network are the intersections between roads and road endpoints, and the edges are road segments between intersections and road endpoints. The network is undirected.
9. **Pennsylvania Roads** – This is the road network of Pennsylvania in the United States of America. The nodes of the network are the intersections between roads and road endpoints, and the edges are road segments between intersections and road endpoints. The network is undirected.
10. **Texas Roads** – This is the road network of Texas in the United States of America. The nodes of the network are the intersections between roads and road endpoints, and the edges are road segments between intersections and road endpoints. The network is undirected.
11. **Pokec** – This is the friendship directed network from the Slovak social network Pokec. Nodes are users of Pokec, and directed edges represent friendships.
12. **Web Google** – This is a network of web pages connected by hyperlinks. The data was released in 2002 by Google as part of the Google Programming Contest.

In Fig. 6, we present visual interpretation guidelines for the first four statistical moments, this well known graphs can help us to understand the results presented in the tables to follow. In Fig. 6a we show that if the empirical data and the model for comparation have the same mean, they actually are centered and can be viewed as the plot is presented, meanwhile if the variance is diferent as in Fig. 6b shows the shape of the curve can be slightly different. For the Fig. 6c and Fig. 6d particular cases are presented for example how to imagine if the means are different but the variance are the same, and what if the variance and the means are different, cases that actually aren‘t presented in our results. For the other hand something important for us is how to interpret the Skewness and the Kurtosis, the Fig. 6e and Fig. 6f help us to get an idea of how the model is fitting the empirical data by understanding the values presented in the tables Table 3 and Table 5. For example if the value of the skewness is negative is means that the fitting model is geting to the right of the values of the empirical data (ED), and getting left for the possitive value, for the kurtosis the positive or negative value of the results can be interpret as how long the tails are getting from the ED. If the kurtosis is positive the tail are shorter, and if they are negative the tails are longer.

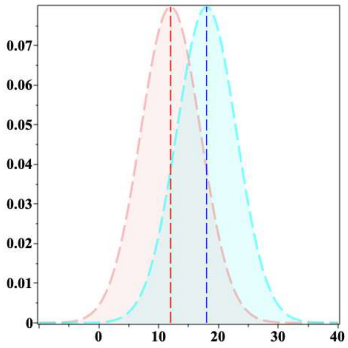
In Tables 2 and 3, we present the CM outcomes for the 12 empirical networks under both power-law exponents, γ and γ_{tail}. The results show that, in most instances, FERGMs consistently surpass standard power-law models, as indicated by their uniformly lower values for each of the first four statistical moments. In Table 2, we present the numerical results for the CM, reporting the reduced temperature, reduced chemical potential, average node degree, relative error, and residual standard deviation for both the primary power-law exponent γ and the exponent below γ_{tail}. We demonstrate that FERGMs consistently outperform traditional power-law models in most cases, as evidenced by their lower RSD.



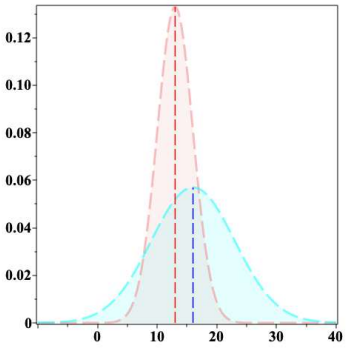
(a) Same Means, Same Variance.



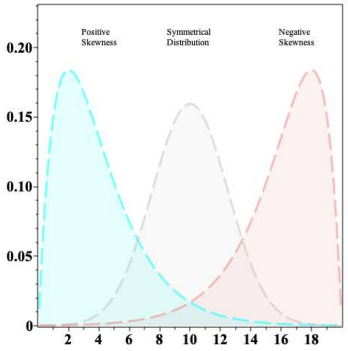
(b) Same Means, Different Variance



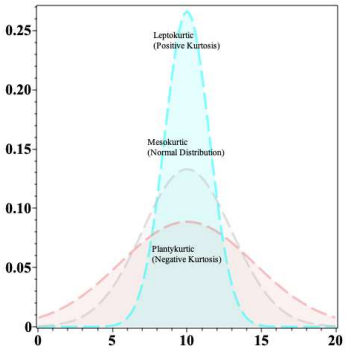
(c) Different Means, Same Variance.



(d) Different Means, Different Variance.



(e) Skewness Comparison.



(f) Kurtosis Comparison.

Fig. 6: We present here the visual interpretation guidelines for the first four statistical moments for reference. To understand how to read the information in the tables, we use this reference plot, which helps us interpret the results.

In Table 3, we present the first moment given by Eq. (69), the second central moment in Eq. (70), the third standardized moment in Eq. (71), and the fourth standardized moment in Eq. (72). From Table 3 we observe that almost all skewness values for the FERGMs are positive, indicating that they are shifted to the right of the ED, while all skewness values for the PDL are negative, indicating that they lie to the left of the ED. Considering the magnitude of these values, we can see that the FERGMs are closer to the ED than the PLD, even though they deviate in opposite directions. Finally, Table 3 shows that the kurtosis is positive for almost all values, and that it is lower for the FERGMs than for the PLD.

In Fig. 7, the FERGMs degree distribution $P(k)$ from Eq. (49) for the CM is plotted as a function of the degree using the exponent γ, with a log-log version of the same data shown in the inset. Numerical results for the CM are presented. The inset of the main plot displays the degree frequency as a function of the degree on a log-log scale, illustrating the power-law exponent γ. Empirical data are depicted as red solid diamonds; FERGM with $P(k)$ given by Eq. (49) is represented by blue solid squares; PLD with $P(k) \propto k^{-\gamma}$ is marked by green solid circles. From the figures, we observe that for examples such as Amazon (MDS), DBLP, and Google Hyperlinks, the FERGMs visually reproduce the distribution of the ED more accurately than the PLD. Other examples, such as Orkut and Pokec, indicate that this model with this configuration is not a good visual match for these real-world networks, even though the numerical results suggest it provides a better fit. For infrastructure networks such as Texas Roads, California Roads, and Pennsylvania Roads, the FERGMs and the PLD model yield very similar results, without any visually noticeable differences. Cases as Youtube Friendships, Brightkite ans Gowalla are visually more aceptable with the PLD Model.

As for Fig. 8, the degree distribution is plotted as a function of degree for the CM, using the tail exponent γ_{tail}, with the inset once more presenting the data on a log-log scale. The ED are shown as red solid diamonds; the FERGMs with $P(k)$ specified by Eq. (49) is depicted by blue solid squares; and the power-law distribution $P(k) \propto k^{-\gamma}$, defined in Appendix E, Sec. (3.3), is represented by green solid circles. For this set of results we can observe that the PLD lost almost all visually predictions of the distributions of the real-world networks, and results as Amazon and DBLP for the FERGMS can still be consider a better visually representations of the distributions and for the case of infrastructure networks the roads networks as California, Pennsylvania and Texas cannot be modeled more effectively by FERGMS.

Table 2: We show in this table the numerical results for the CM, we demonstrate that FERGMs consistently outperform traditional power-law models in most cases, as evidenced by their lower RSD. We can se that the relative error for the PLD can exhibit results from 15% to 768%, and the RSD is in most cases lower for FERGMs.

Network name	PLC	Reduced	Reduced chemical	Parameter	Average node degree		Relative error	RSD	
	γ	temperature, τ	potential, μ^*	$e^{\mu^*}/10^5$	ED, FERGMs	PLD	PLD	FERGMs	PLD
Skitter	1.61077	13.70868	13.70868	8.986	13.0809	66.3457	407.19 %	0.00947	0.01468
	2.29100	12.91997	12.91997	4.083	13.0809	2.69364	79.407 %	0.01201	0.02606
com-amazon	1.69122	12.79796	12.79796	3.614	5.52986	9.75713	76.444 %	0.00460	0.02005
	3.58100	11.66621	11.66621	1.165	5.52986	1.17348	78.779 %	0.00449	0.03725
com-dblp	1.72001	12.51551	12.51551	2.725	6.62209	7.67029	15.828 %	0.00604	0.02093
	3.26100	11.50891	11.50891	0.995	6.62209	1.25559	81.039 %	0.01272	0.04113
com-orkut	1.27235	13.55056	13.55056	7.672	76.2813	662.275	768.20 %	0.00380	0.00380
	1.27235	13.55056	13.55056	7.672	76.2813	662.275	768.20 %	0.00380	0.01136
com-youtube	2.42900	13.34194	13.34194	6.227	5.26505	2.12330	59.671 %	0.02269	0.00922
	2.14100	13.53964	13.53964	7.589	5.26505	3.92924	25.371 %	0.02189	0.00569
loc-brightkite	1.90579	10.46189	10.46189	0.349	7.35309	6.05793	17.613 %	0.01650	0.00968
	2.48100	10.00847	10.00847	0.222	7.35309	1.94193	73.590 %	0.01800	0.01736
loc-gowalla	1.73070	11.64046	11.64046	1.136	9.66806	23.0961	138.89 %	0.01277	0.01154
	2.65100	10.86703	10.86703	0.524	9.66806	1.67229	82.702 %	0.01509	0.02409
roadNet-CA	2.05571	14.78872	14.78872	2.646	2.81559	1.92019	31.801 %	0.11093	0.19640
	8.99100	13.69175	13.69175	8.835	2.81559	1.00207	64.409 %	0.09347	0.28935
roadNet-PA	2.05282	14.19363	14.19363	1.459	2.83413	1.79152	36.787 %	0.13029	0.22744
	8.99100	13.09401	13.09401	4.860	2.83413	1.00207	64.642 %	0.11064	0.33117
roadNet-TX	2.07371	14.42962	14.42962	1.848	2.78518	1.90076	31.754 %	0.11473	0.19591
	8.90100	13.35156	13.35156	6.287	2.78518	1.00221	64.016 %	0.09900	0.28559
soc-pokec †	1.40480	13.13103	13.13103	5.043	37.5092	168.51	349.25 %	0.00794	0.01081
	3.06100	11.47228	11.47228	0.960	37.5092	1.33668	96.436 %	0.00924	0.03343
web-Google †	1.61721	13.14853	13.14853	5.132	11.6592	32.4436	178.26 %	0.00949	0.01275
	2.73100	12.13860	12.13860	1.869	11.6592	1.57532	86.488 %	0.01178	0.02817

Table 3: We show in this table the numerical results for the CM, showing the four moments. We demonstrate that FERGMs consistently outperform traditional power-law models in most cases, as evidenced by their lower results for each of the first four statistical moments.

