Renormalisation of Inhomogeneous Random Graphs

Luca Avena [1]
Diego Garlaschelli [2]
Rajat Subhra Hazra [3]
Frank den Hollander [3]

July 13, 2026

Abstract

We consider inhomogeneous random graphs in which vertices are assigned i.i.d. random weights, pairs of distinct vertices are connected by an edge independently with a probability that is a bi-variate function of the weights of the vertices, and single vertices are connected to themselves by a self-loop independently with a probability that is a univariate function of the weight of the vertex. We apply a renormalisation transformation in which vertices are aggregated into groups of equal size according to a greedy algorithm, namely, distinct groups of aggregated vertices are connected by an aggregated edge if and only if there is at least one edge connecting two constituent vertices across the groups, while a group of aggregated vertices is connected to itself by an aggregated self-loop if and only if there is at least one self-loop at an internal vertex or one edge connecting a pair of distinct internal vertices. We analyse what happens when the renormalisation transformation is iterated. In particular, we show that, starting from appropriately scaled connection functions, the iterated renormalised graphs converge to a two-parameter family of random graphs, acting as an attractor in a universality class. We consider a light-tailed regime, for which the scaling limit is a homogeneous Erdős–Rényi random graph, and a heavy-tailed regime, for which the scaling limit is an inhomogeneous random graph with stable infinite-mean random weights and an exponential disconnection function. Different scalings are needed for the two regimes. Which of the two regimes prevails depends on the choice of the connection functions and the choice of the law of the random weights.

MSC 2020: 05C80; 60C05; 60K35; 60K37;

Key words and phrases. Inhomogeneous random graphs, stable law random graphs, random weights, aggregation, network renormalisation, iteration, universality.

Acknowledgment. LA & RSH & FdH were supported through NWO Gravitation Grant NETWORKS-024.002.003. DG was supported through the project *Network renormalization: from theoretical physics to the resilience of societies*, file number NWA.1418.24.029 of the research programme NWA L3 – Innovative projects within routes 2024, and the project *Redefining renormalization for complex networks*, file number OCENW.M.24.039, both of which are partly financed by NWO (see `https://doi.org/10.61686/AOIJP05368` and `https://doi.org/10.61686/PBSEC42210`).

[1]Dipartimento di Matematica e Informatica 'Ulisse Dini', Università degli Studi di Firenze, Viale Giovanni Battista Morgagni, 67/a, 50134 Firenze, Italy.
`luca.avena@unifi.it`

[2]IMT Scuola di Alti Studi, Lucca, Italy & Lorentz Institute for Theoretical Physics, Leiden University, The Netherlands.
`diego.garlaschelli@imtlucca.it`

[3]Mathematical Institute, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands.
`{r.s.hazra,denholla}@math.leidenuniv.nl`

1 Introduction and main results

The goal of this paper is to carry out a *renormalisation analysis* of *inhomogeneous random graphs with random weights*. The renormalisation map considered in this paper acts on the vertices and the weight-dependent connection probabilities in the form of an *aggregation algorithm*. The target is to identify *universal attractors* for the iterates of the renormalisation map. Section 1.1 provides background and motivation. Section 1.2 describes the model. Section 1.3 identifies the renormalisation map. Section 1.4 states the main theorems. Section 1.5 offers a discussion of the main theorems and lists open problems. Section 1.6 gives an outline of the remainder of the paper, which provides technical definitions and proofs of key statements.

1.1 Background and motivation

Renormalisation. Coarse-graining of a model of *interacting particle systems* is a way to aggregate information about the structure and the evolution of the system. With the help of *renormalisation transformations* that sum out over variables on a microscopic scale while appropriately rescaling the model parameters, information is obtained about the system on a mesoscopic scale. Iteration of the transformation defines a *renormalisation flow* in the space of parameters and ultimately leads to a description of the system on a macroscopic scale [1, 2].

Renormalisation plays an important role in the analysis of phase transitions and critical phenomena in equilibrium statistical physics [3]. Since parsimonious microscopic models of physical systems are ultimately based on *effective* rather than *fundamental* degrees of freedom, renormalisation theory aims to ensure that the macroscopic properties predicted by such models do not depend sensitively on microscopic details. In this setting, identifying possible *attractors* of microscopic models under the renormalisation flow is crucial: microscopic differences between models that flow to the same attractor are macroscopically irrelevant.

Attractors are typically identified as *fixed points*, i.e., appropriately defined invariants, of the renormalisation transformation [1, 2, 3]. Under the action of this transformation, it is essential that the flow of parameters keeps tracks of the system within an appropriate class of target models. Generally, it is hard to work out the details, and consequently many arguments in the literature are approximative and mathematically non-rigorous. For instance, renormalisation can throw the system *out* of the class of target models, while approximations presume that nonetheless the system remains close to this class in some sense [4]. Properly ensuring that fixed points also act as attractors along the renormalisation flow is therefore a mathematically delicate task.

Renormalisation of networks. Renormalisation of *complex networks* [5], especially in the presence of non-homogeneities and randomness, represents an open challenge and remains largely unexplored from a mathematical perspective. As for physical systems, in many empirical networks there is no unique natural resolution for the constituent vertices: for instance, individuals may be grouped into households, cities or countries, web pages into domains, neurons into regions of the brain, and firms into industries or sectors. How the representation of large-scale social, technological, biological and economic networks transforms under a change of resolution, and what microscopic network properties are irrelevant for the behaviour of the system at coarser scales, including the processes that run on it, is therefore an important and fascinating question.

Most of the current approaches to network renormalisation are concerned with the identification of meaningful iterative coarse-graining rules that, given a single input graph, produce

an aggregated graph that is maximally informative about the original graph under a chosen criterion. For instance, *geometric renormalisation* [6] proceeds by first finding an optimal embedding of the vertices of the input graph onto some metric space, and then aggregating groups of vertices that are found within a given mutual distance from one another. Edges between constituent vertices are aggregated onto edges between aggregated vertices, and the geometric embedding of the reduced graph is informative about the embedding of the original graph. By contrast, *Laplacian renormalisation* [7] focusses on the Laplacian matrix of the original graph, which acts as the generator of a discrete diffusion process, and proceeds by averaging out the 'fast eigenmodes', which leads to a reduced Laplacian matrix that mimics the same diffusion process as if it were running on an aggregated network. In [8], a Laplacian-based renormalisation scheme is introduced that is similar in spirit and uses Markov intertwining dualities to construct multi-resolution schemes and wavelet transforms on finite undirected graphs.

Random graph models. A different approach, which is our main interest here, focuses on how *random graph models*, along with their underlying parameters, transform under network aggregation. This approach is known as *ensemble* or *multiscale* network renormalisation [9]. The basic idea is that, given a random graph model on an initial set of graphs, together with a deterministic aggregation rule projecting each graph in the initial set onto a reduced graph in some output set, an aggregated random graph model on the output set is induced. In general, the aggregated graph may have a difficult law, with parameters that are complicated functions of the parameters of the original law. Multiscale network renormalisation looks for *invariant random graph ensembles* under a given aggregation rule, i.e., random graphs that remain governed by the original law of the graph, only up to a rescaling of its defining parameters. Such invariant ensembles represent the graph analogue of stable random variables.

If restricted to the class of random graphs with independent edges, the search for an invariant ensemble under a given coarse-graining rule is relatively straightforward and leads to an inhomogeneous random graph model where vertices are assigned positive real weights and a connection between two vertices is established with a probability that is a specific function of the weights of the two vertices [9]. Under a greedy coarse-graining procedure, whereby an aggregate vertex is assigned a weight equal to the sum of the weights of the constituent vertices while two aggregate vertices are connected whenever at least one microscopic connection exists between their constituents, the invariant probability of *disconnection* between two vertices can be identified as a decreasing exponential of the product of the weights of the two vertices [9] (in the case of directed graphs, appropriate generalisations can be made [10]). This specific connection function has been considered previously in the mathematical literature [11, 12, 13, 14, 15], albeit with different purposes. The identification of this connection function with an invariant under a renormalisation transformation opens up new avenues.

Flow of the law of the weights. While most of the prior research focussed on the case where the vertex weights are either deterministic variables (representing e.g. some empirically observable amount of activity, mass, fitness, population size or capacity) or random variables drawn from convenient probability distributions, the network renormalisation perspective brings in the notion of the *flow of the law of the weights* [9]. In particular, just like the connection probability of an inhomogeneous random graph model transforms under coarse-graining, also the law of the weights transforms, and is given by the convolution of the laws of the weights to produce the weights of aggregated vertices. The identification of a *joint*

fixed point, i.e., the combination of an invariant connection function along with an invariant law of the weights, therefore becomes relevant. If vertices are aggregated into groups of equal size, then an invariant law is necessarily either a delta-like distribution (with vertices having equal weights at any aggregation level) or a one-sided stable-law distribution (with vertices having weights distributed with asymptotic power-law decay).

The *former* fixed point corresponds to the homogeneous Erdős–Rényi random graph where, irrespectively of the form of the connection function, the value of the connection probability is the same for all pairs of vertices (the value itself flowing appropriately under renormalisation). Instead, the *latter* fixed point corresponds, in the regime of infinite-mean fitness, to an extremely inhomogeneous invariant model, also called the Multi Scale Model (MSM) [9], which is different from any other inhomogeneous random graph model with independent edges and finite-mean weights [14]. In particular, previous mathematical investigations of the MSM have highlighted that, under an appropriate scaling of a global parameter tuning the overall edge density, the degree distribution has a universal asymptotic inverse-square decay [16], the local clustering coefficient remains finite even in the sparse regime where the edge density vanishes [17], the fraction of isolated vertices and other non-self-averaging network properties converge to random variables [17], and the spectral properties of the adjacency matrix are quite exotic, with multiple largest eigenvalues scaling like the square root of the number of vertices and the corresponding eigenvector entries being characterised by log-periodicity and complex scaling exponents [18].

Attractors. Motivated by the above results, in the present paper we address for the first time the natural question whether the aforementioned invariant graph ensembles are also attractors of the renormalisation flow for generic random graph models. In particular, we ask whether, starting from arbitrary inhomogeneous random graphs with independent edges, random weights, and weight-dependent connection probabilities, the iteration of the greedy aggregation rule progressively transforms the original model into one of the known fixed points. Note that the coarse-grained network records the existence of a connection between any two aggregate vertices, not the identity of the microscopic connections between the constituent vertices themselves. It therefore naturally represents the original system at a lower resolution level. As illustrated in the remainder of this section and proven rigorously in the subsequent sections, we establish conditions (referred to as light-tailed and heavy-tailed regimes, respectively) under which a random graph model is attracted either to the Erdős–Rényi random graph or to the invariant model with one-sided stable-law weights and exponential disconnection probability. Although we are able to only partially identify the basins of attraction of the attractors, our results highlight that the iteration of a fundamental coarse-graining procedure (representing a change in the resolution at which random graphs are observed) can naturally bring the network close to one of two very different models, characterised by full homogeneity or extreme heterogeneity, respectively.

1.2 Model

The object of interest in this paper is $G = (V, E)$, the *inhomogeneous random graph* with vertex set $V = [n] = \{1, \dots, n\}$ and edge set $E \subset [n] \times [n]$ generated as follows [9], [5]. Let $X = \{X_i\}_{i \in [n]}$ be i.i.d. weights taking values in $\mathbb{R}_+ = [0, \infty)$ drawn from a common probability distribution. Choose functions

$$f^{\neq} \colon \mathbb{R}_+ \times \mathbb{R}_+ \to [0, 1), \qquad f^{=} \colon \mathbb{R}_+ \to [0, 1),$$

called *connection functions*. Connect the unordered pair of vertices $\{i, j\}$, $i \neq j$, with probability $f^{\neq}_{ij}[X] = f^{\neq}(X_i, X_j)$ independently for different pairs of vertices, and connect the vertex i with itself with probability $f^{=}_{i}[X] = f^{=}(X_i)$ independently for different vertices. Note that we allow for single edges (when $i \neq j$) and single self-loops (when $i = j$). We do *not allow for multiple edges or multiple self-loops*. Because edges are undirected, we require $f^{\neq}$ to be symmetric.

Note that G has *two types of randomness*, one coming from X associated with the vertices and one coming from the edges drawn according to $f^{\neq}, f^{=}$ *conditional* on X. Sometimes we write $G[X]$ to exhibit that G depends on X. Throughout the sequel, $\mathbb{E}_X$ denotes expectation with respect to X.

1.3 Renormalisation

Suppose that $n = m^L$ for some $m \in \mathbb{N} \setminus \{1\}$ and $L \in \mathbb{N}$. Partition the n vertices in G into disjoint groups of m vertices each, and *aggregate* the vertices in each group into a single vertex, referred to as a level-1 vertex. Construct an *aggregated graph* $G^{(1)}$ consisting of $n_1 \equiv n/m$ level-1 vertices, labelled by $[n/m]$, by drawing an aggregated edge between two level-1 vertices $I^{(1)}$ and $J^{(1)}$ *if and only if* there is a pair of vertices $i \in I^{(1)}$ and $j \in J^{(1)}$ in G that is connected by an edge in G. For $I^{(1)} = J^{(1)}$ we allow $i = j$, so that an aggregated self-loop at $I^{(1)}$ is drawn *if and only if* either an edge between a pair of distinct constituent vertices $i \neq j$ in $I^{(1)}$ is present or a self-loop at any constituent vertex $i \in I^{(1)}$ is present (= 'greedy connection algorithm'). Note that the greedy connection rule has a simple interpretation: two aggregate vertices are connected whenever at least one microscopic connection exists between their constituents. Thus, the coarse-grained edge records the existence of a connection between two aggregate vertices, but not the identity of the microscopic connections.

