Clustering in a hyperbolic model of complex networks.

Nikolaos Fountoulakis[*1], Pim van der Hoorn[†2], Tobias Müller[‡3], and Markus Schepers[§3]

[1] School of Mathematics, University of Birmingham, United Kingdom.
[2]Department of Physics, Northeastern University, United States.
[3]Bernoulli Institute, University of Groningen, The Netherlands.

February 27, 2019

Abstract

In this paper we consider the clustering coefficient, and clustering function in a random graph model proposed by Krioukov et al. in 2010. In this model, nodes are chosen randomly inside a disk in the hyperbolic plane and two nodes are connected if they are at most a certain hyperbolic distance from each other. It has been previously shown that this model has various properties associated with complex networks, including a power-law degree distribution, "short distances" and a non-vanishing clustering coefficient. The model is specified using three parameters: the number of nodes n, which we think of as going to infinity, and $\alpha, \nu > 0$, which we think of as constant. Roughly speaking α controls the power law exponent of the degree sequence and ν the average degree.

Here we show that the clustering coefficient tends in probability **Pim:** Should we state here that convergence is in the stronger notion of expectation (L_1)? to a constant γ that we give explicitly as a closed form expression in terms of α, ν and certain special functions. This improves over earlier work by Gugelmann et al., who proved that the clustering coefficient remains bounded away from zero with high probability, but left open the issue of convergence to a limiting constant. Similarly, we are able to show that $c(k)$, the average clustering coefficient over all vertices of degree exactly k, tends in probability to a limit $\gamma(k)$ given explicitly as a closed form expression in terms of α, ν and certain special functions. We are able to extend this last result also to sequences $(k_n)_n$ where k_n grows as a function of n. Our results contradict a prediction of Krioukov et al., which stated that the limiting values $\gamma(k)$ should scale with k^{-1} as we let k grow. We find that this is true only when $\alpha > \frac{3}{4}$, while $\gamma(k) = \Theta(k^{2-4\alpha})$ if $\alpha < \frac{3}{4}$ and $\gamma(k) = \Theta(\log(k)/k)$ if $\alpha = \frac{3}{4}$.

*Research partially supported by the Alan Turing Institute, grant no. EP/N510129/1.
†Research partially supported by ARO Grant W911NF-16-1-0391 and W911NF-17-1-0491
‡Research partially supported by NWO grants 639.032.529 and 612.001.409.
§Research partially supported by NWO grant 639.032.529.

Contents

Pim: We need to fix the missing references. In addition, I have noticed the use of UK English grammar, which I do not often encounter in papers I read. Is there a specific reason for using this?

1 Introduction and main results

In this paper, we will consider clustering in a model of random graphs that involves points taken randomly in the hyperbolic plane. This model was introduced by Krioukov, Papadopoulos, Kitsak, Vahdat and Boguñá [12] in 2010 - we abbreviate it as *the KPKVB model*. We should however note that the model also goes by several other names in the literature, including *hyperbolic random geometric graphs* and *random hyperbolic graphs*. Krioukov et al. suggested this model as a suitable model for complex networks. It exhibits the three main characteristics usually associated with complex networks: a power-law degree distribution, a non-vanishing clustering coefficient and short graph distances.

Pim: I would try to refrain from making assumptions about the reader. We start with the definition of the KPKVB model. As mentioned, its nodes are situated in the hyperbolic plane $\mathbb{H}$, which is a surface with constant Gaussian curvature -1. This surface has several convenient representations (i.e. coordinate maps), including the Poincaré halfplane model, the Poincaré disk model and the Klein disk model. A gentle introduction to Gaussian curvature, hyperbolic geometry and these representations of the hyperbolic plane can be found in [**?**]. Throughout this paper we will be working with a representation of the hyperbolic plane using *hyperbolic polar coordinates*, sometimes called the *native representation*. That is, a point $p \in \mathbb{H}$ is represented as (r, θ), where r is the hyperbolic distance between p and the origin O and θ as the angle between the line segment Op and the positive x-axis. Here, when mentioning "the origin" and the angle between the line segment and the positive x-axis, we think of $\mathbb{H}$ embedded as the Poincaré disk in the ordinary euclidean plane.

The KPKVB model has three parameters: the number of vertices n, which we think of as going to infinity, and $\alpha > \frac{1}{2}$, $\nu > 0$ which we think of as fixed. Given n, α, ν we define $R = 2\log(n/\nu)$. Then the hyperbolic random graph $G(n; \alpha, \nu)$ is defined as follows:

- The vertex set is given by n i.i.d. points $u_1, \dots, u_n$ denoted in polar coordinates (r_i, θ_i), where the angular coordinate θ is chosen uniformly from $(-\pi, \pi]$ while the radial coordinate r is sampled independently according to the cumulative distribution function

$$F_{\alpha,R}(r) = \begin{cases} 0 & \text{if } r < 0 \\ \frac{\cosh(\alpha r)-1}{\cosh(\alpha R)-1} & \text{if } 0 \le r \le R \\ 1 & \text{if } r > R \end{cases} \tag{1}$$

- Any two vertices $u_i = (r_i, \theta_i)$ and $u_j = (r_j, \theta_j)$ are adjacent if and only if $d_{\mathbb{H}}(u_i, u_j) \le R$, where $d_{\mathbb{H}}$ denotes the distance in the hyperbolic plane. We will frequently be using that, by the hyperbolic law of cosines, $d_{\mathbb{H}}(u_i, u_j) \le R$ is equivalent to

$$\cosh(r_i)\cosh(r_j) - \sinh(r_i)\sinh(r_j)\cos(|\theta_i - \theta_j|_{2\pi}) \le \cosh(R),$$

  where $|a|_b = \min(|a|, b - |a|)$ for $-b \le a \le b$.

Figure 1 shows a computer simulation of $G(n; \alpha, \nu)$.

As observed by Krioukov et al [12], and proved rigorously by Gugelmann et al. [11], the degree sequence of the KPKVB model follows a power-law with exponent $2\alpha + 1$. Gugelmann et al [11] also showed that the average degree converges in probability to the constant $8\nu\alpha^2/\pi(2\alpha - 1)^2$, and they showed that the (local) clustering coefficient is non-vanishing in the sense that it is bounded below by a positive constant a.a.s. Here, and in the rest of the paper, for a sequence $(E_n)_n$ of events, E_n *asymptotically almost surely* (a.a.s.) means that $\mathbb{P}(E_n) \to 1$.

Apart from the degree sequence and clustering, the third main characteristic associated with complex networks, "short distances", has also been established in the literature. In [1] it is shown

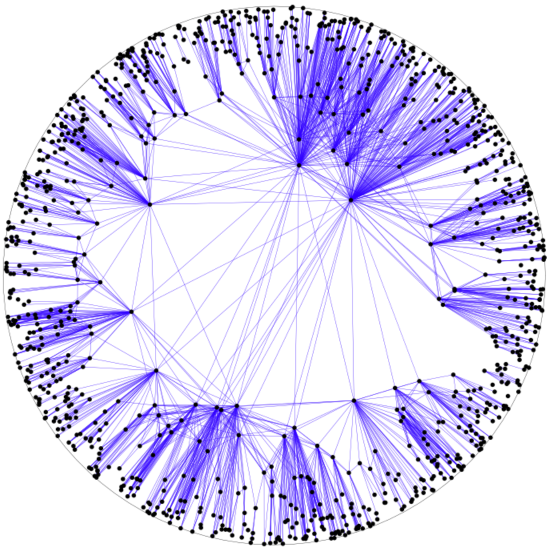

Figure 1: Simulation $G(n;\alpha,\nu)$ with $\alpha = 0.9$, $\nu = 0.2$ and $n = 5000$.

that for $\alpha < 1$ the largest component is what is called an *ultra-small world*: if we randomly sample two vertices of the graph then, a.a.s., conditional on them being in the same component, their graph distance is of order $\log\log n$. In [?] and [?] a.a.s. polylogarithmic upper and lower bounds on the graph diameter of the largest component are shown, and in [?], these were sharpened to show that $\log n$ is the correct order of the diameter.

Earlier work of the first and third authors with Bode [3] and of the first and third authors [8] has established the "threshold for a giant component": when $\alpha < 1$ then there always is a unique component of size linear in n no matter how small ν (and hence the average degree) is; when $\alpha > 1$ all components are sublinear no matter the value of ν; and when $\alpha = 1$ then there is a critical value ν_c such that for $\nu < \nu_c$ all components are sublinear and for $\nu > \nu_c$ there is a unique linear sized component (all of these statements holding a.a.s.). Whether or not there is a giant component when $\alpha = 1$ and $\nu = \nu_c$ remains an open problem. In [?] an [?], Kiwi and Mitsche considered the size of the second largest component and showed that when $\alpha \in (1/2, 1)$, a.a.s., the second largest component has polylogarithmic order with exponent $1/(\alpha - 1/2)$.

In another paper of the first and third authors with Bode [?] it was shown that $\alpha = 1/2$ is the threshold for connectivity: for $\alpha < 1/2$ the graph is a.a.s. connected, for $\alpha > 1/2$ the graph is a.a.s. disconnected and when $\alpha = 1/2$ the probability of being connected tends to a continuous, nondecreasing function of ν which is identically one for $\nu \geq \pi$ and strictly less than one for $\nu < \pi$. Friedrich and Krohmer [2] studied the size of the largest clique as well as the number of cliques of a given size. Boguña et al. [?] and Bläsius et al. [?] considered fitting the KPKVB model to data using maximum likelihood estimation. Kiwi and Mitsche [?] studied the spectral gap and related properties, and Bläsius et al. [?] considered the treewidth and related parameters of the KPKVB model. Recently Owada and Yogeshwaran [?] considered subgraph counts, and in particular established a central limit theorem for the number of copies of a fixed tree T in $G(n;\alpha,\nu)$, subject to some restrictions on the parameter α.

Clustering

In the literature there are unfortunately two distinct, rival definitions of the *clustering coefficient*. One of those, sometimes called the *global* clustering coefficient, is defined as three times the ratio the number of triangles to the number of paths of length two in the graph. Results for this version of the clustering coefficient in the KPKVB model were obtained by Candellero and the first author [4] and for the evolution of graphs on more general spaces with negative curvature by the first author in [**?**].

In this paper we will study the other notion of clustering, the one which is also considered by Krioukov et al. [12] and Gugelmann et al. [11]. It is sometimes called the *local* clustering coefficient, although we should point out that Gugelmann et al. actually call it the global clustering coefficient in their paper. For a graph G and a vertex $v \in V(G)$ we define the clustering coefficient *of* v as:

$$c(v) := \begin{cases} \frac{1}{\binom{\deg(v)}{2}} \sum_{u,w \sim v} 1_{\{uw \in E(G)\}} & \text{if } \deg(v) \geq 2, \\ 0 & \text{otherwise,} \end{cases}$$

where $E(G)$ denotes the edge set of G and $\deg(v)$ is the degree of vertex v. That is, provided v has degree at least two, $c(v)$ equals the number of edges that are actually present between the neighbours **Pim:** I think it is neighbors in US English vs neighbours in UK English. Is there any specific reason to us the UK version? of v divided by the number of edges that could possibly be present between the neighbours given the degree of v. The clustering coefficient of G is now defined as the average of $c(v)$ over all vertices v:

$$c(G) := \frac{1}{|V(G)|} \sum_{v \in V(G)} c(v).$$

As mentioned above, Gugelmann et al. [11], have established that $c(G(n; \alpha, \nu))$ is non-vanishing a.a.s., but they left open the question of convergence. Theorem 1.1 below establishes that the clustering coefficient indeed converges in probability to a constant γ that we give explicitly as a closed form expression involving α, ν and several classical special functions.

In addition to the clustering coefficient, we shall also be interested in the *clustering function*. This assigns to each integer k the value

$$c(k; G) := \begin{cases} \frac{1}{N_k} \sum_{\substack{v \in V(G), \\ \deg(v) = k}} c(v) & \text{if } N_k \geq 1 \\ 0 & \text{else.} \end{cases},$$

where N_k denotes the number of vertices of degree exactly k in G. In other words, the clustering function assigns to the integer k the average of the local clustering coefficient over all vertices of degree k. We remark that, while it might seem natural to consider $c(k)$ to be "undefined" when $N_k = 0$, we prefer to use the above definition for technical convenience. This way $c(k; G(n; \alpha, \nu))$ is a plain vanilla random variable and we can for instance compute its moments without any issues.

A general expression of the clustering function for KPKVB random graphs is given in [12, Equation (59)]. The authors conjecture that as k tends to infinity, the clustering function decays as k^{-1}. They based this prediction on observations (Figure 8 in [12]) in experiments on the infrastructure of the Internet obtained in [5]. Despite these interesting observations and the attention the KPKVB model has generated since then, the behaviour of the clustering function in KPKVB random graphs had not been completely determined. In particular it has not been established whether it converges as $n \to \infty$ to some suitable limit function and, how $c(k; G)$ scales with k. Theorems 1.2, 1.3 and Proposition 1.4 below settle these questions. Theorem 1.2 shows that for each fixed k, the value $c(k; G(n; \alpha, \nu))$ converges in probability to a constant $\gamma(k)$ that we again give explicitly as a closed form expression involving α, ν and several classical special functions. Theorem 1.3 extends this result to growing sequences satisfying $k \ll n^{1/(2\alpha+1)}$. Propostion 1.4 clarifies the asymptotics of the limiting function $\gamma(k)$, as $k \to \infty$. This depends on the parameter α, and $\gamma(k)$ only scales with k^{-1} when $\alpha > 3/4$, which corresponds to the exponent of the degree distribution exceeding $5/2$. So in particular our findings disprove the abovementioned conjecture of Krioukov et al. [12].

Notation

In the statement of our main results, and throughout the rest of the paper, we will use the following notations. We set

$$\xi := \frac{4\alpha\nu}{\pi(2\alpha-1)}.$$

We write $\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t$ for the gamma function, $\Gamma^+(a,b) := \int_b^\infty t^{a-1}e^{-t}\mathrm{d}t$ for the upper incomplete gamma function, $B(a,b) := \int_0^1 u^{a-1}(1-u)^{b-1}\mathrm{d}u = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ for the beta function and $B^-(x,a,b) := \int_0^x u^{a-1}(1-u)^{b-1}\mathrm{d}u$ the lower incomplete beta function. We write $U(a,b,z)$ for the hypergeometric U-function (also called Tricomi's confluent hypergeometric function), which has the integral representation

$$U(a,b,z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}dt,$$

see [7, p.255 Equation (2)], and let $G^{m,\ell}_{p,q}\left(z\middle|\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right)$ denote Meijer's G-Function [14], see Appendix A for more details.

For a sequence $(X_n)_n$ of random variables, we write $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$ to denote that X_n converges in probability to X, as $n\to\infty$.

1.1 Main results

1.1.1 The clustering coefficient

Our first result is the following.

Theorem 1.1. *Let $\alpha > \frac{1}{2}$, $\nu > 0$ be fixed. Writing $G_n := G(n;\alpha,\nu)$, we have*

$$c(G_n) \xrightarrow[n\to\infty]{\mathbb{P}} \gamma,$$

where γ is defined for $\alpha \neq 1$ as

$$\begin{aligned}\gamma = {} & \frac{2+4\alpha+13\alpha^2-34\alpha^3-12\alpha^4+24\alpha^5}{16(\alpha-1)^2\alpha(\alpha+1)(2\alpha+1)} + \frac{2^{-1-4\alpha}}{(\alpha-1)^2}\\
&+ \frac{(\alpha-1/2)(B(2\alpha,2\alpha+1)+B^-(1/2;1+2\alpha,-2+2\alpha))}{2(\alpha-1)(3\alpha-1)}\\
&+ \frac{\xi^{2\alpha}\left(\Gamma^+(1-2\alpha,\xi)+\Gamma^+(-2\alpha,\xi)\right)}{4(\alpha-1)}\\
&+ \frac{\xi^{2\alpha+2}\alpha(\alpha-1/2)^2\left(\Gamma^+(-2\alpha-1,\xi)+\Gamma^+(-2\alpha-2,\xi)\right)}{2(\alpha-1)^2}\\
&- \frac{\xi^{2\alpha+1}\alpha(2\alpha-1)\left(\Gamma^+(-2\alpha,\xi)+\Gamma^+(-2\alpha-1,\xi)\right)}{(\alpha-1)}\\
&- \frac{\xi^{6\alpha-2}2^{-4\alpha}(3\alpha-1)\left(\Gamma^+(-6\alpha+3,\xi)+\Gamma^+(-6\alpha+2,\xi)\right)}{(\alpha-1)^2}\\
&- \frac{\xi^{6\alpha-2}(\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)\left(\Gamma^+(-6\alpha+3,\xi)+\Gamma^+(-6\alpha+2,\xi)\right)}{(\alpha-1)}\\
&- \frac{e^{-\xi}\Gamma(2\alpha+1)\left(U(2\alpha+1,1-2\alpha,\xi)+U(2\alpha+1,2-2\alpha,\xi)\right)}{4(\alpha-1)}\\
&+ \frac{\xi^{6\alpha-2}\Gamma(2\alpha+1)\left(G^{3,0}_{2,3}\left(\xi\middle|\begin{matrix}1,3-2\alpha\\ 3-4\alpha,-6\alpha+2,0\end{matrix}\right)+G^{3,0}_{2,3}\left(\xi\middle|\begin{matrix}1,3-2\alpha\\ 3-4\alpha,-6\alpha+3,0\end{matrix}\right)\right)}{4(\alpha-1)},\end{aligned}$$

and for $\alpha = 1$ as the $\alpha\to 1$ limit of the above expression.

A plot of γ can be found in Figure 2.

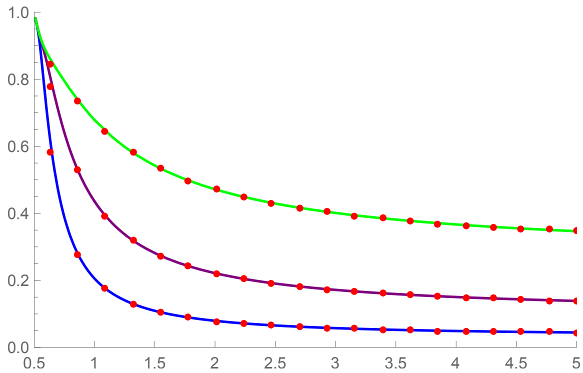

Figure 2: Plot of γ for α varying from 0.5 to 5 on the horizontal axis and for $\nu = \frac{1}{2}$ (blue), $\nu = 1$ (purple), $\nu = 2$ (green); simulations (red dots) with $n = 10000$ and 100 repetitions.

In the above expression for γ, a factor $\alpha - 1$ occurs in the denominator of each term, but we will see that this corresponds to a removable singularity. We have not been able to find a closed form expression in terms of known functions in the case when $\alpha = 1$, but in Section 3.2.4 we do provide an explicit expression involving integrals.

1.1.2 The clustering function

Our second result is on the clustering function for constant k.

Theorem 1.2. *Let $\alpha > \frac{1}{2}$, $\nu > 0$ and $k \geq 2$ be fixed. Writing $G_n := G(n; \alpha, \nu)$, we have*

$$c(k; G_n) \xrightarrow[n\to\infty]{\mathbb{P}} \gamma(k),$$

where $\gamma(k)$ is defined for $\alpha \neq 1$ as

$$\begin{aligned}\gamma(k) =& \frac{1}{8\alpha(\alpha-1)\Gamma^+(k-2\alpha,\xi)}\Bigg(-\Gamma^+(k-2\alpha,\xi) - 2\frac{\alpha(\alpha-1/2)^2\xi^2\Gamma^+(k-2\alpha-2,\xi)}{(\alpha-1)}\\ &+8\alpha(\alpha-1/2)\xi\Gamma^+(k-2\alpha-1,\xi)\\ &+4\xi^{4\alpha-2}\Gamma^+(k-6\alpha+2,\xi)\left(\frac{2^{-4\alpha}(3\alpha-1)}{(\alpha-1)} + (\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)\right)\\ &+\xi^{k-2\alpha}\Gamma(2\alpha+1)e^{-\xi}U(2\alpha+1,1+k-2\alpha,\xi)\\ &-\xi^{4\alpha-2}\Gamma(2\alpha+1)G^{3,0}_{2,3}\left(\begin{matrix}1,3-2\alpha\\ 3-4\alpha,-6\alpha+k+2,0\end{matrix}\middle|\xi\right)\Bigg)\end{aligned}$$

and for $\alpha = 1$ as the $\alpha \to 1$ limit of the above expression.

A plot of $\gamma(k)$ can be found in Figure 3. Again, we remark that the above expression for $\gamma(k)$ appears to have a singularity at $\alpha = 1$, but this will turn out to be a removable singularity. Again we have not been able to find a closed form expression in terms of known functions in the case when $\alpha = 1$, but in Section 3.2.4 we do provide an explicit expression involving integrals.

Theorem 1.2 in fact generalises to increasing sequences $(k_n)_n$.

Theorem 1.3. *Let $\alpha > \frac{1}{2}, \nu > 0$ be fixed and let k_n be a sequence satisfying $1 \ll k_n \ll n^{1/(2\alpha+1)}$. Then, writing $G_n := G(n; \alpha, \nu)$, we have*

$$\mathbb{E}[|c(k_n; G_n) - \gamma(k_n)|] = o(\gamma(k_n)),$$

as $n \to \infty$, where $\gamma(.)$ is as in Theorem 1.2.

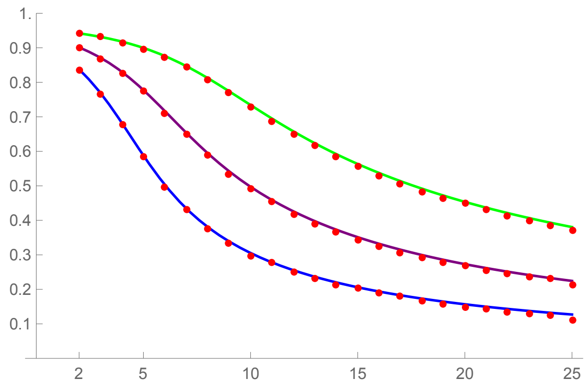

Figure 3: Plot $\gamma(k)$ for k varying from 2 to 25 on the horizontal axis, for $\alpha = 0.8$ and $\nu = \frac{1}{2}$ (blue), $\nu = 1$ (purple), $\nu = 2$ (green); simulations (red dots) with $n = 10000$ and 100 repetitions.

In particular, Theorem 1.3 implies that

$$\frac{c(k_n; G_n)}{\gamma(k_n)} \xrightarrow[n\to\infty]{L^1} 1,$$

so that the clustering function of the KPKVB model scales as $\gamma(k)$ as k grows.

1.1.3 Scaling of $\gamma(k)$

To clarify the scaling behaviour of $\gamma(k)$ with k we offer the following result.

Proposition 1.4. *As $k \to \infty$, we have*

$$\gamma(k) = (1 + o(1)) \cdot \begin{cases} \frac{8\alpha\nu}{\pi(4\alpha-3)} \cdot k^{-1} & \text{if } \alpha > \frac{3}{4}, \\ \frac{6\nu}{\pi} \cdot \frac{\log(k)}{k} & \text{if } \alpha = \frac{3}{4}, \\ c_\alpha \cdot k^{2-4\alpha} & \text{if } \frac{1}{2} < \alpha < \frac{3}{4}, \end{cases},$$

where $c_\alpha := \left(\frac{3\alpha-1}{2^{4\alpha+1}\alpha(\alpha-1)^2} + \frac{(\alpha-\frac{1}{2})B^-(\frac{1}{2},2\alpha+1,2\alpha-2)}{2(\alpha-1)\alpha} - \frac{B(2\alpha,3\alpha-4)}{4(\alpha-1)}\right) \cdot \xi^{4\alpha-2}$.

In particular this contradicts the scaling conjectured in [12] for $\alpha \leq \frac{3}{4}$, and confirms it for $\alpha > \frac{3}{4}$.

We remark that simultaneously and independently Stegehuis, van der Hofstad and van Leeuwaarden [16] found a similar, but less detailed result on the $k \to \infty$ scaling of the clustering function in the KPKVB model.

Pim: I really think we should add a subsection where we discuss the results and their implications. Especially the addition $\log(k)$ term for degree distribution exponent $5/2$, and the fact that $n^{1/(2\alpha_1)}$ is the best scaling possible for $c(k_n; G_n)$.

1.2 Outline of the paper

In the next section we will recall some useful tools from the literature and define a series of auxiliary random graph models that will be used in the proofs. In particular, we relate in a series of steps the KPKVB model to an infinite percolation model G_∞ that was used in previous work of the first and third authors [8] on the largest component of the KPKVB model. The value of the limiting constant γ, respectively limiting clustering function $\gamma(k)$, correspond to the probability that two randomly chosen neighbours of a "typical point" in this infinite model are themselves neighbours, respectively the probability of this event conditional on the typical point having exactly k neighbours. These probabilities can be expressed as certain integrals, which we solve explicitly in Section 3. In the same section we also prove Proposition 1.4, on the asymptotics of $\gamma(k)$. We then proceed to prove Theorems 1.1 and 1.2 by relating said probabilities for the typical point of the infinite model to the corresponding clustering coefficient/function in the original KPKVB random graph, using the Campbell-Mecke formula and some other, relatively straightforward considerations.

The remaining sections are devoted to the proof of Theorem 1.3, which turns out to be a lot more involved.

2 Preliminaries

In this section we recall some definitions and tools that we will need in our proofs.

2.1 The infinite limit model G_∞

We start by recalling the definition of the infinite limit model from [8]. Let $\mathcal{P} = \mathcal{P}_{\alpha,\nu}$ be a Poisson point process on $\mathbb{R}^2$ with intensity function $f = f_{\alpha,\nu}$ given by

$$f(x, y) = \frac{\alpha\nu}{\pi} e^{-\alpha y} \cdot 1_{\{y>0\}}. \tag{2}$$

The *infinite limit model* $G_\infty = G_\infty(\alpha, \nu)$ has vertex set $\mathcal{P}$ and edge set such that

$$pp' \in E(G_\infty) \iff |x - x'| \le e^{\frac{y+y'}{2}},$$

for $p = (x, y), p' = (x', y') \in \mathcal{P}$.

For any point $p \in \mathcal{R}$, we write $\mathcal{B}_\infty(p)$ to denote the *ball* around p, i.e.

$$\mathcal{B}_\infty(p) = \{p' \in \mathbb{R} \times [0, \infty) : |x - x'| \le e^{\frac{y+y'}{2}}\}. \tag{3}$$

With this notation we then have that $\mathcal{B}_\infty(p) \cap \mathcal{P}$ denotes the set of neighbors of a vertex $p \in G_\infty$. We will denote the intensity measure of the Poisson process $\mathcal{P}$ by $\mu = \mu_{\alpha,\nu}$, i.e. for every Borel-measurable subset $S \subseteq \mathbb{R}^2$ we have $\mu(S) = \int_S f(x, y)\, dx\, dy$

2.2 The finite box model $G_{\mathbf{box}}$

For the definition of the finite graph, recall that in the definition of the KPKVB model we set $R = 2\log(n/\nu)$. We consider the box $\mathcal{R} = (-\frac{\pi}{2}e^{R/2}, \frac{\pi}{2}e^{R/2}] \times (0, R]$ in $\mathbb{R}^2$. Then the *finite box model* $G_{\text{box}} := G_{\text{box}}(n; \alpha, \nu)$ has vertex set $\mathcal{V} := \mathcal{P} \cap \mathcal{R}$ and edges set such that

$$pp' \in E(G_{\text{box}}(n; \alpha, \nu)) \iff |x - x'|_{\pi e^{R/2}} \le e^{\frac{y+y'}{2}},$$

where $|x|_r = \min(|x|, r - |x|)$ for $-r \le x \le r$. Using $|.|_{\pi e^{R/2}}$ instead of $|.|$ results in the left and right boundaries of the box $\mathcal{R}$ getting identified, which in particular makes the model invariant under horizontal shifts and reflections in vertical lines. The graph G_{box} can thus be seen as a subgraph of G_∞ induced on $\mathcal{V}_n$, with some additional edges caused by the identification of the boundaries.