Network name	Power-Law Coefficient γ	First Moment Expected Value/10^{-7}		Second Moment Variance		Third Moment Skewness		Fourth Moment Kurtosis/10^5	
		FERGM	PLD	FERGM	PLD	FERGM	PLD	FERGM	PLD
Skitter	1.61077	5.886	5.886	0.000089	0.000215	4251.50	-9792.05	0.59993	2.10613
	2.29100	5.886	5.886	0.000144	0.000679	3215.36	-10272.9	0.41454	2.23136
com-amazon	1.69122	29.86	29.86	0.000021	0.000402	1968.256	-7785.039	0.87927	1.79206
	3.58100	29.86	20.20	0.000020	0.001387	-2389.533	-10321.93	0.50183	2.32226
com-dblp	1.72001	31.53	31.53	0.000036	0.000438	1439.969	-5203.879	0.25700	0.92023
	3.26100	31.53	31.53	0.000161	0.001691	1366.408	-5481.610	0.21570	0.97415
com-orkut	1.27235	3.217	3.217	0.000014	0.000129	-1052.078	-8292.787	0.08419	1.55311
	1.27235	3.217	3.217	0.000014	0.000129	-1052.078	-8292.787	0.08419	1.55311
com-youtube	2.42900	8.806	8.806	0.000514	0.000085	8242.911	-9926.672	1.74895	2.14551
	2.14100	8.806	8.806	0.000479	0.000032	8395.548	-9937.868	1.78748	2.14809
loc-brightkite	1.90579	171.7	171.7	0.000272	0.000093	7286.169	-10304.63	1.49179	2.23936
	2.48100	171.7	171.7	0.000324	0.000301	6714.828	-10289.96	1.33335	2.23619
loc-gowalla	1.73070	50.84	50.84	0.000163	0.000133	5976.821	-10070.77	1.10985	2.17856
	2.65100	50.84	50.84	0.000227	0.000580	4746.506	-10273.14	0.80300	2.23207
roadNet-CA	2.05571	5.088	5.088	0.012306	0.038574	19.14147	-4.830455	0.00031	0.00031
	8.99100	5.088	5.088	0.008737	0.083727	18.86191	-20.14131	0.00038	0.00048
roadNet-PA	2.05282	9.190	9.190	0.016977	0.051733	10.40701	-3.316731	0.00008	0.00010
	8.99100	9.190	9.190	0.012242	0.109678	9.874648	-13.27716	0.00011	0.00021
roadNet-TX	2.07371	7.246	7.246	0.013164	0.038381	18.67345	-2.509578	0.00034	0.00030
	8.90100	7.246	7.246	0.009801	0.081563	16.89331	-18.79923	0.00037	0.00047
soc-pokec †	1.40480	6.122	6.122	0.000063	0.000116	2774.998	-9776.171	0.35404	2.04540
	3.06100	6.122	6.122	0.000085	0.001117	1687.201	-10840.98	0.20169	2.38440
web-Google †	1.61721	11.41	11.41	0.000090	0.000162	5027.033	-9731.367	0.91325	2.07987
	2.73100	11.41	11.41	0.000138	0.000794	3930.578	-10387.38	0.59988	2.25981

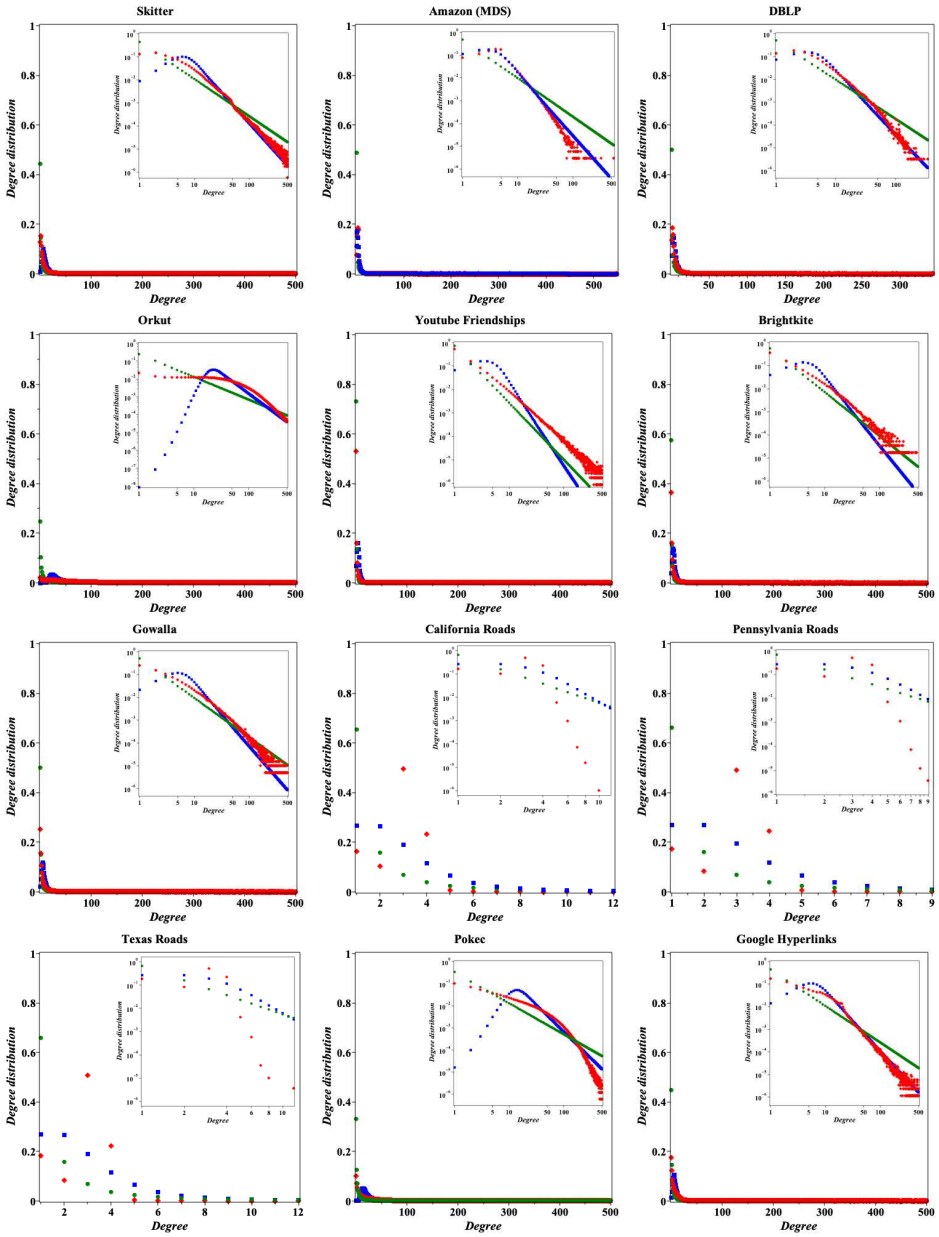


Fig. 7: We present the numerical results for the CM, the degree distribution versus degree is depicted in the log-log scale insert in main graph plot for the Power-law coeficient γ. The Empirical data is represented by red solid daimonds. The FERGMs is represented by blue solid boxes, and finally the PLD with $P(k) \propto k^{-\gamma}$ is represented by the green solid circles.

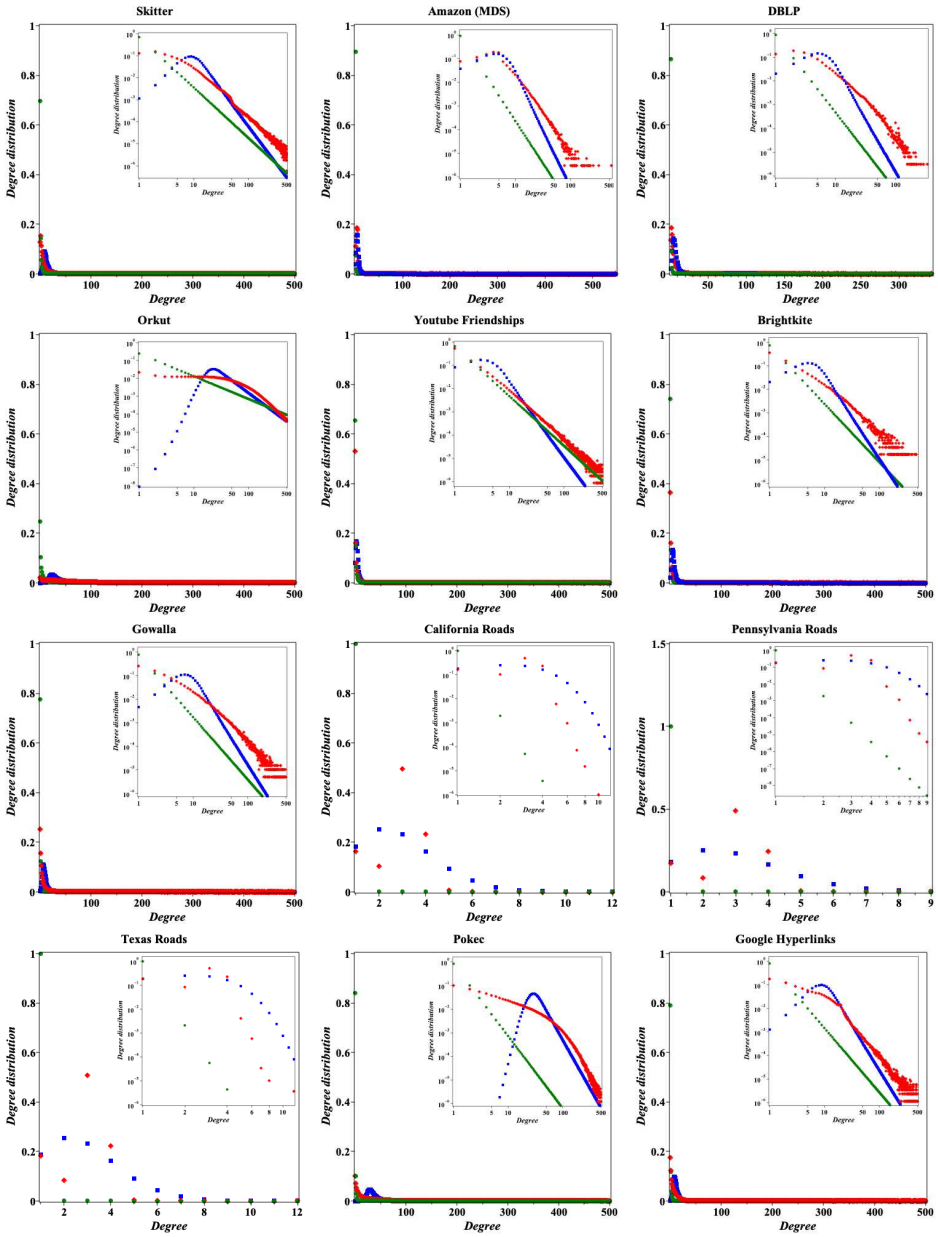


Fig. 8: We present the numerical results for the CM. The degree distribution vs degree is depicted in the log-log scale insert in main graph plot for the Power-law coeficient γ_{tail}. The Empirical data is represented by red solid daimonds. The FERGMs is represented by blue solid boxes, and finally the PLD with $P(k) \propto k^{-\gamma}$ is represented by the green solid circles.