Renormalisation transformation. We view $G^{(1)}$ as a *level-1 renormalisation* of G, which is again an inhomogeneous random graph, on $n_1 = n/m$ level-1 vertices. Clearly, in $G^{(1)}$ the pair of level-1 vertices $\{I^{(1)}, J^{(1)}\}$ with $I^{(1)} \neq J^{(1)}$ is connected with probability

$$f^{\neq}_{I^{(1)},J^{(1)}}[X] = 1 - \prod_{i \in I^{(1)},\, j \in J^{(1)}} \big(1 - f^{\neq}(X_i, X_j)\big),$$

while the level-1 vertex $I^{(1)}$ is connected to itself by a single self-loop with probability

$$f^{=}_{I^{(1)},I^{(1)}}[X] = 1 - \prod_{\substack{i,j \in I^{(1)} \\ i<j}} \big(1 - f^{\neq}(X_i, X_j)\big) \prod_{i \in I^{(1)}} \big(1 - f^{=}(X_i)\big),$$

where the restriction $i < j$ is inserted in order to avoid double counting of the internal pairs. (Ordering of the vertices in $I^{(1)}$ can be done arbitrarily because of exchangeability in the product.) Clearly, in $G^{(1)}$ different pairs of level-1 vertices are connected independently. Note that $G^{(1)}$ is a *deterministic* function of G, obtained via the greedy connection algorithm. All the randomness (weights and connections) sits in G *alone*.

It is convenient to write

$$f^{\neq}(x, y) = 1 - \mathrm{e}^{-g^{\neq}(x,y)}, \qquad f^{=}(x) = 1 - \mathrm{e}^{-\frac{1}{2}[g^{\neq}(x,x) + g^{=}(x)]},$$

with $g^{\neq}\colon \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and $g^{=}\colon \mathbb{R}_+ \to \mathbb{R}_+$. With this change of variables the above relations can be *rewritten in the compact form*

$$g^{\neq}_{I^{(1)},J^{(1)}}[X] = \sum_{i\in I^{(1)},\, j\in J^{(1)}} g^{\neq}(X_i, X_j), \quad I^{(1)} \neq J^{(1)}, \qquad g^{=}_{I^{(1)}}[X] = \sum_{i\in I^{(1)}} g^{=}(X_i).$$

The latter shows that $g^{\neq}$ and $g^{=}$ renormalise *autonomously*, which is an important simplification.

Note that the renormalisation transformation is such that, even when we start from connection functions that depend on the weights of the vertices only, i.e., $g^{\neq}_{ij}[X] = g^{\neq}(X_i, X_j)$ and $g^{=}_{i}[X] = g^{=}(X_i)$ for level-0 vertices i, j, this property *does not carry over* to the aggregated connection functions. Indeed, $g^{\neq}_{I^{(1)},J^{(1)}}[X]$ and $g^{=}_{I^{(1)}}[X]$ in general depend on the two sets $X_{I^{(1)}} \equiv \{X_i\}_{i\in I^{(1)}}$ and $X_{J^{(1)}} \equiv \{X_j\}_{j\in J^{(1)}}$ of weights of the constituent vertices of the aggregated vertices $I^{(1)}$ and $J^{(1)}$. Thus, the transformation *does not preserve* the class of connection functions that depend on the scalar weights only, a situation that is *typical for renormalisation schemes*. One of our goals is to find out whether or not particular connection functions exist that remain inside the class, and whether these act as attractors for the iterated transformation. We will see that the answer is yes.

The following was proved in [9] in a simpler case ($\eta = 0$) and, for completeness, will be reproved in Section 3.1 in full generality.

Lemma 1.1. [Invariance property for the connection probability] *For every $\delta \in (0,\infty)$ and $\eta \in [0,\infty)$, the choice*

$$g^{\neq}_{\delta}(x,y) = \delta xy, \qquad g^{=}_{\eta}(x) = \eta x,$$

is invariant under addition of the weights in the aggregation. In other words, the renormalised graph has the same connection functions as the original graph, provided the weight of an aggregated vertex equals the sum of the weights of the constituent vertices, i.e., the aggregated vertex $I^{(1)}$ has weight $\bar{X}_{I^{(1)}} = \sum_{i\in I^{(1)}} X_i$.

In Section 3.1 we will also check the following invariance property.

Lemma 1.2. [Invariance property for positive stable weights] *Let $\gamma \in (0,1)$, and let $Y_1, \dots, Y_m$ be i.i.d. standard positive γ-stable random variables, with Laplace transform*

$$\mathbb{E}[\mathrm{e}^{-\lambda Y_1}] = \mathrm{e}^{-\lambda^\gamma}, \qquad \lambda \geq 0.$$

Then

$$m^{-1/\gamma} \sum_{i=1}^{m} Y_i \stackrel{d}{=} Y_1.$$

Consequently, for every $\delta \in (0,\infty)$ and $\eta \in [0,\infty)$, the inhomogeneous random graph with connection functions

$$g^{\neq}_{\delta}(x,y) = \delta xy, \qquad g^{=}_{\eta}(x) = \eta x,$$

and i.i.d. standard positive γ-stable weights is invariant in distribution under the renormalisation transformation, provided the aggregated weights are rescaled by $m^{-1/\gamma}$.

The restriction $\gamma \in (0,1)$ is natural because we work with positive one-sided stable weights. Stable laws with index $\gamma \in [1,2)$ do not give non-degenerate positive one-sided weights of this form.

Lemma 1.3. [Invariance property for deterministic weights] *For every $\delta \in (0,\infty)$, $\eta \in [0,\infty)$ and $w \in (0,\infty)$, the deterministic weight law concentrated at w is invariant under averaging of weights: if $W_1,\dots,W_m$ are deterministic random variables equal to w, then*

$$m^{-1}\sum_{i=1}^{m} W_i \stackrel{d}{=} W_1,$$

where in this case equivalence in distribution coincides with equality in value. Combined with the invariant connection functions $g^{\neq}_{\delta}(x,y) = \delta xy$ and $g^{=}_{\eta}(x,x) = \eta x$ identified in Lemma 1.1, this gives a two-parameter homogeneous Erdős–Rényi random graph with edge and self-loop connection probability equal to

$$p^{\neq} = 1 - \mathrm{e}^{-\delta w^2}, \qquad p^{=} = 1 - \mathrm{e}^{-\frac{1}{2}(\delta w^2 + \eta w)}.$$

The above random graph is invariant in distribution under the renormalisation transformation, provided the aggregated weights are rescaled by m^{-1}.

Thus, we have two different *two-parameter families* of connection functions and weights that are invariant under the renormalisation map when weights are *added* and *rescaled*. We will show that each acts as an *attractor* for iterations of the renormalisation map under appropriate scaling of the weights. As illustrated by the different expressions for the transformations of the random variables Y_i and W_i considered in Lemma 1.2 and 1.3, respectively, the required scaling of the weights is different for the two attractors.

Iteration of the renormalisation transformation. The renormalisation can be iterated, leading to a sequence of random graphs $\{G^{(\ell)}[X]\}_{0\le\ell\le L}$ in which vertices get more and more aggregated, at each step into groups of equal size m. The number of vertices at level ℓ is therefore $n_\ell = n/m^\ell$. Note that $G^{(\ell)}$ is a *deterministic* function of G, obtained by applying the greedy connection algorithm ℓ times. Again, all the randomness (weights and connections) sits in G *alone*.

Note that the iteration preserves the dependence of the aggregated connection functions on the full set of microscopic weights inside each aggregate vertex. Indeed, in $G^{(\ell)}$ the pair of level-ℓ vertices $\{I^{(\ell)}, J^{(\ell)}\}$ with $I^{(\ell)} \neq J^{(\ell)}$ is connected with probability

$$f^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = 1 - \prod_{i\in I^{(\ell)},\, j\in J^{(\ell)}} \big(1 - f^{\neq}(X_i,X_j)\big),$$

while the level-ℓ vertex $I^{(\ell)}$ is connected to itself by a single self-loop with probability

$$f^{=}_{I^{(\ell)},I^{(\ell)}}[X] = 1 - \prod_{\substack{i,j\in I^{(\ell)}\\ i<j}} \big(1 - f^{\neq}(X_i,X_j)\big) \prod_{i\in I^{(\ell)}} \big(1 - f^{=}(X_i)\big).$$

This is compactly written as

$$g^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = \sum_{i\in I^{(\ell)},\, j\in J^{(\ell)}} g(X_i,X_j), \qquad g^{=}_{I^{(\ell)}}[X] = \sum_{i\in I^{(\ell)}} g^{=}(X_i).$$

Since $n = m^L$, a total of L iterations is possible, after which all vertices have been aggregated into a *single vertex* at level L, i.e., $n_L = 1$. At every level ℓ with $0 < \ell \le L$ the

connection probabilities a priori depend on *all* the weights whose labels fall in the aggregated vertex sets, not just on the *sum* of these weights. Only for the special case where $g = g_\delta$ does this reduction occur. Nonetheless, we may expect that after many iterations some sort of *homogenisation* takes place, such that this reduction occurs at least *approximately*, in the spirit of abstract renormalisation theory. The goal of the current paper is to identify choices of the connection functions for which such a simplification occurs after appropriate scaling. We will see that light-tailed regimes and heavy-tailed regimes lead to different scalings and different scaling limits.

Although we might loosen the restriction of the renormalisation being done in groups of vertices of equal size, this setup is easy and maintains exchangeability. When passing to the limit $\ell \to \infty$, we need that also $L \to \infty$. In order to achieve this, we may fix an infinite sequence of weights $\{X_i\}_{i\in\mathbb{N}}$ beforehand and embed $\{X_i\}_{i\in[n]}$ into this sequence for every n. We define $N_{L-\ell} = n_\ell = m^{L-\ell}$ and think of the setting where $N_{L-\ell} \gg 1$, so that even after ℓ iterations the renormalised graph still consists of a large number of level-ℓ vertices. In what follows, n does not play much of a role. We only need it to be large enough. Our focus will be on what the renormalised random graph looks like for *large but finite* ℓ, even though our main theorems are phrased with $\ell \to \infty$ to get sharp and universal limiting results. Throughout the paper m is kept fixed.

1.4 Main theorems

We apply the renormalisation transformation ℓ times and investigate what happens when $\ell \to \infty$. In order to end up with a non-degenerate graph limit, we need to start from a graph in which the edges and the loops are *diluted* in an ℓ-dependent way. Here and below, by dilution we mean that the microscopic (= level-0) connection intensities are allowed to depend on the number ℓ of renormalisation steps and tend to zero as $\ell \to \infty$. This should not be confused with sparsity in the sense of bounded average degree. Rather, for each target scale ℓ we start from a microscopic graph whose connection function is weakened by an ℓ-dependent factor. More explicitly, if

$$f^{\neq}(x,y) = 1 - \mathrm{e}^{-g^{\neq}(x,y)}, \qquad f^{=}(x) = 1 - \mathrm{e}^{-\frac{1}{2}[g^{\neq}(x,x)+g^{=}(x)]},$$

then the ℓ-diluted version has rescaled connection functions

$$\mathfrak{f}_\ell^{\neq}(x,y) = 1 - \mathrm{e}^{-\mathfrak{g}_\ell^{\neq}(x,y)}, \qquad \mathfrak{f}_\ell^{=}(x) = 1 - \mathrm{e}^{-\frac{1}{2}[\mathfrak{g}_\ell^{\neq}(x,x)+\mathfrak{g}_\ell^{=}(x)]},$$

where

$$\mathfrak{g}_\ell^{\neq}(x,y) = a_\ell^{\neq} g^{\neq}(x,y), \qquad \mathfrak{g}_\ell^{=}(x) = a_\ell^{=} g^{=}(x),$$

for some appropriately chosen $a_\ell^{\neq} \downarrow 0$ and $a_\ell^{=} \downarrow 0$ as $\ell \to \infty$. For instance, for large ℓ and bounded x, y, we have

$$\mathfrak{f}_\ell^{\neq}(x,y) \approx a_\ell^{\neq} g^{\neq}(x,y),$$

and each microscopic (= level-0) edge is individually unlikely. The reason for this scaling is that a level-ℓ edge represents the union of $m^{2\ell}$ possible microscopic edges between two blocks of size m^ℓ. Without weakening the microscopic connection probabilities, the greedy renormalisation would make the level-ℓ graph complete in the limit as $\ell \to \infty$. The dilution is therefore the *scaling window* in which the renormalised graph has a non-trivial limit.

Light-tailed regime. Write $P(G^{(\ell)}[X])$ to denote the law of the *inhomogeneous random graph* $G^{(\ell)}[X]$ that is obtained after carrying out ℓ renormalisations of the graph with *rescaled connection functions* $\mathfrak{g}_\ell^{\neq}$ and $\mathfrak{g}_\ell^{=}$ given by

$$\mathfrak{g}_\ell^{\neq} = (m^{2\ell})^{-1}\, g^{\neq}, \qquad \mathfrak{g}_\ell^{=} = (m^{\ell})^{-1}\, g^{=},$$

conditional on the weights X (= 'quenched setting'). This law lives on the set of all graphs with $N_{L-\ell}$ vertices and no multiple edges or multiple loops. Write $P(G_{\ell,p})$ to denote the law of

- the *homogeneous Erdős–Rényi random graph* $G_{\ell,p}$ with $N_{L-\ell}$ vertices, edge probability $p^{\neq} \in (0,1)$, loop probability $p^{=} \in (0,1)$, and $p = (p^{\neq}, p^{=})$.