Similar to the infinite graph, for a point $p \in \mathcal{R}$ we define the ball $\mathcal{B}_{\text{box}}(p)$ as

$$\mathcal{B}_{\text{box}}(p) = \left\{p' \in \mathcal{R} : |x - x'|_{\pi e^{R/2}} \le e^{\frac{y+y'}{2}}\right\}. \tag{4}$$

2.3 The Poissonized KPKVB model $G_{\mathbf{Po}}$

Imagine that we have an infinite supply of i.i.d. points $p_1, p_2, \ldots$ in the hyperbolic plane $\mathbb{H}$ chosen according to the (α, R)-quasi uniform distribution. In the standard KPKVB random graph $G(n; \alpha, \nu)$ we take $p_1, \ldots, p_n$ as our vertex set and add edges between points at hyperbolic distance at most $R = 2\log(n/\nu)$. In the *Poissonized* KPKVB random graph $G_{\mathrm{Po}} := G_{\mathrm{Po}}(n;\alpha,\nu)$, we instead take $N \stackrel{\mathrm{d}}{=} \mathrm{Po}(n)$, a Poisson random variable with mean n, independent of our i.i.d. sequence of points and let the vertex set be $p_1, \dots, p_N$ and add edges according to the same rule as before. Equivalently, we could say that the vertex set consists of the points of a Poisson point process with intensity function ng, where g denotes the probability density of the (α, R)-quasi uniform distribution. That is,

$$g(r,\theta) = \frac{\alpha \sinh(\alpha r)}{2\pi(\cosh(\alpha R) - 1)} \cdot 1_{\{0 \le r \le R, -\pi < \theta \le \pi\}}.$$

Working with the Poissonized model has the advantage that when we take two disjoint regions A, B then the number of points in A and the number of points in B are independent Poisson-distributed random variables. As we will see, and as is to be expected, switching to the Poissonized model does not significantly alter the limiting behaviour of the clustering coefficient and function.

2.4 Coupling G_{Po} and G_{box}

The following lemmas from [8] establish a useful coupling between the Poissonized KPKVB random graph and the finite box model and relate the edge sets of the two graphs.

Lemma 2.1 ([8, Lemma 27]). *Let $\mathcal{V}$ denote the vertex set of $G_{Po}(n;\alpha,\nu)$ and $\mathcal{W}$ the vertex set of $G_{box}(n;\alpha,\nu)$. Define the map $\Psi : [0,R] \times (-\pi,\pi] \to \mathcal{R}$ by*

$$\Psi(r,\theta) = \left(\theta\frac{e^{R/2}}{2}, R - r\right). \tag{5}$$

Then there exists a coupling such that, a.a.s., $\mathcal{W} = \Psi[\mathcal{V}]$.

In the remainder of this paper we will write $\mathcal{B}(p)$ to denote the image under Ψ of the ball of hyperbolic radius R around the point $\Psi^{-1}(p)$, i.e.

$$\mathcal{B}(p) := \Psi\left[\left\{u \in \mathbb{H} : d_{\mathbb{H}}(\Psi^{-1}(p), u), d_{\mathbb{H}}(O,u) \le R\right\}\right].$$

Under the map Ψ, a point $p = (x,y) \in \mathcal{R}$ corresponds to $u := \Psi^{-1}(p) = (2e^{-R/2}x, R - y)$.

By the hyperbolic rule of cosines, for two points $p = (x,y) = \Psi((r,\theta)), p' = (x',y') = \Psi((r',\theta')) \in \mathcal{R}$ we have that $p' \in \mathcal{B}(p)$ iff. either $r + r' \le R$ or $r + r' > R$ and

$$\cosh r \cosh r' - \sinh r \sinh r' \cos(|\theta - \theta'|_{2\pi}) \le \cosh(R),$$

This can be rephrased as $p' \in \mathcal{B}(p)$ iff. either $y + y' \ge R$ or $y + y' < R$ and

$$|x - x'|_{\pi e^{R/2}} \le \Phi(r,r') := \frac{1}{2}e^{R/2}\arccos\left(\frac{\cosh r \cosh r' - \cosh R_n}{\sinh r \sinh r'}\right). \tag{6}$$

The following lemma provides useful bounds on the function $\Phi(r,r')$.

Lemma 2.2 ([8, Lemma 28]). *There exists a constant $K > 0$ such that, for every $\varepsilon > 0$ and for R sufficiently large, the following holds. For every $r, r' \in [\varepsilon R, R]$ with $r + r' > R$ we have that*

$$e^{\frac{1}{2}(y+y')} - Ke^{\frac{3}{2}(y+y')-R} \le \Phi(r,r') \le e^{\frac{1}{2}(y+y')} + Ke^{\frac{3}{2}(y+y')-R}, \tag{7}$$

where $y := R - r, y' := R - r'$. Moreover:

$$\Phi(r,r') \ge e^{\frac{1}{2}(y+y')} \quad \text{if} \quad r, r' < R - K. \tag{8}$$

A key consequence of Lemma 2.2 is that the coupling from Lemma 2.1 preserves edges between points whose heights are not too large.

Lemma 2.3 ([8, Lemma 30]). *On the coupling space of Lemma 2.1 the following holds a.a.s.:*

1. *for any two points $p, p' \in \mathcal{W}$ with $y, y' \leq R/2$, we have*

$$pp' \in E(G_{box}) \Rightarrow \Psi^{-1}(p)\Psi^{-1}(p') \in E(G_{Po}),$$

2. *for any two points $p, p' \in \mathcal{W}$ with $y, y' \leq R/4$, we have that*

$$pp' \in E(G_{box}) \iff \Psi^{-1}(p)\Psi^{-1}(p') \in E(G_{Po}).$$

2.5 The Campbell-Mecke formula

Pim: We might want to give some more results for Poisson point processes and Poisson random variables. Such as at the end of Section 2.1 in your recent paper on Hamiltonian cycles.

A very useful tool for analyzing subgraph counts, and their generalizations, in the setting of the Poissonized random geometric graphs, and in particular the Poissonized KPKVB model and the box model is the *Campbell-Mecke formula*. We use the incarnation below, which can be found in the monograph [15], as Theorem 1.6.

Pim: I found it extremely hard to see the implications of this theorem. I therefore strongly recommend using more explicit version, for instance as in the previous version of the paper, or explicitly state the consequences of this result which will be used here.

Theorem 2.4 ([15]). *Let $\mathcal{Q}$ be a Poisson process on $\mathbb{R}^d$ with intensity function g, and suppose that $\lambda := \int g < \infty$. Suppose that $h(\mathcal{Y}, \mathcal{X})$ is a bounded measurable function, defined on pairs $(\mathcal{Y}, \mathcal{X})$ with $\mathcal{Y} \subseteq \mathcal{X} \subseteq \mathbb{R}^d$ and $\mathcal{X}$ finite, such that $h(\mathcal{Y}, \mathcal{X}) = 0$ whenever $|\mathcal{Y}| \neq j$ (for some $j \in \mathbb{N}$). Then*

$$\mathbb{E} \sum_{\mathcal{Y} \subseteq \mathcal{Q}} h(\mathcal{Y}, \mathcal{Q}) = \frac{\lambda^j}{j!} \cdot \mathbb{E}h(\{Y_1, \ldots, Y_j\}, \{Y_1, \ldots, Y_j\} \cup \mathcal{Q}),$$

where the Y_i are i.i.d. random variables that are independent of $\mathcal{P}$ and have common probability density function g/λ.

3 Clustering and the degree of the typical point in G_∞

As mentioned earlier, we plan to make use of the Campbell-Mecke formula for comparing the clustering coefficient and function of G_{Po} with certain quantities associated with G_∞. We will be considering the Poisson process $\mathcal{P}$ to which we add one additinal point $(0, y)$ on the y-axis. In some computations the height y will be fixed, but eventually we shall take it exponentially distributed with parameter α, and independent of $\mathcal{P}$. We refer to $(0, y)$ as "the typical point".

To provide some intuïtion for this definition and name, note that we can alternatively view $\mathcal{P}$ as follows. We take a constant intensity Poisson process on $\mathbb{R}$ corresponding to the x-coordinates, and to each point we attach a random "mark", corresponding to the y-coordinate, where the marks are i.i.d. exponentially distributed with parameter α.

Since $c(G)$ is defined as an average over all vertices of the graph, it is not immediately obvious how to meaningfully define a corresponding notion for infinite graphs, and similarly for the clustering function, the degree sequence, etc. We can however without any issues speak of the (expected) clustering coefficient of the typical point, or the expected clustering coefficient given that it has degree k, or the distribution of the degree of the typical point. (All considered in the graph obtained from G_∞ by adding the typical point to its vertex set.)

If $p = (x, y) \in \mathbb{R} \times [0, \infty)$ is a point, not necessarily part of the Poisson process, then we will write

$$\mu(y) = \mu(p) := \mu(\mathcal{B}_\infty(p)).$$

Integrating the intensity function of $\mathcal{P}$ over $\mathcal{B}_\infty(p)$ gives

$$\begin{aligned}\mu(y) &= \int_{\mathcal{B}_\infty(p)} f(x',y')\,dx'\,dy' = \int_0^\infty \int_{-e^{(y+y')/2}}^{e^{(y+y')/2}} \tfrac{\alpha\nu}{\pi} e^{-\alpha y'}\,dx'\,dy' \\ &= \int_0^\infty 2e^{(y+y')/2} \tfrac{\alpha\nu}{\pi} e^{-\alpha y'}\,dy' = \tfrac{2\alpha\nu e^{y/2}}{\pi} \int_0^\infty e^{(\frac{1}{2}-\alpha)y'}\,dy' \\ &= \tfrac{2\alpha\nu e^{y/2}}{\pi(\alpha-\frac{1}{2})} = \xi e^{y/2}.\end{aligned}$$

3.1 The degree of the typical point

Before considering clustering we briefly investigate the distribution of the degree of the typical point. For $p=(x,y)\in\mathbb{R}\times[0,\infty)$ we define

$$\rho(p,k) := \mathbb{P}\left(\mathrm{Po}(\mu(y)) = k\right), \tag{9}$$

where $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with mean λ. We will often write $\rho(y,k)$ instead of $\rho(p,k)$.

Let the random variable D denote the degree of the typical point. Since the typical point has a height that is independent of the Poisson process and exponential(α)-distributed:

$$p_k := \mathbb{P}(D=k) = \int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,dy. \tag{10}$$

Using the transformation of variables $z=\xi e^{\frac{y}{2}}$ (so $dy=\frac{2}{z}dz$), we compute

$$\begin{aligned}p_k &= \frac{1}{k!}\int_0^\infty \left(\xi e^{\frac{y}{2}}\right)^k e^{-\xi e^{\frac{y}{2}}} \alpha e^{-\alpha y}\,dy \\ &= \frac{\alpha\xi^{2\alpha}}{k!}\int_0^\infty \left(\xi e^{\frac{y}{2}}\right)^{k-2\alpha} e^{-\xi e^{\frac{y}{2}}}\,dy \\ &= \frac{2\alpha\xi^{2\alpha}}{k!}\int_\xi^\infty z^{k-2\alpha-1}e^{-z}\,dz \\ &= \frac{2\alpha\xi^{2\alpha}\Gamma^+(k-2\alpha,\xi)}{k!},\end{aligned}$$

where we recall that Γ denotes the Gamma-function and Γ^+ the upper incomplete Gamma-function. Note that, unsurprisingly, this is identical to the expression Gugelmann et al. [11] gave for the limiting degree distribution of $G(n;\alpha,\nu)$. Using Stirling's approximation to the gamma function, we find that

$$p_k \sim 2\alpha\xi^{2\alpha}k^{-(2\alpha+1)} \quad \text{as } k\to\infty. \tag{11}$$

By a similar computation we have the following result, which will be useful later on. For any $\beta>0$, as $k\to\infty$

$$\int_0^\infty e^{\beta y}\rho(y,k)\alpha e^{-\alpha y}\,dy \sim 2\alpha\xi^{2(\beta+\alpha)}k^{-2(\beta+\alpha)-1}. \tag{12}$$

3.2 The expected clustering coefficient and function of the typical point

Let the random variable C denote the clustering coefficient of the typical point $(0,y)$, in the graph obtained from G_∞ by adding $(0,y)$. We *define*

$$\gamma := \mathbb{E}C, \quad \gamma(k) := \mathbb{E}\left(C|D=k\right).$$

(Where we take the expection over both the Poisson point process $\mathcal{P}$ and $y \overset{\mathrm{d}}{=} \exp(\alpha)$, independent of the Poisson process $\mathcal{P}$.) We shall show shortly that these take on the values stated in Theorem 1.1 and 1.2.

For any fixed value $y_0 > 0$, the set of points inside $\mathcal{B}_\infty((0,y_0))$ is a Poisson process with intensity $f \cdot 1_{\mathcal{B}_\infty((0,y_0))}$. As $\mu(\mathcal{B}_\infty((0,y_0))) = \mu(y_0) = \xi e^{y_0/2} < \infty$, this can be described alternatively by the experiment where we first pick $N \stackrel{\mathrm{d}}{=} \mathrm{Po}(\mu(y_0))$ then and we take N i.i.d. points in $\mathcal{B}_\infty((0,y_0))$ according to the probability density $f \cdot 1_{\mathcal{B}_\infty((0,y_0))}/\mu(y_0)$. (That is, the intesity function of the Poisson point process, but set to zero outside of $\mathcal{B}_\infty((0,y_0))$ and renormalized in such a way that it integrates to one.) Hence, if we condition on the event that y takes on some fixed value y_0 and that there are exactly k points of $\mathcal{P}$ inside $\mathcal{B}_\infty((0,y_0))$, then those k points behave like k i.i.d. points in $\mathcal{B}_\infty((0,y_0))$ chosen according to the mentioned renormalized probability density function. This shows that, for every $k \geq 2$:

$$\mathbb{E}\left(C|D=k, y=y_0\right) = \frac{1}{\binom{k}{2}}\mathbb{E}\left(\sum_{1\leq i<j\leq k} 1_{\{u_i \in \mathcal{B}_\infty(u_j)\}}\right) = \mathbb{E}\left[1_{\{u_1\in\mathcal{B}_\infty(u_2)\}}\right],$$

where $u_1, \ldots, u_n$ are i.i.d. points in $\mathcal{B}_\infty((0,y_0))$ with the above mentioned density. Note that this does not depend on the value of k. For notational convenience, we'll write

$$P(y_0) := \mathbb{E}\left[1_{\{u_1\in\mathcal{B}_\infty(u_2)\}}\right],$$

with u_1, u_2 as above.

We now observe that

$$\gamma(k) = \mathbb{E}(C|D=k) = \int_0^\infty \mathbb{E}\left(C|D=k, y=y_0\right) g_k(y_0)\, dy_0,$$

where g_k denotes the density of y *conditional on* $D = k$. That is,

$$g_k(y_0) = \frac{\rho(y_0,k)\alpha e^{-\alpha y_0}}{\int_0^\infty \rho(t,k)\alpha e^{-\alpha t}\, dt} = \frac{1}{p_k}\cdot \rho(y_0,k)\alpha e^{-\alpha y_0}.$$

Hence,

$$\gamma(k) = \frac{1}{p_k}\cdot \int_0^\infty P(y_0)\rho(y_0,k)\alpha e^{-\alpha y_0}\, dy_0. \tag{13}$$

This also gives

$$\begin{aligned}
\gamma &= \mathbb{E}C = \sum_{k\geq 2}\mathbb{E}(C|D=k)\mathbb{P}(D=k) \\
&= \int_0^\infty P(y_0)\left(\sum_{k=2}^\infty \rho(y_0,k)\right)\alpha e^{-\alpha y_0}\, dy_0 \qquad (14)\\
&= \int_0^\infty P(y_0)\left(1-\rho(y_0,0)-\rho(y_0,1)\right)\alpha e^{-\alpha y_0}\, dy_0.
\end{aligned}$$

A key step is to derive the following explicit expression for $P(y)$.

Lemma 3.1. *If $\alpha \neq 1$, then*

$$\begin{aligned}
P(y) = &-\frac{1}{8(\alpha-1)\alpha} + \frac{(\alpha-1/2)e^{-\frac{1}{2}y}}{\alpha-1} - \frac{(\alpha-1/2)^2 e^{-y}}{4(\alpha-1)^2} \\
&+ (e^{-\frac{1}{2}y})^{4\alpha-2}\left(\frac{2^{-4\alpha-1}(3\alpha-1)}{\alpha(\alpha-1)^2} + \frac{(\alpha-1/2)B^-(1/2; 1+2\alpha, -2+2\alpha)}{2(\alpha-1)\alpha}\right) \\
&+ \frac{(1-e^{-\frac{1}{2}y})^{2\alpha}}{8(\alpha-1)\alpha} - \frac{(e^{-\frac{1}{2}y})^{4\alpha-2}B^-(1-e^{-\frac{1}{2}y}; 2\alpha, 3-4\alpha)}{4(\alpha-1)}
\end{aligned}$$

We will prove this lemma in a sequence of steps.

Recall that $P(y_0)$ is the probability that $u_1 = (x_1,y_1), u_2 = (x_2,y_2)$ are neighbours in G_∞, where u_1, u_2 are i.i.d. with probability density $f \cdot 1_{\mathcal{B}_\infty((0,y_0))}/\mu(y_0)$. In particular

$$\begin{aligned}\mathbb{P}(y_i > t) &= \frac{\nu\alpha}{\pi\mu(y_0)}\int_t^\infty \int_{-e^{(s+y_0)/2}}^{e^{(s+y_0)/2}} e^{-\alpha s}\, ds \\ &= \frac{\nu\alpha}{\pi\mu(y_0)}\int_t^\infty 2e^{(s+y_0)/2}\cdot e^{-\alpha s}\, ds \\ &= \frac{2\nu\alpha e^{y_0/2}}{\pi\xi e^{y_0/2}(\alpha-\frac{1}{2})}\cdot e^{(\frac{1}{2}-\alpha)t} \\ &= e^{(\frac{1}{2}-\alpha)t},\end{aligned}$$

using that $\mu(y_0) = \xi e^{y_0/2} = \left(\frac{2\alpha\nu}{\pi(\alpha-\frac{1}{2})}\right)e^{y_0/2}$. Thus, y_1, y_2 are exponentially distributed with parameter $\alpha - \frac{1}{2}$. Now note that, for each $t > 0$, the probability density $f \cdot 1_{\mathcal{B}_\infty((0,y_0))}/\mu(y_0)$ is constant on $[-e^{(t+y_0)/2}, e^{(t+y_0)/2}]\times\{t\}$ and it is vanishes on $(-\infty, -e^{(t+y_0)/2})\times\{t\}\cup(e^{(t+y_0)/2},\infty)\times\{t\}$.

Hence, given the height y_i of u_i, the x-coordinate of u_i is uniform in $[-e^{\frac{1}{2}(y+y_i)}, e^{\frac{1}{2}(y+y_i)}]$. With this in mind we define $P(y_0, y_1, y_2)$ to be the probability that $(0, y_0), (x_1, y_1), (x_2, y_2)$ satisfy $|x_1 - x_2| \leq e^{(y_1+y_2)/2}$, where x_1 and x_2 are independent uniform random variables in, respectively, $[-e^{\frac{1}{2}(y_0+y_1)}, e^{\frac{1}{2}(y_0+y_1)}]$ and $[-e^{\frac{1}{2}(y_0+y_2)}, e^{\frac{1}{2}(y_0+y_2)}]$. We have that

$$P(y_0) = (\alpha - 1/2)^2 \int_0^\infty \int_0^\infty P(y_0, y_1, y_2) e^{-(\alpha-1/2)(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}y_1. \tag{15}$$

3.2.1 Determining $P(y_0, y_1, y_2)$

To compute the integral (15) it will be convenient to use the change of variable $z_i = e^{-y_i/2}$, for $i = 0, 1, 2$. We will write $y_i(z_i)$ to stress the dependence between y_i and z_i. The following result completely characterizes $P(y_0, y_1, y_2)$.

Lemma 3.2.

$$P(y_0(z_0), y_1(z_1), y_2(z_2)) = \begin{cases} 1 & \textit{if } z_0 \geq z_1 + z_2, z_0 > z_1 > z_2, \\ 1 - G(z_0, z_1, z_2) & \textit{if } z_0 < z_1 + z_2, z_0 > z_1 > z_2, \\ \frac{z_0}{z_1} & \textit{if } z_1 \geq z_0 + z_2, z_1 > \max(z_0, z_2), \\ \frac{z_0}{z_1}\left(1 - G(z_1, z_0, z_2)\right) & \textit{if } z_1 < z_0 + z_2, z_1 > \max(z_0, z_2), \end{cases}$$

where

$$G(a, b, c) = \frac{1}{4}\left(b^{-1}c + bc^{-1} + a^2 b^{-1} c^{-1} + 2 - 2ab^{-1} - 2ac^{-1}\right)$$

We split the proof of this lemma into a couple of smaller pieces. We begin with the following lemma.

Lemma 3.3. *Let $z_i = e^{-y_i/2}$, $i = 0, 1, 2$. If $y_0 < y_1 < y_2$ (or equivalently $z_0 > z_1 > z_2$), then*

$$P(y_0(z_0), y_1(z_1), y_2(z_2)) = \begin{cases} 1, & \textit{if } z_0 \geq z_1 + z_2, \\ 1 - G(z_0, z_1, z_2), & \textit{if } z_0 < z_1 + z_2 \end{cases}$$

Proof. Note that $P(y_0, y_1, y_2)$ is the probability that x_2 falls into the interval $[x_1 - e^{(y_1+y_2)/2}, x_1 + e^{(y_1+y_2)/2}]$, as well as into the interval $[-e^{(y_0+y_2)/2}, e^{(y_0+y_2)/2}]$. By symmetry considerations, we can take x_1 uniformly at random from $[0, e^{y_0/2+y_1/2}]$ as opposed to $[-e^{y_0/2+y_1/2}, e^{y_0/2+y_1/2}]$. Figure 4 show the intersection of the intervals (red line) for two different cases for $x_1 \leq e^{(y_0+y_1)/2}$.

Since $y_0 < y_1 < y_2$ we have that $e^{(y_1+y_2)/2} > e^{(y_0+y_2)/2}$ and so, when $x_1 \geq 0$, the "right half" of the interval $[-e^{(y_0+y_2)/2}, e^{(y_0+y_2)/2}]$ is always covered by the interval $[x_1 - e^{(y_1+y_2)/2}, x_1 + e^{(y_1+y_2)/2}]$. If $e^{(y_1+y_2)/2} - e^{(y_0+y_1)/2} \geq e^{(y_0+y_2)/2}$ then the "left half" is always covered as well. In other words:

$$e^{(y_1+y_2)/2} - e^{(y_0+y_1)/2} \geq e^{(y_0+y_2)/2} \Rightarrow P(y_0, y_1, y_2) = 1.$$

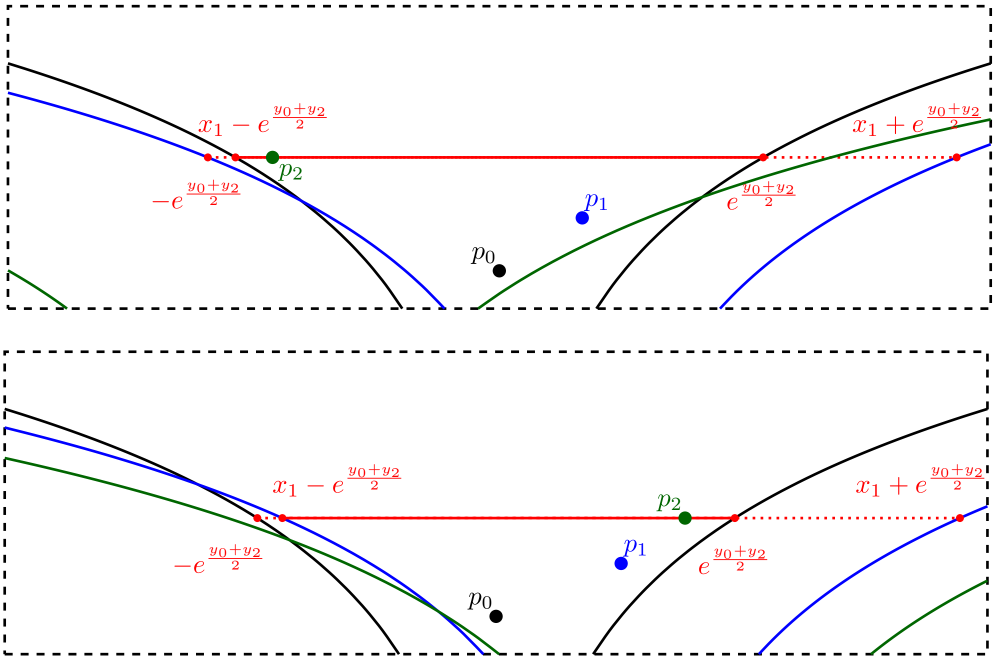

Figure 4: Situation for the intersections of the connection intervals considered in Lemma 3.3, with $y_0 < y_1 < y_2$ fixed and for different cases of $0 \le x_1 \le e^{(y_0+y_1)/2}$. The top figure shows the case where $0 \le x_1 \le e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2}$, while the bottom one shows the case $x_1 > e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2}$. The solid red line indicates the range for x_2 such that the points p_0, p_1 and p_2 form a triangle. The boundaries of their neighborhoods are shown in, respectively, black, blue and green.

Now consider the case where $e^{(y_1+y_2)/2} - e^{(y_0+y_1)/2} < e^{(y_0+y_2)/2}$. Then, if $x_1 \in [0, e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2}]$ the whole interval $[-e^{(y_0+y_2)/2}, e^{(y_0+y_2)/2}]$ is still covered so that p_0, p_1 and p_2 form a triangle. If, on the other hand $e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2} < x_1 \le e^{(y_0+y_1)/2}$ then the probability that $|x_2 - x_1| \le e^{(y_1+y_2)/2}$ equals

$$1 - \frac{x_1 - (e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2})}{2e^{(y_0+y_2)/2}}.$$

Hence, when $e^{(y_1+y_2)/2} - e^{(y_0+y_1)/2} < e^{(y_0+y_2)/2}$ we have

$$\begin{aligned}
P(y_0, y_1, y_2) &= \frac{e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2}}{e^{(y_0+y_1)/2}} \\
&\quad + \int_{e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2}}^{e^{(y_0+y_1)/2}} \left(1 - \frac{x_1 - (e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2})}{2e^{(y_0+y_2)/2}}\right) \cdot \frac{1}{e^{(y_0+y_1)/2}} \, dx_1 \\
&= 1 - \frac{1}{2e^{y_0+y_1/2+y_2/2}} \int_0^{e^{(y_0+y_1)/2} + e^{(y_0+y_2)/2} - e^{(y_1+y_2)/2}} x_1 \, dx_1 \\
&= 1 - \frac{\left(e^{(y_0+y_1)/2} + e^{(y_0+y_2)/2} - e^{(y_1+y_2)/2}\right)^2}{4e^{y_0+y_1/2+y_2/2}},
\end{aligned}$$

At this point it is convenient to rewrite everything in terms of $z_i := e^{-y_i/2}$. Note that $y_0 < y_1 < y_2$ if and only if $z_0 > z_1 > z_2$ while the condition $e^{(y_1+y_2)/2} - e^{(y_0+y_1)/2} < e^{(y_0+y_2)/2}$ becomes

$$e^{(y_1+y_2)/2} - e^{(y_0+y_1)/2} < e^{(y_0+y_2)/2} \Leftrightarrow z_1^{-1} z_2^{-1} < z_0^{-1} z_1^{-1} + z_0^{-1} z_2^{-1} \Leftrightarrow z_0 < z_1 + z_2.$$

We now conclude that

$$P(y_0(z_0), y_1(z_1), y_2(z_2)) = 1 \quad \text{if} \quad z_0 > z_1 > z_2 \text{ and } z_0 \ge z_1 + z_2$$

while for $z_0 > z_1 > z_2$ and $z_0 < z_1 + z_2$

$$\begin{aligned} P(y_0, y_1, y_2) &= 1 - \frac{z_0^2 z_1 z_2}{4} \cdot \left(z_0^{-1} z_1^{-1} + z_0^{-1} z_2^{-1} - z_1^{-1} z_2^{-1}\right)^2 \\ &= 1 - \frac{1}{4}\left(z_1^{-1} z_2 + z_1 z_2^{-1} + z_0^2 z_1^{-1} z_2^{-1} + 2 - 2z_0 z_1^{-1} - 2z_0 z_2^{-1}\right), \end{aligned}$$

which finishes the proof. □

The previous lemma covers the case when $y_0 < y_1 < y_2$. We now leverage it to take care of the other cases as well.