In Tables 4 and 5, we present the GCM results for the 12 empirical networks, considering both power-law exponents, γ and γ_{tail}. In Table 4 provides the numerical results for the GCM, including the reduced temperature, reduced chemical potential, mean node degree, relative error, and residual standard deviation for the two power-law regimes, reported first for γ and then for γ_{tail}. We show that, in most cases, FERGMs systematically outperform standard power-law models, as indicated by their lower RSD values. In Table 5, we present the numerical results for the GCM. In this table, we report the first moment in Eq. (69), the second central moment in Eq. (70), the third standardized moment in Eq. (71), and the fourth standardized moment in Eq. (72) we show that FERGMs generally surpass conventional power-law models. This is reflected in their consistently lower values across each of the first four statistical moments.

In Fig. 9, the degree distribution versus degree for GCM is plotted on a log-log scale, and the corresponding power-law coefficient γ is indicated in the main panel. Empirical data are represented by red solid diamonds; blue solid squares denote the power-law distribution (PLD) with $P(k) \propto k^{-\gamma}$; and green solid circles are also shown. In this results we can observe FERGMs has a rapidly decay to zero in a mayority of cases campare to PLD. Where PLD mantain the tail shape that most of the networks had. For this cases the visually more accurate case presented can be Brightkite and Gowalla where the upper left sector of the plot is better fit, which give more importance by the time to fitting the network. against all odds, this is one of those cases where generalizing a model does not necessarily guarantee better results than those obtained by a more particular case.

In Fig. 10, the relationship between degree and its distribution for GCM is shown, characterized by the tail power-law coefficient γ_{tail}, while the inset presents the same data on a log-log scale. Empirical data are represented by red solid diamonds; the FERGM results with $P(k)$ specified in Eq. (55) are shown as blue solid boxes; and the power-law distribution $P(k) \propto k^{-\gamma}$, defined in Sec. (3.3), is depicted as green solid circles. As we already see in previous figures the PLD lost almost all power of visually model the distributions of the real-world networks, and as in previous figure the FERGMs outperforms PLD by better fitting the upper-left sector of the log-log plot wich give more importance given that the rest of values thend to zero.

The purpose of this setup is to evaluate how changes in the power-law coefficient influence both the body and the tail portions of the distribution. By contrasting the power-law distribution model with the FERGMs, we assess whether these approaches can sufficiently capture the structural characteristics of the network. This assessment relies on comparing statistical moments, examining the relative error in predicting the mean node degree, and analyzing the residual standard deviation. Since the CM has not previously been used to model empirical networks in earlier research, we now apply it to assess how well it captures real-world networks. In addition, we compare its outcomes with those of the GCM, which is presented in this work as a novel extension.

Table 4: We show in this table the numerical results for the GCM. We demonstrate that FERGMs consistently outperform traditional power-law models in most cases, as evidenced by their lower RSD. We can se that the relative error for the PLD can exhibit results from 15% to 768%, and the RSD is in most cases lower for FERGMs.

Network name	PLC	Reduced	Reduced chemical	Parameter	Average node degree		Relative error	RSD	
	γ	temperature, τ	potential, μ^*	e^{μ^*}	ED, FERGMs	PLD	PLD	FERGMs	PLD
Skitter	1.61077	0.00168	-3.83914	0.02151	13.0809	66.3457	407.19 %	0.00541	0.01468
	2.29100	0.00168	-3.83914	0.02151	13.0809	2.69364	79.407 %	0.00541	0.02606
com-amazon	1.69122	0.00008	-0.19744	0.82082	5.52986	9.75713	76.444 %	0.01176	0.02005
	3.58100	0.00008	-0.19744	0.82082	5.52986	1.17348	78.779 %	0.01176	0.03725
com-dblp	1.72001	0.00008	0.20327	1.22540	6.62209	7.67029	15.828 %	0.00896	0.02093
	3.26100	0.00008	0.20327	1.22540	6.62209	1.25559	81.039 %	0.00896	0.04113
com-orkut	1.27235	0.00106	-1.80251	0.16488	76.2813	662.275	768.20 %	0.00673	0.01136
	1.27235	0.00106	-1.80251	0.16488	76.2813	662.275	768.20 %	0.00673	0.01136
com-youtube	2.42900	0.00065	-3.69457	0.02485	5.26505	2.12330	59.671 %	0.01143	0.00922
	2.14100	0.00065	-3.69457	0.02485	5.26505	3.92924	25.371 %	0.01143	0.00569
loc-brightkite	1.90579	0.00124	-0.81708	0.44171	7.35309	6.05793	17.613 %	0.00545	0.00968
	2.48100	0.00124	-0.81708	0.44171	7.35309	1.94193	73.590 %	0.00545	0.01736
loc-gowalla	1.73070	0.00230	-2.36749	0.09371	9.66806	23.0961	138.89 %	0.00277	0.01154
	2.65100	0.00230	-2.36749	0.09371	9.66806	1.67229	82.702 %	0.00277	0.02409
roadNet-CA	2.05571	0.00001	-1.42069	0.24154	2.81559	1.92019	31.801 %	0.13619	0.19640
	8.99100	0.00001	-1.42069	0.24154	2.81559	1.00207	64.409 %	0.13619	0.28935
roadNet-PA	2.05282	0.00002	-1.43697	0.23764	2.83413	1.79152	36.787 %	0.15813	0.22744
	8.99100	0.00002	-1.43697	0.23764	2.83413	1.00207	64.642 %	0.15813	0.33117
roadNet-TX	2.07371	0.00002	-1.42340	0.24089	2.78518	1.90076	31.754 %	0.13742	0.19591
	8.90100	0.00002	-1.42340	0.24089	2.78518	1.00221	64.016 %	0.13742	0.28559
soc-pokec †	1.40480	0.00076	-1.67589	0.18714	37.5092	168.512	349.25 %	0.00327	0.01081
	3.06100	0.00076	-1.67589	0.18714	37.5092	1.33668	96.436 %	0.00327	0.03343
web-Google †	1.61721	0.00028	-1.51023	0.22085	11.6592	32.4436	178.26 %	0.00298	0.01275
	2.73100	0.00028	-1.51023	0.22085	11.6592	1.57532	86.488 %	0.00298	0.02817

Table 5: We show in this table the numerical results for the GCM, showing the four moments. We demonstrate that FERGMs consistently outperform traditional power-law models in most cases, as evidenced by their lower results for each of the first four statistical moments.

Network name	Power-Law Coefficient γ	First Moment Expected Value/10^{-7}		Second Moment Variance		Third Moment Skewness		Fourth Moment Kurtosis/10^5	
		FERGM	PLD	FERGM	PLD	FERGM	PLD	FERGM	PLD
skitter	1.61077	5.886	5.886	0.000029	0.000215	-2080.397	-9792.056	0.67023	2.10613
	2.29100	5.886	5.886	0.000029	0.000679	-2080.397	-10272.99	0.67023	2.23136
com-amazon	1.69122	29.86	29.86	0.000138	0.000402	-3077.101	-7785.039	1.11198	1.79206
	3.58100	29.86	29.86	0.000138	0.001387	-3077.101	-10321.93	1.11198	2.32226
com-dblp	1.72001	31.53	31.53	0.000080	0.000438	-1999.876	-5203.879	0.44618	0.92023
	3.26100	31.53	31.53	0.000080	0.001691	-1999.876	-5481.610	0.44618	0.97415
com-orkut	1.27235	3.217	3.217	0.000045	0.000129	-7104.622	-8292.787	1.19853	1.55311
	1.27235	3.217	3.217	0.000045	0.000129	-7104.622	-8292.787	1.19853	1.55311
com-youtube	2.42900	8.806	8.806	0.000130	0.000085	9723.572	-9926.672	2.08008	2.14551
	2.14100	8.806	8.806	0.000130	0.000032	9723.572	-9937.868	2.08008	2.14809
loc-brightkite	1.90579	171.7	171.7	0.000029	0.000093	8477.308	-10304.63	1.68111	2.23936
	2.48100	171.7	171.7	0.000029	0.000301	8477.308	-10289.96	1.68111	2.23619
loc-gowalla	1.73070	50.84	50.84	0.000007	0.000133	4276.395	-10070.77	0.51785	2.17856
	2.65100	50.84	50.84	0.000007	0.000580	4276.395	-10273.14	0.51785	2.23207
roadNet-CA	2.05571	5.088	5.088	0.018549	0.038574	16.70914	-4.830455	0.00031	0.00031
	8.99100	5.088	5.088	0.018549	0.083727	16.70914	-20.14131	0.00031	0.00048
roadNet-PA	2.05282	9.190	9.190	0.025005	0.051733	10.43173	-3.316731	0.00008	0.00010
	8.99100	9.190	9.190	0.025005	0.109678	10.43173	-13.27716	0.00008	0.00021
roadNet-TX	2.07371	7.246	7.246	0.018886	0.038381	18.87830	-2.509578	0.00034	0.00030
	8.90100	7.246	7.246	0.018886	0.081563	18.87830	-18.79923	0.00034	0.00047
soc-pokec †	1.40480	6.122	6.122	0.000010	0.000111	-8260.712	-9776.171	1.66392	2.04540
	3.06100	6.122	6.122	0.000010	0.001116	-8260.712	-10840.98	1.66392	2.38440
web-Google †	1.61721	11.41	11.41	0.000008	0.000162	-2533.962	-9731.367	0.52856	2.07987
	2.73100	11.41	11.41	0.000008	0.000794	-2533.962	-10387.38	0.52856	2.25981