Let

$$\Delta\Big(P(G^{(\ell)}[X]), P(G_{\ell,p})\Big) = S^{\neq}\Big(P(G^{(\ell)}[X]) \;\Big|\; P(G_{\ell,p})\Big) + S^{=}\Big(P(G^{(\ell)}[X]) \;\Big|\; P(G_{\ell,p})\Big),$$

where $S^{\neq}$ and $S^{=}$ are the specific relative entropies for edges and self-loops defined in Section 2.2.

Theorem 1.4. [Averaged homogeneous attractor] *Suppose that*

$$(\square) \qquad \begin{aligned} &0 < \mathbb{E}_X\big[g^{\neq}(X_1, X_1')\big] < \infty,\\ &0 \le \mathbb{E}_X\big[g^{=}(X_1)\big] < \infty,\\ &k^{-2}\sum_{i=1}^{k} g^{\neq}(X_i, X_i) \xrightarrow[k\to\infty]{\mathbb{P}} 0, \end{aligned}$$

where X_1' is an independent copy of X_1 and $(X_i)_{i\in\mathbb{N}}$ are i.i.d. copies of X_1. Put

$$q^{\neq} = \mathbb{E}_X\big[g^{\neq}(X_1, X_1')\big], \qquad q^{=} = \mathbb{E}_X\big[g^{=}(X_1)\big],$$

and

$$p^{\neq} = 1 - \mathrm{e}^{-q^{\neq}}, \qquad p^{=} = 1 - \mathrm{e}^{-\frac{1}{2}(q^{\neq}+q^{=})}.$$

Then

$$\lim_{\ell\to\infty} \sup_{L\ge\ell} \mathbb{E}_X\left[\Delta\Big(P(G^{(\ell)}[X]), P(G_{\ell,p})\Big)\right] = 0$$

with $p = (p^{\neq}, p^{=})$.

Corollary 1.5. [Product kernel with finite first effective moment] *Suppose that, for some $B, \beta \in (0,\infty)$ and $C, K^{\neq}, K^{=} < \infty \in [0,\infty)$,*

$$\big|g^{\neq}(x,y) - B(xy)^{\beta}\big| \le K^{\neq}, \quad x,y \in \mathbb{R}_+, \qquad \big|g^{=}(x) - Cx^{\beta}\big| \le K^{=}, \quad x \in \mathbb{R}_+.$$

Suppose further that

$$0 < \mathbb{E}_X[X_1^{\beta}] < \infty.$$

Then Theorem 1.4 applies. In particular, if

$$\mathbb{P}(X_1 > x) \sim Ax^{-\alpha}, \qquad x \to \infty,$$

then the product connection function is attracted to the homogeneous Erdős–Rényi random graph whenever $\alpha > \beta$.

Heavy-tailed regime. Fix $\gamma \in (0,1)$. Write $P(G^{(\ell)})$ to denote the law of the random graph $G^{(\ell)}$ that is obtained after carrying out ℓ renormalisations of the graph with *rescaled connection functions* $\mathfrak{g}_\ell^{\neq}$ and $\mathfrak{g}_\ell^{=}$ given by

$$\mathfrak{g}_\ell^{\neq} = (m^{2\ell})^{-1/\gamma} g^{\neq}, \qquad \mathfrak{g}_\ell^{=} = (m^{\ell})^{-1/\gamma}\, g^{=},$$

unconditional on the weights X (= 'annealed setting'). Write $P(G_{\ell,\gamma,\delta,\eta})$ to denote the law of

- the *inhomogeneous stable law random graph* $G_{\ell,\gamma,\delta,\eta}$ with $N_{L-\ell}$ vertices, with weights given by i.i.d. standard γ-stable law random variables $\{Y_I\}_{I\in[N_{L-\ell}]}$, and with connection functions $g_\delta^{\neq}$ and $g_\eta^{=}$.

The standard γ-stable random variable Y_1 on $\mathbb{R}_+$ has Laplace transform $\mathbb{E}\big[\mathrm{e}^{-\lambda Y_1}\big] = \mathrm{e}^{-\lambda^\gamma}$, $\lambda \in \mathbb{R}_+$. It has tail probabilities $\mathbb{P}(Y_1 > x) \sim x^{-\gamma}/\Gamma(1-\gamma)$, $x \to \infty$, and $\mathbb{P}(Y_1 \le x) = o(1)\,\mathrm{e}^{-x^{-\gamma}}$, $x \downarrow 0$. See [19, Section XIII.6] for more details.

Theorem 1.6. [Averaged stable-law attractor] *Suppose that*

$$(\blacksquare) \qquad \begin{aligned} &\mathbb{P}(X_1 > x) \sim Ax^{-\alpha}, \quad x \to \infty, \\ &\big|g^{\neq}(x,y) - B(xy)^{\beta}\big| \le K^{\neq}, \quad x,y \in \mathbb{R}_+, \\ &\big|g^{=}(x) - Cx^{\beta}\big| \le K^{=}, \quad x \in \mathbb{R}_+, \end{aligned}$$

with $\alpha \in (0,\infty)$ a tail exponent, $\beta \in (0,\infty)$ a growth exponent, and $A,B,C,K^{\neq},K^{=} \in (0,\infty)$ constants. Suppose, in addition, that the law of X_1 is absolutely continuous. If $\alpha < \beta$, then there exists a coupling of (X,Y) via a sequence U of i.i.d. UNIF$[0,1]$ *random variables, written $(X(U),Y(U))$, such that*

$$\lim_{\ell\to\infty} \sup_{L\ge\ell} \mathbb{E}_U\Big[\Delta\Big(P(G^{(\ell)}[X(U)]), P(G_{\ell,\gamma,\delta,\eta}[Y(U)])\Big)\Big] = 0$$

with $\gamma = \frac{\alpha}{\beta}$, $\delta = B\big(A\Gamma(1-\gamma)\big)^{2/\gamma}$ and $\eta = C\big(A\Gamma(1-\gamma)\big)^{1/\gamma}$.

1.5 Discussion and open problems

1. To the best of our knowledge, the present paper is the first to provide a fully rigorous renormalisation analysis of a large class of inhomogeneous random graphs with random weights, both in a light-tailed regime and in a heavy-tailed regime. Universality classes and domains of attraction are identified in both regimes, exhibiting different scaling behaviour. We focus on a particular algorithm for the renormalisation, namely, the 'greedy connection algorithm', although the tools we use in our analysis are flexible (see items 9–10 below). From a physical perspective, identifying previously known *fixed points* of a renormalisation transformation as *attractors* of the associated flow is a powerful step forward, because it provides an explicit and natural mechanism (namely, the iterated change of resolution level) according to which effective aggregate models of networks arise, irrespectively of microscopic details that are irrelevant at the macroscopic scale.

2. Theorem 1.4 establishes universality for the renormalisation in the light-tailed regime: after renormalisation and appropriate scaling of the connection functions, $G^{(\ell)}[X]$ and $G_{\ell,p}$ are close to each other in the sense of specific relative entropy when $\ell \to \infty$, uniformly in $L \ge \ell$, provided

$$q^{\neq} = \mathbb{E}_X[g^{\neq}(X_1,X_1')], \qquad q^{=} = \mathbb{E}_X[g^{=}(X_1)],$$

and

$$p^{\neq} = 1 - \mathrm{e}^{-q^{\neq}}, \qquad p^{=} = 1 - \mathrm{e}^{-\frac{1}{2}(q^{\neq}+q^{=})}.$$

The term $q^{\neq}$ comes from level-0 edges inside a block, while the term $q^{=}$ comes from level-0 self-loops. The scaling limit is *homogeneous*, i.e., there are *no* weights involved, and the claim holds in expectation. Both Theorem 1.4 and Theorem 1.6 are averaged relative-entropy statements. In Theorem 1.4, the law of the target graph is deterministic, because the limiting Erdős–Rényi graph has deterministic edge and loop probabilities. Convergence to zero of the specific relative entropy means that around a *typical edge* and a *typical vertex* the probability laws are asymptotically the same. The domain of attraction given by (□) is large: it includes arbitrary g combined with bounded X_1, as well as g bounded by a bi-variate polynomial and X_1 having a finite moment of an appropriate order, as illustrated in Corollary 1.5.

3. Theorem 1.6 establishes the same in the heavy-tailed regime, but with a scaling that depends on the tail exponent of the distribution of X_1 and the growth exponent of $g^{\neq}$, $g^{=}$, and with a scaling limit that is *inhomogeneous*. The limit is invariant in the sense of Lemma 1.1 and has stable law random weights with exponent γ. In Theorem 1.6, the target graph law is itself random through the stable weights. The coupling in Theorem 1.6 is used to realise the original and limiting weight families on a common probability space. For product kernels $g^{\neq}(x,y) = B(xy)^{\beta} + O(1)$ and $g^{=}(x) = Cx^{\beta} + O(1)$, the case $\alpha > \beta$ is covered by Theorem 1.4, as noted in Corollary 1.5, while the case $\alpha < \beta$ is covered by Theorem 1.6. Thus, the borderline between the homogeneous and the stable-law attractor is governed by the finiteness or infiniteness of the first effective moment $\mathbb{E}[X_1^{\beta}]$. The domain of attraction given by (■) is fair, but far from exhaustive. It remains open to obtain a full classification in the heavy-tailed regime. For instance, what happens when $g^{\neq}(x,y) = B(x+y)^{\beta_1} + O(1)$, $x, y \to \infty$, and $g^{=}(x) = Cx^{\beta_2} + O(1)$, $x \to \infty$?

4. Previous research has shown that several random graph models defined by the same (random) vertex weights but different connection functions are either asymptotically *equivalent* [20] or mutually *contiguous* [21] under conditions that ultimately require the finiteness of certain moments of the weights. For instance, the three models defined by the connection probabilities

$$f_1^{\neq}(x,y) = \kappa(x,y) \wedge 1, \qquad f_2^{\neq}(x,y) = \frac{\kappa(x,y)}{1+\kappa(x,y)}, \qquad f_3^{\neq}(x,y) = 1 - \mathrm{e}^{-\kappa(x,y)},$$

for the same kernel $\kappa(x,y)$, are asymptotically equivalent or mutually contiguous when certain conditions are met [20, 21]. In our setting, the corresponding connection functions are

$$g_1^{\neq}(x,y) = -\log[1 - \kappa(x,y) \wedge 1], \qquad g_2^{\neq}(x,y) = \log[1+\kappa(x,y)], \qquad g_3^{\neq}(x,y) = \kappa(x,y),$$

with the convention $-\log 0 = \infty$. Thus the relevant object for the renormalisation is not the connection probability f itself, but the connection function $g^{\neq} = -\log(1 - f^{\neq})$.

- In the light-tailed regime, if the corresponding intensities satisfy (□), then Theorem 1.4 applies to each of these models and gives a homogeneous Erdős–Rényi attractor, with parameter $p^{\neq}$ determined by the corresponding value of

  $$\mathbb{E}_X[g_i^{\neq}(X_1, X_1')].$$

  For the model defined by $f_1^{\neq}$, this requires in particular that $\kappa(x,y) < 1$ on the relevant support, or at least that $-\log[1 - \kappa(X_1, X_1') \wedge 1]$ be integrable. Otherwise $g_1^{\neq}$ may be infinite with positive probability and Theorem 1.4 does not apply.

- In the heavy-tailed regime, the moment assumptions that guarantee asymptotic equivalence or mutual contiguity may fail, and models that are comparable at level-0 may lie in different domains of attraction under the renormalisation map. For instance, take

$$\kappa(x,y) = B(xy)^\beta, \qquad B, \beta \in (0,\infty),$$

and assume that $\mathbb{P}(X_1 > x) \sim Ax^{-\alpha}$ as $x \to \infty$. For the model $f_2^{\neq}$, the associated connection function is $g_2^{\neq}(x,y) = \log\bigl(1+B(xy)^\beta\bigr)$, which satisfies the integrability condition in ($\square$). Indeed,

$$\log\bigl(1 + B(X_1X_1')^\beta\bigr) \le C + \beta\log(1+X_1) + \beta\log(1+X_1')$$

for a constant $C < \infty$, and $\mathbb{E}_X[\log(1+X_1)] < \infty$ for regularly varying X_1 with tail exponent $\alpha > 0$. Similarly,

$$\mathbb{E}_X\bigl[\log(1 + BX_1^{2\beta})\bigr] < \infty.$$

Hence $f_2^{\neq}$ falls within the light-tailed theorem (when $\mathbb{E}_X[g_2^{=}(X)] < \infty$) and flows to a homogeneous Erdős–Rényi attractor. For the model $f_3^{\neq}$, however, the associated connection function is

$$g_3^{\neq}(x,y) = B(xy)^\beta.$$

When $\alpha < \beta$, the random variable X_1^β has tail exponent $\gamma = \alpha/\beta \in (0,1)$, and the hypotheses (■) of Theorem 1.6 are satisfied (when the condition for $g_3^{=}$ is in force). Therefore $f_3^{\neq}$ is attracted to the inhomogeneous γ-stable attractor. This shows that $f_2^{\neq}$ and $f_3^{\neq}$, although built from the same kernel κ, need not have the same attractor in the heavy-tailed regime.

- In the particular case where $\kappa(x,y) = \delta xy$, the three models in item 4 have special meanings:
  - $f_1^{\neq}(x,y) = \delta xy \wedge 1$ is the Chung–Lu model [22],
  - $f_2^{\neq}(x,y) = \frac{\delta xy}{1+\delta xy}$ is the canonical (or maximum-entropy) configuration model [23],
  - $f_3^{\neq}(x,y) = 1 - \mathrm{e}^{-\delta xy}$ is the invariant model in Lemma 1.1.

  Thus, in the light-tailed regime, these models may belong to the same homogeneous universality class, whereas in the heavy-tailed regime they may be separated into different domains of attraction.