Proof of Lemma 3.2. Let $y_i > 0$ and $z_i = e^{-y_i/2}$, $i = 0, 1, 2$. Lemma 3.3 gives the expression for $P(y_0(z_0), y_1(z_1), y_2(z_2))$ in the case $y_0 < y_1 < y_2$, or equivalently $z_0 > z_1 > z_2$, i.e. the first two lines in the claim of Lemma 3.2. To analyze the other cases we shall express $P(y_1, y_0, y_2)$ and $P(y_1, y_2, y_0)$ in terms of $P(y_0, y_1, y_2)$ and z_i. For this we note that we can view $P(y_0, y_1, y_2)$ as a 2-fold integral of the indicator function

$$h(x_0, x_1, x_2) := \mathbb{1}_{\left\{|x_0 - x_1| < e^{(y_0+y_1)/2}, |x_0 - x_2| < e^{(y_0+y_2)/2}, |x_1 - x_2| < e^{(y_1+y_2)/2}\right\}},$$

where x_0 was set to zero, without loss of generality, and the other two x_i are uniform random variables on $[-e^{(y_0)+y_i)/2}, e^{(y_0+y_i)/2}]$. When we consider the probability $P(y_1, y_0, y_2)$, this is the 2-fold integral of $h(x_0, 0, x_2)$ so that

$$\begin{aligned} P(y_1, y_0, y_2) &= \frac{1}{2e^{(y_1+y_0)/2}} \cdot \frac{1}{2e^{(y_1+y_2)/2}} \iint_{\mathbb{R}} h(x_0, 0, x_2)\, \mathrm{d}x_0\, \mathrm{d}x_2 \\ &= \frac{e^{y_0/2}}{e^{y_1/2}} \frac{1}{2e^{(y_0+y_1)/2}} \frac{1}{2e^{(y_0+y_2)/2}} \iint_{\mathbb{R}} h(0, x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \\ &= \frac{e^{y_0/2}}{e^{y_1/2}} P(y_0, y_1, y_2) = \frac{z_1}{z_0} P(y_0, y_1, y_2). \end{aligned}$$

Finally we note that $h(x_0, 0, x_2) = h(x_2, 0, x_0)$ from which we conclude that

$$P(y_0, y_1, y_2) = (z_0/z_1)\, P(y_1, y_0, y_2) = (z_0/z_1)\, P(y_1, y_2, y_0). \tag{16}$$

To complete the proof for the other cases we note that since $P(y_0, y_1, y_2)$ is symmetric in y_1 and y_2, we can assume, without loss of generality, that $y_1 < y_2$. Then, there are two more orderings of y_0, y_1, y_2, namely $y_1 < y_0 < y_2$ and $y_1 < y_2 < y_0$, which can be summarized as $y_1 < \min(y_0, y_2)$, or equivalently $z_1 > \max(z_0, z_2)$. For $y_1 < y_0 < y_2$ and $y_1 < y_2 < y_0$ we can apply Lemma 3.3 to obtain $P(y_1, y_0, y_2) = P(y_1, y_2, y_0)$ which happen to agree due to the symmetry in the last two arguments of the expression found in Lemma 3.3. The expression for $P(y_0, y_1, y_2)$ then follows from (16). □

3.2.2 Integrating over y_1, y_2

Now that we have established the expression for $P(y_0, y_1, y_2)$ we can proceed to compute $P(y_0)$ by integrating over y_1, y_2. We however start with the following observation.

Lemma 3.4. *The function $\alpha \mapsto P_\alpha(y_0)$ is continuous for all $\alpha > \frac{1}{2}$.*

Proof. This follows from the theorem of dominated convergence: Let $\alpha > \frac{1}{2}$ and $(\alpha_n)_{n\in\mathbb{N}}$ a sequence of real numbers converging to α, so we can assume $|\alpha_n - \alpha| < \epsilon := \frac{\alpha - 1/2}{2}$. This means that $-\epsilon < \alpha_n - \alpha < \epsilon$, i.e. $\frac{\alpha - 1/2}{2} < \alpha_n - 1/2 < \frac{3\alpha - 3/2}{2}$. Define

$$f_n(y_1, y_2) = P(y_0, y_1, y_2)(\alpha_n - 1/2)^2 e^{-(\alpha_n - 1/2)(y_1+y_2)}.$$

As the function $x \mapsto x^2$ is increasing in x for $x > 0$ and the function $x \mapsto e^{-(y_1+y_2)x}$ is decreasing in x and $P(y_0, y_1, y_2) \in [0, 1]$, it holds that

$$|f_n(y_1, y_2)| \le \left(\frac{3\alpha - 3/2}{2}\right)^2 e^{-(y_1+y_2)\frac{\alpha-1/2}{2}}$$

which is integrable over $\mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0}$ (with integral equalling $(6\alpha-3)^2/(2\alpha-1)^2$). Application of the theorem of dominated convergence yields that $P_{\alpha_n}(y_0) \to P_\alpha(y_0)$ which gives the claim as the sequence $(\alpha_n)_n$ was arbitrary. □

Due to this lemma we can first assume $\alpha \notin \{\frac{3}{4}, 1\}$, compute $P(y_0)$ and then obtain the values of $P(y_0)$ at the remaining two points by taking the corresponding limit in α. This strategy is executed below. It involves the computation of several integrals which are involved and will take up a few pages. The proof is structured using headers, to aid the reader.

Note that when writing $P(y_0)$ as an integral, see equation (15), by symmetry in the integration variables y_1 and y_2, we can assume that $y_1 < y_2$ in which case either y_0 or y_1 is the smallest height. This gives half the value of $P(y_0)$ and hence

$$P(y_0) = 2(I_1(y_0) + I_2(y_0)),$$

where I_1 and I_2 are given by:

$$I_1(y_0) := \int_{0<y_0<y_1<y_2} P(y_0, y_1, y_2) \cdot (\alpha - 1/2)^2 e^{-(\alpha-1/2)(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}y_1$$
$$I_2(y_0) := \int_{0<y_1<y_0,y_2} P(y_0, y_1, y_2) \cdot (\alpha - 1/2)^2 e^{-(\alpha-1/2)(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}y_1$$

We proceed with computing each of these two integrals, each of which is split in two parts. The final expressions of those four integrals can be found in (17), (22), (23) and (25).

Computing $I_1(y_0)$ Applying the change of variables $z_i := e^{-y_i/2}$ and Lemma 3.2 gives

$$\begin{aligned}
I_1(y_0) &= 4(\alpha - 1/2)^2 \cdot \int_{z_0>z_1>z_2>0} P(y_0, y_1(z), y_2(z)) z_1^{2\alpha-2} z_2^{2\alpha-2} \mathrm{d}z_2 \mathrm{d}z_1 \\
&= 4(\alpha - 1/2)^2 \cdot \Bigg(\int_{z_0>z_1>z_2>0} 1 \cdot z_1^{2\alpha-2} z_2^{2\alpha-2} \mathrm{d}z_2 \mathrm{d}z_1 \\
&\quad - \int_{\substack{z_0>z_1>z_2>0,\\ z_0<z_1+z_2}} G(z_0, z_1, z_2) \cdot z_1^{2\alpha-2} z_2^{2\alpha-2} \mathrm{d}z_2 \mathrm{d}z_1 \Bigg) \\
&=: 4(\alpha - 1/2)^2 (I_{11}(y_0) - I_{12}(y_0)).
\end{aligned}$$

The integral $I_{11}(y_0)$ is easily obtained:

$$\begin{aligned}
I_{11}(y_0) &= \int_0^{z_0} \int_0^{z_1} z_1^{2\alpha-2} z_2^{2\alpha-2} \mathrm{d}z_2 \mathrm{d}z_1 = \int_0^{z_0} z_1^{2\alpha-2} \left[\frac{z_2^{2\alpha-1}}{2\alpha-1}\right]_0^{z_1} \mathrm{d}z_1 \\
&= \frac{1}{2\alpha-1} \cdot \int_0^{z_0} z_1^{4\alpha-3} \mathrm{d}z_1 = \frac{1}{2(2\alpha-1)^2} \cdot z_0^{4\alpha-2}. \qquad (17)
\end{aligned}$$

To deal with I_{12} we note that $G(z_0, z_1, z_2)$ is a linear combination of monomials of the form $z_0^a z_1^b z_2^c$ with $a, b, c \in \{-1, 0, 1, 2\}$ and $a + b + c = 0$. Let us consider the integral $J_{(a,b,c)}(z_0)$ defined by

$$J_{a,b,c}(z_0) := z_0^a \int_{\substack{z_0>z_1>z_2>0,\\ z_0<z_1+z_2}} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} \mathrm{d}z_2 \mathrm{d}z_1. \qquad (18)$$

and note that

$$I_{1,2}(y_0) = \frac{1}{4}(J_{0,-1,1}(z_0)+J_{0,1,-1}(z_0)+J_{2,-1,-1}(z_0)+2J_{0,0,0}(z_0)-2J_{1,-1,0}(z_0)-2J_{1,0,-1}(z_0)). \quad (19)$$

Next we compute $J_{a,b,c}(z_0)$.

$$\begin{aligned}
J_{a,b,c} &= z_0^a \int_{z_0/2}^{z_0} \int_{z_0-z_1}^{z_1} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} \mathrm{d}z_2 \mathrm{d}z_1 = z_0^a \int_{z_0/2}^{z_0} z_1^{b+2\alpha-2} \left[\frac{z_2^{c+2\alpha-1}}{c+2\alpha-1}\right]_{z_0-z_1}^{z_1} \mathrm{d}z_1 \\
&= \frac{z_0^a}{c+2\alpha-1} \cdot \left(\int_{z_0/2}^{z_0} z_1^{b+c+4\alpha-3}\,\mathrm{d}z_1 - \int_{z_0/2}^{z_0} z_1^{b+2\alpha-2}(z_0-z_1)^{c+2\alpha-1}\mathrm{d}z_1\right) \\
&= \frac{z_0^{a+b+c+4\alpha-2}(1-(1/2)^{b+c+4\alpha-2})}{(c+2\alpha-1)(b+c+4\alpha-2)} \\
&\quad - \frac{z_0^{a+b+c+4\alpha-3}}{c+2\alpha-1} \int_{z_0/2}^{z_0} (z_1/z_0)^{b+2\alpha-2} (1-(z_1/z_0))^{c+2\alpha-1}\,\mathrm{d}z_1 \\
&= \frac{z_0^{4\alpha-2}(1-(1/2)^{b+c+4\alpha-2})}{(c+2\alpha-1)(b+c+4\alpha-2)} - \frac{z_0^{4\alpha-2}}{c+2\alpha-1} \int_{1/2}^{1} u^{b+2\alpha-2}(1-u)^{c+2\alpha-1}\,\mathrm{d}u \\
&= \frac{z_0^{4\alpha-2}(1-(1/2)^{b+c+4\alpha-2})}{(c+2\alpha-1)(b+c+4\alpha-2)} - \frac{z_0^{4\alpha-2}}{c+2\alpha-1} B^-(1/2; c+2\alpha, b+2\alpha-1),
\end{aligned}$$

where we've used the substitution $u := z_1/z_0$ giving $z_0\mathrm{d}u = \mathrm{d}z_1$ in the penultimate line and B^- denotes the (lower) incomplete beta-function. Note that since $c \geq -1$, $-a \in \{0,-1,-2\}$ and by our assumption $\alpha \notin \{\frac{3}{4}, 1\}$, the denominators that occur during the integration are all non-zero.

Plugging this back into (19) gives

$$\begin{aligned}
I_{1,2}(y_0) &= \frac{z_0^{4\alpha-2}(1-(1/2)^{4\alpha-2})}{32\alpha(\alpha-1/2)} - \frac{z_0^{4\alpha-2}}{8\alpha} B^-(1/2; 1+2\alpha, 2\alpha-2) \\
&\quad + \frac{z_0^{4\alpha-2}(1-(1/2)^{4\alpha-2})}{32(\alpha-1)(\alpha-1/2)} - \frac{z_0^{4\alpha-2}}{4(2\alpha-2)} B^-(1/2; 2\alpha-1, 2\alpha) \\
&\quad + \frac{z_0^{4\alpha-2}(1-(1/2)^{4\alpha-4})}{32(\alpha-1)^2} - \frac{z_0^{4\alpha-2}}{4(2\alpha-2)} B^-(1/2; -1+2\alpha, 2\alpha-2) \\
&\quad + \frac{z_0^{4\alpha-2}(1-(1/2)^{4\alpha-2})}{16(\alpha-1/2)^2} - \frac{z_0^{4\alpha-2}}{2(2\alpha-1)} B^-(1/2; 2\alpha, 2\alpha-1) \\
&\quad - \frac{z_0^{4\alpha-2}(1-(1/2)^{4\alpha-3})}{16(\alpha-1/2)(\alpha-3/4)} + \frac{z_0^{4\alpha-2}}{2(2\alpha-1)} B^-(1/2; 2\alpha, 2\alpha-2) \\
&\quad - \frac{z_0^{4\alpha-2}(1-(1/2)^{4\alpha-3})}{16(\alpha-1)(\alpha-3/4)} + \frac{z_0^{4\alpha-2}}{2(2\alpha-2)} B^-(1/2; -1+2\alpha, 2\alpha-1) \\
&= \frac{\left(\frac{3}{64} - \frac{3}{16}2^{-4\alpha} + \alpha(-\frac{41}{128} + \frac{13}{16}2^{-4\alpha}) + \alpha^2(\frac{5}{8} - \frac{3}{4}2^{-4\alpha}) - \frac{15}{32}\alpha^3 + \frac{1}{8}a^4\right) z_0^{4\alpha-2}}{4(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)\alpha} \\
&\quad + \frac{z_0^{4\alpha-2}}{8(\alpha-1)\alpha(2\alpha-1)}(4(\alpha-1)\alpha(B^-(1/2; 2\alpha, 2\alpha-2) - B^-(1/2; 2\alpha, 2\alpha-1)) \\
&\quad - (2\alpha-1)\alpha(B^-(1/2; 2\alpha-1, 2\alpha-2) + B^-(1/2; 2\alpha-1, 2\alpha) - 2B^-(1/2; 2\alpha-1, 2\alpha-1)) \\
&\quad - (2\alpha-1)(\alpha-1)B^-(1/2; 1+2\alpha, 2\alpha-2)) \\
&= \frac{\left(\frac{3}{64} - \frac{3}{16}2^{-4\alpha} + \alpha(-\frac{41}{128} + \frac{13}{16}2^{-4\alpha}) + \alpha^2(\frac{5}{8} - \frac{3}{4}2^{-4\alpha}) - \frac{15}{32}\alpha^3 + \frac{1}{8}a^4\right) z_0^{4\alpha-2}}{4(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)\alpha} \\
&\quad + \frac{z_0^{4\alpha-2}}{8(\alpha-1)\alpha(2\alpha-1)}(4(\alpha-1)\alpha B^-(1/2; 2\alpha+1, 2\alpha-2) \\
&\quad - (2\alpha-1)\alpha B^-(1/2; 2\alpha+1, 2\alpha-2)
\end{aligned}$$

$$- (2\alpha-1)(\alpha-1)B^{-}(1/2; 2\alpha+1, 2\alpha-2)).$$

For the last step we use the identities

$$B^{-}(z;a,b) - B^{-}(z;a,b+1) = B^{-}(z;a+1,b), \tag{20}$$
$$B^{-}(z;a,b) + B^{-}(z;a,b+2) - 2B^{-}(z;a,b+1) = B^{-}(z;a+2,b). \tag{21}$$

to obtain

$$\begin{aligned} I_{1,2}(y_0) = &\frac{\left(\frac{3}{64} - \frac{3}{16}2^{-4\alpha} + \alpha(-\frac{41}{128} + \frac{13}{16}2^{-4\alpha}) + \alpha^2(\frac{5}{8} - \frac{3}{4}2^{-4\alpha}) - \frac{15}{32}\alpha^3 + \frac{1}{8}a^4\right) z_0^{4\alpha-2}}{4(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)\alpha} \\ &- \frac{z_0^{4\alpha-2}B^{-}(1/2; 2\alpha+1, 2\alpha-2)}{8(\alpha-1)\alpha(2\alpha-1)} \end{aligned} \tag{22}$$

Computing $I_2(y_0)$ We will follow a similar strategy as for $I_1(y_0)$. First, using the change of variables $z_i := e^{-y_i/2}$ we get

$$\begin{aligned} I_2(y_0) &= 4(\alpha-1/2)^2 \cdot \int_{1>z_1>z_2,z_0>0} P(y_0(z_0), y_1(z_1), y_2(z_2)) z_1^{2\alpha-2} z_2^{2\alpha-2} dz_2 dz_1 \\ &= 4(\alpha-1/2)^2 \cdot \Bigg(\int_{1>z_1>z_0,z_2>0} z_0 z_1^{2\alpha-3} z_2^{2\alpha-2} dz_2 dz_1 \\ &\quad - \int_{\substack{1>z_1>z_0,z_2>0 \\ z_1<z_0+z_2}} G(z_1, z_0, z_2) z_0 z_1^{2\alpha-3} z_2^{2\alpha-2}\, dz_2 dz_1 \Bigg) \\ &=: 4(\alpha-1/2)^2 (I_{21}(y_0) - I_{22}(y_0)). \end{aligned}$$

We proceed with the easy integral:

$$\begin{aligned} I_{21}(y_0) &= z_0 \int_{1>z_1>\max(z_2,z_0); z_0,z_2>0} z_1^{2\alpha-3} z_2^{2\alpha-2} dz_2 dz_1 = z_0 \int_{z_0}^{1} \int_0^{z_1} z_1^{2\alpha-3} z_2^{2\alpha-2} dz_2 dz_1 \\ &= z_0 \int_{z_0}^{1} \left[\frac{z_2^{2\alpha-1}}{2\alpha-1} \right]_0^{z_1} z_1^{2\alpha-3} dz_1 = \frac{z_0}{2\alpha-1} \int_{z_0}^{1} z_1^{4\alpha-4} dz_1 = \frac{z_0 - z_0^{4\alpha-2}}{(4\alpha-3)(2\alpha-1)}. \end{aligned} \tag{23}$$

We note that the denominators above are non-zero as $\alpha > \frac{1}{2}$ and $\alpha \neq \frac{3}{4}$.

To deal with $I_{22}(y_0)$ we consider the integral function

$$J'_{a,b,c}(z_0) := z_0^a \int_{\substack{1>z_1>\max(z_0,z_2); z_0,z_2>0 \\ z_1<z_0+z_2}} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} dz_2 dz_1$$

and note that

$$\begin{aligned} I_{2,2}(y_0) = &\frac{1}{4}\left(J'_{0,-1,1}(z_0) + J'_{2,-1,-1}(z_0) + J'_{0,1,-1}(z_0)\right) \\ &+ \frac{1}{2}\left(J'_{1,-1,0}(z_0) - J'_{0,0,0}(z_0) - J'_{1,0,-1}(z_0)\right). \end{aligned} \tag{24}$$

We know compute $J'_{a,b,c}(z_0)$

$$\begin{aligned} J'_{a,b,c}(z_0) &= z_0^a \int_{z_0}^{1} \int_{z_1-z_0}^{z_1} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} dz_2 dz_1 \\ &= z_0^a \int_{z_0}^{1} \frac{1}{c+2\alpha-1} z_1^{b+2\alpha-2} (z_1^{c+2\alpha-1} - (z_1-z_0)^{c+2\alpha-1}) dz_1 \\ &= z_0^a \int_{z_0}^{1} \frac{1}{c+2\alpha-1} z_1^{b+c+4\alpha-3} dz_1 - z_0^a \int_{z_0}^{1} \frac{1}{c+2\alpha-1} z_1^{b+2\alpha-2} (z_1-z_0)^{c+2\alpha-1} dz_1 \end{aligned}$$

$$= z_0^a \frac{1}{(c+2\alpha-1)(b+c+4\alpha-2)}(1-z_0^{b+c+4\alpha-2})$$
$$- \frac{z_0^a}{c+2\alpha-1} z_0^{b+c+4\alpha-2} B^-(1-z_0; c+2\alpha, -b-c-4\alpha+2)$$
$$= \frac{z_0^a - z_0^{4\alpha-2}}{(c+2\alpha-1)(b+c+4\alpha-2)} - \frac{z_0^{4\alpha-2} B^-(1-z_0; c+2\alpha, -b-c-4\alpha+2)}{c+2\alpha-1}.$$

Here we used that for $x \in \mathbb{R}, y > -1$ (note that as $c \geq -1$, it holds that $c + 2\alpha - 1 > -1$):

$$\begin{aligned}
\int_{z_0}^{1} z_1^x (z_1 - z_0)^y \mathrm{d}z_1 &= \int_0^{1-z_0} (s+z_0)^x s^y \mathrm{d}s \\
&= z_0^{x+y} \int_0^{1-z_0} ((s/z_0)+1)^x (s/z_0)^y \mathrm{d}s \\
&= z_0^{x+y+1} \int_0^{1/z_0-1} (t+1)^x t^y \mathrm{d}t \\
&= z_0^{x+y+1} \int_0^{1-z_0} u^y (1-u)^{-(x+y+2)} \mathrm{d}u \\
&= z_0^{x+y+1} B^-(1-z_0; y+1, -x-y-1).
\end{aligned}$$

As $c \geq -1$ and $-a \in \{0, -1, -2\}$ and by our assumption $\alpha \notin \{\frac{3}{4}\}$, the denominators that occur during the computations above are non-zero.

Plugging the expression for $J'_{a,b,c}(z_0)$ back into (24) we get,

$$\begin{aligned}
I_{2,2}(y_0) = &\frac{1 - z_0^{4\alpha-2}}{32\alpha(\alpha-1/2)} - \frac{z_0^{4\alpha-2} B^-(1-z_0; 1+2\alpha, -4\alpha+2)}{8\alpha} \\
&+ \frac{z_0^2 - z_0^{4\alpha-2}}{32(\alpha-1)^2} - \frac{z_0^{4\alpha-2} B^-(1-z_0; -1+2\alpha, -4\alpha+4)}{8(\alpha-1)} \\
&+ \frac{1 - z_0^{4\alpha-2}}{32(\alpha-1)(\alpha-1/2)} - \frac{z_0^{4\alpha-2} B^-(1-z_0; -1+2\alpha, -4\alpha+2)}{8(\alpha-1)} \\
&+ \frac{z_0 - z_0^{4\alpha-2}}{16(\alpha-1/2)(\alpha-3/4)} - \frac{z_0^{4\alpha-2} B^-(1-z_0; 2\alpha, -4\alpha+3)}{4(\alpha-1/2)} \\
&- \frac{1 - z_0^{4\alpha-2}}{16(\alpha-1/2)^2} + \frac{z_0^{4\alpha-2} B^-(1-z_0; 2\alpha, -4\alpha+2)}{4(\alpha-1/2)} \\
&- \frac{z_0 - z_0^{4\alpha-2}}{16(\alpha-1)(\alpha-3/4)} + \frac{z_0^{4\alpha-2} B^-(1-z_0; -1+2\alpha, -4\alpha+3)}{4(\alpha-1)}.
\end{aligned}$$

Using some algebra and the identities (20) and (21) this can be reduced to

$$\begin{aligned}
I_{2,2}(y_0) = &\frac{1}{64\alpha(\alpha-1/2)^2(\alpha-1)} - \frac{(1-z_0)^{2\alpha}}{64\alpha(\alpha-1/2)^2(\alpha-1)} - \frac{z_0}{8(\alpha-1/2)(\alpha-1)(4\alpha-3)} \\
&+ \frac{z_0^2}{32(\alpha-1)^2} + \frac{(-6+25\alpha-48\alpha^2+44\alpha^3-16\alpha^4) z_0^{4\alpha-2}}{512\alpha(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)} \\
&+ \frac{z_0^{4\alpha-2} B^-(1-z_0; 2\alpha, 3-4\alpha)}{32(\alpha-1)(\alpha-1/2)^2}.
\end{aligned} \tag{25}$$

Combining the results for $I_1(y_0)$ and $I_2(y_0)$ Combining the results for $I_{11}(y_0), I_{12}(y_0), I_{21}(y_0)$ and $I_{22}(y_0)$ we get, after some algebra, an explicit expression for $P(y_0)$ as a linear combination of terms of the form z_0^u, $(1-z_0)^u$ and $z_0^u B^-(1-z_0; a, b)$:

$$P(y_0) = 2(I_1 + I_2) = 8(\alpha - 1/2)^2 (I_{1,1} - I_{1,2} + I_{2,1} - I_{2,2})$$

$$
\begin{aligned}
=&8(\alpha-1/2)^2\Bigg(\frac{1}{2(2\alpha-1)^2}z_0^{4\alpha-2}\\
&-\frac{\left(\frac{3}{64}-\frac{3}{16}2^{-4\alpha}+\alpha(-\frac{41}{128}+\frac{13}{16}2^{-4\alpha})+\alpha^2(\frac{5}{8}-\frac{3}{4}2^{-4\alpha})-\frac{15}{32}\alpha^3+\frac{1}{8}a^4\right)z_0^{4\alpha-2}}{4(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)\alpha}\\
&+\frac{z_0^{4\alpha-2}B^-(1/2;2\alpha+1,2\alpha-2)}{8(\alpha-1)\alpha(2\alpha-1)}+\frac{z_0-z_0^{4\alpha-2}}{(4\alpha-3)(2\alpha-1)}\\
&-\frac{1}{64\alpha(\alpha-1/2)^2(\alpha-1)}+\frac{(1-z_0)^{2\alpha}}{64\alpha(\alpha-1/2)^2(\alpha-1)}+\frac{z_0}{8(\alpha-1/2)(\alpha-1)(4\alpha-3)}\\
&-\frac{z_0^2}{32(\alpha-1)^2}-\frac{(-6+25\alpha-48\alpha^2+44\alpha^3-16\alpha^4)z_0^{4\alpha-2}}{512\alpha(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)}\\
&-\frac{z_0^{4\alpha-2}B^-(1-z_0;2\alpha,3-4\alpha)}{32(\alpha-1)(\alpha-1/2)^2}\Bigg)\\
=&-\frac{1}{8(\alpha-1)\alpha}+\frac{(\alpha-1/2)z_0}{\alpha-1}-\frac{(\alpha-1/2)^2z_0^2}{4(\alpha-1)^2}\\
&+z_0^{-2+4\alpha}\left(\frac{2^{-4\alpha-1}(3\alpha-1)}{\alpha(\alpha-1)^2}+\frac{(\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)}{2(\alpha-1)\alpha}\right)\\
&+\frac{(1-z_0)^{2\alpha}}{8(\alpha-1)\alpha}-\frac{z_0^{4\alpha-2}B^-(1-z_0;2\alpha,3-4\alpha)}{4(\alpha-1)}
\end{aligned}
$$

Observe that the above expression only contains terms of the form $\alpha-1$ in the denominator. The only expression of the form $\alpha-3/4$ is in the lower incomplete beta-function $B^-(1-z_0;2\alpha,3-4\alpha)$ which appears twice in the expression for $P(y_0)$.