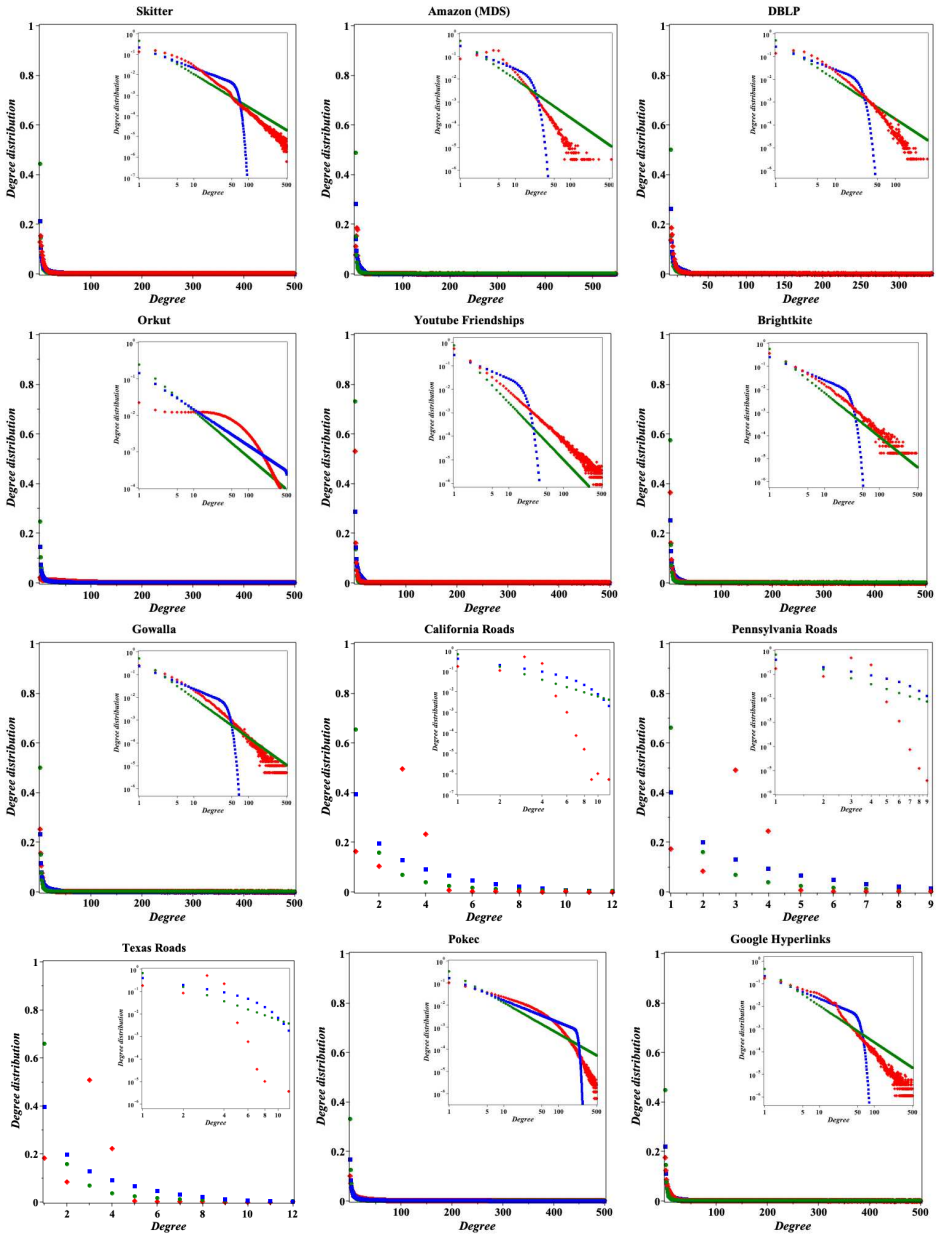


Fig. 9: We present the numerical results for the GMC. The degree distribution vs degree is depicted in the log-log scale and insert in main graph plot for the Power-law coeficient γ. The Empirical data is represented by red solid daimonds. The FERGMs is represented by blue solid boxes, and finally the PLD with $P(k) \propto k^{-\gamma}$ is represented by the green solid circles.

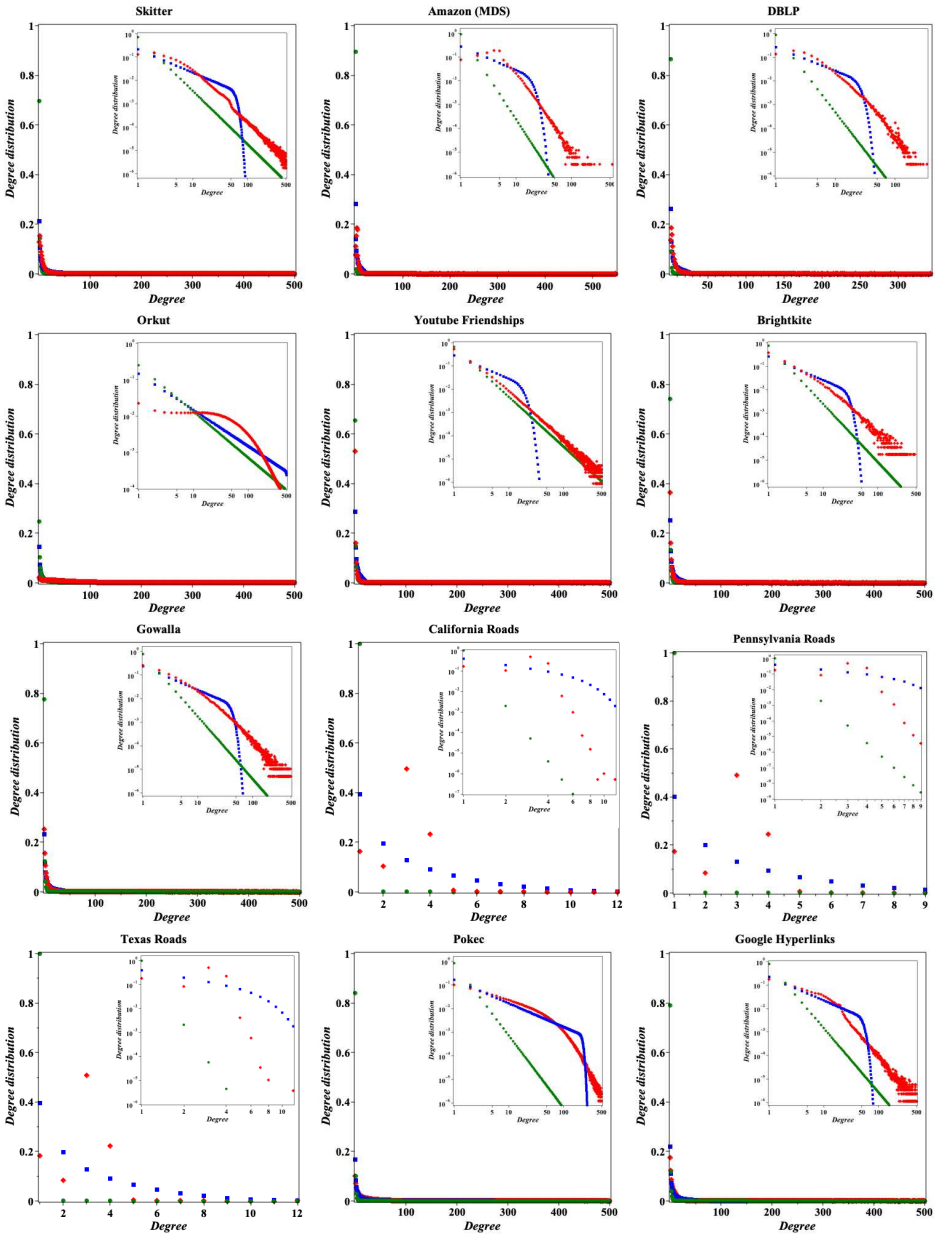


Fig. 10: For purely illustrative purposes, we present the numerical results for the GMC. The degree distribution vs degree is depicted in the log-log scale and insert in main graph plot for the Power-law coeficient $\gamma_{\mathbf{tail}}$. The Empirical data is represented by red solid daimonds. The FERGMs is represented by blue solid boxes, and finally the PLD with $P(k) \propto k^{-\gamma}$ is represented by the green solid circles.

5 Discussion and Conclusion

In this work, we present a unified statistical–mechanical framework for modeling empirical networks using FERGMs. We propose two complementary formulations: the Configuration Model (CM) with bounded fitness, $0 \leq \varepsilon \leq \varepsilon_0$, and the Generalized Configuration Model (GCM) with unbounded fitness, $0 \leq \varepsilon \leq \infty$. These two formulations capture different physical regimes of network formation, each arising from distinct constraints on the available resources. We have derived exact and approximate analytical expressions for degree distributions across multiple temperature regimes. For the CM in the high-energy limit ($e^{\beta\mu} \gg 1$, $e^{\beta\varepsilon} \gg 1$), the degree distribution takes the form of Eq. (49), expressed through the lower incomplete γ function. For the GCM in the sparse network regime ($e^{\mu^*} \ll 1$), we obtained a closed-form solution, Eq. (55), revealing a Poisson-like structure at low temperatures, smoothly transitioning to broader distributions as the temperature increases.

Through systematic comparison with 12 diverse real-world networks—spanning technological (Skitter, web-Google), social (Orkut, YouTube, Brightkite, Gowalla, Pokec), collaborative (DBLP), commercial (Amazon), and infrastructure (California, Pennsylvania, Texas road networks) domains—we demonstrate that FERGMs consistently outperform traditional power-law models in reproducing observed degree distributions, as Fig. 11 shows.