5. Another notion of equivalence between different models has been established in the context of Exponential Random Graphs (ERGs) [24], which are defined as random graph ensembles with a probability distribution that maximises Shannon entropy [23] while enforcing a desired set of constraints in terms of expected network properties. In particular, it has been shown [25, 26] that, in the dense regime, various ERGs defined through different constraints are asymptotically indistinguishable from the Erdős–Rényi model (which is itself an ERG, defined through a given expected link density). Those results indicate that it is hard to perturb the homogeneous Erdős–Rényi model in the dense regime via the addition of other constraints. Under the different lens of renormalisation, our results show that: on one hand, the dense Erdős–Rényi model is indeed a homogeneous attractor for a large class of microscopic models

(Theorem 1.4); on the other hand, a different dense graph, clearly distinguishable from the Erdős–Rényi one, exists and acts as an inhomogeneous attractor for a separate class of models (Theorem 1.6).

6. The scaling of the connection functions has a simple probabilistic interpretation. In the light-tailed regime, no single microscopic (= level-0) vertex dominates a block. Many small contributions average out, and the coarse-grained graph is effectively homogeneous. This corresponds to a law of large numbers universality class: microscopic heterogeneity is washed out under aggregation. In the heavy-tailed regime, however, a few large weights may dominate a block. Aggregation, even under the greedy rule, does not erase this heterogeneity. Instead, the effective block weights converge to a stable law, and the coarse-grained graph remains inhomogeneous. The word *dilute* should be understood in this sense. It refers to weakening of individual microscopic connection probabilities, and not to sparsity in the sense of bounded average degree. In the light-tailed regime, a typical microscopic connection probability is of order $m^{-2\ell}$. Since the original graph has $n = m^L$ vertices, the microscopic expected degree is typically of order $nm^{-2\ell} = m^{L-2\ell}$, up to constants depending on the connection function and the weight distribution. Hence the microscopic graph has bounded average degree only in the regime $L - 2\ell = O(1)$. In general, the microscopic graph is diluted because individual edge probabilities are small, but it need not be sparse. A concrete example is obtained by taking, in the light-tailed regime, the rescaled invariant edge probability

$$\mathfrak{f}_\ell^{\neq}(x, y) = 1 - \exp\left(-\frac{\delta xy}{m^{2\ell}}\right).$$

Then each pair of vertices has only a small probability of being connected. However, between two level-ℓ blocks $I^{(\ell)}, J^{(\ell)}$ the effective connection intensity is

$$\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}) = \frac{\delta}{m^{2\ell}} \sum_{i \in I^{(\ell)}, j \in J^{(\ell)}} X_i X_j = \delta \left(\frac{1}{m^\ell} \sum_{i \in I^{(\ell)}} X_i\right) \left(\frac{1}{m^\ell} \sum_{j \in J^{(\ell)}} X_j\right).$$

If X_1 has finite mean, then this product converges to $\delta(\mathbb{E}[X_1])^2$, and the coarse-grained graph becomes a homogeneous Erdős–Rényi random graph with edge probability

$$1 - \exp(-\delta(\mathbb{E}[X_1])^2).$$

This example shows that the dilution is not merely a technical device: it is the scale at which many weak microscopic interactions combine into a macroscopic interaction of order one.

7. The scaling of the connection functions $g^{\neq}, g^{=}$ to the connection functions $\mathfrak{g}_\ell^{\neq}, \mathfrak{g}_\ell^{=}$ is important in order to end up with a *non-trivial graph limit*. Indeed, under the aggregation, weights 'accumulate' and the graph becomes 'denser'. By starting with a dilute graph where the edges have a density of order $m^{-2\ell}$, respectively, $(m^{-2\ell})^{1/\gamma}$, and the loops have a density of order $m^{-\ell}$, respectively, $(m^{-\ell})^{1/\gamma}$, the initial graph is *pre-engineered* so that after ℓ renormalisations it acquires an edge density and a loop density of order 1. Consequently, the whole renormalisation flow acts in the space of *dense graphs*, from dilute to non-dilute. Without the rescaling, the graph limit would be the complete graph, which is not interesting.

8. Relative entropy satisfies the contraction property: $S(TP \mid TQ) \leq S(P \mid Q)$ for any two probability measures P, Q to which an aggregation map T is applied, i.e., a surjection [27]. It therefore follows that $\ell \mapsto S(P(G^{(\ell)}[X]) \mid P(\bar{G}^{(\ell)}[\bar{X}]))$ is non-increasing for any two starting

graphs $G[X], \bar{G}[\bar{X}]$. Indeed, for fixed weights X, the map sending $G^{(\ell-1)}[X]$ to $G^{(\ell)}[X]$ is a deterministic map on graph configurations. Hence the law $P(G^{(\ell)}[X])$ is the push-forward of $P(G^{(\ell-1)}[X])$ under this map. Therefore relative entropy is a natural quantity to analyse under renormalisation.

9. It is interesting to pursue other types of aggregation as well. For instance, we may connect two level-1 vertices if and only if *every vertex in one group is connected to every vertex in the other group* (= 'lazy connection algorithm'). In that case the connection probabilities renormalise as

$$f^{\neq}_{I^{(1)},J^{(1)}}[X] = \prod_{i\in I^{(1)},\, j\in J^{(1)}} f^{\neq}(X_i,X_j), \qquad I^{(1)}\neq J^{(1)},$$
$$f^{=}_{I^{(1)},I^{(1)}}[X] = \prod_{\substack{i,j\in I^{(1)}\\ i<j}} f^{\neq}(X_i,X_j) \prod_{i\in I^{(1)}} f^{=}(X_i),$$

The proper change of variables is

$$f^{\neq}(x,y) = \mathrm{e}^{-g^{\neq}(x,y)}, \qquad f^{=}(x) = \mathrm{e}^{-\frac{1}{2}[g^{\neq}(x,x)+g^{=}(x)]}.$$

Iteration of the renormalisation gives

$$g^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = \sum_{i\in I^{(\ell)},\, j\in J^{(\ell)}} g^{\neq}(X_i,X_j), \qquad g^{=}_{I^{(\ell)},I^{(\ell)}}[X] = \sum_{i\in I^{(\ell)}} g^{=}(X_i).$$

The invariant connection functions are

$$g^{\neq}_{\delta}(x,y) = \delta xy, \qquad g^{=}_{\eta}(x) = \eta x.$$

In the present paper we will not pursue such generalisations, even though the arguments developed in Sections 2–3 can in principle be carried over.

10. Another example, not based on an aggregation algorithm, is where we connect two level-1 vertices with a probability that equals the *maximum of the connection probabilities of the constituent pairs* in the two groups. In that case the connection probabilities renormalise as

$$f^{\neq}_{I^{(1)},J^{(1)}}[X] = \max_{i\in I^{(1)},\, j\in J^{(1)}} f^{\neq}(X_i,X_j), \qquad I^{(1)}\neq J^{(1)},$$
$$f^{=}_{I^{(1)}}[X] = \max_{i\in I^{(1)}} f^{=}(X_i).$$

Iteration of the renormalisation gives

$$f^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = \max_{i\in I^{(\ell)},\, j\in J^{(\ell)}} f^{\neq}(X_i,X_j), \qquad I^{(\ell)}\neq J^{(\ell)},$$
$$f^{=}_{I^{(\ell)}}[X] = \max_{i\in I^{(1)}} f^{=}(X_i).$$

If $f^{\neq}, f^{=}$ are non-decreasing in both arguments, then

$$\max_{i\in I^{(\ell)},\, j\in J^{(\ell)}} f^{\neq}(X_i,X_j) = f^{\neq}(\tilde{X}_{I^{(\ell)}},\tilde{X}_{J^{(\ell)}}), \qquad I^{(\ell)}\neq J^{(\ell)},$$
$$\max_{i\in I^{(\ell)}} f^{=}(X_i) = f^{=}(\tilde{X}_{I^{(\ell)}}),$$

with $\tilde{X}_{I^{(\ell)}} = \max_{i \in I^{(\ell)}} X_i$. In other words, the aggregation no longer sums the weights, but maximises the weights.

For heavy-tailed weights the difference between sums and maxima of constituent weights should be small, so that different models should exhibit similar scaling behaviour. More in general, it is interesting to compare different renormalisation maps applied to the same starting graph.

11. What happens when the weights are dependent? A natural extension is to assume that the infinite sequence $(X_i)_{i \geq 1}$ is exchangeable, in the sense that its joint law is invariant under finite permutations. By de Finetti's theorem, this amounts to weights that are conditionally i.i.d. Their dependence may affect the attractors of the renormalisation. Again we may ask what happens for multiscale stochastic block models, configuration-type models, or weights arising from more general infinitely divisible laws.

1.6 Outline of the remainder

Section 2 analyses the renormalisation transformation by identifying its iterates and establishing bounds on the relative entropy of the laws of the renormalised connection functions. Section 3 provides the proofs of the invariance lemma and the two scaling theorems.

2 Renormalisation transformation

In Section 2.1 we identify how the renormalisation transformation iterates. In Section 2.2 we estimate specific relative entropies. In Section 2.3 we estimate perturbations of specific relative entropies. The results will be used in Section 3 to prove our main theorems. Appendix A collects a few basic facts about stable laws, domains of attraction and coupling.

2.1 Iteration

The renormalisation transformation defined in Section 1.3 can be iterated L times when $n = m^L$. After L iterations the renormalised graph $G^{(L)}$ consists of a *single vertex*. For $0 < \ell \leq L$, aggregate level-$(\ell - 1)$ vertices into disjoint groups of size m each, referred to as level-(ℓ) vertices, to obtain an aggregated graph $G^{(\ell)}$ with connection function $f^{(\ell)}$ obtained iteratively as

$$f^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = 1 - \prod_{I^{(\ell-1)} \in I^{(\ell)},\, J^{(\ell-1)} \in J^{(\ell)}} \big(1 - f^{\neq}_{I^{(\ell-1)},J^{(\ell-1)}}[X]\big), \quad I^{(\ell)} \neq J^{(\ell)},$$

$$f^{=}_{I^{(\ell)},I^{(\ell)}}[X] = 1 - \prod_{\substack{I^{(\ell-1)},J^{(\ell-1)} \in I^{(\ell)} \\ I^{(\ell-1)} < J^{(\ell-1)}}} \big(1 - f^{\neq}_{I^{(\ell-1)},J^{(\ell-1)}}[X]\big) \prod_{I^{(\ell-1)} \in I^{(\ell)}} \big(1 - f^{=}_{I^{(\ell-1)}}[X]\big),$$

which after the change of variables reads

$$g^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = \sum_{I^{(\ell-1)} \in I^{(\ell)},\, J^{(\ell-1)} \in J^{(\ell)}} g^{\neq}_{I^{(\ell-1)},J^{(\ell-1)}}[X],$$

$$g^{=}_{I^{(\ell)}}[X] = \sum_{I^{(\ell-1)} \in I^{(\ell)}} g^{=}_{I^{(\ell-1)}}[X].$$

When the right-hand sides of the above expressions are recursively expanded down to the microscopic level (= level-0) where $I^{(0)} = i$ and $J^{(0)} = j$, we obtain the iterated transformation equations shown in Section 1.3.

As already noted, while the connection functions $g^{\neq}_{i,j}[X] = g^{\neq}(X_i, X_j)$ at level $\ell = 0$ are assumed to depend on X only through the two vertex weights X_i and X_j, this is in general not true for levels $\ell > 0$, where we can instead write

$$g^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = g^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}), \quad X_{I^{(\ell)}} \equiv \{X_i\}_{i\in I^{(\ell)}}$$

which depends on the fulls sets $X_{I^{(\ell)}}$, $X_{J^{(\ell)}}$ of microscopic weights of the constituent vertices of the aggregate vertices $I^{(\ell)}$ and $J^{(\ell)}$.

For the special case $g^{\neq} = g^{\neq}_{\delta}$, $g^{=} = g^{=}_{\eta}$, iteration gives

$$g^{\neq}_{\delta}(X_{I^{(\ell)}}, X_{J^{(\ell)}}) = \delta \bar{X}_{I^{(\ell)}} \bar{X}_{J^{(\ell)}}, \qquad g^{=}_{\eta}(X_{I^{(\ell)}}) = \eta \bar{X}_{I^{(\ell)}},$$

with $\bar{X}_{I^{(\ell)}} = \sum_{i\in I^{(\ell)}} X_i$. For this invariant case, the connection functions $g^{\neq}_{I^{(\ell)},J^{(\ell)}}[X]$, $g^{=}_{I^{(\ell)}}[X]$ therefore still depend only on the two (renormalised) weights $\bar{X}_{I^{(\ell)}}$, $\bar{X}_{J^{(\ell)}}$ of the end-point aggregated vertices $I^{(\ell)}$ and $J^{(\ell)}$.