The case of $\boldsymbol{\alpha=3/4}$

Note that the factor $\alpha-\frac{3}{4}$ does not occur in any denominator of the previously obtained expression. For the lower incomplete beta function, the last argument $3-4\alpha$ is zero for $\alpha=\frac{3}{4}$, however as $z_0<1$ the integration domain of the lower incomplete beta function does not touch the singularity at $t=1$ (note $B^-(1-z_0;2\alpha;3-4\alpha)=\int_0^{1-z_0}t^{2\alpha-1}(1-t)^{2-4\alpha}$). Therefore, the previous expression holds for this case as well.

3.2.3 Computing γ and $\gamma(k)$

Now that we have an expression for $P(y_0)$ we can compute $\gamma,\gamma(k)$ by integrating over y_0 and prove that they equal the expressions given in, respectively, Theorem 1.1 and Theorem 1.2.

We define

$$
I^{(k)}:=\int_0^\infty P(y)\alpha e^{-\alpha y}\rho(y,k)\,\mathrm{d}y=\int_0^\infty P(y)\alpha e^{-\alpha y}\frac{\left(\xi e^{y/2}\right)^k}{k!}e^{-\xi e^{y/2}}\,\mathrm{d}y
$$

and

$$
J:=\int_0^\infty P(y)\alpha e^{-\alpha y}\,\mathrm{d}y.
$$

Then, recalling (14) and (13), we have

$$
\gamma=J-I^{(1)}-I^{(2)}\quad\text{and}\quad\gamma(k)=\frac{I^{(k)}}{p_k}.
$$

We will thus compute J and $I^{(k)}$. It will be helpful to change coordinates to $z:=e^{-y/2}$. This yields

$$
J=2\alpha\int_0^1 P(y)z^{2\alpha-1}\,\mathrm{d}z,
$$

and

$$I^{(k)} = \frac{2\alpha\xi^k}{k!} \cdot \int_0^1 P(y(z)) \cdot z^{2\alpha-(k+1)} e^{-\xi z^{-1}} \, \mathrm{d}z.$$

We shall be assuming $\alpha \neq 1$. We observe from Lemma 3.1 that for $\alpha \neq 1$, $P(y(z))$ is in fact a linear combination of terms of the form z^u, $(1-z)^u$ and $z_0^u B^-(1-z_0, v, w)$.

To compute J we observe that, by partial integration,

$$\begin{aligned}\int_0^1 z^{u+2\alpha-1} B^-(1-z; v, w) \, \mathrm{d}z &= \left[\frac{z^{u+2\alpha}}{u+2\alpha} B^-(1-z; v, w)\right]_0^1 + \frac{1}{u+2\alpha} \int_0^1 z^{u+2\alpha+w-1}(1-z)^{v-1} \, \mathrm{d}z \\ &= \frac{1}{u+2\alpha} B(u+w+2\alpha, v)\end{aligned}$$

where we used that $\frac{\partial}{\partial z} B^-(1-z; v, w) = -z^{w-1}(1-z)^{v-1}$. This takes care of the two integrands involving the Beta function in $P(y)$. The other integrals are easily computed and yield the following expression for J (note that it only depends on α but not on ν)

$$\begin{aligned}J = &\frac{2+4\alpha+13\alpha^2-34\alpha^3-12\alpha^4+24\alpha^5}{16(\alpha-1)^2\alpha(\alpha+1)(2\alpha+1)} + \frac{2^{-1-4\alpha}}{(\alpha-1)^2} \\ &+ \frac{(\alpha-1/2)(B(2\alpha, 2\alpha+1) + B^-(1/2; 1+2\alpha, -2+2\alpha))}{2(\alpha-1)(3\alpha-1)}\end{aligned}$$

We proceed to work out $I^{(k)}$. For this we will compute the integrals involving terms in $P(y(z))$ of the form z^u, $(1-z)^u$ and $B(1-z, v, w)$ separately. We first point out that for any $0 \leq a < b \leq 1$

$$\begin{aligned}\int_a^b z^{u+2\alpha-(k+1)} e^{-\xi z^{-1}} \, \mathrm{d}z &= \xi^{u+2\alpha-k} \int_{\xi/b}^{\xi/a} t^{k-1-2\alpha-u} e^{-t} \, \mathrm{d}t \\ &= \xi^{u+2\alpha-k} \left(\Gamma^+(k-2\alpha-u; \xi/b) - \Gamma^+(k-2\alpha-u; \xi/a)\right),\end{aligned}$$

In particular

$$\int_0^1 z^{u+2\alpha-k-1} e^{-\xi z^{-1}} \, \mathrm{d}z = \xi^{u+2\alpha-k} \Gamma^+(k-2\alpha-u; \xi) \tag{26}$$

where Γ^+ denotes the (upper) incomplete gamma function, and we've used the subsitution $t = \xi/z$ which gives $\mathrm{d}z = -\xi_{\alpha,\nu} t^{-2} \, \mathrm{d}t$. (And of course it is understood that $\xi/0 = \infty$). This takes care of the integrals of all terms in $P(y(z))$ of the form z^u.

Next we will consider the integrals over the terms in $P(y(z))$ of the form $(1-z)^u$. For this we need the hypergeometric U-function (also called Tricomi's confluent hypergeometric function), which has the integral representation

$$U(a, b, z) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} \, \mathrm{d}t$$

which holds for $a, b, z \in \mathbb{C}$, $b \notin \mathbb{Z}_{\leq 0}$, $Re(a), Re(z) > 0$, see [7, p.255]. Applying the change of variables $t = \frac{1-s}{s}$ (i.e. $\mathrm{d}t = -s^{-2} \, \mathrm{d}s$ and $s = \frac{1}{t+1}$) yields

$$U(a, b, z) = \frac{e^z}{\Gamma(a)} \int_0^1 s^{-b} (1-s)^{a-1} e^{-z/s} ds$$

Plugging in $a = 2\alpha+1 > 0$, $b = -2\alpha+k+1$, $z = \xi > 0$, then gives

$$\int_0^1 z_0^{2\alpha-k-1} e^{-\xi/z_0} (1-z_0)^{2\alpha} dz_0 = \Gamma(2\alpha+1) e^{-\xi} U(2\alpha+1, 1+k-2\alpha, \xi) \tag{27}$$

Finally we need to deal with the terms in $P(y(z))$ that involve the incomplete Beta-function. Let $a, c \in \mathbb{R}$, $\xi, b > 0$ positive real numbers. Using the integral definition of the incomplete beta function, the change of variables $s = 1-t$ gives:

$$\int_0^1 z^a e^{-\xi/z} B^-(1-z; b, c) \, \mathrm{d}z = \int_0^1 z^a e^{-\xi/z} \int_0^{1-z} t^{b-1}(1-t)^{c-1} \, \mathrm{d}t \, \mathrm{d}z$$

$$= \int_0^1 z^a e^{-\xi/z} \int_z^1 s^{c-1}(1-s)^{b-1}\, \mathrm{d}s\, \mathrm{d}z$$

Then changing the order of integration and using the substitution $u = \xi/z$ and recognizing the upper incomplete Γ-function yields

$$\begin{aligned}
&\int_0^1 z^a e^{-\xi/z} \int_z^1 s^{c-1}(1-s)^{b-1}\, \mathrm{d}s\, \mathrm{d}z \\
&\quad = \int_0^1 \int_0^s z^a e^{-\xi/z}\, \mathrm{d}z\, s^{c-1}(1-s)^{b-1}\, \mathrm{d}s \\
&\quad = \int_0^1 \int_{\xi/s}^\infty \xi^{a+1} u^{-a-2} e^{-u}\, \mathrm{d}u\, s^{c-1}(1-s)^{b-1}\, \mathrm{d}s \\
&\quad = \xi^{a+1} \int_0^1 \Gamma^+(-a-1, \xi/s) s^{c-1}(1-s)^{b-1}\, \mathrm{d}s. \qquad (28)
\end{aligned}$$

To compute this last integral we make use of the fact that the incomplete Γ-function has a representation in terms of Meijer's G-function (see Lemma A.1 in Appendix A)

$$\Gamma^+(-a-1, \xi/s) = G^{2,0}_{1,2}\left(\begin{matrix} 1 \\ -a-1, 0 \end{matrix} \middle| \frac{\xi}{s} \right),$$

which holds for any $a \in \mathbb{R}$ and $s > 0$ (that for a fixed second argument, the upper incomplete Γ-function is entire in the first argument, see [10, pp. 899, 1032ff.]). We can now evaluate the integral in (28) using several identities for Meijer's G-function. First, inserting the expression for the incomplete Gamma-function into (28) gives

$$\xi^{a+1} \int_0^1 s^{c-1}(1-s)^{b-1} G^{2,0}_{1,2}\left(\begin{matrix} 1 \\ -a-1, 0 \end{matrix} \middle| \frac{\xi}{s} \right) \mathrm{d}s$$

Next we apply the inversion identity for Meijer's G-function (see [7, p. 209, 5.3.1.(9))]) to get

$$\xi^{a+1} \int_0^1 s^{c-1}(1-s)^{b-1} G^{0,2}_{2,1}\left(\begin{matrix} 2+a, 1 \\ 0 \end{matrix} \middle| \frac{s}{\xi} \right) \mathrm{d}s$$

This expression is actually the Euler transform of Meijer's G-function (see [7, p. 214, 5.5.2.(5)]) and (as the conditions $2+1 < 2(0+2)$ and $|\arg(\xi^{-1})| < \frac{\pi}{2}$ (as $\xi > 0$) and $1-c-b < 1-c$ (as $b > 0$) are satisfied) it equals

$$\xi^{a+1} \Gamma(b) G^{0,3}_{3,2}\left(\begin{matrix} 1-c, 2+a, 1 \\ 0, 1-c-b \end{matrix} \middle| \xi^{-1} \right)$$

Using again the inversion identity for Meijer's G-function we now get

$$\xi^{a+1} \Gamma(b) G^{3,0}_{2,3}\left(\begin{matrix} 1, b+c \\ c, -1-a, 0 \end{matrix} \middle| \xi \right)$$

Finally, plugging in $a = 6\alpha - k - 3$, $b = 2\alpha$, $c = 3 - 4\alpha$ we obtain

$$\int_0^1 z^a e^{-\xi/z} B^-(1-z; b, c)\, \mathrm{d}z = \xi^{6\alpha-k-2} \Gamma(2\alpha) G^{3,0}_{2,3}\left(\begin{matrix} 1, 3-2\alpha \\ 3-4\alpha, -6\alpha+k+2, 0 \end{matrix} \middle| \xi \right) \qquad (29)$$

Using equation (26), (27) and (29) we get

$$\begin{aligned}
I^{(k)} = \frac{\xi^{2\alpha}}{4k!(\alpha-1)} \Bigg(&-\Gamma^+(k-2\alpha, \xi) - 2\frac{\alpha(\alpha-1/2)^2 \xi^2 \Gamma^+(k-2\alpha-2, \xi)}{(\alpha-1)} \\
&+ 8\alpha(\alpha-1/2)\xi\Gamma^+(k-2\alpha-1, \xi)
\end{aligned}$$

$$
\begin{aligned}
&+4\xi^{4\alpha-2}\Gamma^{+}(k-6\alpha+2,\xi)\left(\frac{2^{-4\alpha}(3\alpha-1)}{(\alpha-1)}+(\alpha-1/2)B^{-}(1/2;1+2\alpha,-2+2\alpha)\right)\\
&+\xi^{k-2\alpha}\Gamma(2\alpha+1)e^{-\xi}U(2\alpha+1,1+k-2\alpha,\xi)\\
&-\xi^{4\alpha-2}\Gamma(2\alpha+1)G^{3,0}_{2,3}\left(\left.\begin{matrix}1,3-2\alpha\\3-4\alpha,-6\alpha+k+2,0\end{matrix}\right|\xi\right)\Bigg)
\end{aligned}
$$

With the expressions for J and $I^{(k)}$ and using $\Gamma^{*}(q,z)=\Gamma^{+}(q+1,z)+\Gamma^{+}(q,z)$ we now obtain, after some algebra, the expression for γ

$$
\begin{aligned}
\gamma &= J-I^{(0)}-I^{(1)}\\
&=\frac{2+4\alpha+13\alpha^2-34\alpha^3-12\alpha^4+24\alpha^5}{16(\alpha-1)^2\alpha(\alpha+1)(2\alpha+1)}+\frac{2^{-1-4\alpha}}{(\alpha-1)^2}\\
&+\frac{(\alpha-1/2)(B(2\alpha,2\alpha+1)+B^{-}(1/2;1+2\alpha,-2+2\alpha))}{2(\alpha-1)(3\alpha-1)}\\
&-\frac{\xi^{2\alpha}}{4(\alpha-1)}\Bigg(-\Gamma^{+}(-2\alpha,\xi)-2\frac{\alpha(\alpha-1/2)^2\xi^2\Gamma^{+}(-2\alpha-2,\xi)}{(\alpha-1)}\\
&\quad+8\alpha(\alpha-1/2)\xi\Gamma^{+}(-2\alpha-1,\xi)\\
&\quad+4\xi^{4\alpha-2}\Gamma^{+}(-6\alpha+2,\xi)\left(\frac{2^{-4\alpha}(3\alpha-1)}{(\alpha-1)}+(\alpha-1/2)B^{-}(1/2;1+2\alpha,-2+2\alpha)\right)\\
&\quad+\xi^{-2\alpha}\Gamma(2\alpha+1)e^{-\xi}U(2\alpha+1,1-2\alpha,\xi)\\
&\quad-\xi^{4\alpha-2}\Gamma(2\alpha+1)G^{3,0}_{2,3}\left(\left.\begin{matrix}1,3-2\alpha\\3-4\alpha,-6\alpha+2,0\end{matrix}\right|\xi\right)\Bigg)\\
&-\frac{\xi^{2\alpha}}{4(\alpha-1)}\Bigg(-\Gamma^{+}(1-2\alpha,\xi)-2\frac{\alpha(\alpha-1/2)^2\xi^2\Gamma^{+}(-2\alpha-1,\xi)}{(\alpha-1)}\\
&\quad+8\alpha(\alpha-1/2)\xi\Gamma^{+}(1-2\alpha-1,\xi)\\
&\quad+4\xi^{4\alpha-2}\Gamma^{+}(1-6\alpha+2,\xi)\left(\frac{2^{-4\alpha}(3\alpha-1)}{(\alpha-1)}+(\alpha-1/2)B^{-}(1/2;1+2\alpha,-2+2\alpha)\right)\\
&\quad+\xi^{1-2\alpha}\Gamma(2\alpha+1)e^{-\xi}U(2\alpha+1,2-2\alpha,\xi)\\
&\quad-\xi^{4\alpha-2}\Gamma(2\alpha+1)G^{3,0}_{2,3}\left(\left.\begin{matrix}1,3-2\alpha\\3-4\alpha,-6\alpha+3,0\end{matrix}\right|\xi\right)\Bigg)\\
&=\frac{2+4\alpha+13\alpha^2-34\alpha^3-12\alpha^4+24\alpha^5}{16(\alpha-1)^2\alpha(\alpha+1)(2\alpha+1)}+\frac{2^{-1-4\alpha}}{(\alpha-1)^2}\\
&+\frac{(\alpha-1/2)(B(2\alpha,2\alpha+1)+B^{-}(1/2;1+2\alpha,-2+2\alpha))}{2(\alpha-1)(3\alpha-1)}\\
&+\frac{\xi^{2\alpha}\Gamma^{*}(-2\alpha,\xi)}{4(\alpha-1)}+\frac{\xi^{2\alpha+2}\alpha(\alpha-1/2)^2\Gamma^{*}(-2\alpha-2,\xi)}{2(\alpha-1)^2}\\
&-\frac{\xi^{2\alpha+1}\alpha(2\alpha-1)\Gamma^{*}(-2\alpha-1,\xi)}{(\alpha-1)}-\frac{\xi^{6\alpha-2}2^{-4\alpha}(3\alpha-1)\Gamma^{*}(-6\alpha+2,\xi)}{(\alpha-1)^2}\\
&-\frac{\xi^{6\alpha-2}(\alpha-1/2)B^{-}(1/2;1+2\alpha,-2+2\alpha)\Gamma^{*}(-6\alpha+2,\xi)}{(\alpha-1)}\\
&-\frac{e^{-\xi}\Gamma(2\alpha+1)\left(U(2\alpha+1,1-2\alpha,\xi)+U(2\alpha+1,2-2\alpha,\xi)\right)}{4(\alpha-1)}\\
&+\frac{\xi^{6\alpha-2}\Gamma(2\alpha+1)\left(G^{3,0}_{2,3}\left(\xi\left|\begin{matrix}1,3-2\alpha\\3-4\alpha,-6\alpha+2,0\end{matrix}\right.\right)+G^{3,0}_{2,3}\left(\xi\left|\begin{matrix}1,3-2\alpha\\3-4\alpha,-6\alpha+3,0\end{matrix}\right.\right)\right)}{4(\alpha-1)}.
\end{aligned}
$$

and note that this equals the expression in Theorem 1.1.

Similarly, we get

$$\begin{aligned}
\gamma(k) &= \frac{I^{(k)}}{p_k} \\
&= \frac{1}{8\alpha(\alpha-1)\Gamma^+(k-2\alpha,\xi)}\Bigg(-\Gamma^+(k-2\alpha,\xi) - 2\frac{\alpha(\alpha-1/2)^2\xi^2\Gamma^+(k-2\alpha-2,\xi)}{(\alpha-1)} \\
&\quad +8\alpha(\alpha-1/2)\xi\Gamma^+(k-2\alpha-1,\xi) \\
&\quad +4\xi^{4\alpha-2}\Gamma^+(k-6\alpha+2,\xi)\left(\frac{2^{-4\alpha}(3\alpha-1)}{(\alpha-1)} + (\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)\right) \\
&\quad +\xi^{k-2\alpha}\Gamma(2\alpha+1)e^{-\xi}U(2\alpha+1,1+k-2\alpha,\xi) \\
&\quad -\xi^{4\alpha-2}\Gamma(2\alpha+1)G^{3,0}_{2,3}\left(\begin{matrix}1,3-2\alpha\\ 3-4\alpha,-6\alpha+k+2,0\end{matrix}\middle|\xi\right)\Bigg),
\end{aligned}$$

which equals the expression in Theorem 1.2.

3.2.4 Explicit expressions for $\gamma, \gamma(k)$ when $\alpha = 1$.

Although we've already established that $\gamma, \gamma(k)$ can be obtained at $\alpha = 1$ by taking the $\alpha \to 1$ limit of the expression obtained for $\alpha = 1$, it is still helpful to derive an alternate, more explicit expression. This is what we will do in the current section. We will prove

Proposition 3.5. *If $\alpha = 1$ then*

$$\begin{aligned}
\gamma &= \frac{575-12\pi^2}{576} + \frac{\eta^4(7+\pi^2)\Gamma^*(-4,\eta)}{4} \\
&\quad - \frac{1}{2}\int_0^1 (1-4z+3z^3)\log(1-z)(z+\eta)e^{-\eta/z}\,\mathrm{d}z \\
&\quad - \int_0^1 \mathrm{Li}_2(z)(z^3+\eta z^2)e^{-\eta/z}\,\mathrm{d}z,
\end{aligned}$$

and

$$\begin{aligned}
\gamma(k) &= \frac{9\eta^3}{2k!}\Gamma^+(k-3,\eta) - \frac{\xi^4}{k!}\frac{7+\pi^2}{4}\Gamma^+(k-4,\eta) \\
&\quad + \frac{\eta^k}{2k!}\int_0^1 (1-4z+3z^2)\ln(1-z)z^{1-k}e^{-\eta/z}\,\mathrm{d}z \\
&\quad + \frac{\eta^k}{k!}\int_0^1 z^{3-k}\,\mathrm{Li}_2(z)e^{-\eta/z}\,\mathrm{d}z,
\end{aligned}$$

with $\eta = 4\nu/\pi$ and $\mathrm{Li}_2(z) = \sum_{t=1}^\infty z^t/t^2$, the dilogarithm function.

Naturally, the proof proceeds by proving the analogue of Lemma 3.1:

Lemma 3.6. *If $\alpha = 1$, then*

$$P(y) = \frac{9}{4}e^{-\frac{1}{2}y} + \frac{1-4e^{-\frac{1}{2}y}+3e^{-y}}{4}\ln(1-e^{-\frac{1}{2}y}) - \frac{7+\pi^2}{8}e^{-y} + \frac{1}{2}e^{-y}\,\mathrm{Li}_2(e^{-y})$$

where $\mathrm{Li}_2(z) = \int_0^z \frac{\ln(1-t)}{t}dt$ is the dipolylogarithm function.

Proof. **Tobias:** Warning! this has not been properly proofread by me, or anyone else. We want to compute the limit $\lim_{\alpha\to 1} P_\alpha(y_0(z_0))$. For $\alpha \neq 1$, we label the terms as follows:

$P_\alpha(y_0(z_0))$

$$= \frac{1}{\alpha - 1}\left(s_1(\alpha, z_0) + s_2(\alpha, z_0) + \frac{1}{\alpha - 1}(s_3(\alpha, z_0) + s_4(\alpha, z_0)) + s_5(\alpha, z_0) + s_6(\alpha, z_0) + s_7(\alpha, z_0)\right)$$

where

$$\begin{aligned}
s_1(\alpha, z_0) &= -\frac{1}{8\alpha} \\
s_2(\alpha, z_0) &= (\alpha - 1/2)z_0 \\
s_3(\alpha, z_0) &= -\frac{(\alpha - 1/2)^2 z_0^2}{4} \\
s_4(\alpha, z_0) &= z_0^{-2+4\alpha}\frac{2^{-4\alpha-1}(3\alpha - 1)}{\alpha} \\
s_5(\alpha, z_0) &= z_0^{-2+4\alpha}\frac{(\alpha - 1/2)B^-(1/2; 1 + 2\alpha, -2 + 2\alpha)}{2\alpha} \\
s_6(\alpha, z_0) &= \frac{(1 - z_0)^{2\alpha}}{8\alpha} \\
s_7(\alpha, z_0) &= -\frac{z_0^{4\alpha-2}B^-(1 - z_0; 2\alpha, 3 - 4\alpha)}{4}
\end{aligned}$$

Now, we consider the functions $s_i(\alpha) = s_i(\alpha, z_0)$ as functions of α only and compute their Taylor expansion at $\alpha = 1$, for $i \in \{1, 2, 5, 6, 7\}$ up to linear and for $i \in \{3, 4\}$ up to quadratic order, i.e. we write $s_i(\alpha) = s_i(1) + s_i'(1)(\alpha - 1) + o(\alpha - 1)$ for $i \in \{1, 2, 5, 6, 7\}$ and $s_i(\alpha) = s_i(1) + s_i'(1)(\alpha - 1) + \frac{s_i''(1)}{2}(\alpha - 1)^2 + o((\alpha - 1)^2)$ for $i \in \{3, 4\}$. Using these expansions, we can rewrite

$$\begin{aligned}
P(y_0(z_0)) = \frac{1}{\alpha - 1}&\left(\sum_{i\in\{1,2,5,6,7\}} s_i(1) + \sum_{i\in\{1,2,5,6,7\}} s_i'(1)(\alpha - 1) + o(\alpha - 1)\right. \\
&\left.+\frac{1}{\alpha - 1}(s_3(1) + s_4(1) + (s_3'(1) + s_4'(1))(\alpha - 1) + \frac{1}{2}(s_3''(1) + s_4''(1))(\alpha - 1)^2 + o((\alpha - 1)^2))\right)
\end{aligned}$$

In order to continue, we compute:

$$\begin{aligned}
s_1(\alpha) &= -\frac{1}{8} + \frac{1}{8}(\alpha - 1) + o(\alpha - 1) \\
s_2(\alpha) &= \frac{1}{2}z_0 + z_0(\alpha - 1) + o(\alpha - 1) \\
s_3(\alpha) &= -\frac{1}{16}z_0^2 - \frac{1}{4}z_0^2(\alpha - 1) - \frac{1}{4}z_0^2(\alpha - 1)^2 + o((\alpha - 1)^2) \\
s_4(\alpha) &= \frac{1}{16}z_0^2 + \frac{z_0^2}{4}\left(\frac{1}{8} + \ln\frac{z_0}{2}\right)(\alpha - 1) \\
&\quad + \frac{z_0^2}{8}\left(4\left(\ln\frac{z_0}{2}\right)^2 + \ln\frac{z_0}{2} - \frac{1}{4}\right)(\alpha - 1)^2 + o((\alpha - 1)^2) \\
s_5(\alpha) &= \frac{z_0^2}{4}B^-(1/2; 3, 0) + z_0^2\left(\left(\ln(z_0) + \frac{1}{4}\right)B^-(1/2; 3, 0)\right. \\
&\quad \left.+1/2\int_0^{\frac{1}{2}} \ln(t)t^2(1 - t)^{-1} + \ln(1 - t)t^2(1 - t)^{-1}dt\right)(\alpha - 1) + o(\alpha - 1) \\
s_6(\alpha) &= \frac{(1 - z_0)^2}{8} + \frac{(1 - z_0)^2}{4}(\ln(1 - z_0) - 1/2)(\alpha - 1) + o(\alpha - 1) \\
s_7(\alpha) &= -\frac{z_0^2}{4}B^-(1 - z_0; 2, -1) - z_0^2\left(\ln(z_0)B^-(1 - z_0; 2, -1)\right. \\
&\quad \left.+\int_0^{1-z_0} 1/2\ln(t)t(1 - t)^{-2} - t\ln(1 - t)(1 - t)^{-2}dt\right)(\alpha - 1) + o(\alpha - 1)
\end{aligned}$$

Based on this we see that

$$s_3(1) + s_4(1) = -\frac{1}{16}z_0^2 + \frac{1}{16}z_0^2 = 0$$

and

$$\begin{aligned}
&\sum_{i\in\{1,2,5,6,7\}} s_i(1) + s_3'(1) + s_4'(1) \\
&= -\frac{1}{8} + \frac{1}{2}z_0 - \frac{1}{4}z_0^2 + \frac{z_0^2}{32} + \frac{z_0^2}{4}\ln(\frac{z_0}{2}) + \frac{z_0^2}{4}B^-(1/2;3,0) + \frac{(1-z_0)^2}{8} - \frac{z_0^2}{4}B^-(1-z_0;2,-1) \\
&= -\frac{1}{8} + \frac{1}{2}z_0 - \frac{1}{4}z_0^2 + \frac{z_0^2}{32} + \frac{z_0^2}{4}\ln(z_0) - \frac{z_0^2}{4}\ln 2 - \frac{5z_0^2}{32} + \frac{z_0^2}{4}\ln 2 \\
&\qquad + \frac{1}{8} - \frac{z_0}{4} + \frac{z_0^2}{8} + \frac{z_0^2}{4} - \frac{z_0}{4} - \frac{z_0^2}{4}\ln z_0 \\
&= 0
\end{aligned}$$

Finally, it follows that

$$P(y_0(z_0)) = \sum_{i\in\{1,2,5,6,7\}} s_i'(1) + \frac{1}{2}(s_3''(1) + s_4''(1)) + o(1)$$