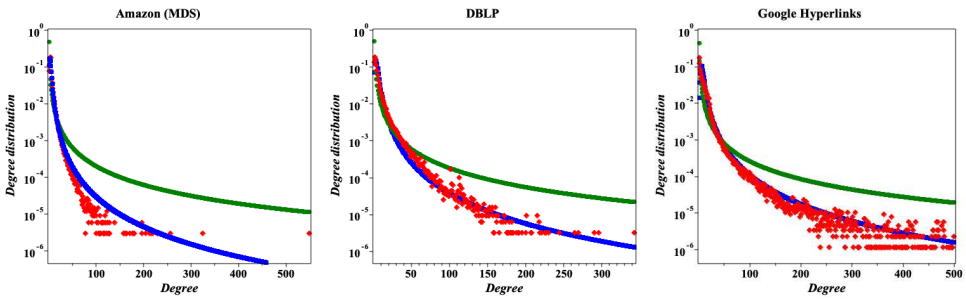


Fig. 11: We present the cases of success for CM: the degree distribution vs degree is depicted in the log-plot for the Power-law coefficient γ. The Empirical data is represented by red solid daimonds. The FERGMs is represented by blue solid boxes, and finally the PLD with $P(k) \propto k^{-\gamma}$ is represented by the green solid circles.

As shown, in Tables 2-5, FERGMs achieve zero relative error in average node degree by construction, while power-law distributions exhibit errors ranging from 15% to 768%. The residual standard deviations are consistently lower for FERGMs, particularly for social and technological networks.

Our results contribute to the growing body of evidence challenging the universality of scale-free networks [25, 27]. Instead of following a strict power-law degree distribution $P(k) \propto k^{-\gamma}$ over all scales, most empirical networks are, at best, only "weakly scale-free," showing power-law-like behavior only within restricted degree intervals. The FERGM framework naturally captures this phenomenon via the temperature-dependent interaction between fitness distributions and connection probabilities.

The FERGM framework offers several practical advantages. For network reconstruction, given partial observations such as the average node degree values or clustering coefficients, FERGMs generate ensemble-averaged network properties with quantified uncertainties. For network design, understanding temperature-fitness relationships enables the engineering of networks with desired topological properties through controlled fitness distributions. More broadly, our work exemplifies the deep connection between network science and equilibrium statistical mechanics. The Shannon-Gibbs entropy maximization principle naturally leads to ERGMs, while the grand canonical ensemble formulation enables the treatment of networks with variable link numbers. The fitness variable plays a role analogous to energy in physical systems, with the density of states $\rho(\varepsilon)$ determining accessible microstates. This study introduces FERGMs as a rigorous, analytically tractable, and empirically validated framework for characterizing the structural properties of real-world networks. By formulating network formation within statistical-mechanical principles and incorporating intrinsic vertex fitness as a hidden variable, we propose an alternative to the canonical scale-free framework that better captures the diversity of observed degree distributions.

The analytical tractability of our approach facilitates a clear separation of topological (fitness distribution $\rho(\varepsilon)$) and thermodynamic (temperature, chemical potential) contributions to network structure. This decomposition enables the systematic study of how microscopic node properties (fitness) and macroscopic constraints (temperature, average connectivity) jointly determine emergent network topology.

The temperature–fitness duality demonstrates that network topology arises from a subtle equilibrium between entropic disorder and fitness-driven structural organization. Distinct classes of networks—ranging from highly structured infrastructure networks to heterogeneous social networks—occupy specific regions within this thermodynamic landscape, thereby accounting for systematic deviations from otherwise universal scaling laws.

Our demonstration that FERGMs outperform traditional power-law models across a wide range of network types, together with our derivation of analytical solutions in several thermodynamic regimes, establishes this framework as a valuable tool for the network science community. As complex networks remain central to understanding biological, social, and technological phenomena, frameworks like FERGMs—combining theoretical sophistication with strong empirical performance—will become increasingly vital for advancing both fundamental theory and real-world applications.

To sum up, FERGMs provide a powerful perspective for interpreting the world as a network of interdependent relationships. By integrating the concept of fitness

and allowing for node-level attributes, these models yield a refined picture of network organization. Their applications span diverse fields, from social science to marketing, offering important insights into the behavior and evolution of interconnected systems. Although challenges persist, the capacity of FERGMs to expose latent structures in networks and deepen our comprehension of the world around us is undeniable.

Several limitations of the current framework suggest directions for future research, as: (i) While global clustering serves as a fitting parameter, our formulation does not fully constrain local clustering patterns or higher-order motif distributions. Extensions incorporating triadic closure mechanisms or community structure could enhance the predictive power for networks with strong mesoscale organization. (ii) Analytical evaluation of degree distributions beyond asymptotic limits involves integrals of hypergeometric functions that can be numerically challenging. The development of efficient approximation schemes, possibly leveraging asymptotic expansions or saddle-point methods, would broaden the practical applicability of our approach. (iii) The current framework describes static network snapshots. Extension to temporal ERGMs (TERGMs) that incorporate fitness-driven evolution while preserving analytical tractability represents an important frontier. Similarly, generalization to directed networks with asymmetric fitness and weighted networks with continuous interaction strengths would significantly expand the model's scope.

Even with these limitations, FERGMs continue to serve as a valuable instrument for deciphering the intricate patterns of connections in our world. As future work further improves these models and broadens their range of applications, the hidden order within networks is likely to emerge more clearly, allowing us to better understand and potentially anticipate the behavior of the complex systems that influence our lives.

Finally, while we demonstrate that exponential fitness distributions $\rho(\varepsilon) \propto e^{\pm\alpha\beta\varepsilon}$ effectively capture observed degree distributions, understanding the mechanistic origins of these distributions—whether they emerge from multiplicative growth processes, optimization principles, or other dynamical rules—remains an open question warranting further investigation.

Appendix A Generating Function

Following [37], we define a generating function as

$$G_0(z) = \sum_k z^k P(k), \tag{A1}$$

where $P(k)$ is the degree distribution (the probability that any given vertex has degree k). Further, all calculations will be confined to the region $0 \leq z \leq 1$.

Once the generating function is known, the degree distribution and its moments can be readily computed:

$$P(k) = \frac{1}{k!}\frac{d^k}{dz^k}G_0(z)\Big|_{z=0}, \tag{A2}$$

$$\langle k^n \rangle = \left(z\frac{d}{dz}\right)^n G_0(z)\Big|_{z=1}. \tag{A3}$$

In particular, this yields

$$\langle k \rangle = G_0'(1), \tag{A4}$$

$$\langle k^2 \rangle = G_0''(1) + G_0'(1). \tag{A5}$$

Furthermore, it is useful to introduce a shorthand notation for the derivatives of the generating function:

$$z_n = \frac{d^n}{dz^n} G_0(z)\Big|_{z=1}. \tag{A6}$$

Then, applying Eq. (A3), we find that $\langle k \rangle = z_1$, $\langle k^2 \rangle = z_2 + z_1$, and so on.

We are now ready to analyze the topological properties of the network. First, we are interested in the degree distribution $P(k)$. To compute $G_0(z)$, we use the generating function formalism for networks with hidden variables developed in [65]. One can write the degree distribution as

$$P(k) = \int g(k \mid \varepsilon)\rho(\varepsilon)d\varepsilon, \tag{A7}$$

where $g(k \mid \varepsilon)$ denotes the propagator, which satisfies the normalization condition $\sum_k g(k \mid \varepsilon) = 1$. By inserting $P(k)$ into Eq. (A1), we obtain

$$G_0(z) = \int d\varepsilon \rho(\varepsilon) \sum_k z^k g(k \mid \varepsilon). \tag{A8}$$

As shown in [65],

$$\ln \sum_k z^k g(k \mid \varepsilon) = N \int d\varepsilon' \rho(\varepsilon') \ln \big(1 - (1-z)p(\varepsilon, \varepsilon')\big). \tag{A9}$$

Using this result in Eq. (A8), we obtain

$$G_0(z) = \int d\varepsilon \rho(\varepsilon) \exp\Big[N \int d\varepsilon' \rho(\varepsilon') \ln \big(1 - (1-z)p(\varepsilon, \varepsilon')\big)\Big]. \tag{A10}$$

In the limit of $N \gg 1$, one can approximate the generating function as Eq. (35) where $\bar{k}(\varepsilon)$ is the expected degree of the node with the hidden variable ε defined by Eq. (25). The computation of the degree distribution $P(k)$ is defined by Eq. (36).

A direct calculation yields the following expression: $z_n = \langle \bar{k}^n \rangle$, where $\langle \bar{k}^n \rangle$ denotes the n-th moment of the node degree:

$$\langle \bar{k}^n \rangle = \int \bar{k}^n(\varepsilon)\rho(\varepsilon)d\varepsilon. \tag{A11}$$

By applying Eq. (A3), the n-th moment of the degree distribution, $\langle k^n \rangle$, can be determined provided that the corresponding moments of the hidden variables are known, $\langle \overline{k}^n \rangle$. In particular, we obtain, $\langle k^2 \rangle = \langle \overline{k}^2 \rangle + \langle k \rangle$.

Appendix B Degree Distribution

We are now prepared to investigate the network's topological characteristics. Our initial focus is on the degree distribution, $P(k)$. To this end, we employ the generating function method described in Appendix (A). For sparse networks, the generating function is specified by Eq. (35), where

$$\overline{k}(\varepsilon) = (N-1) \int_0^{\varepsilon_0} p(\varepsilon, \varepsilon') \rho(\varepsilon') d\varepsilon', \tag{B12}$$

it denotes the expected degree of the vertex associated with the hidden variable ε; see Eq. (53).

The calculation of the degree distribution gives

$$P(k) = \frac{1}{k!} \int_0^{\varepsilon_0} e^{-\overline{k}(\varepsilon)} \left(\overline{k}(\varepsilon)\right)^k \rho(\varepsilon) \, d\varepsilon. \tag{B13}$$

The degree distribution $P(k)$ describes how the connections in a network are distributed among nodes. Specifically, it gives the probability distribution of the degrees (the number of connections or edges) across all nodes in the network.