2.2 Relative entropy

To measure how close a graph measure $P_f(G)$ is to another graph measure $P_{f'}(G)$ on the same set $\mathcal{G}_N$ of graphs with N vertices, we consider the relative entropy

$$H(P_f \mid P_{f'}) = \sum_{G\in\mathcal{G}_N} P_f(G) \log \frac{P_f(G)}{P_{f'}(G)}$$

For a random graph with $\binom{N}{2}$ independent edges and N independent self-loops, the graph measure factorises into $\binom{N}{2} + N = \binom{N+1}{2}$ Bernoulli variables. Consequently, the relative entropy separates additively into such variables as

$$H(P_f \mid P_{f'}) = \sum_{1\le i\le j\le N} \left[f_{ij} \log\left(\frac{f_{ij}}{f'_{ij}}\right) + (1-f_{ij}) \log\left(\frac{1-f_{ij}}{1-f'_{ij}}\right)\right]$$

(where f_{ij} and f'_{ij} are the connection probabilities between i and j under P_f and $P_{f'}$ respectively). The above expression takes into account both edges between distinct pairs of vertices ($i < j$) and self-loops at individual vertices ($i = j$). To characterise asymptotic convergence to the attractor(s), we could consider the *specific relative entropy* (= relative entropy density)

$$S(P_f \mid P_{f'}) = \binom{N+1}{2}^{-1} \sum_{1\le i\le j\le N} \left[f_{ij} \log\left(\frac{f_{ij}}{f'_{ij}}\right) + (1-f_{ij}) \log\left(\frac{1-f_{ij}}{1-f'_{ij}}\right)\right].$$

Although this would be sufficient to characterise convergence, we note that the self-loop contribution, consisting of N terms in the sum, would be overruled by the overall normalising factor. Therefore, to gain proper control on the convergence of the laws of the edges and of the self-loops separately, we first define two separately normalised relative entropies $S^{\neq}(P_f \mid P_{f'})$ and $S^{=}(P_f \mid P_{f'})$ for edges and self-loops respectively, i.e.,

$$S^{\neq}(P_f \mid P_{f'}) = \binom{N}{2}^{-1} \sum_{1\le i<j\le N} \left[f_{ij} \log\left(\frac{f_{ij}}{f'_{ij}}\right) + (1-f_{ij}) \log\left(\frac{1-f_{ij}}{1-f'_{ij}}\right)\right],$$

$$S^{=}(P_f \mid P_{f'}) = N^{-1} \sum_{1\le i\le N} \left[f_{ii} \log\left(\frac{f_{ii}}{f'_{ii}}\right) + (1-f_{ii}) \log\left(\frac{1-f_{ii}}{1-f'_{ii}}\right)\right],$$

and afterwards combine them into the single quantity

$$\Delta(P_f, P_{f'}) = S^{\neq}(P_f \mid P_{f'}) + S^{=}(P_f \mid P_{f'}).$$

Note that, while $\Delta(P_f, P_{f'})$ is not a relative entropy *per se*, its vanishing is sufficient to ensure the separate convergence of the law of the edges and the law of the self-loops encoded in P_f to the corresponding laws encoded in $P_{f'}$.

In terms of $g_{ij} = -\log(1 - f_{ij})$ and $g'_{ij} = -\log(1 - f'_{ij})$ we have the following estimate. In what follows we use the following bounds, valid for $U, V \in (0, 1)$:

$$(*) \qquad (1-U)\log\left(\frac{1-U}{1-V}\right) = (1-U)\log\left(1 + \frac{V-U}{1-V}\right) \begin{cases} \leq (1-U)\,\frac{V-U}{1-V}, \\ \geq V - U. \end{cases}$$

Lemma 2.1. *$0 \leq H^{\neq}(P_f \mid P_{f'}) \leq d^{\neq}(g, g')$ and $0 \leq H^{=}(P_f \mid P_{f'}) \leq d^{=}(g, g')$, where*

$$d^{\neq}(g, g') = \binom{N}{2}^{-1} \sum_{1 \leq i < j \leq N} \zeta(g_{ij}, g'_{ij}), \qquad d^{=}(g, g') = N^{-1} \sum_{1 \leq i \leq N} \zeta(g_{ii}, g'_{ii}),$$

with

$$(\heartsuit_1) \qquad \zeta(a, b) = [1 \wedge (a-b)]^2\, \frac{1}{b}\, 1_{\{a>b\}} + \mathrm{e}^{-a}(b-a)\, 1_{\{b>a\}}, \qquad a, b \in \mathbb{R}_+.$$

Proof. The lower bounds are trivial by Jensen's inequality and convexity of the function $x \mapsto x \log x$. To get the upper bounds, define

$$\chi = f \log\left(\frac{f}{f'}\right) + (1-f)\log\left(\frac{1-f}{1-f'}\right) = (1 - \mathrm{e}^{-g})\log\frac{(1-\mathrm{e}^{-g})}{(1-\mathrm{e}^{-g'})} + \mathrm{e}^{-g}(g' - g),$$

where the indices ij and ii are suppressed to simplify the notation. Note that the two terms in the right-hand side have opposite signs. Abbreviate $U = \mathrm{e}^{-g}$ and $V = \mathrm{e}^{-g'}$. For $g > g'$, estimate $V - U = V(1 - \frac{U}{V}) = V(1 - \mathrm{e}^{-(g-g')}) \leq V[1 \wedge (g - g')]$, to get from $(*)$ that

$$\chi \leq [1 \wedge (g-g')] \left[\frac{1-U}{1-V}\, V - U\right] = [1 \wedge (g-g')]\, \frac{V-U}{1-V} \leq [1 \wedge (g-g')]^2\, \frac{V}{1-V}.$$

Since $\log(\frac{1}{V}) = \log(1 + \frac{1-V}{V}) \leq \frac{1-V}{V}$, it follows that $\frac{V}{1-V} \leq \frac{1}{g'}$, and so $\chi \leq [1 \wedge (g-g')]^2\, \frac{1}{g'}$. For $g < g'$, use $(*)$ to estimate

$$\begin{aligned} \chi &\leq \frac{1-U}{1-V}\,(V-U) + \mathrm{e}^{-g}(g'-g) = -\frac{1-\mathrm{e}^{-g}}{1-\mathrm{e}^{-g'}}\,(\mathrm{e}^{-g} - \mathrm{e}^{-g'}) + \mathrm{e}^{-g}(g'-g) \\ &= \mathrm{e}^{-g}\left[-\frac{1-\mathrm{e}^{-g}}{1-\mathrm{e}^{-g'}}\,(1 - \mathrm{e}^{-(g'-g)}) + (g'-g)\right]. \end{aligned}$$

The term in square brackets is bounded above by $g' - g$, because the first term is non-positive. Hence

$$\chi \leq \mathrm{e}^{-g}(g' - g),$$

which is the second term in $\zeta(g, g')$. □

2.3 Perturbation of relative entropy

Lemma 2.2. $|H^{\neq}(P_f \mid P_{f'}) - H^{\neq}(P_{\tilde f} \mid P_{f'})| \leq d^{\neq}(g,\tilde g,g')$ *and* $|H^{=}(P_f \mid P_{f'}) - H^{=}(P_{\tilde f} \mid P_{f'})| \leq d^{=}(g,\tilde g,g')$, *where*

$$d^{\neq}(g,\tilde g,g') = \binom{N}{2}^{-1} \sum_{1\leq i<j\leq N} \zeta(g_{ij},\tilde g_{ij},g'_{ij}),$$
$$d^{=}(g,\tilde g,g') = N^{-1} \sum_{1\leq i\leq N} \zeta(g_{ii},\tilde g_{ii},g'_{ii}),$$

with

$$(\heartsuit_2) \qquad \zeta(a,b,c) = \frac{1}{c}\Big(\mathrm{e}^{-(a\wedge c)}\,[1\wedge|a-c|] \wedge \mathrm{e}^{-(b\wedge c)}\,[1\wedge|b-c|]\Big) + \mathrm{e}^{-(a\wedge b)}\Big([1\wedge|a-b|]+|a-c|+|b-c|\Big), \qquad a,b,c\in\mathbb{R}_+.$$

Proof. Again we suppress the indices ij and ii. Define

$$\tilde\chi = \tilde f \log\left(\frac{\tilde f}{f'}\right) + (1-\tilde f)\log\left(\frac{1-\tilde f}{1-f'}\right) = (1-\mathrm{e}^{-\tilde g})\log\frac{(1-\mathrm{e}^{-\tilde g})}{(1-\mathrm{e}^{-g'})} + \mathrm{e}^{-\tilde g}(g'-\tilde g).$$

Abbreviate $\tilde U = \mathrm{e}^{-\tilde g}$. Use $(*)$ to estimate

$$(V-U) - (1-\tilde U)\,\frac{V-\tilde U}{1-V} \leq \chi - \tilde\chi - [U(g'-g) - \tilde U(g'-\tilde g)] \leq (1-U)\,\frac{V-U}{1-V} - (V-\tilde U).$$

The upper bound equals $\frac{(V-U)^2}{1-V} + (\tilde U - U)$, the lower bound equals $-\frac{(V-\tilde U)^2}{1-V} + (\tilde U - U)$. Hence

$$|\chi-\tilde\chi| \leq \frac{(V-U)^2 \wedge (V-\tilde U)^2}{1-V} + |U-\tilde U| + |U(g'-g) - \tilde U(g'-\tilde g)|.$$

Since $\frac{V}{1-V} \leq \frac{1}{g'}$, the first term is bounded from above by $\frac{1}{g'}[|V-U| \wedge |V-\tilde U|]$, where $|V-U| \leq (U\vee V)\,[1\wedge|g-g'|]$ and $|V-\tilde U| \leq (\tilde U\vee V)\,[1\wedge|\tilde g-g'|]$. Moreover, the second term is bounded from above by $(U\vee\tilde U)\,[1\wedge|g-\tilde g|]$ and the third term by $(U\vee\tilde U)[|g-g'|+|\tilde g-g'|]$. □

3 Proof of the invariance lemmas and the scaling theorems

In this section we prove Lemmas 1.1–1.2 (Section 3.1), Theorem 1.4 (Section 3.2) and Theorem 1.6 (Section 3.4).

3.1 Proof of the invariance lemmas

Proof of Lemma 1.1. For the case $g^{\neq} = g^{\neq}_\delta$ and $g^{=} \equiv 0$, we have for $I^{(1)} \neq J^{(1)}$,

$$g^{\neq}_{I^{(1)},J^{(1)}}[X] = \sum_{i\in I^{(1)},\, j\in J^{(1)}} g^{\neq}_\delta(X_i,X_j) = \sum_{i\in I^{(1)},\, j\in J^{(1)}} \delta X_i X_j = \delta \bar X_{I^{(1)}} \bar X_{J^{(1)}} = g_\delta(\bar X_{I^{(1)}}, \bar X_{J^{(1)}})$$

with $\bar X_{I^{(1)}} = \sum_{i\in I^{(1)}} X_i$ and $\bar X_{J^{(1)}} = \sum_{i\in J^{(1)}} X_i$, and

$$g^{=}_{I^{(1)}}[X] = 2\sum_{\substack{i,\,j\in I^{(1)}\\ i<j}} \delta X_i X_j + \sum_{i\in I^{(1)}} \delta X_i^2 = \sum_{i,\,j\in I^{(1)}} \delta X_i X_j = \delta \bar X^2_{I^{(1)}} = g_\delta(\bar X_{I^{(1)}}),$$

where we use that $\bar{X}^2_{I^{(1)}} = (\sum_{i\in I^{(1)}} X_i)^2 = \sum_{i\in I^{(1)}} X_i^2 + 2\sum_{i,\,j\in I^{(1)}, i<j} X_iX_j$. This settles the claim in Lemma 1.1 for $\eta = 0$. Similarly, for the case $g^{=} = g^{=}_{\eta}$,

$$g^{=}_{I^{(1)}}[X] = \sum_{i\in I^{(1)}} g^{=}_{\eta}(X_i) = \sum_{i\in I^{(1)}} \eta X_i = \eta \bar{X}_{I^{(1)}} = g_{\eta}(\bar{X}_{I^{(1)}}) \qquad \forall I^{(1)},$$

settling the claim for $\eta > 0$ as well. □

Proof of Lemma 1.2. For $\lambda \geq 0$,

$$\mathbb{E}\left[\exp\left\{-\lambda m^{-1/\gamma}\sum_{i=1}^{m} Y_i\right\}\right] = \prod_{i=1}^{m} \mathbb{E}\left[\mathrm{e}^{-\lambda m^{-1/\gamma} Y_i}\right] = \mathrm{e}^{-m(\lambda m^{-1/\gamma})^{\gamma}} = \mathrm{e}^{-\lambda^{\gamma}}.$$

Hence $m^{-1/\gamma}\sum_{i=1}^{m} Y_i$ has the same law as Y_1. The graph-level invariance follows by applying Lemma 1.1: after aggregation, the rescaled block weight has again the same stable law, while the invariant connection functions $g^{\neq}_{\delta}(x,y) = \delta xy$ and $g^{=}_{\eta}(x) = \eta x$ keep the same form. This settles the claim in Lemma 1.2. □

Proof of Lemma 1.3. Let all the weights be equal to w. Under aggregation into blocks of size m, the summed weight of a block is mw, and after the deterministic rescaling by m^{-1} the block weight is again w. For the invariant connection functions

$$g^{\neq}_{\delta}(x,y) = \delta xy, \qquad g^{=}_{\eta}(x) = \eta x,$$

the corresponding edge and loop densities at weight w are

$$q^{\neq} = \delta w^2, \qquad q^{=} = \eta w.$$

Hence the associated homogeneous graph has edge probability $p^{\neq} = 1 - \mathrm{e}^{-q^{\neq}}$ and self-loop probability $p^{=} = 1 - \mathrm{e}^{-\frac{1}{2}(q^{\neq}+q^{=})}$. Thus, the law of the deterministic weights is invariant under rescaling by m^{-1}, and the law of the resulting random graph belongs to the two-parameter family of Erdős–Rényi random graphs. This settles the claim in Lemma 1.3. □

Note that, for $w = 1$, the proof of Lemma 1.3 can be seen as a particular case of the proof of Lemma 1.2 with $\gamma = 1$, because the Laplace transform of a deterministic variable W_i equal to $w = 1$ is $\mathbb{E}[\mathrm{e}^{-\lambda W_i}] = \mathrm{e}^{-\lambda}$.