Therefore, the desired value of $\lim_{\alpha\to 1} P(y_0(z_0))$ is given by

$$\begin{aligned}
&\sum_{i\in\{1,2,5,6,7\}} s_i'(1) + \frac{1}{2}(s_3''(1) + s_4''(1)) \\
&= \frac{1}{8} + z_0 - \frac{z_0^2}{4} + \frac{z_0^2}{8}(4(\ln\frac{z_0}{2})^2 + \ln\frac{z_0}{2} - \frac{1}{4}) + \frac{(1-z_0)^2}{4}(\ln(1-z_0) - 1/2) \\
&\quad + z_0^2\left(\left(\ln(z_0) + \frac{1}{4}\right)B^-(1/2;3,0) + 1/2\int_0^{\frac{1}{2}} \ln(t)t^2(1-t)^{-1} + \ln(1-t)t^2(1-t)^{-1}dt\right) \\
&\quad - z_0^2\left(\ln(z_0)B^-(1-z_0;2,-1) + \int_0^{1-z_0} 1/2\ln(t)t(1-t)^{-2} - t\ln(1-t)(1-t)^{-2}dt\right) \\
&= \frac{1}{8} + z_0 - \frac{z_0^2}{4} + \frac{z_0^2}{2}(\ln\frac{z_0}{2})^2 + \frac{z_0^2}{8}\ln\frac{z_0}{2} - \frac{z_0^2}{32} \\
&\quad - \frac{5}{8}z_0^2\ln(z_0) + z_0^2\ln(z_0)\ln 2 - \frac{5z_0^2}{32} + \frac{z_0^2\ln 2}{4} \\
&\quad + z_0^2/2\int_0^{\frac{1}{2}} \ln(t)t^2(1-t)^{-1} + \ln(1-t)t^2(1-t)^{-1}\,\mathrm{d}t \\
&\quad + \frac{(1-z_0)^2}{4}\ln(1-z_0) - \frac{1}{8} + \frac{z_0}{4} - \frac{z_0^2}{8} \\
&\quad + z_0^2\ln(z_0) - z_0\ln z_0 - z_0^2(\ln z_0)^2 - z_0^2\int_0^{1-z_0} 1/2\ln(t)t(1-t)^{-2} - t\ln(1-t)(1-t)^{-2}dt \\
&= \frac{5}{4}z_0 - \frac{9}{16}z_0^2 + \frac{z_0^2}{2}(\ln\frac{z_0}{2})^2 + \frac{z_0^2}{8}\ln\frac{z_0}{2} + \frac{(1-z_0)^2}{4}\ln(1-z_0) \\
&\quad + \frac{3}{8}z_0^2\ln(z_0) + z_0^2\ln(z_0)\ln 2 + \frac{z_0^2\ln 2}{4} + z_0^2/2\int_0^{\frac{1}{2}} \ln(t)t^2(1-t)^{-1} + \ln(1-t)t^2(1-t)^{-1}dt \\
&\quad - z_0\ln z_0 - z_0^2(\ln z_0)^2 - z_0^2\int_0^{1-z_0} 1/2\ln(t)t(1-t)^{-2} - t\ln(1-t)(1-t)^{-2}dt \\
&= \frac{5}{4}z_0 - \frac{9}{16}z_0^2 + \frac{z_0^2}{2}(\ln\frac{z_0}{2})^2 + \frac{z_0^2}{8}\ln\frac{z_0}{2} + \frac{(1-z_0)^2}{4}\ln(1-z_0)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{3}{8}z_0^2\ln(z_0)+z_0^2\ln(z_0)\ln 2+\frac{z_0^2\ln 2}{4}+z_0^2/2(11/8-1/4\ln 2-3/2\ln(2)^2-\mathrm{Li}_2(1/2))\\
&-z_0\ln z_0-z_0^2(\ln z_0)^2+z_0(1+\frac{1}{2}(2-z_0)\ln(z_0)+\frac{1}{2}z_0\ln(z_0)^2-\frac{1}{2}(1-z_0)\ln(1-z_0)\\
&+\frac{1}{2}z_0\,\mathrm{Li}_2(z_0))-z_0^2-\frac{1}{2}z_0^2\,\mathrm{Li}_2(1)\\
=&\frac{9}{4}z_0-\frac{25}{16}z_0^2+\frac{z_0^2}{2}(\ln\frac{z_0}{2})^2+\frac{z_0^2}{8}\ln\frac{z_0}{2}+\frac{(1-z_0)^2}{4}\ln(1-z_0)\\
&-\frac{1}{8}z_0^2\ln(z_0)+z_0^2\ln(z_0)\ln 2+\frac{z_0^2\ln 2}{4}+z_0^2/2(11/8-1/4\ln 2-3/2\ln(2)^2\\
&-\mathrm{Li}_2(1/2)-\mathrm{Li}_2(1)+\mathrm{Li}_2(z_0))-\frac{1}{2}z_0^2(\ln z_0)^2-\frac{1}{2}z_0(1-z_0)\ln(1-z_0)
\end{aligned}
$$

where we used that

$$
z_0^2/2\int_0^{\frac{1}{2}}\ln(t)t^2(1-t)^{-1}+\ln(1-t)t^2(1-t)^{-1}\,\mathrm{d}t=11/8-1/4\ln 2-3/2\ln(2)^2-\mathrm{Li}_2(1/2),
$$

and

$$
\begin{aligned}
&z_0^2\int_0^{1-z_0}1/2\ln(t)t(1-t)^{-2}-t\ln(1-t)(1-t)^{-2}dt\\
=&-\frac{1}{z_0}(1+\frac{1}{2}(2-z_0)\ln(z_0)+\frac{1}{2}z_0\ln(z_0)^2-\frac{1}{2}(1-z_0)\ln(1-z_0)+\frac{1}{2}z_0Li_2(z_0))+1+\frac{1}{2}\,\mathrm{Li}_2(1).
\end{aligned}
$$

By expanding the squares and collecting terms, the last expression can be simplified to

$$
\begin{aligned}
&\frac{9}{4}z_0+\frac{1-4z_0+3z_0^2}{4}\ln(1-z_0)+z_0^2\left(-7/8-\frac{\ln(2)^2+2\,\mathrm{Li}_2(1/2)+2\,\mathrm{Li}_2(1)}{4}\right)+\frac{1}{2}z_0^2\,\mathrm{Li}_2(z)\\
=&\frac{9}{4}z_0+\frac{1-4z_0+3z_0^2}{4}\ln(1-z_0)-\frac{7+\pi^2}{8}z_0^2+\frac{1}{2}z_0^2\,\mathrm{Li}_2(z)
\end{aligned}
$$

which finishes the computation. □

Proof of Proposition 3.5: It suffices to find the value of J and $I^{(k)}$ at $\alpha=1$. We can do this by computing the integrals with the expression for $P(y)$ that we found for $\alpha=1$, i.e.

$$
\begin{aligned}
J&=2\alpha\int_0^1\left(\frac{9}{4}z+\frac{1-4z+3z^2}{4}\ln(1-z)-\frac{7+\pi^2}{8}z^2+\frac{1}{2}z^2\,\mathrm{Li}_2(z)\right)z^{2\alpha-1}\,\mathrm{d}z\\
&=\frac{575-12\pi^2}{576}
\end{aligned}
$$

and

$$
\begin{aligned}
I^{(k)}&=\frac{2\alpha\xi^k}{k!}\int_0^1\left(\frac{9}{4}z+\frac{1-4z+3z^2}{4}\ln(1-z)-\frac{7+\pi^2}{8}z^2+\frac{1}{2}z^2\,\mathrm{Li}_2(z)\right)z^{2\alpha-k-1}e^{-\xi/z}\,\mathrm{d}z\\
&=\frac{2\eta^k}{k!}\int_0^1\left(\frac{9}{4}z+\frac{1-4z+3z^2}{4}\ln(1-z)-\frac{7+\pi^2}{8}z^2+\frac{1}{2}z^2\,\mathrm{Li}_2(z)\right)z^{1-k}e^{-\eta/z}\,\mathrm{d}z\\
&=\frac{9\eta^k}{2k!}\eta^{3-k}\Gamma^+(k-3,\eta)-\frac{\eta^k}{k!}\frac{7+\pi^2}{4}\eta^{4-k}\Gamma^+(k-4,\eta)\\
&\quad+\frac{\eta^k}{2k!}\int_0^1(1-4z+3z^2)\ln(1-z)z^{1-k}e^{-\eta/z}dz+\frac{\eta^k}{k!}\int_0^1 z^{3-k}\,\mathrm{Li}_2(z)e^{-\eta/z}\,\mathrm{d}z\\
&=\frac{9\eta^3}{2k!}\Gamma^+(k-3,\eta)-\frac{\eta^4}{k!}\frac{7+\pi^2}{4}\Gamma^+(k-4,\eta)
\end{aligned}
$$

$$+\frac{\eta^k}{2k!}\int_0^1 (1-4z+3z^2)\ln(1-z)z^{1-k}e^{-\eta/z}dz + \frac{\eta^k}{k!}\int_0^1 z^{3-k}\,\mathrm{Li}_2(z)e^{-\eta/z}\,\mathrm{d}z$$

where $\eta = \frac{4\nu}{\pi}$ and and $\mathrm{Li}_2(z) = \sum_{t=1}^{\infty} z^t/t^2$, the dilogarithm function. Plugging this into (14) and (13) yields the expressions in the statement of the proposition. ■

3.3 The proof of Proposition 1.4

Instead of extracting the scaling of $\gamma(k)$ from its explicit expression, it turns out to be more convenient to derive it directly. Recall that

$$\gamma(k) = \frac{\int_0^\infty \rho(y,k)P(y)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}.$$

The asymptotic behavior for the denominator follows from (11). Hence, the main term to consider is the numerator

$$\int_0^\infty P(y)\,\rho(y,k)\alpha e^{-\alpha y}\,dy,$$

and in particular the function $P(y)$. We therefore start with establishing the asymptotic behavior of the latter. First we combine (11) and (12) to obtain the following scaling result

$$\frac{\int_0^\infty e^{-\beta y}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y} \sim \xi_{\alpha,\nu}^{2\beta}k^{-2\beta}. \tag{30}$$

Proposition 3.7 (Asymptotic behavior of $P(y)$)**.** *Let $\alpha > \frac{1}{2}$, $\nu > 0$ and c_α as defined in Proposition 1.4 Then, as $y \to \infty$,*

1. *for $\frac{1}{2} < \alpha < \frac{3}{4}$,*
$$P(y) \sim e^{-\frac{y}{2}(4\alpha-2)}c_\alpha \xi^{4\alpha-2},$$
2. *for $\alpha = \frac{3}{4}$,*
$$P(y) \sim \frac{y}{2}e^{-\frac{y}{2}},$$
3. *and for $\alpha > \frac{3}{4}$,*
$$P(y) \sim e^{-\frac{y}{2}}\frac{\alpha-\frac{1}{2}}{\alpha-\frac{3}{4}}.$$

Proof. We shall deal with each of the three cases for α separately.

Proof for $1/2 < \alpha < 3/4$ By Lemma 3.1 we get that

$$\begin{aligned} e^{(4\alpha-2)\frac{y}{2}}P(y) &= \frac{2^{-4\alpha-1}(3\alpha-1)}{\alpha(\alpha-1)^2} + \frac{(\alpha-\frac{1}{2})B^-(\frac{1}{2};1+2\alpha,-2+2\alpha)}{2(\alpha-1)\alpha} - \frac{B^-(1-e^{-\frac{y}{2}};2\alpha,3-4\alpha)}{4(\alpha-1)} \\ &+ \frac{e^{(4\alpha-2)\frac{y}{2}}}{8(\alpha-1)\alpha}\left((1-e^{-\frac{y}{2}})^{2\alpha}-1\right) + \frac{\alpha-\frac{1}{2}}{\alpha-1}e^{(4\alpha-3)\frac{y}{2}} - \frac{(\alpha-\frac{1}{2})^2}{4(\alpha-1)^2}e^{4(\alpha-1)\frac{y}{2}}. \end{aligned}$$

Because for any $b < 1$, $B^-(1-z,a,b)$ converges to $B^-(a,b) < \infty$ as $z \to 0$, we get that as $y \to \infty$, the first line is asymptotically equivalent to

$$\frac{3\alpha-1}{2^{4\alpha+1}\alpha(\alpha-1)^2} + \frac{(\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)}{2(\alpha-1)\alpha} - \frac{B(2\alpha,3-4\alpha)}{4(\alpha-1)} = c_\alpha\xi^{-(4\alpha-2)},$$

with c_α as defined in Proposition (1.4). The proof now follows since for $1/2 < \alpha < 3/4$, the remaining three terms go to zero as $y \to \infty$.

Proof for $\alpha = 3/4$ Similar to the previous case we use Lemma 3.1 to obtain

$$\begin{aligned}\frac{2}{y}e^{\frac{y}{2}}P(y) &= \frac{2}{y}\frac{B^-(1-e^{-\frac{y}{2}},2\alpha,3-4\alpha)}{4(\alpha-1)} \\ &+ \frac{2}{y}\frac{e^{\frac{y}{2}}\left((1-e^{-\frac{y}{2}})^{2\alpha}-1\right)}{8(\alpha-1)\alpha} + \frac{(\alpha-\frac{1}{2})}{(\alpha-1)y} - \frac{(\alpha-\frac{1}{2})^2e^{-\frac{y}{2}}}{4(\alpha-1)^2y} \\ &+ \frac{2}{y}\left(\frac{2^{-4\alpha-1}(3\alpha-1)}{\alpha(\alpha-1)^2} + \frac{(\alpha-\frac{1}{2})B^-(\frac{1}{2};1+2\alpha,-2+2\alpha)}{2(\alpha-1)\alpha}\right)\end{aligned}$$

First we note that as $y \to \infty$,

$$e^{\frac{y}{2}}\left(\left(1-e^{-\frac{y}{2}}\right)^{2\alpha}-1\right) \sim -2\alpha, \tag{31}$$

which implies that

$$\lim_{y\to\infty}\frac{2}{y}\frac{e^{\frac{y}{2}}\left((1-e^{-\frac{y}{2}})^{2\alpha}-1\right)}{8(\alpha-1)\alpha} = 0.$$

We can now conclude that all terms in $\frac{2}{y}e^{\frac{y}{2}}P(y)$ except the first one are $o(1)$ as $y \to \infty$. By writing $z = e^{-\frac{y}{2}}$ we can rewrite the first term as

$$\frac{2}{y}\frac{B^-(1-e^{-\frac{y}{2}},2\alpha,3-4\alpha)}{4(\alpha-1)} = -\frac{1}{\log(z)}\frac{B^-(1-z,2\alpha,3-4\alpha)}{4(\alpha-1)}.$$

Since $B^-(1-z,2\alpha,0) \sim -\log(z)$ as $z \to \infty$, see Lemma B.1, it now follows that for $\alpha = 3/4$,

$$\lim_{y\to\infty}\frac{2}{y}\frac{B^-(1-e^{-\frac{y}{2}},2\alpha,3-4\alpha)}{4(\alpha-1)} = \lim_{z\to 0} -\frac{1}{\log(z)}\frac{B^-(1-z,2\alpha,3-4\alpha)}{4(\alpha-1)} = 1.$$

We therefore conclude that

$$P(y) \sim \frac{y}{2}e^{-\frac{y}{2}},$$

as $y \to \infty$.

Proof for $\alpha > 3/4$ We first deal with the case $\alpha = 1$. Here it follows from Proposition 3.5 that

$$\begin{aligned}e^{y/2}P(y) &= \frac{9}{4} + \frac{e^{y/2}\log(1-e^{-y/2})}{4} \\ &\quad -\log(1-e^{-y/2}) + e^{-y/2}\left(\frac{3}{4}\log(1-e^{-y/2}) - \frac{7+\pi^2}{8} + \frac{1}{2}\mathrm{Li}_2(e^{-y})\right) \\ &= 2 + \left(\frac{e^{y/2}\log(1-e^{-y/2})}{4} + 1\right) \\ &\quad -\log(1-e^{-y/2}) + e^{-y/2}\left(\frac{3}{4}\log(1-e^{-y/2}) - \frac{7+\pi^2}{8} + \frac{1}{2}\mathrm{Li}_2(e^{-y})\right)\end{aligned}$$

The last two terms are $o(1)$ as $y \to \infty$, while $2 = (\alpha - 1/2)/(\alpha - 3/4)$ for $\alpha = 1$.

Now we will deal with the case $\alpha > 3/4$ and $\alpha \neq 1$. For simplicity we write

$$Q_\alpha := \frac{2^{-4\alpha-1}(3\alpha-1)}{\alpha(\alpha-1)^2} + \frac{(\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)}{2(\alpha-1)\alpha}.$$

Then, by Lemma 3.1 we get

$$e^{y/2}P(y) = \frac{\alpha-\frac{1}{2}}{\alpha-1} + \frac{e^{\frac{y}{2}}}{8(\alpha-1)\alpha}\left(\left(1-e^{-\frac{y}{2}}\right)^{2\alpha}-1\right)$$

$$
\begin{aligned}
&- e^{-(4\alpha-3)\frac{y}{2}} \frac{B^-(1-e^{-\frac{1}{2}y}; 2\alpha, 3-4\alpha)}{4(\alpha-1)} \\
&+ e^{-(4\alpha-3)\frac{y}{2}} Q_\alpha + \frac{(\alpha-\frac{1}{2})^2}{4(\alpha-1)^2} e^{-\frac{y}{2}}.
\end{aligned}
$$

The first term is constant while the last two terms go to zero as $y \to \infty$. We will therefore focus on the remaining two terms. For the first we have, see (31)

$$
\frac{e^{\frac{y}{2}}}{8(\alpha-1)\alpha} \left(\left(1 - e^{-\frac{y}{2}}\right)^{2\alpha} - 1 \right) \sim \frac{-2\alpha}{8(\alpha-1)\alpha} = -\frac{1}{4(\alpha-1)},
$$

as $y \to \infty$. Finally, writing $z = e^{-\frac{y}{2}}$ we get that

$$
e^{-(4\alpha-3)\frac{y}{2}} B^-(1-e^{-\frac{1}{2}y}; 2\alpha, 3-4\alpha) = z^{4\alpha-3} B^-(1-z, 2\alpha, 3-4\alpha).
$$

Therefore it follows, see Lemma B.1, that

$$
\begin{aligned}
\lim_{y\to\infty} -e^{-(4\alpha-3)\frac{y}{2}} \frac{B^-(1-e^{-\frac{1}{2}y}; 2\alpha, 3-4\alpha)}{4(\alpha-1)} &= \lim_{z\to 0} z^{4\alpha-3} \frac{B^-(1-z, 2\alpha, 3-4\alpha)}{4(\alpha-1)} \\
&= \frac{1}{4(\alpha-1)(4\alpha-3)}.
\end{aligned}
$$

We conclude that as $y \to \infty$

$$
e^{y/2} P(y) \sim \frac{\alpha-\frac{1}{2}}{\alpha-1} - \frac{1}{4(\alpha-1)} - \frac{1}{4(\alpha-1)(4\alpha-3)} = \frac{1-3\alpha+2\alpha^2}{(\alpha-1)(\alpha-\frac{3}{4})} = \frac{\alpha-\frac{1}{2}}{\alpha-\frac{3}{4}},
$$

which finishes the proof. □

With the asymptotic behavior of $P(y)$ we are almost ready to prove Proposition 1.4. First we will prove a result that will allow us to limit the values of y, when performing the integration. For this, fix some $C > 0$ and define

$$
a(k)^{\pm} = 2\log\left(\frac{k \pm C\sqrt{k\log(k)}}{\xi} \vee 1 \right).
$$

We will show that, as $k \to \infty$,

$$
\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y} \, \mathrm{d}y = (1 + o(1)) \int_{a(k)^-}^{a(k)^+} P(y)\rho(y,k)\alpha e^{-\alpha y} \, \mathrm{d}y. \tag{32}
$$

To establish (32) recall that $\mu(y) = \xi e^{\frac{y}{2}}$ and consider $\rho(y,k) = \mathbb{P}(\mathrm{Po}(\mu(y)) = k)$ as a function of y. Then, since $\mu'(y) = \mu(y)/2$, we get that

$$
\frac{\partial \rho(y,k)}{\partial y} = \frac{1}{2}\left(k - \mu(y)\right)\rho(y,k),
$$

which implies that $\rho(y,k)$ attains its maximum at $\mu(y) = k$. Moreover we see that the derivative is strictly positive when $\mu(y) < k$ and strictly negative when $\mu(k) > k$. Since $\mu(a(k)^-) < k$ and $\mu(a(k)^+) > k$, we conclude that $\rho(y,k)$, as a function of y, is strictly increasing on $[0, a(k)^-]$ and strictly decreasing on $[a(k)^+, \infty)$. Hence, using that $P(y) \le 1$,

$$
\begin{aligned}
&\int_{\mathbb{R}_+ \setminus [a(k)^-, a(k)^+]} P(y)\rho(y,k)\alpha e^{-\alpha y} \, \mathrm{d}y \\
&\qquad \le \int_0^{a(k)^-} \rho(y,k)\alpha e^{-\alpha y} \, \mathrm{d}y + \int_{a(k)^+}^\infty \rho(y,k)\alpha e^{-\alpha y} x \, \mathrm{d}y
\end{aligned}
$$

$$\le \rho(a(k)^-,k)\int_0^{a(k)^-} e^{-\alpha y}\, \mathrm{d}y + \rho(a(k)^+,k)\int_{a(k)^+}^{\infty} e^{-\alpha y}\, \mathrm{d}y$$
$$= O(1)\left(\rho(a(k)^-,k)+\rho(a(k)^+,k)\right),$$

as $k\to\infty$. Next we show that

$$\rho(a(k)^{\pm},k) = O\left(k^{-(1+C^2)/2}\right).$$

Since the arguments are almost completely identical, we give the prove for $a(k)^+$.

Using Stirling's formula $k! \sim \sqrt{2\pi} k^{k+1/2} e^{-k}$ as $k\to\infty$ we write

$$\begin{aligned}\rho(a(k)^+,k) &= \frac{\mu(a(k)^+)^k}{k!} e^{-\mu(a(k)^+)}\\ &\sim (2\pi)^{-1/2} k^{-1/2}\left(\frac{\mu(a(k)^+)}{k}\right)^k e^{-(\mu(a(k)^+-k)}\\ &= (2\pi)^{-1/2}k^{-1/2} e^{-k\left(\frac{\mu(a(k)^+)}{k}-1-\log\left(\frac{\mu(a(k)^+)}{k}\right)\right)}.\end{aligned}$$

Since

$$\frac{\mu(a(k)^+)}{k} = 1 + C\sqrt{\frac{\log(k)}{k}}$$

and $x-\log(1+x)\sim x^2/2$ as $x\to 0$, we get

$$\begin{aligned}\rho(a(k)^+,k) &\sim \sqrt{2\pi} k^{-1/2} e^{-k\left(C\sqrt{\frac{\log(k)}{k}}-\log\left(1+C\sqrt{\frac{\log(k)}{k}}\right)\right)}\\ &\sim (2\pi)^{-1/2}k^{-1/2} e^{-\frac{k\left(C\sqrt{\frac{\log(k)}{k}}\right)^2}{2}}\\ &= O\left(k^{-(1+C^2)/2}\right).\end{aligned}$$

Since $C>0$ can be chosen arbitrarily large we conclude that

$$\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y = (1+o(1))\int_{a^-(k)}^{a^+(k)} P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y,$$

as $k\to\infty$. Note that this implies that if $P(y)=h(y)(1+o(1))$ as $y\to\infty$, then

$$\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \sim \int_0^\infty h(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y, \tag{33}$$

as $y\to\infty$.

We now proceed with the proof of Proposition 1.4, which is split over the different cases for α. First we combine (11) and (12) to obtain the following scaling result

$$\frac{\int_0^\infty e^{-\beta y}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y} \sim \xi_{\alpha,\nu}^{2\beta} k^{-2\beta}. \tag{34}$$

Proof when $1/2<\alpha<3/4$ By Proposition 3.7 and (33) it follows that as $k\to\infty$,

$$\gamma(k)\sim c_\alpha \xi^{-(4\alpha-2)}\frac{\int_0^\infty e^{-(4\alpha-2)y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho_y(k)\alpha e^{-\alpha y}\,\mathrm{d}y} \sim c_\alpha k^{-4\alpha+2}.$$

where the last line is due to (34) with $\beta = 2\alpha-1$.

Proof when $\alpha = 3/4$ Similar to the previous case Proposition 3.7 and (33) imply that as $k \to \infty$

$$\gamma(k) = \frac{\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho_y(k)\alpha e^{-\alpha y}\,\mathrm{d}y} \sim \frac{\int_0^\infty \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho_y(k)\alpha e^{-\alpha y}\,\mathrm{d}y}.$$

However, the final step does not follow immediately from (34) because of the additional logarithmic term. To prove the result we first show that

$$\int_{a(k)^-}^{a(k)^+} P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \sim \int_{a(k)^-}^{a(k)^+} \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y. \tag{35}$$

For this we establish an upper bound for the left hand side

$$\int_{a(k)^-}^{a(k)^+} \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \leq \frac{a(k)^+}{2}\int_{a(k)^-}^{a(k)^+} e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y$$

and similarly, a lower bound

$$\int_{a(k)^-}^{a(k)^+} \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \geq \frac{a(k)^-}{2}\int_{a(k)^-}^{a(k)^+} e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y$$

Now observe that

$$\frac{a(k)^\pm}{2} = \log\left(\frac{k \pm \sqrt{k\log(k)}}{\xi_{\alpha,\nu}}\right) \sim \log(k)$$

and therefore it follows that

$$\limsup_{k\to\infty} \frac{\int_{a(k)^-}^{a(k)^+} \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\log(k)\int_{a(k)^-}^{a(k)^+} e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y} \leq 1.$$

and

$$\liminf_{k\to\infty} \frac{\int_{a(k)^-}^{a(k)^+} \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\log(k)\int_{a(k)^-}^{a(k)^+} e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y} \geq 1.$$

This proofs (35).

Next we note that by (34) with $\beta = 1/2$ we have

$$\frac{\int_0^\infty e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y} \sim \xi k^{-1}.$$

Therefore, since by (32),

$$\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \sim \int_{a(k)^-}^{a(k)^+} P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y$$

it follows from (35) that as $k \to \infty$,

$$\begin{aligned}\gamma(k) &\sim \frac{\int_0^\infty \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}\\ &\sim \log(k)\frac{\int_0^\infty e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y} \sim \xi\log(k)k^{-1} = \frac{6\nu}{\pi}\log(k)k^{-1},\end{aligned}$$

when $\alpha = 3/4$.

Proof when $\alpha > 3/4$ Again, by Proposition 3.7, equation (33) and (34) with $\beta = 1/2$, it follows that as $k \to \infty$,

$$\gamma(k) \sim \frac{\alpha - \frac{1}{2}}{\alpha - \frac{3}{4}} \frac{\int_0^\infty e^{-y/2} \rho(y,k) \alpha e^{-\alpha y}\, \mathrm{d}y}{\int_0^\infty \rho_y(k) \alpha e^{-\alpha y}\, \mathrm{d}y} \sim \frac{\alpha - \frac{1}{2}}{\alpha - \frac{3}{4}} \xi k^{-1} = \frac{8\alpha\nu}{\pi(4\alpha - 3)}.$$

4 Proofs of Theorem 1.1 and Theorem 1.2

We will first derive Theorem 1.2. It will turn out that Theorem 1.1 has a quick derivation assuming Theorem 1.2.

4.1 The proof of Theorem 1.2

Tobias: This is still in need of a serious polish

We first show we can restrict attention to the Poissonized instead of the standard KPKVB model.

Lemma 4.1. *Let $\alpha > \frac{1}{2}, \nu > 0, k \geq 2$ be fixed and write $G_n := G(n;\alpha,\nu), G_{n,Po} := G_{Po}(n;\alpha,\nu)$. Then*

$$|c(k;G_n) - c(k;G_{n,Po})| \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proof. We use a sprinkling argument. We observe that $G((1+\epsilon)n;\alpha,(1+\epsilon)\nu), G(n;\alpha,\nu)$ and $G_{\mathrm{Po}}(n;\alpha,\nu)$ all live on the same disk (with radius $R = 2\log\left(\frac{n}{\nu}\right)$). We consider the coupling where we have an infinite supply of i.i.d. points $u_1, u_2, \dots$ chosen according to the (α, R)-quasi uniform distribution, the vertices of $G((1+\epsilon)n;\alpha,(1+\epsilon)\nu)$ are $u_1,\dots,u_{(1+\epsilon)n}$; the vertices of $G(n;\alpha,\nu)$ are $u_1,\dots,u_n$ and the vertices of $G_{\mathrm{Po}}(n;\alpha,\nu)$ are $u_1,\dots,u_N$ with $N \stackrel{\mathrm{d}}{=} \mathrm{Po}(n)$ independent of $u_1, u_2, \dots$.