B.1 Case $e^{\beta\mu} \gg 1$, $e^{\beta\varepsilon} \gg 1$

We can express both the expected vertex degree and the average node degree in terms of the reduced chemical potential, defined as $\mu^* = \beta\mu$, and the new hidden parameter $\mu^* x = \beta\varepsilon$ as follows

$$\overline{k}(x) = \frac{N-1}{2\sinh(\alpha\mu^*/2)} \left(e^{\alpha\mu^*/2} \, {}_2F_1\left(1, \alpha; 1+\alpha; -e^{\mu^* x}\right) - e^{-\alpha\mu^*/2} \, {}_2F_1\left(1, \alpha; 1+\alpha; -e^{\mu^*(x-1)}\right)\right), \tag{B14}$$

$$\langle k \rangle = \frac{(N-1)}{4\sinh^2(\alpha\mu^*/2)} \Big[e^{\alpha\mu^*} \, {}_3F_2\left(1, \alpha, \alpha; 1+\alpha, 1+\alpha; -e^{\mu^*}\right) - 2 \, {}_3F_2\left(1, \alpha, \alpha; 1+\alpha, 1+\alpha; -1\right) + e^{-\alpha\mu^*} \, {}_3F_2\left(1, \alpha, \alpha; 1+\alpha, 1+\alpha; -e^{-\mu^*}\right) \Big]. \tag{B15}$$

To investigate the thermodynamic limit, the model must be defined by fixing the parameters α and μ^*. Using the condition $\beta\mu = \mu^* \gg 1$, we find

$$\overline{k}(x) = \frac{N-1}{2\sinh(\alpha\mu^*/2)} e^{\alpha\mu^*/2} \, {}_2F_1\left(1, \alpha; 1+\alpha; -e^{\mu^* x}\right), \tag{B16}$$

$$\langle k\rangle = \frac{(N-1)}{4}\frac{e^{\alpha\mu^*}}{\sinh^2(\alpha\mu^*/2)}\ {}_3F_2\left(1,\alpha,\alpha;1+\alpha,1+\alpha;-e^{\mu^*}\right). \tag{B17}$$

The degree distribution is given by high energies, $\beta\varepsilon = \mu^* x \gg 1$. Using the asymptotic properties of the hypergeometric functions, we obtain

$$\overline{k}(x) = \Lambda e^{-\mu^* x} + \Omega e^{-\alpha\mu^* x}, \tag{B18}$$

$$\langle k\rangle = \Lambda' e^{-\mu^*} + \Omega' e^{-\alpha\mu^*}, \tag{B19}$$

where

$$\Lambda = \frac{(N-1)}{2}\frac{e^{\alpha\mu^*/2}}{\sinh(\alpha\mu^*/2)}\frac{\alpha}{\alpha-1}, \tag{B20}$$

$$\Omega = \frac{(N-1)}{2}\frac{e^{\alpha\mu^*/2}}{\sinh(\alpha\mu^*/2)}\frac{\alpha\pi}{\sin(\alpha\pi)}, \tag{B21}$$

$$\Lambda' = \left[\frac{(N-1)}{4}\frac{e^{\alpha\mu^*}}{\sinh^2(\alpha\mu^*/2)}\alpha^2\right]\left(\frac{1}{(\alpha-1)^2}\right), \tag{B22}$$

$$\Omega' = \left[\frac{(N-1)}{4}\frac{e^{\alpha\mu^*}}{\sinh^2(\alpha\mu^*/2)}\alpha^2\right]\left(\frac{\mu^*\pi}{\sin(\pi\alpha)} + \frac{\pi^2\cdot\cos(\pi\alpha)}{\sin^2(\pi\alpha)}\right). \tag{B23}$$

It is easy to prove that the only terms that make a contribution are the first terms such that we can rewrite Eqs. (B18) - (B19) as

$$\overline{k}(x) = \Lambda e^{-\mu^* x}, \tag{B24}$$

$$\langle k\rangle = \Lambda' e^{-\mu^*}. \tag{B25}$$

With these results, we can express Eq. (35) in the form

$$G_0(z) = -\frac{\alpha e^{-\alpha\mu^*/2}}{2\sinh(\alpha\mu^*/2)}\int_1^{-y_0} e^{(z-1)\Lambda y} y^{-(\alpha+1)}dy, \tag{B26}$$

where $y = e^{-\beta\varepsilon} = e^{\mu^* x}$ and $y_0 = e^{-\beta\mu} = e^{-\mu^*}$. And integrating, it is easy to confirm that $G_0(1) = 1$. By applying another change of variables, we can rewrite Eq. (B26)

$$G_0(z) = -\frac{\alpha e^{-\alpha\mu^*/2}\left((z-1)\Lambda\right)^\alpha}{2\sinh(\alpha\mu^*/2)}\int_{w_0}^{w_1} e^{-w} y^{-(\alpha+1)}dy, \tag{B27}$$

where $-w = (z-1)\Lambda y$. Upon performing the integration, we obtain

$$G_0(z) = -\frac{\alpha e^{-\alpha\mu^*/2}\left((z-1)\Lambda\right)^\alpha}{2\sinh(\alpha\mu^*/2)}\left[\Gamma(-\alpha,w_0) - \Gamma(-\alpha,w_1)\right], \tag{B28}$$

where $\Gamma(a,x)$ is the upper incomplete gamma function. To verify the correctness of our result, we evaluate $G_0(1)$ and $G_0^{''}(1)$. This gives $G_0(1)=1$ and $G_0^{''}(1)=\langle k\rangle$, in agreement with expectations.

To determine the degree distribution, we insert the expression of Eq. (A2) into the generating function. Carrying out the differentiation, we find

$$P(k)=-\frac{\alpha e^{-\alpha\mu^*/2}}{2\sinh(\alpha\mu^*/2)}\Lambda^{\alpha}\frac{1}{k!}\left[\Gamma(k-\alpha,\Lambda)-\Gamma\left(k-\alpha,\Lambda e^{-\mu^*}\right)\right], \tag{B29}$$

where Eq. (D55) has been employed to determine the degree distribution.

Appendix C Generalized Configuration Model

For the GCM case where the upper limit of the energy isn't cut off. Using the generation function defined in Eq. (42), where

$$\overline{k}(\varepsilon)=(N-1)\int_0^{\infty}p(\varepsilon,\varepsilon')\rho(\varepsilon')d\varepsilon', \tag{C30}$$

is the expected vertex degree for the node associated with the hidden variable ε as defined in Eq. (53). Evaluating the degree distribution leads to

$$P(k)=\frac{1}{k!}\int_0^{\infty}e^{-\bar{k}(\varepsilon)}\left(\bar{k}(\varepsilon)\right)^k\rho(\varepsilon)d\varepsilon. \tag{C31}$$

C.1 Arbitrary reduced temperature τ and $e^{\mu^*}\ll 1$

In the limit where $e^{\beta\mu}\ll 1$, the expressions for $\bar{k}(\varepsilon)$ and $\langle k\rangle$ can be approximated as follows:

$$\bar{k}(\varepsilon)\approx(N-1)\frac{\tau}{1+\tau}e^{\beta(\mu-\varepsilon)}, \tag{C32}$$

$$\langle k\rangle\approx(N-1)\frac{\tau^2}{(1+\tau)^2}e^{\beta\mu}. \tag{C33}$$

Using the asymptotic form of the generalized hypergeometric function, in conjunction with Eqs. (D53) and (D58), we can rewrite Eq.(35) as:

$$G_0(z)=\tau\beta\int_0^{\infty}e^{\Lambda^+e^{-\beta\varepsilon}}e^{-\tau\beta\varepsilon}d\varepsilon, \tag{C34}$$

where

$$\Lambda=(z-1)(N-1)\frac{\tau}{\tau+1},\quad\text{and}\quad\Lambda^+=\Lambda e^{\beta\mu}, \tag{C35}$$

by introducing the substitutions $y=e^{-\beta\varepsilon}$ and $b=e^{\beta\mu}$ in Eq. (C34), we obtain

$$G_0(z)=\tau\int_0^1 e^{\Lambda by}y^{\tau-1}dy, \tag{C36}$$

we identify this as the lower incomplete gamma function by setting $-s = \Lambda b$ and $sy = u$, which gives

$$G_0(z) = \frac{\tau}{s^\tau} \int_0^s e^{-u} u^{\tau-1} du = \frac{\tau}{s^\tau} \gamma(\tau, s). \tag{C37}$$

By substituting these results into Eq. (A2), we obtain

$$P(k) = \frac{\tau \Lambda^{-\tau} \gamma(k+\tau, \Lambda)}{k!}, \tag{C38}$$

where

$$\Lambda = \frac{(1+\tau)\langle k \rangle}{\tau}. \tag{C39}$$

A similar approach can be reached by solving the integral $P(k)$ directly in Eq. (C31) using the limiting case for $\overline{k}(\varepsilon)$ in Eq. (C32) and employing some of these changes of variables, as we do it in this section.