3.2 Proof of the light-tailed scaling theorem

Proof. The iteration formula in Section 2.1 tells us that

$$g^{\neq}_{I^{(\ell)},J^{(\ell)}}[X] = g^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}) = \sum_{i\in I^{(\ell)},\, j\in J^{(\ell)}} g^{\neq}(X_i, X_j), \qquad I^{(\ell)} \neq J^{(\ell)},$$

$$g^{=}_{I^{(\ell)}}[X] = g^{=}(X_{I^{(\ell)}}) = \sum_{i\in I^{(\ell)}} g^{=}(X_i).$$

In the light-tailed regime we use the rescaled connection functions

$$\mathfrak{g}^{\neq}_{\ell} = m^{-2\ell} g^{\neq}, \qquad \mathfrak{g}^{=}_{\ell} = m^{-\ell} g^{=}.$$

With this rescaling included, the ℓ-renormalised edge function is

$$\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}) = m^{-2\ell} \sum_{i \in I^{(\ell)},\, j \in J^{(\ell)}} g^{\neq}(X_i, X_j),$$

while the ℓ-renormalised self-loop function is

$$\mathfrak{g}_\ell^{=}(X_{I^{(\ell)}}, X_{I^{(\ell)}}) = m^{-\ell} \sum_{i \in I^{(\ell)}} g^{=}(X_i).$$

Put

$$q^{\neq} = \mathbb{E}_X[g^{\neq}(X_1, X_1')], \qquad q^{=} = \mathbb{E}_X[g^{=}(X_1)],$$

where X_1' is an independent copy of X_1. By assumption ($\square$), $q^{\neq}$ is finite and strictly positive, while $q^{=}$ is finite and non-negative. Hence $\frac{1}{2}(q^{\neq} + q^{=}) > 0$. For fixed $I^{(\ell)} \neq J^{(\ell)}$, the two-sample law of large numbers gives

$$m^{-2\ell} \sum_{i \in I^{(\ell)},\, j \in J^{(\ell)}} g^{\neq}(X_i, X_j) \longrightarrow q^{\neq} \text{ in } L^1.$$

Similarly, for a fixed $I^{(\ell)}$, the law of large numbers for U-statistics gives

$$m^{-2\ell} \sum_{\substack{i,j \in I^{(\ell)} \\ i \neq j}} g^{\neq}(X_i, X_j) \longrightarrow q^{\neq} \text{ in } L^1,$$

while, by the last assumption in ($\square$),

$$m^{-2\ell} \sum_{i \in I^{(\ell)}} g^{\neq}(X_i, X_i) \longrightarrow 0$$

in probability. The ordinary law of large numbers gives

$$m^{-\ell} \sum_{i \in I^{(\ell)}} g^{=}(X_i) \longrightarrow q^{=} \text{ in } L^1.$$

Hence

$$\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}) \longrightarrow q^{\neq}, \qquad \mathfrak{g}_\ell^{=}(X_{I^{(\ell)}}) \longrightarrow q^{=}$$

in probability.

We next estimate the specific relative entropies. By exchangeability of the level-ℓ blocks, it is enough to consider fixed representative blocks $I^{(\ell)}, J^{(\ell)}$, with $I^{(\ell)} \neq J^{(\ell)}$, for the edge contribution, and a fixed representative block $I^{(\ell)}$ for the self-loop contribution. Recall that

$$p^{\neq} = 1 - \mathrm{e}^{-q^{\neq}}, \qquad p^{=} = 1 - \mathrm{e}^{-\frac{1}{2}(q^{\neq} + q^{=})},$$

or

$$-\log(1 - p^{\neq}) = q^{\neq}, \qquad -\log(1 - p^{=}) = \frac{1}{2}(q^{\neq} + q^{=}).$$

Using Lemma 2.1, we obtain

$$\mathbb{E}_X\left[S^{\neq}\big(P(G^{(\ell)}[X]) \mid P(G_{\ell,p})\big)\right] \leq c_\ell^{\neq}, \qquad \mathbb{E}_X\left[S^{=}\big(P(G^{(\ell)}[X]) \mid P(G_{\ell,p})\big)\right] \leq c_\ell^{=},$$

where

$$c_\ell^{\neq} = \mathbb{E}_X\left[\zeta\Big(\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}), q^{\neq}\Big)\right],$$
$$c_\ell^{=} = \mathbb{E}_X\left[\zeta\Big(\tfrac{1}{2}\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{I^{(\ell)}}) + \tfrac{1}{2}\mathfrak{g}_\ell^{=}(X_{I^{(\ell)}}), \tfrac{1}{2}(q^{\neq} + q^{=})\Big)\right].$$

By $(\heartsuit_1)$, $\zeta(a,b) \le (\frac{1}{b}+1)|a-b|$, $a,b>0$. Therefore

$$c_\ell^{\neq} \le \left(\frac{1}{q^{\neq}} + 1\right)\mathbb{E}_X\left[\Big|\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{J^{(\ell)}}) - q^{\neq}\Big|\right] \longrightarrow 0,$$

and

$$c_\ell^{=} \le \left(\frac{2}{q^{\neq}+q^{=}} + 1\right)\mathbb{E}_X\left[\Big|\tfrac{1}{2}\mathfrak{g}_\ell^{\neq}(X_{I^{(\ell)}}, X_{I^{(\ell)}}) + \tfrac{1}{2}\mathfrak{g}_\ell^{=}(X_{I^{(\ell)}}) - \tfrac{1}{2}(q^{\neq}+q^{=})\Big|\right] \longrightarrow 0.$$

The bounds do not depend on L, apart from the requirement $L \ge \ell$ ensuring that level-ℓ blocks exist. Hence

$$\lim_{\ell\to\infty}\sup_{L\ge\ell}\mathbb{E}_X\left[\Delta\Big(P(G^{(\ell)}[X]), P(G_{\ell,p})\Big)\right] = 0.$$

This settles the claim in Theorem 1.4. □

3.3 Proof of Corollary 1.5

Proof. By assumption,

$$g^{\neq}(x,y) = B(xy)^\beta + O(1), \qquad g^{=}(x) = Cx^\beta + O(1).$$

Since $\mathbb{E}_X[X_1^\beta] < \infty$, we have

$$\mathbb{E}_X[g^{\neq}(X_1, X_1')] < \infty, \qquad \mathbb{E}_X[g^{=}(X_1)] < \infty.$$

Thus, the first two conditions in $(\square)$ are satisfied, with

$$q^{\neq} = \mathbb{E}_X[g^{\neq}(X_1, X_1')], \qquad q^{=} = \mathbb{E}_X[g^{=}(X_1)].$$

It remains to verify the last condition in $(\square)$, which we do for a generic level-ℓ block $I^{(\ell)}$ of $k \equiv m^\ell$ microscopic vertices as $\ell \to \infty$. Put $W_i = X_i^\beta$. Then $\mathbb{E}_X[W_1] < \infty$. For a fixed level-$\ell$ block $I^{(\ell)}$,

$$m^{-2\ell}\sum_{i\in I^{(\ell)}} X_i^{2\beta} = m^{-2\ell}\sum_{i\in I^{(\ell)}} W_i^2 \le \left(m^{-\ell}\max_{i\in I^{(\ell)}} W_i\right)\left(m^{-\ell}\sum_{i\in I^{(\ell)}} W_i\right).$$

By the law of large numbers,

$$m^{-\ell}\sum_{i\in I^{(\ell)}} W_i \longrightarrow \mathbb{E}_X[W_1], \qquad m\to\infty,$$

in probability. Moreover,

$$m^{-\ell}\max_{i\in I^{(\ell)}} W_i \longrightarrow 0, \qquad m\to\infty,$$

in probability. Indeed, for every $\varepsilon > 0$,

$$\mathbb{P}\left(m^{-\ell}\max_{i\in I^{(\ell)}} W_i > \varepsilon\right) \le m^\ell\,\mathbb{P}(W_1 > \varepsilon m^\ell) \longrightarrow 0, \qquad m\to\infty,$$

because $\mathbb{E}_X[W_1] < \infty$ implies

$$t\,\mathbb{P}(W_1 > t) \longrightarrow 0, \qquad t \to \infty.$$

Therefore

$$m^{-2\ell} \sum_{i \in I^{(\ell)}} X_i^{2\beta} \longrightarrow 0$$

in probability.

The $O(1)$-term in $g^{\neq}(x,x)$ contributes at most $O(m^{\ell})$, and hence also vanishes. Thus,

$$m^{-2\ell} \sum_{i \in I^{(\ell)}} g^{\neq}(X_i, X_i) \longrightarrow 0$$

in probability. Hence Theorem 1.4 applies.

Finally, if

$$\mathbb{P}(X_1 > x) \sim Ax^{-\alpha}, \qquad x \to \infty,$$

then $\mathbb{E}_X[X_1^{\beta}] < \infty$ precisely when $\alpha > \beta$. This proves the final claim. □

3.4 Proof of the heavy-tailed scaling theorem

Proof. The proof proceeds via two key lemmas (Lemmas 3.1–3.2 below). Along the way we use some standard facts listed in Appendix A. Throughout the proof, (■) is in force. The key parameters are $\gamma = \frac{\alpha}{\beta} \in (0,1)$, $\delta = B(A\Gamma(1-\gamma))^{2/\gamma}$ and $\eta = C(A\Gamma(1-\gamma))^{1/\gamma}$.

Define weights $\hat{X}^{(\ell)} = \{\hat{X}_I^{(\ell)}\}_{I \in [N_{L-\ell}]}$ by putting

$$\hat{X}_I^{(\ell)} = (m^{\ell})^{-1/\gamma} \sum_{i \in I} X_i^{\beta},$$

and $\hat{Y} = \{Y_I\}_{I \in [N_{L-\ell}]}$ by picking i.i.d. copies of Y_1, the standard γ-stable law random variable. In Appendix A we show how the weights $X, Y, \hat{X}, \hat{Y}$ can be coupled via a sequence U of i.i.d. UNIF$[0,1]$ random variables, which we suppress from the notation.

- Let $\hat{G}^{(\ell)}[\hat{X}]$ denote the inhomogeneous random graph with invariant connection functions $g_{\delta}^{\neq}$, $g_{\eta}^{=}$ and weights $\hat{X}^{(\ell)}$. From Lemma A.1 we know that the connection probabilities in $\hat{G}^{(\ell)}[\hat{X}]$ are *almost the same* as those in $G^{(\ell)}[X]$, the original graph after ℓ renormalisations.
- Let $\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}]$ denote the inhomogeneous random graph with invariant connection functions $g_{\delta}^{\neq}$, $g_{\eta}^{=}$ and weights $\hat{Y}$. From Lemma A.3 we know that the connection probabilities in $\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}]$ are *the same* as those in $G_{\ell,\gamma,\delta,\eta}[Y]$, the limiting graph in Theorem 1.6.

Our first key lemma shows that $\hat{G}^{(\ell)}[\hat{X}(U)]$ can be approximated by $G^{(\ell)}[X(U)]$ as $\ell \to \infty$.

Lemma 3.1. *Uniformly in $L \geq \ell$,*

$$\lim_{\ell \to \infty} \mathbb{E}_U \Big[\Big| S^{\neq}\Big(P(G^{(\ell)}[X]) \;\Big|\; P(G_{\ell,\gamma,\delta,\eta}[Y])\Big) - S^{\neq}\Big(P(\hat{G}^{(\ell)}[\hat{X}]) \;\Big|\; P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\Big) \Big| \Big] = 0,$$

$$\lim_{\ell \to \infty} \mathbb{E}_U \Big[\Big| S^{=}\Big(P(G^{(\ell)}[X]) \;\Big|\; P(G_{\ell,\gamma,\delta,\eta}[Y])\Big) - S^{=}\Big(P(\hat{G}^{(\ell)}[\hat{X}]) \;\Big|\; P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\Big) \Big| \Big] = 0.$$

Proof. For $L \geq \ell$, write $[N_{L-\ell}]$ for the set of level-ℓ vertices. For $I, J \in [N_{L-\ell}]$, define

$$a_\ell^{\neq}(I,J) = m^{-2\ell/\gamma} \sum_{i \in I,\, j \in J} g^{\neq}(X_i, X_j), \qquad b_\ell^{\neq}(I,J) = B \hat{X}_I^{(\ell)} \hat{X}_J^{(\ell)}.$$

For $I \in [N_{L-\ell}]$, define the genuine self-loop contribution

$$a_\ell^{=}(I) = m^{-\ell/\gamma} \sum_{i \in I} g^{=}(X_i), \qquad b_\ell^{=}(I) = C \hat{X}_I^{(\ell)}.$$

Recall that, under the present parametrisation,

$$f^{=}(x) = 1 - \mathrm{e}^{-\frac{1}{2}[g^{\neq}(x,x) + g^{=}(x)]}.$$

Consequently, the Bernoulli density of a self-loop at a level-ℓ block I is given by the following

$$A_\ell^{=}(I) = \tfrac{1}{2}\big[a_\ell^{\neq}(I,I) + a_\ell^{=}(I)\big].$$

Similarly, for the approximating graph and the limiting stable-law graph, put

$$B_\ell^{=}(I) = \tfrac{1}{2}\big[b_\ell^{\neq}(I,I) + b_\ell^{=}(I)\big],$$

and

$$C_\ell^{=}(I) = \tfrac{1}{2}\big[c_\ell^{\neq}(I,I) + c_\ell^{=}(I)\big],$$

where

$$c_\ell^{\neq}(I,J) = \delta \hat{Y}_I \hat{Y}_J, \qquad c_\ell^{=}(I) = \eta \hat{Y}_I.$$