We imagine the situation where we remove εn vertices from $G((1+\epsilon)n;\alpha,(1+\epsilon)\nu)$ to obtain $G(n;\alpha,\nu)$ and we remove (or add, in the event that $N > (1+\varepsilon)n$) $|(1+\varepsilon)n - N|$ vertices to obtain $G_{\mathrm{Po}}(n;\alpha,\nu)$. Note that, by Chebyschev, $\mathbb{P}(|N-n| > \varepsilon n) = o(1)$. So, a.a.s., in both cases we remove no more than $2\varepsilon n$ vertices. Now let X, respectively X_{Po}, denote the number of vertices that are either themselves removed or have at least one neighbour that is removed.

Let $\overline{D}$ denote the average degree of $G((1+\epsilon)n;\alpha,(1+\epsilon)\nu)$. By the results of Gugelmann et al. [11] $\overline{D}$ converges in probability to a certain constant. In particular, there is a constant C such that $\overline{D} \leq C$ a.a.s.

Now note that, under the coupling described above:

$$\mathbb{E}\left(X | \overline{D} \leq C\right), \mathbb{E}\left(X_{\mathrm{Po}} | \overline{D} \leq C, |N-n| < \varepsilon n\right) \leq (1+C)2\varepsilon n.$$

Markov's inequality gives

$$\mathbb{P}(X \geq \sqrt{\varepsilon} n), \mathbb{P}(X_{\mathrm{Po}} \geq \sqrt{\varepsilon} n) = O(\sqrt{\varepsilon}).$$

By the results of Gugelmann et al. on the degree sequence the number of vertices of degree exactly k in $G(n;\alpha,\nu)$ equals $(1+o(1))np_k$ a.a.s. In particular it is $\Omega(n)$ a.a.s.

Since $c(k;G)$ is the average of $c(v)$ over all v of degree k, and $0 \leq c(v) \leq 1$, we have

$$\mathbb{P}(|c(k;G(n;\alpha,\nu)) - c(k;G((1+\epsilon)n;\alpha,(1+\epsilon)\nu))| > \sqrt{\varepsilon}) = O(\sqrt{\varepsilon}),$$

and simlarly for G_{Po}. In particular also $\mathbb{P}\left(|c(k;G(n;\alpha,\nu)) - c(k;G_{\mathrm{Po}}(n;\alpha,\nu))| > \sqrt{\varepsilon}\right) = O(\sqrt{\varepsilon})$. Sending $\varepsilon \searrow 0$ proves the lemma. **Pim:** I think we need to be a bit more precise here and set $\varepsilon = \varepsilon_n := Qn^{-1}$. ∎

Lemma 4.2. *Let $N_1(H)$ denote the number of vertices with height at least $H \in [0,R]$ in the finite box model (with n vertices).*

For all $\epsilon > 0$, there is $H = H_\epsilon$ such that $\mathbb{E}N_1(H) \leq \epsilon n$.

Proof. By linearity of expectation and the probability mass of a centered disk,

$$\mathbb{E}N_1(H) = n\mathbb{P}(B_{R-H}(0)) \leq ne^{-\alpha H}$$

Now, choose H large enough such that $e^{-\alpha H} \leq \epsilon$. $\square$

Lemma 4.3. *Let $N_2(H)$ denote the number of vertices with no neighbour with height above $H \in [0, R]$ in the finite box model (with n vertices).*

For all $\epsilon > 0$, there is $H = H_\epsilon$ such that $\mathbb{E}N_2 \leq \epsilon n$.

Proof. Let $\epsilon > 0$. By Lemma 4.2, there is H_1 such that the expected number of vertices with height at least H_1 is upper bounded by $\frac{\epsilon}{2}n$. So, if we write $N_{\geq H_1}$ for the number of vertices with height at least H_1 and $N_{\leq H_1}(H)$ for the number of vertices with height at most H_1 and no neighbour with height above H, then

$$N_2(H) \leq N_{\geq H_1} + N_{\leq H_1}(H)$$

Let $\mu(H)$ denote the expected number of vertices in the neighbourhood ball above H of a vertex with height H_1. For fixed H_1, $\mu(H)$ is a continuous function which is decreasing in H and with $\mu(R) = 0$. Therefore, there is H_2 large enough such that $1 - e^{-\mu(H_2)} \leq \frac{\epsilon}{2}$. Hence,

$$\mathbb{E}N_{\leq H_1}(H_2) \leq n(1 - e^{-\mu(H_2)}) \leq n\frac{\epsilon}{2}$$

Finally, $\mathbb{E}N_2(H_2) \leq \mathbb{E}N_{\geq H_1} + \mathbb{E}N_{\leq H_1}(H_2) \leq \epsilon n$. $\square$

Lemma 4.4. *Let $c(k)$ denote the clustering function of the KPKVB random graph for $k \in \mathbb{N}$ and $c_{\leq H}(k)$ denote the average of the clustering coefficient of all vertices with degree k and height $\leq H$ and with no neighbour with height $> H$. Then, $c(k) = c_{\leq H}(k) + O(\sqrt{\epsilon})$ with probability $\geq 1 - O(\sqrt{\epsilon}) + o(1)$.*

Proof. Let $N(k)$ denote the number of vertices with degree k in KPKVB. Let $N_H(k)$ denote the number of vertices with degree k and height $\leq H$ and with no neighbour with height $> H$. From the previous two lemmas, it follows that $\mathbb{E}(n - N_H) \leq \epsilon n$ and from this also that $\mathbb{E}(N(k) - N_H(k)) \leq \epsilon n$. Therefore, by Markov's inequality,

$$\mathbb{P}(n - N_H \geq \sqrt{\epsilon}n) \leq \frac{\mathbb{E}(n - N_H)}{n\sqrt{\epsilon}} \leq \sqrt{\epsilon}$$

Consider

$$c(k) - c_{\leq H}(k) = \frac{1}{N(k)}\sum c(v) - \frac{1}{N_H(k)}\sum c(v) \leq \frac{N(k) - N_H(k)}{N(k)}$$

So, finally

$$\mathbb{P}(c - c_{\leq H} = O(\sqrt{\epsilon})) \geq 1 - O(\sqrt{\epsilon})$$

$\square$

Lemma 4.5. *Let $c_{\leq H}(k)$ denote the average over all clustering coefficients of all vertices with degree k and height $\leq H$ and with no neighbour with height $> H$ in the poissonized KPKVB model.*

Let $\Delta_H(y)$ denote the probability that two neighbours of $(0, y)$ are adjacent. Let $\tau(y, k)$ be the Poisson probability that $(0, y)$ has degree k. Let $\rho(y)$ denote the truncated exponential density. **Pim:** *These notations are not in sync with the rest of the paper.* *Then, $c_{\leq H}(k) = (1 + o(1)) \int_0^H \int_0^H \int_0^H \Delta_H(y)\tau(y, k)\rho(y_1)\rho(y_2)\rho(y)dy_1dy_2dy$ a.a.s.* **Pim:** *I do not understand how a.a.s. is proven.*

Proof. First of all, as for $c_{\leq H}(k)$ all vertices involved have height $\leq H$, we can perform the computations in the box model: As H is constant, for R large enough, vertices with radial coordinates at least $R - H$ will have radial coordinate $\geq \frac{3}{4}R$, such that by Lemma 2.3, edges exist between such vertices in the disk if and only if they exist in the box.

Let $N_{trig}(k, H)$ denote the number of tuples (v, uw) for pairwise distinct vertices v, u, w such that v, u, w form a triangle, the radial coordinates of u, v, w are all at least $R-H$ and v has degree k (where H is the cut-off from the first two lemmas). So, we have $c_{\leq H}(k) = \frac{N_{trig}(k,H)}{N_H}$.

Let $M_{trig}(k, H)$ denote the number of tuples (v, u, w) for pairwise distinct vertices v, u, w such that v, u, w form a triangle, the radial coordinates of u, v, w are all at least $R-H$ and v has degree k. Note that $N_{trig}(k, H) = \frac{1}{2}M_{trig}(k, H)$. $M_{trig}(k, H)$ can be written as a sum over 3-tuples:

$$M_{trig}(k, H) = \sum_{(v,u,w)\in V^3_{\neq}} h(V, v, u, w)$$

where h is the indicator function that v, u, w form a triangle, have radial coordinates at least $R-H$ and v has degree k. Using Mecke's formula, we obtain that

$$\mathbb{E}M_{trig}(k, H) = \int_{E_R}\int_{E_R}\int_{E_R} \mathbb{E}h(V, v, u, w)dvdudw$$

Pim: What is E_R and what is D_H?

From there, it follows that

$$\mathbb{E}c_{\leq H}(k) = \frac{1}{\mathbb{P}(D_H = k)}\int_0^H\int_0^H\int_0^H \Delta_H(y)\tau(y, k)\rho(y_1)\rho(y_2)dy_1dy_2dy$$

For $H \to \infty$, this converges to the integral obtained in the paper.

In order to complete the proof, we need to show that $\mathbb{E}c_{\leq H}(k)^2 = (1 + o(1))(\mathbb{E}c_{\leq H}(k))^2$. For this, it is sufficient to show that $\mathbb{E}N^2_{trig} = (1 + o(1))(\mathbb{E}N_{trig})^2$. In the box model, we know that two vertices with height at most H and horizontal distance $> 2e^H$ cannot be adjacent and their neighbourhood balls are disjoint (w.r.t. the connection rule of the box model). If the neighbourhood balls are disjoint, then the clustering coefficients of the vertices are independent. The horizontal coordinate is uniform in the width of the box which is $\pi e^{\frac{R}{2}}$. Therefore, the probability that two vertices have horizontal distance $\leq 2e^H$ is $\leq \frac{2e^H}{\pi e^{\frac{R}{2}}} = o(1)$. □

Pim: Where is the proof of Theorem 1.2?

4.2 The proof of Theorem 1.1

Proof of Theorem 1.1: We first note that

$$\gamma = \mathbb{E}C = \sum_{k\geq 2}\mathbb{E}(C|D = k)\mathbb{P}(D = k) = \sum_{k\geq 2}\gamma(k)p_k.$$

Writing N_k for the number of vertices of degree k, we have the similar relation

$$c(G) = \sum_{k\geq 2} c(k; G)\cdot(N_k/n).$$

Theorem 1.2 establishes that $c(k; G(n; \alpha, \nu))$ converges to $\gamma(k)$ in probability for any fixed k. The results of Gugelmann et al. [11] on the degree sequence establish that N_k/n converges to p_k in probability, for any fixed k. The result now follows in a straightforward manner from these observations. ■

5 Overview of the proof strategy for $k \to \infty$

Pim: References need to be included once Section 1 is updated.

The proof of Theorem 1.3 follows the same strategy as outlined in Section 2. However, the fact that $k = k_n \to \infty$ as $n \to \infty$, introduces additional technical challenges. For example, the coupling we use becomes less exact so that we can no longer use Lemma 2.3 to conclude that triangle counts in G_{Po} and G_{box} are asymptotically equivalent. In this section we explain the challenges with each step and give a detailed overview of the structure for the proof of Theorem 1.3 using intermediate results for each of the steps. Since we are ultimately interested in recovering the scaling of $c(k_n; G_n)$, which Theorem 1.3 claims is $\gamma(k_n)$, we need to show that each step only introduces error terms that are of smaller order, i.e. that are $o(\gamma(k_n))$. To this end we define the scaling function

$$s(k) = \begin{cases} k^{-(4\alpha-2)} & \text{if } \frac{1}{2} < \alpha < \frac{3}{4}, \\ \log(k)k^{-1} & \text{if } \alpha = \frac{3}{4}, \\ k^{-1} & \text{if } \alpha > \frac{3}{4}, \end{cases} \tag{36}$$

so that $\gamma(k) = \Theta(s(k))$ as $k \to \infty$. We will end this section with the proof of Theorem 1.3, based on the intermediate results.

Remark 5.1 (Diverging k_n)**.** *Throughout the remainder of this paper $\{k_n\}_{n\geq 1}$ will always denote a sequence of integers satisfying $k_n \to \infty$ and $k_n = o\left(n^{\frac{1}{2\alpha+1}}\right)$, as $n \to \infty$.*

We start with introducing a slightly modified version of the local clustering function, which will be convenient for computations later,

$$c^*(k; G) = \frac{1}{\mathbb{E}[N_k]} \sum_{\substack{v \in V(G) \\ \deg(v) = k}} c(v). \tag{37}$$

Notice that the only difference between $c(k; G)$ and $c^*(k; G)$ is that we replace $N_G(k)$ by its expectation $\mathbb{E}[N_G(k)]$. The advantage is that now, the only randomness is in triangle counting. In addition, note that since $\mathbb{E}[N_G(k)] > 0$ a case distinction for N_k is no longer needed for $c^*(k; G)$. It is however still relevant since we are eventually interested in $c(k; G)$. Following the notational convention, throughout the remainder of this paper we write $c^*(k; G_{\mathrm{Po}})$ and $c^*(k; G_{\mathrm{box}})$ to denote the modified local clustering function in G_{Po} and $G_{\mathcal{P},n}(\alpha, \nu)$, respectively.

Figure 5 shows a schematic overview of the proof of Theorem 1.3 based on the different propositions described below, plus the sections in which theses propositions are proved. Observe that the order in which the intermediate results are proved is reversed with respect to the natural order of reasoning. This does not create any circular logic, since each intermediate result is independent of the others. We choose this order because results proved in the later stages are helpful to deal with error terms coming up in proofs at earlier stages and hence help streamline those proofs.

5.1 Adjusted clustering and Poisson nodes in hyperbolic graphs

Recall that the first step for the fixed k case was to show that the transition from the hyperbolic random graph $G_{\mathbb{H},n}$ to the Poisson version G_{Po} did not influence clustering. Here we first make a transition from the local clustering function $c_{\mathbb{H},n}(k)$ to the adjusted version $c^*(k; G_n)$. The following lemma justifies working with this modified version. The proof uses a concentration result for $\mathbb{N}_{\mathbb{H},n}(k_n)$ and full details can be found in Section 9.3.

Lemma 5.1. *As $n \to \infty$,*

$$\mathbb{E}\left[|c(k_n; G_n) - c^*(k_n; G_n)|\right] = o\left(s(k_n)\right).$$

We then establish that the modified local clustering function in the hyperbolic model $G_{\mathbb{H},n}(\alpha, \nu)$ behaves similarly to that in the Poisson version $G_{\widetilde{\mathbb{H}},n}(\alpha, \nu)$. This is based on a standard coupling between a Binomial Point Process and Poisson Point Process.

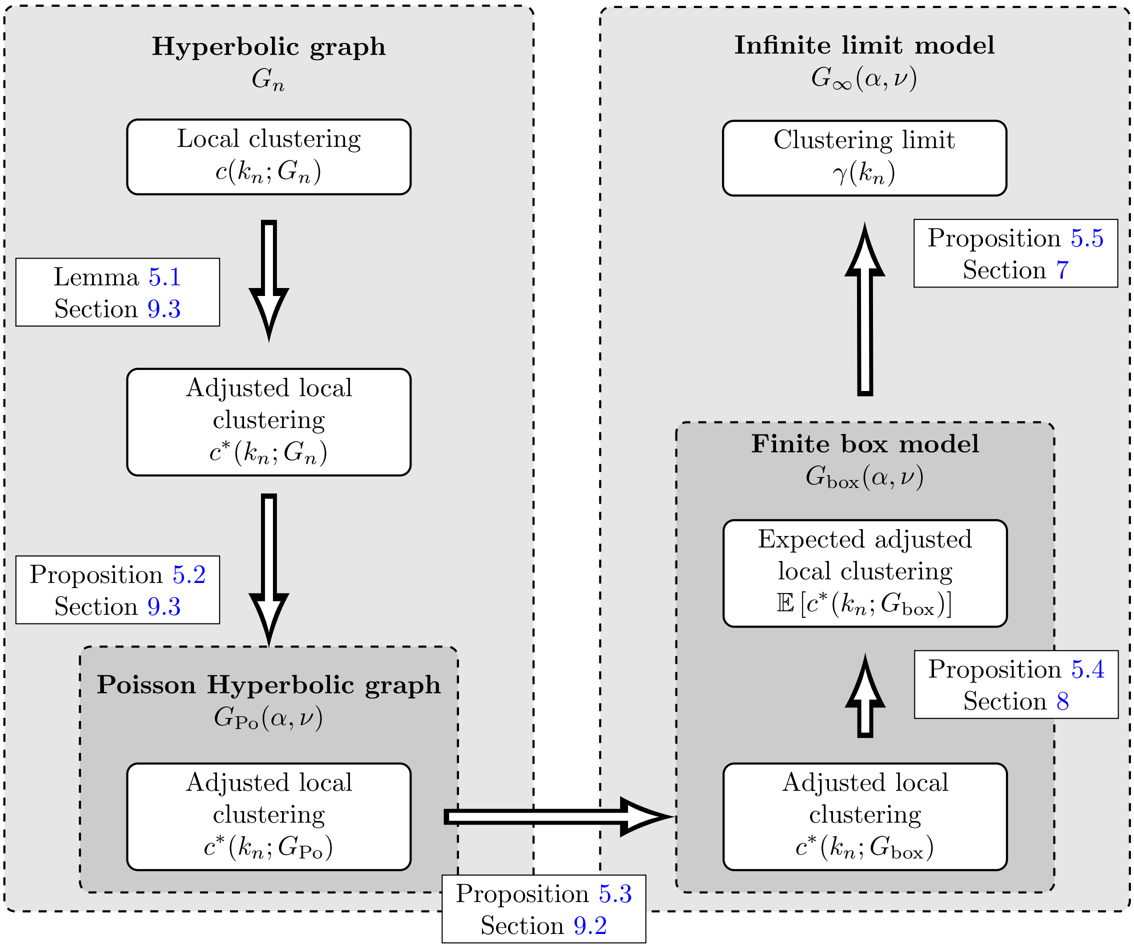

Figure 5: Overview of the proof strategy for Theorem 1.2.

Proposition 5.2. *As $n \to \infty$,*

$$\mathbb{E}\left[|c^*(k_n; G_n) - c^*(k_n; G_{Po})|\right] = o\left(s(k_n)\right).$$

5.2 Coupling of local clustering between G_{Po} and G_{box}

The next step is to show that the modified clustering is preserved under the coupling described in Section 2.4. The proof can be found in Section 9.2. This is one of the key technical challenges we face.

To understand why, recall that the degree k of a node is related to its height y, roughly speaking, by $k \approx \xi e^{y/2}$. Therefore, when k is fixed we have that the heights of nodes with that degree are also fixed, in particular $y < R_n/4$ for large enough n. In addition, the main contribution of triangles would also come from nodes with heights $y' < R_n/4$. This allowed us to use Lemma 2.3 and conclude that the triangles present in the graph G_{Po} where exactly those present in G_{box} and therefore the local clustering function was the same in both models. When $k_n \to \infty$ this is no longer true in general. For instance, suppose $k_n = n^{\frac{1-\varepsilon}{2\alpha+1}}$, for some small $0 < \varepsilon < 1$. Then the relation $k_n \approx \xi e^{y_n/2}$ implies that $y_n \approx \frac{2(1-\varepsilon)}{2\alpha+1}\log(n) - 2\log(\xi)$. Since $R_n/4 = \frac{1}{2}\log(n) - \frac{1}{2}\log(\nu)$ we get that $R_n/4 = o\left(y_n\right)$ for all $\alpha > (3-4\varepsilon)/2$ and hence $y_n > R_n/4$ for large enough n, violating the conditions of Lemma 2.3. However, by carefully analyzing the difference between the adjusted local clustering function in both models we can still make the same conclusion. This is summarized in the following proposition whose proof is found in Section 9.2.

Proposition 5.3 (Coupling result for local clustering)**.** *As $n \to \infty$,*

$$\mathbb{E}\left[|c^*(k_n; G_{Po}) - c^*(k_n; G_{box})|\right] = o\left(s(k_n)\right).$$

Tobias: Maybe we could replace these three with a statement on $|c(k_n; G_n) - c^*(k_n; G_{\text{box}})|$, at least at this point of the paper. This is only the high level description and that is all we need for the "final proof".**Pim:** I would vote for keeping them split, since this allows us to clearly point to the main technical challenge we have to overcome to obtain the final result. These three results together imply that the difference between the local clustering function of the hyperbolic random graph and the modified local clustering function of the finite box graph converges to zero faster than the proposed scaling $\gamma(k_n)$ in Theorem [??]. Hence, to prove this theorem it is enough to prove it for $c^*(k; G_{\text{box}})$.

5.3 From the finite to the infinite model

To compute the limit of the modified local clustering function $c^*(k; G_{\text{box}})$ in the finite graph $G_{\mathcal{P},n}(\alpha, \nu)$ we first prove in Section 8 that it is concentrated around its expectation $\mathbb{E}\left[c^*(k_n; G_{\text{box}})\right]$.

Tobias: "concentration" is a loaded term in probability. I am not sure this use of the word will not be counterintuitive to many. **Pim:** I am not sure. Concentration generally refers to how much a random variable deviates from its expectation. This is exactly what this proposition tells us. I would therefore vote for keeping this terminology, although I would have no problem with replacing it with an other term if someone has a good suggestion.

Proposition 5.4 (Concentration for local clustering function in $G_{\mathcal{P},n}(\alpha, \nu)$)**.** *As $n \to \infty$,*

$$\mathbb{E}\left[|c^*(k_n; G_{box}) - \mathbb{E}\left[c^*(k_n; G_{box})\right]|\right] = o\left(s(k_n)\right).$$

This result is another one of the technical challenges we face when considering $k_n \to \infty$. For the proof, we first identify the specific range of heights that give the main contribution to the triangle count, showing that the triangles coming from nodes with heights outside this range is of smaller order. Then we prove a concentration result for the main term, by carefully analyzing the joint neighborhoods of two nodes whose heights fall into the identified range. The full details are found in Section 8.

Assuming this concentration result, we are left with the task to compute the limit of $\mathbb{E}\left[c^*(k_n; G_{\text{box}})\right]$ as $n \to \infty$ and show that it is equivalent to $\gamma(k_n)$. To accomplish this we move to the infinite limit model G_∞ and show that the difference between the expected value of $c^*(k; G_{\text{box}})$ and $\gamma(k_n)$ goes to zero faster than the proposed scaling in Theorem 1.2.

Proposition 5.5 (Transition to the infinite limit model)**.** *As $n \to \infty$,*

$$|\mathbb{E}\left[c^*(k_n; G_{box})\right] - \gamma(k_n)| = o\left(s(k_n)\right).$$

Tobias: This is only for $c(k)$. Should we not also mean c? **Pim:** I do not think so. We can prove everything for c once we have the result for $c(k_n)$.

Recall that for the finite box model the left and right boundaries of $\mathcal{R}_n$ where identified, so that graph G_{box} contains some additional edge with respect to the induced subgraph of G_∞ on $\mathcal{R}_n$. The prove of Proposition 5.5 therefore relies on analyzing the number of triangles coming from these additional edges and showing that their contribution to the local clustering function are of negligible order, see Section 7.

Remark 5.2 (Notations different graphs)**.** *We will use the subscripts n,* Po*, box and ∞ to identify properties of, respectively, the KPKVB mode G_n, the Poisson version G_{Po}, the finite box model G_{box} and the infinite model G_∞. For example $N_{\text{Po}}(k)$ denotes number of nodes with degree k in G_{Po} and $\rho_{box}(y, k) = \mathbb{P}\left(\text{Po}(\mu(\mathcal{B}_\infty(y))) = k\right)$, i.e. the degree distribution of a typical point in G_{box}.*

5.4 Proof of the main results

We are now ready to prove Theorem 1.3, using the propositions stated in the previous sections.

Proof of Theorem 1.3. First of all, due to cancellation of equal terms we can rewrite

$$\begin{aligned} c(k_n; G_n) - \gamma(k_n) =& c(k_n; G_n) - c^*(k_n; G_n) + c^*(k_n; G_n) - c^*(k_n; G_{\text{Po}}) + c^*(k_n; G_{\text{Po}}) - c^*(k_n; G_{\text{box}}) \\ &+ c^*(k_n; G_{\text{box}}) - \mathbb{E}c^*(k_n; G_{\text{box}}) + \mathbb{E}c^*(k_n; G_{\text{box}}) - \gamma(k_n) \end{aligned}$$

Then, we take absolute values and apply the triangle inequality. By monotonicity of expectation, we can apply it to both sides and obtain

$$\begin{aligned} \mathbb{E}[|c(k_n; G_n) - \gamma(k_n)|] \leq& \mathbb{E}[|c(k_n; G_n) - c^*(k_n; G_n)|] + \mathbb{E}[|c^*(k_n; G_n) - c^*(k_n; G_{\text{Po}})|] \\ &+ \mathbb{E}[|c^*(k_n; G_{\text{Po}}) - c^*(k_n; G_{\text{box}})|] + \mathbb{E}[|c^*(k_n; G_{\text{box}}) - \mathbb{E}c^*(k_n; G_{\text{box}})|] \\ &+ \mathbb{E}[|\mathbb{E}c^*(k_n; G_{\text{box}}) - \gamma(k_n)|] \end{aligned}$$

At this point, the lemmas and propositions presented above in this section can be applied in order to show that all summands are $o\left(\gamma(k_n)\right)$: Lemma 5.1 for the transition to the modified clustering function in the first term, Proposition 5.2 for the Poissonization in the disk in the second term, Proposition 5.3 for the coupling from the disk to the finite box model in the third term, Proposition 5.4 for the concentration in the fourth term and finally Proposition 5.5 for the transition to the infinite limit model. All of this together yields that:

$$\mathbb{E}[|c(k_n; G_n) - \gamma(k_n)|] = o\left(s(k_n)\right) = o\left(\gamma(k_n)\right),$$

i.e. the statement of the theorem.

□

6 Concentration of heights for vertices with degree k

In the proof of Proposition 1.4 we used a result that allowed us to restrict integration over y to the interval $[a(k)^-, a(k)^+]$, with

$$a(k)^{\pm} = 2\log\left(\frac{k \pm C\sqrt{k\log(k)}}{\xi} \vee 1\right).$$

The reason for this was that the integrand included the function $\rho(y,k) = \mathbb{P}\left(\text{Po}(\mu(y)) = k\right)$, where $\text{Po}(\lambda)$ denotes a Poisson random variable with expectation λ and Poisson random variables are well concentrated around their mean, i.e. around heights y for which $\mu(y) \approx k$. Since $\mu(y) = \mu(\mathcal{B}_\infty(y)) = \xi e^{y/2}$, this implies that integration with respect to $\rho(y,k)$ is concentrated around $y \approx 2\log(k/\xi)$.

In the remainder of this paper we will often encounter integrands involving the function $\mathbb{P}\left(\text{Po}(\mu_n(y)) = k\right)$, for some $\mu_n(y)$. In these case we want to be able to restrict our integration around those heights y for which $y \approx 2\log(k_n/\xi)$. We will refer to such results as *concentration of heights arguments*. In this section we establish such results. We start with a concentration of heights lemma for the infinite model G_∞ (Lemma 6.1) and explain in Remark 6.1 how such a result will be used throughout the paper. We then establish similar results for the Poissonized KPKVB and finite box model based on a generalization of Lemma 6.1 obtained in Section 6.2. To apply this lemma to G_{Po} and G_{box} we have to analyze the expected number of nodes in a typical neighborhood in these models, which is done in Sections 6.3 and 6.4.

6.1 Concentration of heights argument for the infinite model

The next lemma states that for a large class of functions $h(y)$ and $k_n \to \infty$, to compute the integral

$$\int_0^\infty \rho(y,k_n)h(y)e^{-\alpha y}\,\mathrm{d}y$$

it is enough to consider integration over a small interval on which $e^{y/2} \approx k_n$, instead of $\mathbb{R}_+$.

Lemma 6.1. *Let $\alpha > \frac{1}{2}$, $\nu > 0$, $\{k_n\}_{n\geq 1}$ be any positive sequence such that $k_n \to \infty$ and $k_n = o(n)$ and let $\ell_n = k_n(1+\epsilon_n)$, with $\epsilon_n \to 0$. In addition, define for any constant $C > 0$,*

$$\lambda_n^{\pm} = (\ell_n \pm C\sqrt{\ell_n \log(\ell_n)}) \wedge \xi, \quad a_n^{\pm} = 2\log\left(\frac{\lambda_n^{\pm}}{\xi}\right).$$

Then the following holds.