C.2 Arbitrary reduced temperature τ and $e^{\beta\mu} \gg 1$

For this case, we use the definition written in Eq. (53) and Eq.(54) for the limit when $\beta\mu \to \infty$. We then approximate the hypergeometric function using Eq. (D54), making use of the facts that in our situation $-1 < \mathbb{R}(c-a-b) \leq 0$ and $|\arg(-z)| < \pi$.

$$\overline{k}(\varepsilon) \approx (N-1) e^{\beta\varepsilon}, \tag{C40}$$

$$\langle k \rangle \approx (N-1). \tag{C41}$$

For the low-temperature regime: $\tau \ll 1$, since the main contribution in the integral for the computation of the generating function is given by the region $0 < \varepsilon < \mu$, one can approximate the expected node degree as follows:

$$\bar{k}(\varepsilon) = (N-1)\mathcal{H}(\mu - \varepsilon), \tag{C42}$$

where $\mathcal{H}(x)$ is the Heaviside step function, and $x = \mu - \varepsilon$. If $\mu < \varepsilon$, then $x < 0$, and if $\mu > \varepsilon$, then $x > 0$,

$$\mathcal{H}(x) := \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases}. \tag{C43}$$

The transition at $\varepsilon = \mu$ becomes sharper as β increases. Straightforward computation when $x > 0$ of the generating function yields

$$G_0(z) = e^{-(1-z)\langle k \rangle}. \tag{C44}$$

Next, applying Eq. (C44) into Eq. (A2) and differentiate, we derive the Poisson

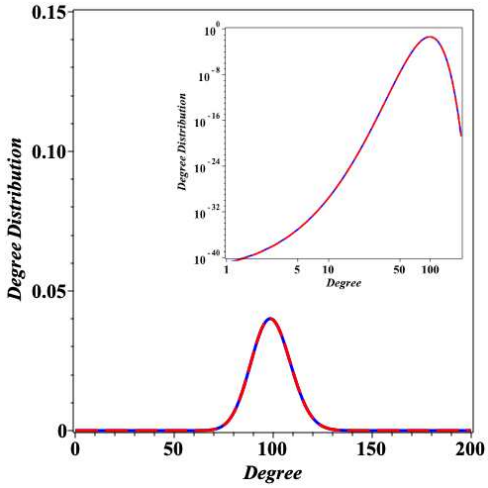


Fig. C1: We present the degree distribution $P(k)$, for the generalized configuration model in the main panel, and the corresponding log–log representation for the same regime in the inset. The blue curve represent the limit case for degree distribution $P(k)$ for arbitrary reduced temperature τ and $e^{\mu^*} \gg 1$. The red dashed curve presents the Poisson distribution where the analytical prediction and the exact numerical result coincide within graphical accuracy. Number of nodes $N = 10^2$. Parameters: $\tau = 0.1$, $\mu^* = 20$.

distribution

$$P(k) = \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!}. \tag{C45}$$

Note that for $x < 0$, the expression becomes undefined. Therefore, in the low-temperature limit with $\mu^* \gg 1$, the network reduces to a classical random graph characterized by a Poisson degree distribution [71]. In Fig. C1, for a network temperature of $\tau = 0.1$, we compare this Poisson distribution with the degree distribution obtained from Eq. (C31). The number of nodes is $N = 10^2$, and the parameters $\tau = 0.1$, $\mu^* = 20$. One observes an excellent match between the two results.

Appendix D Analytical Identities

The definition of a generalized hypergeometric function

$${}_pF_q\left(a_1, ..., a_p; b_1, .., b_q; z\right) = \sum_{n=0}^{\infty} \frac{(a_1)_n \dots (a_p)_n}{(b_1)_n \dots (b_q)_n} \frac{z^n}{n!}, \tag{D46}$$

where $(a_i)_n$ and $(b_i)_n$ are the ascendant factorial (also called the Pochhammer symbol, Pochhammer function, or Pochhammer polynomial).

In this work, we continue to assume that none of the bottom parameters b_i are nonpositive integers. The Pochhammer symbols are given by

$$(x)_n = x\,(x+1)\,(x+2)\ldots(x+n-1)\,, \tag{D47}$$

for the particular case $p = q+1$, with $(a_1) = 1$, $(a_{i\neq 1}) = a$, and $(b_i) = 1+a$, we may express it as

$${}_{q+1}F_q\,(1, a, ..., a; 1+a, .., 1+a; z) = a^q \phi\,(z, q, a)\,, \tag{D48}$$

where $\phi\,(z, q, a)$ is Lerch transcendent function defined as

$$\phi\,(z, q, a) = \sum_{n=0}^{\infty} \frac{z^n}{(n+a)^q}. \tag{D49}$$

The polylogarithm arises as a particular case of the Lerch transcendent function, if we set $k = n+1$, we obtain

$$Li_s\,(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s} = z\phi\,(z, s, 1)\,, \tag{D50}$$

and, it is straightforward to verify that

$${}_3F_2\,(1, 1, 1; 2, 2; z) = z^{-1}\mathrm{Li}_2\,(z)\,. \tag{D51}$$

Another particular case occurs for $a = 1$, $b = 1$, and $c = 2$, and is specified when

$${}_2F_1\,(1, 1; 2; z) = -z^{-1}\ln\,(1-z)\,, \tag{D52}$$

It is straightforward to verify that, in this particular case, the identity relating the generalized hypergeometric functions for $q = 2$ and $q = 1$ can be expressed as:

$${}_3F_2\big(1, a, a; 1+a, 1+a; z\big) = -a^2 \frac{\partial}{\partial a}\Big(\frac{1}{a}\,{}_2F_1\big(1, a; 1+a; z\big)\Big), \tag{D53}$$

another useful identity employed in this work arises in the limit $z \to \infty$, where we can approximate the hypergeometric function by [72]:

$$\begin{aligned} {}_2F_1\,(a, b; c; z) =& \frac{\Gamma\,(c)\,\Gamma\,(b-a)}{\Gamma\,(b)\,\Gamma\,(c-a)}\,(-z)^{-a}\;{}_2F_1\left(a, 1-c+a; 1-b+a; \frac{1}{z}\right) \\ &+ \frac{\Gamma\,(c)\,\Gamma\,(a-b)}{\Gamma\,(a)\,\Gamma\,(c-b)}\,(-z)^{-b}\;{}_2F_1\left(b, 1-c+b; 1-a+b; \frac{1}{z}\right) \qquad (|\arg\,(-z)| < \pi)\,. \end{aligned} \tag{D54}$$

For the derivation of the degree distribution in Eq. (B29)

$$\int_{v_0}^{v_1} e^{-v} v^{k-\alpha-1} dw = \int_0^{v_1} e^{-v} v^{k-\alpha-1} dv - \int_0^{v_0} e^{-v} v^{k-\alpha-1} dv = \gamma\left(k-\alpha, v_1\right) - \gamma\left(k-\alpha, v_0\right)$$
$$= \int_{v_0}^{\infty} e^{-v} v^{k-\alpha-1} dv - \int_{v_1}^{\infty} e^{-v} v^{-\alpha-1} dv = \Gamma\left(k-\alpha, v_0\right) - \Gamma\left(k-\alpha, v_1\right). \tag{D55}$$

A helpfull identity used in this work is [68, 72]:

$${}_2F_1\left(a, b; c; z\right) = {}_2F_1\left(b, a; c; z\right), \tag{D56}$$

and specifically in the case where $b = c$ [69]

$${}_2F_1\left(a, b; b; z\right) = (1-z)^{-a}, \tag{D57}$$

for the specific choice $a = 1$, $b = \alpha$, $c = \alpha + 1$ and $z = -1$ we may apply

$${}_2F_1\left(1, \alpha; 1+\alpha; -1\right) = \frac{1}{2}\alpha\left[\psi\left(\frac{1}{2}\alpha - \frac{1}{2}\right) - \psi\left(\frac{1}{2}\alpha\right)\right], \tag{D58}$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$ for $x \neq 0, -1, -2, \ldots$, which is known as the ψ-function.

References

[1] Barabási, A.-L.: Network Science. Cambridge University Press, Cambridge (2016)

[2] Menczer, F., Fortunato, S., Davis, C.A.: A First Course in Network Science. Cambridge University Press, ??? (2020). https://books.google.com.mx/books?id=q1abxgEACAAJ

[3] Dorogovtsev, S.N., Mendes, J.F.F.: The Nature of Complex Networks. Oxford University Press, ??? (2022). https://doi.org/10.1093/oso/9780199695119.001.0001 . https://doi.org/10.1093/oso/9780199695119.001.0001

[4] Vito Latora, G.R. Vincenzo Nicosia: Complex Networks: Principles, Methods and Applications. Cambridge University Press, ??? (2017)

[5] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex networks: Structure and dynamics. Physics Reports **424**(4), 175–308 (2006) https://doi.org/10.1016/j.physrep.2005.10.009

[6] Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 47–97 (2002) https://doi.org/10.1103/RevModPhys.74.47

[7] Park, J., Newman, M.E.J.: Statistical mechanics of networks. Phys. Rev. E **70**, 066117 (2004) https://doi.org/10.1103/PhysRevE.70.066117

[8] Cimini, G., Squartini, T., Saracco, F., Garlaschelli, D., Gabrielli, A., Caldarelli, G.: The statistical physics of real-world networks. Nature Reviews Physics **1**(1), 58–71 (2019) https://doi.org/10.1038/s42254-018-0002-6

[9] Ghafouri, S., Khasteh, S.H.: A survey on exponential random graph models: an application perspective. PeerJ Computer Science **6**, 269 (2020) https://doi.org/10.7717/peerj-cs.269

[10] Vega Yon, G.G., Slaughter, A., de la Haye, K.: Exponential random graph models for little networks. Social Networks **64**, 225–238 (2021) https://doi.org/10.1016/j.socnet.2020.07.005

[11] Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J.: Everyone's an influencer: quantifying influence on twitter. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. WSDM '11, pp. 65–74. Association for Computing Machinery, New York, NY, USA (2011). https://doi.org/10.1145/1935826.1935845 . https://doi.org/10.1145/1935826.1935845

[12] Robins, G., Pattison, P., Kalish, Y., Lusher, D.: An introduction to exponential random graph (p*) models for social networks. Social Networks **29**(2), 173–191 (2007) https://doi.org/10.1016/j.socnet.2006.08.002 . Special Section: Advances in Exponential Random Graph (p*) Models

[13] Leifeld, P., Cranmer, S.J., Desmarais, B.A.: Temporal exponential random graph models with btergm: Estimation and bootstrap confidence intervals. Journal of Statistical Software **83**(6), 1–36 (2018) https://doi.org/10.18637/jss.v083.i06

[14] Cranmer, S.J., Desmarais, B.A.: Inferential network analysis with exponential random graph models. Political Analysis **19**(1), 66–86 (2011) https://doi.org/10.1093/pan/mpq037

[15] Kronegger, L., Mali, F., Ferligoj, A., Doreian, P.: Collaboration structures in Slovenian scientific communities. Scientometrics **90**(2), 631–647 (2012) https://doi.org/10.1007/s11192-011-0493-8

[16] Brodie, R.J., Ilic, A., Juric, B., Hollebeek, L.: Consumer engagement in a virtual brand community: An exploratory analysis. Journal of Business Research **66**(1), 105–114 (2013) https://doi.org/10.1016/j.jbusres.2011.07.029 . (1)Thought leadership in brand management(2)Health Marketing

[17] Iyengar, R., Bulte, C., Valente, T.W.: Opinion leadership and social contagion in new product diffusion. Marketing Science **30**(2), 195–212 (2011) https://doi.org/10.1287/mksc.1100.0566 https://doi.org/10.1287/mksc.1100.0566