Thus

$$C_\ell^{=}(I) = \tfrac{1}{2}\big[\delta (\hat{Y}_I)^2 + \eta \hat{Y}_I\big].$$

By (■), uniformly in I, J,

$$\begin{aligned} \big|a_\ell^{\neq}(I,J) - b_\ell^{\neq}(I,J)\big| &\leq m^{-2\ell/\gamma} \sum_{i \in I,\, j \in J} \big|g^{\neq}(X_i, X_j) - B X_i^{\beta} X_j^{\beta}\big| \\ &\leq K^{\neq} m^{2\ell(1-1/\gamma)} =: r_\ell^{\neq}. \end{aligned}$$

Since $\gamma < 1$, $r_\ell^{\neq} \to 0$. Similarly,

$$\begin{aligned} \big|a_\ell^{=}(I) - b_\ell^{=}(I)\big| &\leq m^{-\ell/\gamma} \sum_{i \in I} \big|g^{=}(X_i, X_i) - C X_i^{\beta}\big| \\ &\leq K^{=} m^{\ell(1-1/\gamma)} =: r_\ell^{=}, \end{aligned}$$

and $r_\ell^{=} \to 0$. Hence

$$(\sharp_1) \qquad \sup_{I,J} \big|a_\ell^{\neq}(I,J) - b_\ell^{\neq}(I,J)\big| \leq r_\ell^{\neq} \longrightarrow 0,$$

and

$$(\sharp_2) \qquad \sup_{I} \big|a_\ell^{=}(I) - b_\ell^{=}(I)\big| \leq r_\ell^{=} \longrightarrow 0.$$

In particular,

$$(\sharp_2') \qquad \sup_{I} \big|A_\ell^{=}(I) - B_\ell^{=}(I)\big| \leq \tfrac{1}{2}(r_\ell^{\neq} + r_\ell^{=}) =: \rho_\ell^{=} \longrightarrow 0.$$

For fixed representative blocks I, J with $I \neq J$, abbreviate

$$a_\ell^{\neq} = a_\ell^{\neq}(I,J), \qquad b_\ell^{\neq} = b_\ell^{\neq}(I,J), \qquad c_\ell^{\neq} = c_\ell^{\neq}(I,J).$$

For a fixed representative block I, abbreviate

$$A_\ell^{=} = A_\ell^{=}(I), \qquad B_\ell^{=} = B_\ell^{=}(I), \qquad C_\ell^{=} = C_\ell^{=}(I).$$

By Lemma A.1(iii),

$$(\sharp_3) \qquad b_\ell^{\neq} - c_\ell^{\neq} \longrightarrow 0 \quad \text{a.s.}$$

and also

$$b_\ell^{\neq}(I,I) - c_\ell^{\neq}(I,I) \longrightarrow 0, \qquad b_\ell^{=}(I) - c_\ell^{=}(I) \longrightarrow 0 \quad \text{a.s.}$$

Therefore

$$(\sharp_3') \qquad B_\ell^{=} - C_\ell^{=} \longrightarrow 0 \quad \text{a.s.}$$

Combining $(\sharp_1)$–$(\sharp_3')$, we get

$$(\sharp_4) \qquad a_\ell^{\neq} - c_\ell^{\neq} \longrightarrow 0 \quad \text{a.s.}, \qquad A_\ell^{=} - C_\ell^{=} \longrightarrow 0 \quad \text{a.s.}$$

By Lemma 2.2, for edges,

$$\begin{aligned}&\Big|S^{\neq}\big(P(G^{(\ell)}[X]) \mid P(G_{\ell,\gamma,\delta,\eta}[Y])\big) - S^{\neq}\big(P(\hat{G}^{(\ell)}[\hat{X}]) \mid P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\big)\Big| \\ &\qquad \le d^{\neq}(a_\ell^{\neq}, b_\ell^{\neq}, c_\ell^{\neq}).\end{aligned}$$

For self-loops, the same lemma must be applied to the actual Bernoulli self-loop densities. Hence

$$\begin{aligned}&\Big|S^{=}\big(P(G^{(\ell)}[X]) \mid P(G_{\ell,\gamma,\delta,\eta}[Y])\big) - S^{=}\big(P(\hat{G}^{(\ell)}[\hat{X}]) \mid P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\big)\Big| \\ &\qquad \le d^{=}(A_\ell^{=}, B_\ell^{=}, C_\ell^{=}).\end{aligned}$$

Taking expectation with respect to U and using exchangeability of the level-ℓ vertices, we see that it is enough to prove

$$\mathbb{E}_U\big[\zeta(a_\ell^{\neq}, b_\ell^{\neq}, c_\ell^{\neq})\big] \longrightarrow 0$$

and

$$\mathbb{E}_U\big[\zeta(A_\ell^{=}, B_\ell^{=}, C_\ell^{=})\big] \longrightarrow 0.$$

We prove a general estimate. Let (u_ℓ, v_ℓ, w_ℓ) denote either $(a_\ell^{\neq}, b_\ell^{\neq}, c_\ell^{\neq})$ or $(A_\ell^{=}, B_\ell^{=}, C_\ell^{=})$. In both cases,

$$|u_\ell - v_\ell| \le \rho_\ell, \qquad \rho_\ell \to 0,$$

with $\rho_\ell = r_\ell^{\neq}$ for edges and $\rho_\ell = \rho_\ell^{=}$ for loops, while

$$v_\ell - w_\ell \longrightarrow 0 \quad \text{a.s.}$$

We show that

$$\mathbb{E}_U[\zeta(u_\ell, v_\ell, w_\ell)] \longrightarrow 0.$$

Fix $\varepsilon \in (0,1)$ and $M > 1$.

• On the event $\{w_\ell \le \varepsilon\}$. Since $|u_\ell - v_\ell| \le \rho_\ell$ and $\rho_\ell \le 1$ for all large ℓ, the bound $(\heartsuit_2)$ gives, after increasing the constant if necessary,

$$\zeta(u_\ell, v_\ell, w_\ell) \le C\left(w_\ell^{-1} + 1\right) \qquad \text{on } \{w_\ell \le \varepsilon\}.$$

Therefore

$$\mathbb{E}_U\big[\zeta(u_\ell, v_\ell, w_\ell)1_{\{w_\ell \le \varepsilon\}}\big] \le C\,\mathbb{E}_U\big[w_\ell^{-1}1_{\{w_\ell \le \varepsilon\}}\big] + C\,\mathbb{P}_U(w_\ell \le \varepsilon).$$

For the edge term, w_ℓ has the law of $\delta Y_1 Y_2$. For the loop term, w_ℓ has the law of

$$\tfrac{1}{2}\left(\delta Y_1^2 + \eta Y_1\right).$$

In both cases Lemma A.2 gives a finite negative first moment. Since $w_\ell > 0$ a.s., dominated convergence yields

$$\lim_{\varepsilon \downarrow 0} \sup_{\ell} \mathbb{E}_U\big[\zeta(u_\ell, v_\ell, w_\ell)1_{\{w_\ell \le \varepsilon\}}\big] = 0.$$

• On the event $\{\varepsilon < w_\ell \le M\}$. On this event w_ℓ stays in a compact subinterval of $(0, \infty)$. Moreover $|u_\ell - v_\ell| \le \rho_\ell \to 0$ and $v_\ell - w_\ell \to 0$ a.s. Thus (u_ℓ, v_ℓ, w_ℓ) converges a.s. to the diagonal $\{(x, x, x)\colon\ x \in [\varepsilon, M]\}$. The function ζ is continuous near this diagonal and vanishes on it. The same deterministic bound $|u_\ell - v_\ell| \le \rho_\ell$ gives a uniform bound on ζ on this event. Hence bounded convergence gives

$$\mathbb{E}_U\big[\zeta(u_\ell, v_\ell, w_\ell)1_{\{\varepsilon < w_\ell \le M\}}\big] \longrightarrow 0.$$

• On the event $\{w_\ell > M\}$. Define

$$E_{\ell,M} = \left\{|v_\ell - w_\ell| \le \tfrac{1}{4}w_\ell,\ \rho_\ell \le \tfrac{1}{4}w_\ell\right\}.$$

On $E_{\ell,M} \cap \{w_\ell > M\}$, we have

$$\tfrac{3}{4}w_\ell \le v_\ell \le \tfrac{5}{4}w_\ell, \qquad \tfrac{1}{2}w_\ell \le u_\ell \le \tfrac{3}{2}w_\ell,$$

and therefore

$$u_\ell \wedge w_\ell \ge \tfrac{1}{2}w_\ell, \qquad v_\ell \wedge w_\ell \ge \tfrac{3}{4}w_\ell, \qquad u_\ell \wedge v_\ell \ge \tfrac{1}{2}w_\ell.$$

Also,

$$|u_\ell - w_\ell| \le \tfrac{1}{2}w_\ell, \qquad |v_\ell - w_\ell| \le \tfrac{1}{4}w_\ell, \qquad |u_\ell - v_\ell| \le \tfrac{1}{4}w_\ell.$$

Substitution into $(\heartsuit_2)$ gives

$$\zeta(u_\ell, v_\ell, w_\ell) \le C(1 + w_\ell)\mathrm{e}^{-w_\ell/2} \qquad \text{on } E_{\ell,M} \cap \{w_\ell > M\},$$

where $C < \infty$ is independent of ℓ and M. Hence, for $M \ge 2$,

$$\zeta(u_\ell, v_\ell, w_\ell) \le C(1 + M)\mathrm{e}^{-M/2} \qquad \text{on } E_{\ell,M} \cap \{w_\ell > M\}.$$

It remains to control the complement of $E_{\ell,M}$. Since $v_\ell - w_\ell \to 0$ a.s. and $\rho_\ell \to 0$, for fixed M,

$$\mathbb{P}_U\big(E_{\ell,M}^{\mathrm{c}} \cap \{w_\ell > M\}\big) \longrightarrow 0.$$

To pass from this probability estimate to an expectation estimate, first fix $R > M$ and decompose

$$E_{\ell,M}^{\mathrm{c}} \cap \{w_\ell > M\} = \Big(E_{\ell,M}^{\mathrm{c}} \cap \{M < w_\ell \le R\}\Big) \cup \Big(E_{\ell,M}^{\mathrm{c}} \cap \{w_\ell > R\}\Big).$$

On the first event, using again $|u_\ell - v_\ell| \le \rho_\ell$, we see that the function $\zeta(u_\ell, v_\ell, w_\ell)$ is bounded by a constant depending only on R for all large ℓ. Hence its contribution vanishes as $\ell \to \infty$. On the second event, the exponential factors in $(\heartsuit_2)$, together with the bound $|u_\ell - v_\ell| \le \rho_\ell$,

give an upper bound that can be made arbitrarily small by choosing R large and then letting $\ell \to \infty$. Consequently,

$$\lim_{\ell\to\infty} \mathbb{E}_U\big[\zeta(u_\ell, v_\ell, w_\ell)\, 1_{E_{\ell,M}^{\mathrm{c}} \cap \{w_\ell > M\}}\big] = 0.$$

Combining the last display with the estimate on $E_{\ell,M} \cap \{w_\ell > M\}$, we obtain

$$\limsup_{\ell\to\infty} \mathbb{E}_U\big[\zeta(u_\ell, v_\ell, w_\ell) 1_{\{w_\ell > M\}}\big] \le C(1+M)\mathrm{e}^{-M/2}.$$

Letting $M \to \infty$, we see shows that the contribution from $\{w_\ell > M\}$ vanishes.

Combining the three regions, we get

$$\mathbb{E}_U[\zeta(u_\ell, v_\ell, w_\ell)] \longrightarrow 0.$$

Applying this first to $(u_\ell, v_\ell, w_\ell) = (a_\ell^{\neq}, b_\ell^{\neq}, c_\ell^{\neq})$ and then to $(u_\ell, v_\ell, w_\ell) = (A_\ell^{=}, B_\ell^{=}, C_\ell^{=})$ we have settled both the edge and the self-loop estimates. □

Our second key lemma shows that the specific relative entropy of $P(\hat{G}^{(\ell)}[\hat{X}(U)])$ with respect to $P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}(U)])$ vanishes as $\ell \to \infty$ under the coupling.