1. *For any continuous function $h : \mathbb{R}_+ \to \mathbb{R}$, such that $h(y) = O\left(e^{\beta y}\right)$ as $y \to \infty$ for some $\beta < \alpha$,*
$$\int_{\mathbb{R}_+\setminus(a_n^-,a_n^+)} \rho(y,k_n)h(y)\alpha e^{-\alpha y}\,\mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right), \tag{38}$$
as $n \to \infty$.

2. *If in addition, $h(a_n) \sim h(b_n)$ whenever $a_n \sim b_n$, as $n \to \infty$. Then,*
$$\int_0^\infty h(y)\rho(y,k_n)\alpha e^{-\alpha y'}\,\mathrm{d}y' \sim 2\alpha\xi^{2\alpha}h(2\log(k_n/\xi))k_n^{-(2\alpha+1)}, \tag{39}$$
as $n \to \infty$.

Proof.

Proof of the first statement. Recall (see proof of Proposition 1.4) that $\rho(y,k_n)$, as a function of y, is strictly increasing on $[0,a_n^-]$ and strictly decreasing on $[a_n^+,\infty)$. Therefore, by our assumption on $h(y)$,

$$\begin{aligned}
&\int_{\mathbb{R}_+\setminus(a_n^-,a_n^+)} h(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \\
&\quad= O(1)\int_0^{a_n^-} e^{\beta y}\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y + O(1)\int_{a_n^+}^\infty e^{\beta y}\rho(y,k_n)\alpha e^{-\alpha y}x\,\mathrm{d}y \\
&\quad= O(1)\int_0^{a_n^-} \rho(y,k_n)e^{-(\alpha-\beta)y}\,\mathrm{d}y + O(1)\int_{a_n^+}^\infty \rho(y,k_n)e^{-(\alpha-\beta)y}\,\mathrm{d}y \\
&\quad\leq O(1)\,\rho(a_n^-,k_n)\int_0^{a_n^-} e^{-(\alpha-\beta)y}\,\mathrm{d}y + O(1)\,\rho(a_n^+,k_n)\int_{a_n^+}^\infty e^{-(\alpha-\beta)y}\,\mathrm{d}y.
\end{aligned}$$

Since $\alpha - \beta > 0$, we conclude that

$$\int_{\mathbb{R}_+\setminus(a_n^-,a_n^+)} h(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = O(1)\left(\rho(a_n^-,k_n) + \rho(a_n^+,k_n)\right) \tag{40}$$

We shall now bound the terms $\rho(a_n^{\pm},k_n)$, starting with $\rho(a_n^+,k_n)$. Using Stirling's approximation $k! \sim \sqrt{2\pi}k^{k+1/2}e^{-k}$ as $k \to \infty$ we write

$$\rho(a_n^+,k_n) = \frac{\mu(a_n^+)^{k_n}}{k_n!}e^{-\mu(a_n^+)}$$

$$\sim (2\pi)^{-1/2} k_n^{-1/2} \left(\frac{\mu(a_n^+)}{k_n}\right)^{k_n} e^{-(\mu(a_n^+)-k_n)}$$
$$= (2\pi)^{-1/2} k_n^{-1/2} e^{-k_n\left(\frac{\mu(a_n^+)}{k_n}-1-\log\left(\frac{\mu(a_n^+)}{k_n}\right)\right)}.$$

Since

$$\frac{\mu(a_n^+)}{k_n} = \frac{\lambda_n^+}{k_n} = 1+\epsilon_n + C\frac{\kappa_n}{k_n} = 1+\epsilon_n + C\sqrt{\frac{(1+\epsilon_n)\log((1+\epsilon_n)k_n)}{k_n}},$$

and $x - \log(1+x) \sim x^2/2$ as $x \to 0$, we get

$$\begin{aligned}\rho(a_n^+, k_n) &\leq \sqrt{2\pi} k_n^{-1/2} e^{-k_n\left(\epsilon_n + C\frac{\kappa_n}{k_n} - \log\left(1+\epsilon_n + C\frac{\kappa_n}{k_n}\right)\right)} \\ &\sim (2\pi)^{-1/2} k_n^{-1/2} e^{-\frac{k_n(\epsilon_n + C\kappa_n/k_n)^2}{2}} \\ &= O\left(k_n^{-(1+C^2)/2}\right),\end{aligned} \tag{41}$$

where for the last line we used that

$$-k_n \frac{(\epsilon_n + C\kappa_n/k_n)^2}{2} = -\frac{C^2}{2}\log(k_n) + \Theta(1).$$

A similar analysis as above yields

$$\rho(a_n^-, k_n) \leq \Theta(1)\, k_n^{-1/2} e^{-\frac{k_n(\epsilon_n - C\kappa_n/k_n)^2}{2}} = O\left(k_n^{-(1+C^2)/2}\right). \tag{42}$$

Plugging (42) and (41) into (40) yields the result.

Proof of the second statement. By the mean value theorem for definite integrals, there exists a $c_n \in (a_n^-, a_n^+)$ such that

$$\int_{a_n^-}^{a_n^+} h(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = h(c_n)\int_{a_n^+}^{a_n^+} \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y.$$

Since $\int_0^\infty \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = \Theta\left(k_n^{-(2\alpha+1)}\right)$, taking any $C > \sqrt{4\alpha+1}$, (38) implies that

$$\int_{a_n^+}^{a_n^+} \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = (1+o(1))\int_0^\infty \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y,$$

from which we conclude that (see (11)),

$$\int_{a_n^+}^{a_n^+} \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = (1+o(1))2\alpha\xi^{2\alpha}k_n^{-(2\alpha+1)},$$

as $n \to \infty$. Finally, since $c_n \in (a_n^-, a_n^+)$ it follows that

$$\left|\frac{c_n}{2\log(k_n/\xi)} - 1\right| \leq 2C\sqrt{\frac{\log(k_n)}{k_n}},$$

so that $c_n \sim k_n$. Therefore, by assumption on h

$$\int_{a_n^-}^{a_n^+} h(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \sim h(c_n)2\alpha\xi^{2\alpha}k_n^{-(2\alpha+1)} \sim 2\alpha\xi^{2\alpha}h(2\log(k_n/\xi))k_n^{-(2\alpha+1)},$$

as $n \to \infty$. $\square$

Note that we can tune the error in (38) by selecting an appropriately large $C > 0$, i.e. by restrict the function $h(y)$ inside the integral to an appropriate interval around $2\log(k_n/\xi)$. This makes Lemma 6.1 very powerful. Below we list several important corollaries.

Corollary 6.2. *Let $h : \mathbb{R}_+ \to \mathbb{R}$ be any continuous function such that for some $\beta < \alpha$, $h(y) = O\left(e^{\beta y}\right)$ as $y \to \infty$ and $h(a_n) \sim h(b_n)$ whenever $a_n \sim b_n$. Then for any other continuous function $g : \mathbb{R}_+ \to \mathbb{R}$, such that $g(y) \sim h(y)$ as $y \to \infty$*

$$\int_0^\infty g(y)\rho(y,k_n)\alpha e^{-\alpha y'}\,\mathrm{d}y' \sim 2\alpha\xi^{2\alpha}h(2\log(k_n/\xi))k_n^{-(2\alpha+1)}, \tag{43}$$

as $n \to \infty$.

Proof. By assumption, g satisfies the conditions of for the second statement of Lemma 6.1. Since in addition $g(2\log(k_n/\xi)) \sim h(2\log(k_n/\xi))$, the result follows. □

For any positive sequence $a_n \to \infty$ and $C > 0$ we define

$$\mathcal{K}_C(a_n) = \left\{ y \in \mathbb{R}_+ : \frac{a_n - C\sqrt{a_n\log(a_n)}}{\xi} \vee 1 \le e^{\frac{y}{2}} \le \frac{a_n + C\sqrt{a_n\log(a_n)}}{\xi} \right\}. \tag{44}$$

In addition we define

$$\mathcal{K}_{C,n}(a_n) := (-I_n, I_n] \times ((0, R_n] \cap \mathcal{K}_C(a_n)), \tag{45}$$

where $I_n := \frac{\pi}{2}e^{R_n/2}$. Recall that $\mathcal{R} = (-I_n, I_n] \times [0, R]$. The following corollary allows us to bound integrals of function by considering there maximum of $\mathcal{K}_C(k_n)$.

Corollary 6.3. *Let $h_n : \mathbb{R}_+ \to \mathbb{R}_+$ be a sequence of continuous functions which such that for some $s \in \mathbb{R}$ and $\beta < \alpha$, as $n \to \infty$, $h_n(y) = O\left(k_n^s e^{\beta y}\right)$ and $h_n(y) = \Omega(1)$, uniformly on $0 \le y \le R$. Then, as $n \to \infty$,*

$$\int_{\mathcal{R}} h_n(y)\rho(y,k_n)f(x,y)\,\mathrm{d}x\,\mathrm{d}y = (1 + o(1))\, n \int_{\mathcal{K}_C(k_n)} h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y.$$

In particular,

$$\int_{\mathcal{R}} h_n(y)\rho(y,k_n)f(x,y)\,\mathrm{d}x\,\mathrm{d}y = O(1)\, n\, k_n^{-(2\alpha+1)} \max_{y\in\mathcal{K}_C(k_n)} h_n(y),$$

as $n \to \infty$.

Proof. The second result follows immediately from the first. For the first result we note that by the first statement of Lemma 6.1

$$\begin{aligned}\int_{[0,R]\setminus(a_n^-,a_n^+)} h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y &\le O(k_n^s)\int_{\mathbb{R}_+\setminus(a_n^-,a_n^+)} e^{\beta y}\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \\ &= O\left(k_n^{s-(1+C^2)/2}\right).\end{aligned}$$

By assumption on $h_n(y)$,

$$\int_{\mathcal{K}_C(k_n)} h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = O\left(k_n^{s+2\beta}\right)\int_{\mathcal{K}_C(k_n)} \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = O\left(k_n^{s+2\beta-(2\alpha+1)}\right),$$

and

$$\int_{\mathcal{K}_C(k_n)} h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = \Omega(1)\int_{\mathcal{K}_C(k_n)} \rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = \Omega(k_n^{-(2\alpha+1)}).$$

Hence, by taking $C > 0$ such that $(1 + C^2)/2 > \max\{2\alpha + 1 + s, 2\alpha + 1 - \beta\}$ we get that

$$\int_{[0,R]\setminus(a_n^-,a_n^+)} h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y = o(1)\int_{\mathcal{K}_C(k_n)} h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y.$$

The result then follows since

$$\int_{\mathcal{R}} h_n(y)\rho(y,k_n)f(x,y)\,\mathrm{d}x\,\mathrm{d}y = n\int_0^R h_n(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y.$$

$\square$

Corollary 6.4. *Let $h:\mathbb{R}_+\to\mathbb{R}$ be a continuous function which satisfies the conditions of Lemma 6.1 and let h_n be a sequence of functions such that, as $n\to\infty$, $h_n(y)=\Omega(1)$ and $h_n(y)=O\left(k_n^s\right)h(y)\rho(y,k_n)$, uniformly on $0\le y\le R$. Then,*

$$\int_{\mathcal{R}} h_n(y)\rho(y,k_n)f(x,y)\,\mathrm{d}x\,\mathrm{d}y \sim 2\alpha\xi^{2\alpha}\,n\,h_n(2\log(k_n/\xi))k_n^{-(2\alpha+1)}, \tag{46}$$

as $n\to\infty$.

Proof. The result immediately follows by first applying Corollary 6.3 and then using the second statement from Lemma 6.1. $\square$

Remark 6.1 (Concentration of heights argument)**.** *All the above corollaries use the same reasoning, namely that when the integrand contains $h_n(y)\rho(y,k_n)$, for some "nice" functions $h_n(y)$, then the main contribution is determined by the integration over $\mathcal{K}_C(k_n)$. This implies, for instance, that we only need to carefully analyze the functions $h_n(y)$ on $\mathcal{K}_C(y)$, while for a certain class of functions we can even simply replace it with $h_n(2\log(k_n/\xi))$. We will refer collectively to any of these arguments as a* concentration of heights argument*. For example, suppose we know that $h_n(y)=\Omega(1)$ and $h_n(y)=O\left(k_n^4\right)e^{\alpha/2y}$, while $h_n(y)=(1+o(1))k_n^2e^{\alpha/2y}$, uniformly on $\mathcal{K}_C(k_n)$. Then by a concentration of heights argument (in this case Corollary 6.4)*

$$\begin{aligned}\int_{\mathcal{R}} h_n(y)\rho(y,k_n)f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= (1+o(1))2\alpha\xi^{2\alpha}\,n\,k_n^2(k_n/\xi)^{\alpha/2}k_n^{-(2\alpha+1)}\\ &= (1+o(1))2\alpha\xi^{5\alpha/2}\,n\,k_n^{1+5\alpha/2}.\end{aligned}$$

Remark 6.2 (Proof of Proposition 1.4 revisited)**.** *Note that due to Proposition 3.7, the function $P(y)$ from Section 3 satisfies all the necessary conditions in Corollary 6.2. Hence, Proposition 1.4 directly follows from Proposition 3.7 and a concentration of heights argument (second statement of Lemma 6.1).*

Although powerful, the current versions of the concentration of heights arguments are only valid for the function $\rho(y,k_n) := \mathbb{P}\left(\mathrm{Po}\left(\mu\left(\mathcal{B}_\infty\left(y\right)\right)\right)=k_n\right)$, which uses the neighborhoods in the infinite model G_∞. Since we will also be working in the Poissonized KPKVB model G_{Po} and the finite box model G_{box}, we would like to use concentration of heights arguments when the integrand contains either the function $\rho_{\mathrm{Po}}(y,k_n) := \mathbb{P}\left(\mathrm{Po}\left(\mu\left(\mathcal{B}\left(y\right)\right)\right)=k_n\right)$ or $\rho_{\mathrm{box}}(y,k_n) := \mathbb{P}\left(\mathrm{Po}\left(\mu\left(\mathcal{B}_{\mathrm{box}}\left(y\right)\right)\right)=k_n\right)$. The next proposition establishes this result and we will spend the rest of this section on its proof.

Proposition 6.5. *The conclusions from Lemmas 6.1 and Corollaries 6.2, 6.3 and 6.4 still hold if we replace $\rho(y,k_n)$ with either $\rho_{\mathrm{Po}}(y,k_n)$ or $\rho_{box}(y,k_n)$.*

6.2 A more general concentration of heights argument

Note that since all three corollaries follow from Lemma 6.1, Proposition 6.5 follows if we can prove that the conclusions from this lemma holds when we replace $\rho(y,k_n)$ with either $\rho_{\mathrm{Po}}(y,k_n)$ or $\rho_{\mathrm{box}}(y,k_n)$. For this the following, slightly more general version of Lemma 6.1 will be important.

Lemma 6.6. *Let $\alpha>\frac{1}{2},\nu>0$, $k_n\to\infty$ be such that $k_n=o(n)$, $\ell_n=(1+\epsilon_n)k_n$, with $\epsilon_n\to 0$ and define*

$$\lambda_n^{\pm} = (\ell_n \pm C\sqrt{\ell_n\log(\ell_n)})\wedge\xi, \quad \textit{and} \quad a_n^{\pm} = 2\log\left(\frac{\lambda_n^{\pm}}{\xi}\right).$$

In addition, define $\hat{\rho}_n(y,k) = \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n(y)) = k\right)$*, where* $\hat{\mu}_n(y)$ *satisfies,*

$$\hat{\mu}_n(y) = (1+\phi_n(y))\mu(y),$$

where $\phi_n(y)$ *is a continuous differentiable function such that*

i) *For some* $0<\varepsilon<1$, $\lim_{n\to\infty} \sup_{0\le y\le(1-\varepsilon)R_n} |\phi_n(y)| = 0,$

ii) $\lim_{n\to\infty} \sup_{y\in\mathcal{K}_C(\ell_n)} |\phi_n'(y)| = 0.$

Then the following holds.

1. *For any continuous function* $h : \mathbb{R}_+ \to \mathbb{R}$*, such that* $h(y) = O\left(e^{\beta y}\right)$ *as* $y\to\infty$ *for some* $\beta<\alpha$*,*

$$\int_{\mathbb{R}_+\setminus(a_n^-,a_n^+)} \hat{\rho}_n(y,k_n)h(y)\alpha e^{-\alpha y}\,\mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right), \tag{47}$$

as $n\to\infty$*.*

2. *If in addition,* $h(a_n)\sim h(b_n)$ *whenever* $a_n\sim b_n$*, as* $n\to\infty$*. Then,*

$$\int_0^\infty h(y)\hat{\rho}_n(y,k_n)\alpha e^{-\alpha y'}\,\mathrm{d}y' \sim 2\alpha\xi^{2\alpha}h(2\log(k_n/\xi))k_n^{-(2\alpha+1)}, \tag{48}$$

as $n\to\infty$*.*

Proof.

Proof of statement 1. Let $0<\varepsilon<1$ be such that condition ii) holds and take $\varepsilon' = \min\{\varepsilon, 1/3\} < 1/2$. We first show that the integration over $(1-\varepsilon')R \le y \le R$ is negligible.

Since $h(y) = O\left(e^{\beta y}\right)$ we have

$$\begin{aligned}\int_{(1-\varepsilon')R_n}^{R_n} h(y)\hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y &= O\left(1\right)\hat{\rho}_n((1-\varepsilon')R,k_n)e^{-(\alpha-\beta)(1-\varepsilon')R}\\ &= O\left(\hat{\rho}_n((1-\varepsilon')R,k_n)n^{-2(\beta-\alpha)(1-\varepsilon')}\right).\end{aligned}$$

By assumption on $\hat{\mu}_n(y)$ we have that $\hat{\mu}_n((1-\varepsilon')R) = \Theta\left(\mu((1-\varepsilon')R\right) = \Theta\left(n^{2(1-\varepsilon')}\right)$. Since $k_n = o\left(n\right)$, by our choice of ε', $\hat{\mu}_n((1-\varepsilon')R)/k_n = \omega\left(n^{1-2\varepsilon'}\right) \to\infty$ as $n\to\infty$. We can now use Stirling's approximation to bound $\hat{\rho}_n((1-\varepsilon')R,k_n)$ as

$$\begin{aligned}\hat{\rho}_n((1-\varepsilon')R,k_n) &= \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n((1-\varepsilon')R)) = k_n\right)\\ &= \frac{\hat{\mu}_n((1-\varepsilon')R)^{k_n}}{k_n!}e^{-\hat{\mu}_n((1-\varepsilon')R)}\\ &= O\left(1\right)k_n^{-1/2}\left(\frac{\hat{\mu}_n((1-\varepsilon')R)}{k_n}\right)^{k_n} e^{k_n-\hat{\mu}_n((1-\varepsilon')R)}\\ &= O\left(1\right)k_n^{-1/2}e^{k_n\left(1-\frac{\hat{\mu}_n((1-\varepsilon')R)}{k_n}\right)+\log\left(\frac{\hat{\mu}_n((1-\varepsilon')R)}{k_n}\right)}\\ &\le O\left(1\right)k_n^{-1/2}e^{-\hat{\mu}_n((1-\varepsilon')R)/2},\end{aligned}$$

where the last line follows since $1-x+\log(x) \le -x/2$ for large enough x. We conclude that

$$\int_{(1-\varepsilon')R_n}^{R_n} h(y)\hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y = O\left(k_n^{-1/2}n^{-2(\beta-\alpha)(1-\varepsilon')}e^{-n^{2(1-\varepsilon')}}\right) = O\left(k_n^{(1+C^2)/2}\right).$$

We are left to show that for sufficiently large $C > 0$,

$$\int_0^{a_n^-} \hat{\rho}_n(y, k_n) e^{(\beta-\alpha)y} \, \mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right), \tag{49}$$

and

$$\int_{a_n^+}^{(1-\varepsilon')R_n} \hat{\rho}_n(y, k_n) e^{(\beta-\alpha)y} \, \mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right). \tag{50}$$

For simplicity we write $\mu(y) := \mu\left(\mathcal{B}_\infty\left(y\right)\right)$. Now fix some $0 < \delta < 1$ and let n be large enough such that

$$\sup_{0<y\le(1-\varepsilon')R_n} |\phi_n(y)| < \delta, \quad \lambda_n^+(1-\delta) > k_n \quad \text{and} \quad \lambda_n^-(1+\delta) < k_n.$$

Next, recall that the function $\lambda \mapsto \mathbb{P}\left(\mathrm{Po}(\lambda) = k\right)$ is monotonic increasing on $[0, k]$ and monotonic decreasing on $[k, \infty)$. Then since for n large enough we have

$$\hat{\mu}_n(y) = \mu(y)(1 + \phi_n(y)) \ge \mu(y)(1 - \delta) \ge \mu(a_n^+)(1 - \delta) = (\ell_n + C\kappa_n)(1 - \delta) > k_n,$$

it follows that

$$\hat{\rho}_n(y, k_n) = \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n(y)) = k_n\right) \le \mathbb{P}\left(\mathrm{Po}(\mu(y)(1-\delta)) = k_n\right),$$

for all $a_n^+ \le y \le (1-\varepsilon')R_n$. By making the change of variables $z = \mu^{-1}(\mu(y)(1-\delta)) = y + 2\log(1-\delta)$ we then get

$$\begin{aligned}
\int_{a_n^+}^{(1-\varepsilon')R_n} \hat{\rho}_n(y, k_n) e^{(\beta-\alpha)y} \, \mathrm{d}y &\le \int_{a_n^+}^{(1-\varepsilon')R_n} \mathbb{P}\left(\mathrm{Po}(\mu(y)(1-\delta)) = k_n\right) e^{(\beta-\alpha)y} \, \mathrm{d}y \\
&= (1-\delta)^{\beta-\alpha} \int_{a_n^+ + 2\log(1-\delta)}^{(1-\varepsilon')R_n + 2\log(1-\delta)} \mathbb{P}\left(\mathrm{Po}(\mu(z)) = k_n\right) e^{(\beta-\alpha)z} \, \mathrm{d}z \\
&= O\left(1\right) \int_{a_n^+}^{\infty} \rho(z, k_n) e^{(\beta-\alpha)z} \, \mathrm{d}z \\
&= O\left(k_n^{-(1+C^2)/2}\right),
\end{aligned}$$

where the last line is due to Lemma 6.1. This proves (50).

The proof of (49) follows the same line of reasoning. This time we use that for $0 \le y \le a_n^-$,

$$\hat{\mu}_n(y) \le \mu(y)(1 + \delta) \ge \mu(a_n^-)(1 + \delta) = (\ell_n - C\kappa_n)(1 + \delta) < k_n,$$

so that

$$\hat{\rho}_n(y, k_n) = \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n(y)) = k_n\right) \le \mathbb{P}\left(\mathrm{Po}(\mu(y)(1+\delta)) = k_n\right).$$

Making a similar change of variables $z = y + 2\log(1+\delta)$ we then get

$$\begin{aligned}
\int_0^{a_n^-} \hat{\rho}_n(y, k_n) e^{(\beta-\alpha)y} \, \mathrm{d}y &= \int_0^{a_n^-} \mathbb{P}\left(\mathrm{Po}(\mu(y)(1-\delta)) = k_n\right) e^{(\beta-\alpha)y} \, \mathrm{d}y \\
&\le (1+\delta)^{\beta-\alpha} \int_0^{a_n^- + 2\log(1+\delta)} \rho(z, k_n) e^{(\beta-\alpha)z} \, \mathrm{d}z,
\end{aligned}$$

and hence (49) follows by another application of Lemma 6.1.

Proof of statement 2. The proof of the second statement follows the same line of reasoning as above. First, by the mean value theorem for definite integrals

$$\int_{a_n^-}^{a_n^+} \hat{\rho}_n(y, k_n) h(y) e^{-\alpha y} \, \mathrm{d}y = h(c_n) \int_{a_n^-}^{a_n^+} \rho(z, k_n) e^{-\alpha z} \, \mathrm{d}z,$$

for some $c_n \in (a_n^-, a_n^+)$. Since, $c_n \sim 2\log(k_n/\xi)$, by assumption on h, $h(c_n) \sim h(2\log(k_n/\xi))$. Therefore it is enough to show that

$$\int_{a_n^-}^{a_n^+} \hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y = (1+o(1))\int_{a_n^-}^{a_n^+} \rho(z,k_n)e^{-\alpha z}\,\mathrm{d}z.$$

or equivalently,

$$\int_{a_n^-}^{a_n^+} \rho(z,k_n)e^{-\alpha z}\,\mathrm{d}z = (1+o(1))\int_{a_n^-}^{a_n^+} \hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y.$$

Using the change of variable $z = \mu^{-1}(\hat{\mu}_n(y))$ and writing $\hat{a}_n^{\pm} = \hat{\mu}_n^{-1}(\mu(a_n^{\pm}))$, we get

$$\begin{aligned}\int_{a_n^-}^{a_n^+} \rho(z,k_n)e^{-\alpha z}\,\mathrm{d}z &= \int_{a_n^-}^{a_n^+} \mathbb{P}\left(\mathrm{Po}(\mu(z)) = k_n\right)e^{-\alpha z}\,\mathrm{d}z \\ &= \int_{\hat{a}_n^-}^{\hat{a}_n^+} \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n(y)) = k_n\right)e^{-\alpha\mu^{-1}(\hat{\mu}_n(y))}\frac{\hat{\mu}_n'(y)}{\mu'(\mu^{-1}(\hat{\mu}_n(y)))}\,\mathrm{d}y,\end{aligned}$$

where the fraction in the last line follows from the chain rule and the fact that $(\mu^{-1})'(t) = (\mu'(\mu^{-1}(t)))^{-1}$.

Now recall that $\hat{\mu}_n(y) = \mu(y)(1+\phi_n(y))$, with $\phi_n(y)$ satisfying conditions i) and ii). It then follows $\hat{\mu}_n'(y) = (1+o(1))\mu'(y)$ as $n \to \infty$, uniformly on (a_n^-, a_n^+). Next we note that $\mu'(y) = \mu(y)/2$ from which it follows that uniformly on $[a_n^-, a_n^+]$,

$$\frac{\hat{\mu}_n'(y)}{\mu'(\mu^{-1}(\hat{\mu}_n(y)))} = \frac{2\hat{\mu}_n'(y)}{\hat{\mu}_n(y)} = \frac{(1+o(1))2\mu'(y)}{(1+o(1))\mu(y)} = (1+o(1))$$

Therefore,

$$\begin{aligned}\int_{a_n^-}^{a_n^+} \rho(z,k_n)e^{-\alpha z}\,\mathrm{d}z &= (1+o(1))\int_{\hat{a}_n^-}^{\hat{a}_n^+} \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n(y)) = k_n\right)e^{-\alpha y}\,\mathrm{d}y \\ &= (1+o(1))\int_{a_n^-}^{a_n^+} \hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y \\ &= (1+o(1))\int_{a_n^-}^{a_n^+} \hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y\end{aligned}$$

which finishes the proof. □

With this lemma, the proof of Proposition 6.5 follows if we can show that both $\rho_{\mathrm{Po}}(y,k_n)$ or $\rho_{\mathrm{box}}(y,k_n)$ are of the form $\rho(y,k_n)(1+\phi_n(y))$, where $\phi_n(y)$ satisfies the conditions i) and ii). We deal with each of these cases separately.

6.3 Concentration of heights for the finite box model

The following lemma immediately implies that $\mu\left(\mathcal{B}_{\mathrm{box}}\left(y\right)\right)$ satisfies the conditions of Lemma 6.6 and hence that Proposition 6.5 holds for $\rho_{\mathrm{box}}(y,k_n)$.