[18] Nesterov, A.I., Villafuerte, P.H.M., Berman, G.P.: In Response to COVID-19: Configuration Model of the Epidemic Spreading (2020). https://arxiv.org/abs/2007.05496

[19] Lusher, D., Koskine, J., (eds.), G.R.: Exponential Random Graph Models for Social Networks Theory, Methods, and Applications. Structural Analysis in the Social Sciences 35. Cambridge University Press, Cambridge (2012)

[20] Dereich, S., Ortgiese, M.: Robust analysis of preferential attachment models with fitness. Combinatorics, Probability and Computing **23**(3), 386–411 (2014) https://doi.org/10.1017/S0963548314000157

[21] Smolyarenko, I.E., Hoppe, K., Rodgers, G.J.: Network growth model with intrinsic vertex fitness. Phys. Rev. E **88**, 012805 (2013) https://doi.org/10.1103/PhysRevE.88.012805

[22] Bedogne', C., Rodgers, G.J.: Complex growing networks with intrinsic vertex fitness. Phys. Rev. E **74**, 046115 (2006) https://doi.org/10.1103/PhysRevE.74.046115

[23] Caldarelli, G., Capocci, A., De Los Rios, P., Muñoz, M.A.: Scale-free networks from varying vertex intrinsic fitness. Phys. Rev. Lett. **89**, 258702 (2002) https://doi.org/10.1103/PhysRevLett.89.258702

[24] Servedio, V.D.P., Caldarelli, G., Buttà, P.: Vertex intrinsic fitness: How to produce arbitrary scale-free networks. Phys. Rev. E **70**, 056126 (2004) https://doi.org/10.1103/PhysRevE.70.056126

[25] Broido, A.D., Clauset, A.: Scale-free networks are rare. Nat. Commun. **10**(1), 1017 (2019)

[26] Artico, I., Smolyarenko, I., Vinciotti, V., Wit, E.C.: How rare are power-law networks really? Proc Math Phys Eng Sci **476**(2241), 20190742 (2020)

[27] Holme, P.: Rare and everywhere: Perspectives on scale-free networks. Nat. Commun. **10**(1), 1016 (2019)

[28] Newman, M.E.J.: The structure and function of complex networks. SIAM Review **45**(2), 167–256 (2003) https://doi.org/10.1137/s003614450342480

[29] Erdös, P., Rényi, A.: On random graphs i. Publicationes Mathematicae Debrecen **6**, 290–297 (1959)

[30] Barthélemy, M.: Spatial networks. Physics Reports **499**(1), 1–101 (2011) https://doi.org/10.1016/j.physrep.2010.11.002

[31] Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999) https://doi.org/10.1126/science.286.5439.509 https://www.science.org/doi/pdf/10.1126/science.286.5439.509

[32] Barabási, A.L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaborations. Physica A: Statistical

Mechanics and its Applications **311**(3), 590–614 (2002) https://doi.org/10.1016/S0378-4371(02)00736-7

[33] Newman, M.E.J.: The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences **98**(2), 404–409 (2001) https://doi.org/10.1073/pnas.98.2.404 https://www.pnas.org/doi/pdf/10.1073/pnas.98.2.404

[34] Kunegis, J.: KONECT – The Koblenz Network Collection. In: Proc. Int. Conf. on World Wide Web Companion, pp. 1343–1350 (2013). http://dl.acm.org/citation.cfm?id=2488173

[35] Leskovec, J., Krevl, A.: SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data (2014)

[36] Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world'networks. Nature **393**(6684), 440–442 (1998) https://doi.org/10.1038/30918

[37] Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**, 026118 (2001) https://doi.org/10.1103/PhysRevE.64.026118

[38] Squartini, T., Garlaschelli, D.: Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics. Springer, Cham (2017)

[39] Hoorn, P., Lippner, G., Krioukov, D.: Sparse maximum-entropy random graphs with a given power-law degree distribution. Journal of Statistical Physics **173**(3), 806–844 (2018) https://doi.org/10.1007/s10955-017-1887-7

[40] Frank, O.: An exponential family of probability distributions for directed graphs. Journal of the American Statistical Association **76**(373), 58–59 (1981)

[41] Amati, V., Lomi, A., Mira, A.: Social network modeling. Annual Review of Statistics and Its Application **5**(1), 343–369 (2018) https://doi.org/10.1146/annurev-statistics-031017-100746 https://doi.org/10.1146/annurev-statistics-031017-100746

[42] Pol, J.: Introduction to network modeling using exponential random graph models (ergm): Theory and an application using r-project. Computational Economics **54**(3), 845–875 (2019) https://doi.org/10.1007/s10614-018-9853-2

[43] Garlaschelli, D., Loffredo, M.I.: Maximum likelihood: Extracting unbiased information from complex networks. Phys. Rev. E **78**, 015101 (2008) https://doi.org/10.1103/PhysRevE.78.015101

[44] Garlaschelli, D., Ahnert, S.E., Fink, T.M.A., Caldarelli, G.: Low-temperature behaviour of social and economic networks. Entropy **15**(8), 3148–3169 (2013) https://doi.org/10.3390/e15083238

[45] Watts, D.J.: Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton (1999)

[46] Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, ??? (2003). https://doi.org/10.1093/acprof:oso/9780198515906.001.0001 . https://doi.org/10.1093/acprof:oso/9780198515906.001.0001

[47] Barrat, A., Barthélemy, M., Vespignani, A.: Dynamical Processes on Complex Networks. Cambridge University Press, ??? (2008). https://doi.org/10.1017/CBO9780511791383

[48] Caldarelli, G.: Scale-Free Networks: Complex Webs in Nature and Technology. Oxford University Press, ??? (2007). https://doi.org/10.1093/acprof:oso/9780199211517.001.0001 . https://doi.org/10.1093/acprof:oso/9780199211517.001.0001

[49] Barabási, A.-L.: Scale-free networks: A decade and beyond. Science **325**(5939), 412–413 (2009) https://doi.org/10.1126/science.1173299 https://www.science.org/doi/pdf/10.1126/science.1173299

[50] Kwapień, J., Drożdż, S.: Physical approach to complex systems. Physics Reports **515**(3), 115–226 (2012) https://doi.org/10.1016/j.physrep.2012.01.007 . Physical approach to complex systems

[51] Newman, M.: Networks. Oxford University Press, Oxford (2018)

[52] Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences **99**(12), 7821–7826 (2002) https://doi.org/10.1073/pnas.122653799 https://www.pnas.org/doi/pdf/10.1073/pnas.122653799

[53] Voitalov, I., Hoorn, P., Hofstad, R., Krioukov, D.: Scale-free networks well done. Phys. Rev. Res. **1**, 033034 (2019) https://doi.org/10.1103/PhysRevResearch.1.033034

[54] Bianconi, G.: Entropy of network ensembles. Phys. Rev. E **79**, 036114 (2009) https://doi.org/10.1103/PhysRevE.79.036114

[55] Castellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. Rev. Mod. Phys. **81**, 591–646 (2009) https://doi.org/10.1103/RevModPhys.81.591

[56] Barabási, A.-L., Gulbahce, N., Loscalzo, J.: Network medicine: a network-based approach to human disease. Nature Reviews Genetics **12**(1), 56–68 (2011) https://doi.org/10.1038/nrg2918

[57] Strogatz, S.H.: Exploring complex networks. Nature **410**(6825), 268–276 (2001) https://doi.org/10.1038/35065725

[58] Watts, D.J., Dodds, P.S., Newman, M.E.J.: Identity and search in social networks. Science **296**(5571), 1302–1305 (2002) https://doi.org/10.1126/science.1070120 http://science.sciencemag.org/content/296/5571/1302.full.pdf

[59] Newman, M.E.J.: Clustering and preferential attachment in growing networks. Phys. Rev. E **64**, 025102 (2001) https://doi.org/10.1103/PhysRevE.64.025102

[60] Fortunato, S.: Community detection in graphs. Physics Reports **486**(3), 75–174 (2010)

[61] Robledo, O.F., Zhan, X.-X., Hanjalic, A., Wang, H.: Influence of clustering coefficient on network embedding in link prediction. Applied Network Science **7**(1), 35 (2022) https://doi.org/10.1007/s41109-022-00471-1

[62] Barrat, A., Weigt, M.: On the properties of small-world network models. The European Physical Journal B - Condensed Matter and Complex Systems **13**(3), 547–560 (2000) https://doi.org/10.1007/s100510050067

[63] Bollobás, B., Riordan, O.M.: In: Bornholdt, S., Schuster, H.G. (eds.) Mathematical results on scale-free random graphs. Wiley-VCH, Weinheim (2003)

[64] Wang, Y., Ghumare, E., Vandenberghe, R., Dupont, P.: Comparison of Different Generalizations of Clustering Coefficient and Local Efficiency for Weighted Undirected Graphs. Neural Computation **29**(2), 313–331 (2017) https://doi.org/10.1162/NECO_a_00914

[65] Boguñá, M., Pastor-Satorras, R.: Class of correlated random networks with hidden variables. Phys. Rev. E **68**, 036112 (2003) https://doi.org/10.1103/PhysRevE.68.036112

[66] Wilf, H.S.: Generatingfunctionology. A K Peters, Wellesley, MA (2006)

[67] Nesterov, A., Villafuerte, P.H.M.: Critical phenomena in complex networks: from scale-free to random networks. The European Physical Journal B **96**(11), 143 (2023) https://doi.org/10.1140/epjb/s10051-023-00612-0

[68] Erdéley, A., Magnus, W., Oberhettinger, F.: Higher Transcendental Functions, Vol. I. McGraw-Hill, New York, NY, USA (1953)

[69] Olver, F.W.J., Lozier, D.W., Boisvert, R.F., Clark, C.W.: NIST Handbook of Mathematical Functions. Cambridge University Press, Cambridge (2010)

[70] Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I., Queen, N.M.: Integrals and Series: Elementary Functions Vol.1. Taylor & Francis, London (1998)

[71] Dorogovtsev, S.N., Goltsev, A.V., Mendes, J.F.F.: Critical phenomena in complex networks. Rev. Mod. Phys. **80**, 1275–1335 (2008) https://doi.org/10.1103/RevModPhys.80.1275

[72] Abramowitz, M., Stegun, I.A. (eds.): Handbook of Mathematical Functions. Dover, New York (1965)