Lemma 3.2. *Uniformly in $L \ge \ell$,*

$$\lim_{\ell\to\infty} \mathbb{E}_U\left[S^{\neq}\Big(P(\hat{G}^{(\ell)}[\hat{X}]) \;\Big|\; P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\Big)\right] = 0,$$

and

$$\lim_{\ell\to\infty} \mathbb{E}_U\left[S^{=}\Big(P(\hat{G}^{(\ell)}[\hat{X}]) \;\Big|\; P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\Big)\right] = 0.$$

Proof. By Lemma 2.1,

$$S^{\neq}\big(P(\hat{G}^{(\ell)}[\hat{X}]) \mid P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\big) \le d^{\neq}(b_\ell^{\neq}, c^{\neq}),$$

and, for loops,

$$S^{=}\big(P(\hat{G}^{(\ell)}[\hat{X}]) \mid P(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}])\big) \le d^{=}\left(\tfrac{1}{2}b_\ell^{=}, \tfrac{1}{2}c^{=}\right).$$

By exchangeability, it is enough to prove

$$\mathbb{E}_U[\zeta(b_\ell^{\neq}, c^{\neq})] \to 0, \qquad \mathbb{E}_U\left[\zeta\left(\tfrac{1}{2}b_\ell^{=}, \tfrac{1}{2}c^{=}\right)\right] \to 0.$$

Again we treat the edge term only. The same three-event decomposition works. On $\{c^{\neq} \le \varepsilon\}$, use

$$\zeta(a,b) \le \frac{|a-b|}{b} 1_{\{a>b\}} + \mathrm{e}^{-a}(b-a)1_{\{b>a\}} \le \frac{a}{b} + b,$$

together with ($\sharp_3$), positivity of $c^{\neq}$ and Lemma A.2, to conclude that this contribution vanishes as $\varepsilon \downarrow 0$. On $\{\varepsilon < c^{\neq} \le M\}$, bounded convergence applies, because $b_\ell^{\neq} \to c^{\neq}$ a.s. On $\{c^{\neq} > M\}$, the same event $E_{\ell,M}$ and the estimates above imply

$$\zeta(b_\ell^{\neq}, c^{\neq}) \le C(1+M)\,\mathrm{e}^{-M/2} \qquad \text{on } E_{\ell,M} \cap \{c^{\neq} > M\}.$$

Hence the claim follows. □

Combining Lemmas 3.1–3.2, we get the claim in Theorem 1.6. □

A Stable laws, domains of attraction and coupling

In this appendix we collect three lemmas that are needed in the proof of Theorem 1.6. The key parameters are

$$\gamma = \frac{\alpha}{\beta} \in (0,1), \qquad \sigma_\gamma = \big(A\Gamma(1-\gamma)\big)^{1/\gamma}, \qquad \delta = B\sigma_\gamma^2, \qquad \eta = C\sigma_\gamma.$$

► **Convergence and coupling.** The assumption in (■) implies that

$$\mathbb{P}(X_1^\beta > x) \sim Ax^{-\gamma}, \qquad x \to \infty,$$

so that, by the stable domain-of-attraction theorem (see [19, Section XVII.5]),

$$\hat{X}_I^{(\ell)} = m^{-\ell/\gamma} \sum_{i \in I} X_i^\beta \Longrightarrow \sigma_\gamma Y, \qquad \ell \to \infty,$$

where Y is the standard positive γ-stable random variable with Laplace transform $\mathbb{E}[\mathrm{e}^{-\lambda Y}] = \mathrm{e}^{-\lambda^\gamma}$, $\lambda \in \mathbb{R}_+$.

Lemma A.1. *For $\ell \in \mathbb{N}$, let F_ℓ be the distribution function of $\hat{X}_I^{(\ell)}$, and let F be the distribution function of $\sigma_\gamma Y$. Then the following hold:*

(i) F_ℓ and F are continuous distribution functions.

(ii) If U_I, $I \in [N_{L-\ell}]$, are i.i.d. UNIF$[0,1]$*, then*

$$\hat{X}_I^{(\ell)} = F_\ell^{\leftarrow}(U_I), \qquad \hat{Y}_I = \sigma_\gamma^{-1} F^{\leftarrow}(U_I), \qquad I \in [N_{L-\ell}],$$

define two coupled families such that $\{\hat{X}_I^{(\ell)}\}_I$ is i.i.d. with the block-sum law above, while $\{\hat{Y}_I\}_I$ is i.i.d. with the standard positive γ-stable law. Here,

$$G^{\leftarrow}(u) = \inf\{x \geq 0 \colon G(x) \geq u\}$$

denotes the generalised inverse, or quantile function, of a distribution function G.

(iii) For Lebesgue-a.e. $u \in (0,1)$,

$$F_\ell^{\leftarrow}(u) \longrightarrow F^{\leftarrow}(u).$$

Consequently, for every fixed I,

$$\hat{X}_I^{(\ell)} - \sigma_\gamma \hat{Y}_I \longrightarrow 0 \qquad \text{a.s.}$$

Proof. (i) Because X_1 is absolutely continuous, so is X_1^β, and therefore every finite sum of i.i.d. copies of X_1^β has a continuous distribution. This proves the continuity of F_ℓ. The law F is continuous because the positive stable law has a continuous density; see [19, Section XIII.6].

(ii) For a continuous distribution function G, the probability integral transform gives that $G^{\leftarrow}(U)$ has law G when U is UNIF$[0,1]$; see [28]. Applying this to F_ℓ and F gives the claimed marginal laws. Since the variables U_I are independent, both families are internally i.i.d., while the two families are coupled coordinatewise through the same uniforms.

(iii) Since $\hat{X}_I^{(\ell)} \Rightarrow \sigma_\gamma Y$ and F is continuous, the quantile convergence theorem for weak convergence gives

$$F_\ell^{\leftarrow}(u) \longrightarrow F^{\leftarrow}(u)$$

for Lebesgue-a.e. $u \in (0,1)$; see again [28]. Taking $u = U_I(\omega)$ outside the exceptional null set gives the almost sure convergence for every fixed I. □

▶ **Stable law tails.**

Lemma A.2. *Let Y be a standard positive γ-stable random variable with $\gamma \in (0,1)$. Then, for every $r > 0$,*

$$\mathbb{E}[Y^{-r}] = \frac{\Gamma(r/\gamma)}{\gamma\,\Gamma(r)} < \infty.$$

Consequently, if Y_1, Y_2 are i.i.d. copies of Y, then

$$C^{\neq} = \delta Y_1 Y_2, \qquad C^{=} = \tfrac{1}{2}(\delta Y_1^2 + \eta Y_1),$$

satisfy

$$\mathbb{E}[(C^{\neq})^{-1}] < \infty, \qquad \mathbb{E}[(C^{=})^{-1}] < \infty.$$

Proof. For $r > 0$,

$$y^{-r} = \frac{1}{\Gamma(r)} \int_0^\infty \lambda^{r-1} \mathrm{e}^{-\lambda y}\,\mathrm{d}\lambda, \qquad y > 0.$$

Applying Fubini's theorem and using $\mathbb{E}[\mathrm{e}^{-\lambda Y}] = \mathrm{e}^{-\lambda^\gamma}$, we obtain

$$\mathbb{E}[Y^{-r}] = \frac{1}{\Gamma(r)} \int_0^\infty \lambda^{r-1} \mathrm{e}^{-\lambda^\gamma}\,\mathrm{d}\lambda.$$

With the substitution $t = \lambda^\gamma$, this becomes

$$\mathbb{E}[Y^{-r}] = \frac{1}{\gamma\,\Gamma(r)} \int_0^\infty t^{r/\gamma - 1} \mathrm{e}^{-t}\,\mathrm{d}t = \frac{\Gamma(r/\gamma)}{\gamma\,\Gamma(r)}.$$

For $C^{\neq}$, independence gives

$$\mathbb{E}[(C^{\neq})^{-1}] = \delta^{-1}\,\mathbb{E}[Y_1^{-1}]\,\mathbb{E}[Y_2^{-1}] < \infty.$$

For $C^{=}$, since $\eta \geq 0$,

$$C^{=} = \tfrac{1}{2}(\delta Y_1^2 + \eta Y_1) \geq \tfrac{1}{2}\delta Y_1^2.$$

Hence

$$(C^{=})^{-1} \leq \tfrac{1}{2}\delta Y_1^{-2},$$

and the right-hand side has finite expectation. □

▶ **Transcription.**

Lemma A.3. *Let $\hat{G}^{(\ell)}[\hat{X}]$ be the inhomogeneous random graph on $[N_{L-\ell}]$ with edge densities*

$$b_\ell^{\neq}(I,J) = B\hat{X}_I^{(\ell)}\hat{X}_J^{(\ell)}, \qquad I \neq J,$$

and self-loop densities

$$\tfrac{1}{2}[b_\ell^{\neq}(I,I) + b_\ell^{=}(I)] = \tfrac{1}{2}\big[B\big(\hat{X}_I^{(\ell)}\big)^2 + C\hat{X}_I^{(\ell)}\big].$$

Let $\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}]$ be the inhomogeneous stable-law random graph on $[N_{L-\ell}]$ with edge densities

$$c_\ell^{\neq}(I,J) = \delta\hat{Y}_I\hat{Y}_J, \qquad I \neq J,$$

and self-loop densities

$$\tfrac{1}{2}\big[c_\ell^{\neq}(I,I)+c_\ell^{=}(I)\big]=\tfrac{1}{2}\big[\delta(\hat{Y}_I)^2+\eta\hat{Y}_I\big].$$

Then, for every measurable set $\mathcal{A}$ of graphs on $[N_{L-\ell}]$,

$$\mathbb{E}_U\big[P\big(\hat{G}^{(\ell)}[\hat{X}]\in\mathcal{A}\mid U\big)\big]=P\big(\hat{G}^{(\ell)}[\hat{X}]\in\mathcal{A}\big),$$
$$\mathbb{E}_U\big[P\big(\hat{G}_{\ell,\gamma,\delta,\eta}[\hat{Y}]\in\mathcal{A}\mid U\big)\big]=P\big(G_{\ell,\gamma,\delta,\eta}[Y]\in\mathcal{A}\big).$$

In the first line, the law on the right-hand side is the annealed graph law obtained by first sampling an i.i.d. family of block variables with the same law as $\hat{X}_I^{(\ell)}$ and then sampling edges and self-loops conditionally independently. In the second line, the law on the right-hand side is the annealed stable-law graph law.

Proof. Conditional on U, the two graphs are inhomogeneous random graphs with deterministic connection parameters. By Lemma A.1(ii), the random array $\hat{X}^{(\ell)}$ has the same law as an i.i.d. family of block variables

$$m^{-\ell/\gamma}\sum_{i\in I}X_i^{\beta},$$

and the random array $\hat{Y}$ has the same law as an i.i.d. family of standard positive γ-stable variables. Therefore, after integrating the conditional graph law over U, we obtain exactly the annealed law obtained by first sampling the weights and then sampling the edges and self-loops conditionally independently given these weights. □

References

[1] S.-K. Ma, Introduction to the renormalization group, Rev. Mod. Phys. 45:4 (1973) 589–614.

[2] K.G. Wilson, The renormalization group and critical phenomena, Rev. Mod. Phys. 55:3 (1983) 583–600.

[3] A. Pelissetto and E. Vicari, Critical phenomena and renormalization-group theory, Phys. Rep. 368 (2002) 549–727.

[4] A.C.D. van Enter, R. Fernández and A.D. Sokal, Regularity properties and pathologies of position-space renormalization-group transformations: Scope and limitations of Gibbsian theory, J. Stat. Phys. 72 (1993) 879–1167.

[5] A. Gabrielli, D. Garlaschelli, S.P. Patil, M.-Á. Serrano, Network renormalization, Nat. Rev. Phys. 7 (2025) 203–219.

[6] G. García-Pérez, M. Boguñá, M.-Á. Serrano, Multiscale unfolding of real networks by geometric renormalization, Nat. Phys. 14 (2018) 583–589.

[7] P. Villegas, T. Gili, G. Caldarelli, A. Gabrielli, Laplacian renormalization group for heterogeneous networks, Nat. Phys. 19 (2023) 445–450.

[8] L. Avena, F. Castell, A. Gaudillière and C. Mélot, Intertwining wavelets or multiresolution analysis on graphs through random rooted forests, Appl. Comput. Harmon. Anal. 48 (2020) 949–992.

[9] E. Garuccio, M. Lalli, D. Garlaschelli, Multiscale network renormalization: Scale-invariance without geometry, Phys. Rev. Res. 5 (2023) 043101.

[10] M. Lalli, D. Garlaschelli, Geometry-free renormalization of directed networks: scale-invariance and reciprocity, Preprint arXiv:2403.00235 (2024).

[11] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent, Ann. Probab. 25 (1997) 812–854.

[12] I. Norros and H. Reittu, On a conditionally Poissonian graph process, Adv. Appl. Probab. 38 (2006) 59–75.

[13] B. Bollobás, S. Janson, O. Riordan, The phase transition in inhomogeneous random graphs, Random Struct. Algorithms 31 (2007) 3–122.

[14] R. van der Hofstad, Random Graphs and Complex Networks: Volume 1, Cambridge series on statistical and probabilistic mathematics (Cambridge University Press, 2017).

[15] F. Caron, E.B. Fox, Sparse graphs using exchangeable random measures, J. R. Stat. Soc., B: Stat. Methodol. 79 (2017) 1295–1366.

[16] L. Avena, D. Garlaschelli, R.S. Hazra, M. Lalli, Inhomogeneous random graphs with infinite-mean fitness variables. J. Appl. Prob. 63 (2026) 375–400.

[17] A. Catanzaro, R. van der Hofstad, D. Garlaschelli, Clustering without geometry in sparse networks with independent edges, Preprint arXiv:2603.13159 (2026).

[18] A. Catanzaro, R.S. Hazra, D. Garlaschelli, Spectra of random graphs with discrete scale invariance, Preprint arXiv:2509.12407 (2025).

[19] W. Feller, *An Introduction to Probability Theory and Its Applications*, Volume II, Wiley Series in Probability and Mathematical Statistics, Wiley, 1996.

[20] B. Bollobás, S. Janson, O. Riordan, The phase transition in inhomogeneous random graphs, Random Struct. Algorithms 31 (2007) 3–122.

[21] S. Janson, Asymptotic equivalence and contiguity of some random graphs, Random Struct. Algorithms 36 (2010) 26–45.

[22] F.R.K. Chung, L. Lu, *Complex Graphs and Networks*, American Mathematical Society 107 (2006).

[23] T. Squartini, D. Garlaschelli, *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics*, Springer, 2017.

[24] S. Chatterjee, Exponential Random Graph Models. In Large Deviations for Random Graphs: École d'Été de Probabilités de Saint-Flour XLV-2015, 99–117 (Springer International Publishing, Cham, 2017).

[25] S. Bhamidi, G. Bresler and A. Sly, Mixing time of exponential random graphs. In 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS) 803–812 (IEEE, Washington, DC, 2008).

[26] S. Chatterjee, P. Diaconis, Estimating and understanding exponential random graph models, Ann. Statist. 41:5 (2013) 2428–2461.

[27] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (2nd ed.), Wiley-Interscience, 2006.

[28] P. Billingsley, *Probability and Measure* (3rd ed.), Wiley, 1995.