Lemma 6.7. *For all $y > 2\log(\pi/2)$,*

$$\mu\left(\mathcal{B}_{box}\left(p\right)\right) = \mu\left(\mathcal{B}_{\infty}\left(p\right)\right)\left(1-\phi_n(y)\right)$$

where $\phi_n(y) \geq 0$ is given by

$$\phi_n(y) = \left(\frac{\pi}{2}\right)^{-(2\alpha-1)}e^{-(\alpha-\frac{1}{2})(R_n-y)} - \frac{(2\alpha-1)\pi}{4\alpha}\left(\left(\frac{\pi}{2}\right)^{-2\alpha}e^{-(\alpha-\frac{1}{2})(R_n-y)} - e^{-(\alpha-\frac{1}{2})R_n-\frac{y}{2}}\right).$$

On the other hand, if $y \leq 2\log(\pi/2)$ then

$$\mu_{\alpha,\nu}(\mathcal{B}_{box}\left(p\right)) = \mu_{\alpha,\nu}(\mathcal{B}_{\infty}\left(p\right))\left(1-e^{-(\alpha-\frac{1}{2})R_n}\right).$$

Proof. First note that since we have identified the boundaries of $[-\frac{\pi}{2}e^{\frac{R_n}{2}}, \frac{\pi}{2}e^{\frac{R_n}{2}}]$ we can assume, without loss of generality, that $p = (0, y)$. We then have that the boundaries of $\mathcal{B}_{\text{box}}(p)$ are given by the equations $x' = \pm e^{\frac{y+y'}{2}}$, which intersect the left and right boundaries of $[-\frac{\pi}{2}e^{\frac{R_n}{2}}, \frac{\pi}{2}e^{\frac{R_n}{2}}]$ at height

$$h(y) = R_n + 2\log\left(\frac{\pi}{2}\right) - y.$$

Therefore, if $y \le 2\log(\pi/2)$ this intersection occurs above the height R_n of the box $\mathcal{R}_n$ while in the other case the full region of the box above $h(y)$ is connected to p.

We will first consider the case where $y > 2\log(\pi/2)$. Recall that $\mu_{\alpha,\nu}(\mathcal{B}_\infty(p)) = \xi e^{\frac{y}{2}}$ where $\xi = \frac{4\alpha\nu}{(2\alpha-1)\pi}$. Then, after some simple algebra, we have that

$$\begin{aligned}
\mu_{\alpha,\nu}(\mathcal{B}_{\text{box}}(p)) &= \int_0^{h(y)} \int_{-\frac{\pi}{2}e^{\frac{R_n}{2}}}^{\frac{\pi}{2}e^{\frac{R_n}{2}}} \mathbb{1}_{\left\{|x'|\le e^{\frac{y+y'}{2}}\right\}} f_{\alpha,\nu}(x',y')\, dx'\, dy' \\
&\quad + \int_{h(y)}^{R_n} \int_{-\frac{\pi}{2}e^{\frac{R_n}{2}}}^{\frac{\pi}{2}e^{\frac{R_n}{2}}} f_{\alpha,\nu}(x',y')\, dx'\, dy' \\
&= \frac{2\alpha\nu}{\pi} e^{\frac{y}{2}} \int_0^{h(y)} e^{-(\alpha-\frac{1}{2})y'}\, dy' + \alpha\nu e^{\frac{R_n}{2}} \int_{h(y)}^{R_n} e^{-\alpha y'}\, dy' \\
&= \xi e^{\frac{y}{2}} \left(1 - \left(\frac{\pi}{2}\right)^{-(2\alpha-1)} e^{-(\alpha-\frac{1}{2})(R_n-y)}\right) \\
&\quad + \nu e^{\frac{R_n}{2}} \left(\left(\frac{\pi}{2}\right)^{-2\alpha} e^{-\alpha(R_n-y)} - e^{-\alpha R_n}\right) \\
&= \mu_{\alpha,\nu}(\mathcal{B}_\infty(p))\left(1 - \phi_n(y)\right).
\end{aligned}$$

Since, for all $\alpha > \frac{1}{2}$,

$$\left(\frac{\pi}{2}\right)^{-(2\alpha-1)} \ge \frac{(2\alpha-1)\pi}{4\alpha}\left(\frac{\pi}{2}\right)^{-2\alpha}$$

it follows that $\phi_n(y) \ge 0$.

When $y \le 2\log(\pi/2)$ we have

$$\begin{aligned}
\mu_{\alpha,\nu}(\mathcal{B}_{\text{box}}(p)) &= \int_0^{R_n} \int_{-\frac{\pi}{2}e^{\frac{R_n}{2}}}^{\frac{\pi}{2}e^{\frac{R_n}{2}}} \mathbb{1}_{\left\{|x'|\le e^{\frac{y+y'}{2}}\right\}} f_{\alpha,\nu}(x',y')\, dx'\, dy' \\
&= \frac{2\alpha\nu}{\pi} e^{\frac{y}{2}} \int_0^{R_n} e^{-(\alpha-\frac{1}{2})y'}\, dy' \\
&= \mu_{\alpha,\nu}(\mathcal{B}_\infty(p))\left(1 - e^{-(\alpha-\frac{1}{2})R_n}\right).
\end{aligned}$$

□

6.4 Concentration of heights for the KPKVB model

We will now show that a concentration of heights argument also applies to the KPKVB model. Due to the hyperbolic distance formula, the computations are however more involved than for the finite box model. Recall that under the coupling between the hyperbolic random graph and the finite box model, for two points p, p' with $y + y' < R_n$, $p' \in \mathcal{B}(p)$ exactly when $|x - x'|_{\pi e^{R_n/2}} \le \Phi(y, y')$, where we slightly abused notation to write $\Phi(y, y') = \Phi(R - y, R - y')$. In this setting, the coupling lemma (Lemma 2.2) gives that

$$e^{\frac{1}{2}(y+y')} - Ke^{\frac{3}{2}(y+y')-R_n} \le \Phi(y, y') \le e^{\frac{1}{2}(y+y')} + Ke^{\frac{3}{2}(y+y')-R_n},$$

for some constant K. This result enables us to determine the measure of a ball around a given point $p = (0, y)$. Recall that the hyperbolic ball $\mathcal{B}(p)$ is a subset of $\mathcal{R}_n$ and not of the hyperbolic disc $\mathcal{D}_{R_n}$, i.e. the balls $\mathcal{B}(p)$ "live" in the finite box and not the hyperbolic disc. We start with the following preliminary result.

Lemma 6.8. *Let $\Phi(y,y') := \Phi(R-y, R-y')$ be defined as in (6). Then, for any $0 \le \delta < 1$*

$$\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \mu\left(\mathcal{B}_\infty(y)\right)^{-1} \frac{2\nu\alpha}{\pi} \int_0^{(1-\delta)(R-y)} \Phi(y,y') e^{-\alpha y'} = 1.$$

Proof. Recall that $\mu\left(\mathcal{B}_\infty(0,y)\right) = \xi e^{y/2}$ where $\xi = \frac{4\alpha\nu}{\pi(2\alpha-1)}$. Using Lemma 2.2 we have

$$\begin{aligned}
\frac{2\nu\alpha}{\pi} \int_0^{(1-\delta)(R-y)} \Phi(y,y') e^{-\alpha y'} &\le \frac{2\nu\alpha}{\pi} \int_0^{(1-\delta)(R-y)} \left(e^{\frac{y+y'}{2}} + Ke^{\frac{3}{2}(y+y')-R_n}\right) e^{-\alpha y'}\, dy' \\
&= \mu\left(\mathcal{B}_\infty(0,y)\right)\left(1 - e^{-(\alpha-\frac{1}{2})(1-\delta)(R_n-y)}\right) \\
&\quad + \frac{2\nu\alpha}{\pi} K e^{\frac{3y}{2}-R_n} \int_0^{(1-\delta)(R-y)} e^{(\frac{3}{2}-\alpha)y'}\, dy'
\end{aligned}$$

We first compute the integral, which depends on the value of α,

$$\int_0^{(1-\delta)(R-y)} e^{(\frac{3}{2}-\alpha)y'}\, dy' = \begin{cases} \frac{2}{3-2\alpha}\left(e^{(\frac{3}{2}-\alpha)(1-\delta)(R-y)} - 1\right) & \text{if } 1/2 < \alpha < 3/2, \\ (1-\delta)(R-y) & \text{if } \alpha = 3/2, \\ \frac{2}{2\alpha-3}\left(1 - e^{-(\alpha-\frac{3}{2})(1-\delta)(R-y)}\right) & \text{if } \alpha > 3/2. \end{cases}$$

Therefore we get

$$\begin{aligned}
&\frac{2\nu\alpha}{\pi} K e^{\frac{3y}{2}-R_n} \int_0^{(1-\delta)(R-y)} e^{(\frac{3}{2}-\alpha)y'}\, dy' \\
&= \mu\left(\mathcal{B}_\infty(0,y)\right) \begin{cases} \frac{(2\alpha-1)K}{3-2\alpha}\left(e^{-(\alpha-\frac{1}{2})(R_n-y)-(\frac{3}{2}-\alpha)\delta(R-y)} - e^{-(R_n-y)}\right) & \text{if } 1/2 < \alpha < 3/2, \\ \frac{(2\alpha-1)K}{2}(1-\delta)(R-y)e^{-(R_n-y)} & \text{if } \alpha = 3/2, \\ \frac{(2\alpha-1)K}{2\alpha-3}\left(e^{-(R_n-y)} - e^{-(\alpha-\frac{1}{2})(R_n-y)-(\alpha-\frac{3}{2})(R-y)}\right) & \text{if } \alpha > 3/2, \end{cases}
\end{aligned}$$

and hence

$$\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \mu\left(\mathcal{B}_\infty(y)\right)^{-1} \frac{2\nu\alpha}{\pi} K e^{\frac{3y}{2}-R_n} \int_0^{(1-\delta)(R-y)} e^{(\frac{3}{2}-\alpha)y'}\, dy' = 0.$$

Since $\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} e^{-(\alpha-\frac{1}{2})(1-\delta)(R_n-y)} = 0$, we conclude that

$$\limsup_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \mu\left(\mathcal{B}_\infty(y)\right)^{-1} \int_0^{(1-\delta)(R-y)} \Phi(y,y')\alpha e^{-\alpha y'} = 1.$$

The proof that this also holds for the limit infimum immediately follows, by observing that the only difference with the above computations is the change of sign in front of

$$\frac{2\nu\alpha}{\pi} K e^{\frac{3y}{2}-R_n} \int_0^{R_n-y} e^{(\frac{3}{2}-\alpha)y'}\, dy'.$$

□

We can now show that the measure of the balls in the KPKVB model and the infinite model are asymptotically equivalent.

Lemma 6.9. *For any $0 < \varepsilon < 1$*

$$\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \frac{\mu\left(\mathcal{B}(0,y)\right)}{\mu\left(\mathcal{B}_\infty(0,y)\right)} = 1.$$

Proof. We perform the computation of $\mu(\mathcal{B}(0,y))$ by splitting the integration with respect to the height y' into the cases $y' > R_n - y$ and $y' \le R_n - y$,

$$\mu(\mathcal{B}(y)) = \mu(\mathcal{B}(y) \cap \mathcal{R}_n([0, R_n - y))) + \mu(\mathcal{B}(y) \cap \mathcal{R}_n([R_n - y, R_n])).$$

For the first part we have that

$$\mu(\mathcal{B}((0,y)) \cap \mathcal{R}_n[(0, R_n - y)]) = \frac{2\nu\alpha}{\pi} \int_0^{R-y} \Phi(y, y') e^{-\alpha y'} \, \mathrm{d}y'.$$

Hence, by applying Lemma 6.8 with $\delta = 0$ we conclude that

$$\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \mu(\mathcal{B}_\infty(y))^{-1} \mu(\mathcal{B}((0,y)) \cap \mathcal{R}_n[(0, R_n - y)]) = 1.$$

For the second part we observer that $\mathcal{B}((0,y)) \cap \mathcal{R}_n([R_n - y, R_n]) = \mathcal{R}_n([R_n - y, R_n])$. Thus,

$$\begin{aligned}
&\mu(\mathcal{B}((0,y)) \cap \mathcal{R}_n([R_n - y, R_n])) \\
&\quad = \int_{R_n - y}^{R_n} \int_{I_n} f_{\alpha,\nu}(x', y') \, \mathrm{d}x' \, \mathrm{d}y' = \nu\alpha e^{R_n/2} \left(e^{-\alpha(R_n - y)} - e^{-\alpha R_n}\right) \\
&\quad = \mu(\mathcal{B}_\infty(0,y)) \frac{2\alpha - 1}{4\pi} \left(e^{-(\alpha - \frac{1}{2})(R_n - y)} - e^{-(\alpha - \frac{1}{2})R_n - y/2}\right),
\end{aligned} \tag{51}$$

from which we conclude that

$$\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \mu(\mathcal{B}_\infty(y))^{-1} \mu(\mathcal{B}((0,y)) \cap \mathcal{R}_n([R_n - y, R_n])) = 0,$$

which finishes the proof. $\square$

A direct consequence of Lemma 6.9 is that $\mu(\mathcal{B}(y)) = \mu(\mathcal{B}_\infty(y))(1 + \phi_n(y))$, where $\phi_n(y) := \mu(\mathcal{B}(y)) / \mu(\mathcal{B}_\infty(y)) - 1$ satisfies condition i) in Lemma 6.6. To show that condition ii) is also satisfied we need to analyze

$$\phi_n'(y) = \mu(\mathcal{B}_\infty(y))^{-1} \frac{\partial}{\partial y} \mu(\mathcal{B}(y)) - \frac{1}{2} \frac{\mu(\mathcal{B}(y))}{\mu(\mathcal{B}_\infty(y))},$$

where we used that $\frac{\partial}{\partial y} \mu(\mathcal{B}_\infty(y)) = \frac{1}{2} \mu(\mathcal{B}_\infty(y))$. Again, Lemma 6.9 implies that

$$\lim_{n\to\infty} \sup_{0\le y\le(1-\varepsilon)R_n} \frac{1}{2} \frac{\mu(\mathcal{B}(y))}{\mu(\mathcal{B}_\infty(y))} = \frac{1}{2}.$$

The following lemma shows that the same holds for the first term from which we conclude that $\phi_n(y)$ satisfies condition ii) in Lemma 6.6 and hence that Proposition 6.5 also holds for $\rho_{\mathrm{Po}}(y, k_n)$.

Lemma 6.10. *For any* $0 < \varepsilon < 1$,

$$\lim_{n\to\infty} \sup_{0\le y\le(1-\varepsilon)R_n} \mu(\mathcal{B}_\infty(y))^{-1} \frac{\partial}{\partial y} \mu(\mathcal{B}(y)) = \frac{1}{2}.$$

Proof. We again split $\mu(\mathcal{B}(y))$ over the top and bottom part,

$$\mu(\mathcal{B}(y)) = \mu(\mathcal{B}(y) \cap \mathcal{R}_n([0, R_n - y))) + \mu(\mathcal{B}(y) \cap \mathcal{R}_n([R_n - y, R_n])),$$

where

$$\mu(\mathcal{B}(y) \cap \mathcal{R}_n([0, R_n - y))) = \frac{2\alpha\nu}{\pi} \int_0^{R-y} \Phi(y, y') e^{-\alpha y'} \, \mathrm{d}y',$$

with $\Phi(y,y')$ defined as in (6) and, see (51),

$$\mu\left(\mathcal{B}\left(y\right)\cap\mathcal{R}_n([R_n-y,R_n])\right)=\xi e^{y/2}\frac{2\alpha-1}{4\pi}\left(e^{-(\alpha-\frac{1}{2})(R_n-y)}-e^{-(\alpha-\frac{1}{2})R_n-y/2}\right).$$

Taking the derivative of the last expression gives

$$\begin{aligned}
&\frac{\partial}{\partial y}\mu\left(\mathcal{B}\left(y\right)\cap\mathcal{R}_n([R_n-y,R_n])\right)\\
&=\frac{1}{2}\mu\left(\mathcal{B}\left(y\right)\cap\mathcal{R}_n([R_n-y,R_n])\right)+\xi e^{y/2}\frac{2\alpha-1}{4\pi}\left(\left(\alpha-\frac{1}{2}\right)e^{-(\alpha-\frac{1}{2})(R_n-y)}+\frac{1}{2}e^{-(\alpha-\frac{1}{2})R_n-y/2}\right)\\
&=\frac{1}{2}\mu\left(\mathcal{B}\left(y\right)\cap\mathcal{R}_n([R_n-y,R_n])\right)\left(1+\frac{(2\alpha-1)e^{-(\alpha-\frac{1}{2})(R_n-y)}+e^{-(\alpha-\frac{1}{2})R_n-y/2}}{e^{-(\alpha-\frac{1}{2})(R_n-y)}-e^{-(\alpha-\frac{1}{2})R_n-y/2}}\right).
\end{aligned}$$

Since, $\lim_{n\to\infty}\sup_{0<y\leq(1-\varepsilon)R_n}\mu\left(\mathcal{B}_\infty\left(y\right)\right)^{-1}\mu\left(\mathcal{B}\left(y\right)\cap\mathcal{R}_n([R_n-y,R_n])\right)=0$, we are left to show that

$$\lim_{n\to\infty}\sup_{0<y\leq(1-\varepsilon)R_n}\mu\left(\mathcal{B}_\infty\left(y\right)\right)^{-1}\frac{2\alpha\nu}{\pi}\frac{\partial}{\partial y}\int_0^{R-y}\Phi(r,r')e^{-\alpha y'}\,\mathrm{d}y'=\frac{1}{2}. \tag{52}$$

We start with some preliminary computations. For convenience we define

$$\Xi_n(y,y')=1-\frac{\cosh(R-y)\cosh(R-y')-\cosh(R)}{\sinh(R-y)\sinh(R-y')},$$

so that

$$\Phi(R-y,R-y')=\frac{1}{2}e^{R/2}\arccos\left(1-\Xi_n(y,y')\right).$$

Next, following the same calculation as in the proof of [8, Lemma 28], we write

$$\begin{aligned}
\Xi(y,y')&=2e^{-(R-y-y')}\frac{\left(1-e^{y'-y-R}\right)\left(1-e^{y-y'-R}\right)}{\left(1-e^{-2(R-y')}\right)\left(1-e^{-2(R-y)}\right)}\\
&:=2e^{-(R-y-y')}\frac{h_1(y)h_2(y)}{h_3(y')h_3(y)},
\end{aligned}$$

with

$$h_1(y)=1-e^{y'-y-R},\quad h_2(y)=1-e^{y-y'-R}\quad\text{and}\quad h_3(y)=1-e^{-2(R-y)}.$$

We suppressed the dependence on n and, in some cases, on y' for notation convenience.

We make two important observations. First, $\Xi(y,y')$ is an increasing function in both arguments, for $y,y'<R$ and $y+y'<R$. Second, for all $y+y'<R$, $h_1(y)\leq h_3(y')$ and $h_2(y)\leq h_3(y)$, while $h_3(y),h_3(y')<1$, so that

$$2e^{-(R-y-y')}h_1(y)h_2(y)\leq\Xi_n(y,y')\leq 2e^{-(R-y-y')}. \tag{53}$$

In particular, there exists a $0<\delta<1$ such that $1\leq\Xi(y,y')\leq 2$ for all $0<y<R$ and $(1-\delta)(R-y)<y'<R$.

Next, taking the derivative of $\Xi_n(y,y')$ yields,

$$\begin{aligned}
\frac{\partial}{\partial y}\Xi(y,y')&=\Xi(y,y')+2e^{-(R-y-y')}\left(\frac{h_1'(y)h_2(y)}{h_3(y')h_3(y)}+\frac{h_1(y)h_2'(y)}{h_3(y')h_3(y)}-\frac{h_1(y)h_2(y)h_3'(y)}{h_3(y')h_3(y)^2}\right)\\
&=\Xi(y,y')\left(1+\frac{h_1'(y)}{h_1(y)}+\frac{h_2'(y)}{h_2(y)}-\frac{h_3'(y)}{h_3(y)}\right)\\
&:=\Xi(y,y')\left(1+\varphi_n(y,y')\right),
\end{aligned}$$

with

$$\varphi_n(y,y') = \frac{e^{y'-y-R}}{1-e^{y'-y-R}} - \frac{e^{y-y'-R}}{1-e^{y-y'-R}} - \frac{2e^{-2(R-y)}}{1-e^{-2(R-y)}}.$$

Therefore, by the chain rule,

$$\begin{aligned}\frac{\partial}{\partial y}\Phi(R-y,R-y') &= \frac{1}{2}e^{R/2}\frac{1}{\sqrt{1-(1-\Xi(y,y'))^2}}\frac{\partial}{\partial y}\Xi(y,y') \\ &= \frac{\frac{1}{2}e^{R/2}\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}}\left(1+\varphi_n(y,y')\right). \end{aligned} \tag{54}$$

Applying the Leibniz's rule we then get

$$\begin{aligned}&\frac{\partial}{\partial y}\int_0^{R-y}\Phi(y,y')\alpha e^{-\alpha y'}\,\mathrm{d}y' \\ &\quad = -\alpha\Phi(y,R-y)e^{-\alpha(R-y)} + \int_0^{R-y}\frac{\partial}{\partial y}\Phi(y,y')\alpha e^{-\alpha y'}\,\mathrm{d}y' \\ &\quad = -\frac{\pi}{2}e^{-(\alpha-\frac{1}{2})R_n+\alpha y} + \int_0^{R-y}\frac{\frac{1}{2}e^{R/2}\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}}\left(1+\varphi_n(y,y')\right)\alpha e^{-\alpha y'}\,\mathrm{d}y' \\ &\quad = -\frac{\pi}{2}e^{-(\alpha-\frac{1}{2})R_n+\alpha y} + \int_0^{(1-\delta)(R-y)}\frac{\frac{1}{2}e^{R/2}\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}}\left(1+\varphi_n(y,y')\right)\alpha e^{-\alpha y'}\,\mathrm{d}y' \\ &\qquad + \int_{(1-\delta)(R-y)}^{R-y}\frac{\frac{1}{2}e^{R/2}\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}}\left(1+\varphi_n(y,y')\right)\alpha e^{-\alpha y'}\,\mathrm{d}y' \\ &\quad := -I_1(y)+I_2(y)+I_3(y),\end{aligned}$$

with δ such that $1 \le \Xi(y,y') \le 2$ for all $0 < y < R$ and $(1-\delta)(R-y) < y' < R$.

We proceed by showing that

$$\lim_{n\to\infty}\sup_{0<y\le(1-\varepsilon)R_n}\mu\left(\mathcal{B}_\infty(y)\right)^{-1}I_t(y) = 0, \quad \text{for } t=1,3 \tag{55}$$

while

$$\lim_{n\to\infty}\sup_{0\le y\le(1-\varepsilon)R_n}\mu\left(\mathcal{B}_\infty(y)\right)^{-1}\frac{2\nu\alpha}{\pi}I_2(y) = \frac{1}{2}. \tag{56}$$

This then implies (52) and finishes the proof.

Let us first consider $I_1(y)$. Since

$$\lim_{n\to\infty}\sup_{0<y\le(1-\varepsilon)R_n}\mu\left(\mathcal{B}_\infty(y)\right)^{-1}I_1(y) \le \lim_{n\to\infty}\sup_{0<y\le(1-\varepsilon)R_n}\frac{\pi}{2\xi}e^{-(\alpha-\frac{1}{2})(R-y)} = 0.$$

For $I_3(y)$ we first use that $y' < R-y$ to bound $\varphi(y,y')$ as follows,

$$\varphi_n(y,y') \le \frac{e^{y'-y-R}}{1-e^{y'-y-R}} \le \frac{e^{-2y}}{1-e^{-2y}}.$$

This then yields that

$$I_3(y) \le \frac{\alpha}{2}\left(1+\frac{e^{-2y}}{1-e^{-2y}}\right)e^{R/2}\int_{(1-\delta)(R-y)}^{R-y}\frac{\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}}e^{-\alpha y'}\,\mathrm{d}y'.$$

Since for all $1 \le x < 2$,

$$\frac{1}{\sqrt{1-(1-x)^2}} \le \frac{2}{\sqrt{2(2-x)}},$$

and $1 \le \Xi(y,y') \le 2$, for all $(1-\delta)(R-y) \le y\prime < R-y$ and $y < R$, it follows that

$$\begin{aligned}
&\int_{(1-\delta)(R-y)}^{R-y} \frac{\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}} e^{-\alpha y'}\,\mathrm{d}y' \\
&\qquad \le 2\int_{(1-\delta)(R-y)}^{R-y} \frac{\Xi(y,y')}{\sqrt{2(2-\Xi(y,y'))}} e^{-\alpha y'}\,\mathrm{d}y' \\
&\qquad \le 2e^{-\alpha(R-y)} \int_{(1-\delta)(R-y)}^{R-y} \frac{e^{-(R-y-y')}}{\sqrt{1-e^{-(R-y-y')}}} e^{\alpha(R-y-y')}\,\mathrm{d}y'.
\end{aligned}$$

Making the change of variables $z = e^{-(R-y-y')}$ ($\mathrm{d}y' = z^{-1}\,\mathrm{d}z$) we get that

$$\begin{aligned}
2e^{-\alpha(R-y)} \int_{(1-\delta)(R-y)}^{R-y} \frac{e^{-(R-y-y')}}{\sqrt{1-e^{-(R-y-y')}}} e^{\alpha(R-y-y')}\,\mathrm{d}y' &= 2e^{-\alpha(R-y)} \int_{e^{-\delta(R-y)}}^{1} \frac{z^{-\alpha}}{\sqrt{1-z}}\,\mathrm{d}z \\
&\le 2e^{-\alpha(R-y)} \sqrt{1-e^{-\delta(R-y)}} \\
&\le 2e^{-\alpha(R-y)}.
\end{aligned}$$

We therefore conclude that

$$I_3(y) \le \alpha \left(1 + \frac{e^{-2y}}{1-e^{-2y}}\right) e^{-(\alpha-\frac{1}{2})R+\alpha y}.$$

which implies

$$\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \mu\left(\mathcal{B}_\infty(y)\right)^{-1} I_3(y) \le \lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R_n} \alpha \left(1 + \frac{e^{-2y}}{1-e^{-2y}}\right) e^{-(\alpha-\frac{1}{2})(R-y)} = 0.$$

Finally, to show (56) we define

$$\begin{aligned}
\varphi_n^- &:= \sup_{0<y\le(1-\varepsilon)R_n} \inf_{0\le y'\le(1-\delta)(R-y)} \varphi_n(y,y') \\
\varphi_n^+ &:= \sup_{0<y\le(1-\varepsilon)R_n} \sup_{0\le y'\le(1-\delta)(R-y)} \varphi_n(y,y')
\end{aligned}$$

and note that since for $0 \le y' \le (1-\delta)(R-y)$,

$$\frac{e^{-(R+y)}}{1-e^{-(R+y)}} - \frac{e^{-(R-y)}}{1-e^{-(R-y)}} - \frac{2e^{-2(R-y)}}{1-e^{-2(R-y)}} \le \varphi_n(y,y') \le \frac{e^{-\delta(R-y)}}{1-e^{-\delta(R-y)}}.$$

we get $\lim_{n\to\infty}\varphi_n^\pm = 0$. Next, recall that $\Phi(y,y') = \frac{1}{2}e^{R/2}\Xi(y,y')$. Then, since $\Xi(y,y') < 2$ for all $y' < (1-\delta)(R-y)$ and $y < R$, there exists a $K > 0$ such that (see Lemma C.1),

$$\frac{1}{2}\Phi(y,y')\left(1 - \frac{(1+\sqrt{2})\Xi(y,y')}{1+\Xi(y,y')}\right) \le \frac{\frac{1}{2}e^{R/2}\Xi(y,y')}{\sqrt{1-(1-\Xi(y,y'))^2}} \le \frac{1}{2}\Phi(y,y')\left(1 + \frac{(1+K)\Xi(y,y')}{1-\Xi(y,y')}\right)$$

for all $y' < (1-\delta)(R-y)$ and $y < R$. Using that $e^{-(r-y)} \le e^{-(R-y-y')} \le e^{-\delta(R-y)}$, for $0 < y' < (1-\delta)(R-y)$, we get

$$1 - \frac{(1+\sqrt{2})\Xi(y,y')}{1+\Xi(y,y')} \ge 1 - 2(1+\sqrt{2})e^{-(R-y)}$$

and

$$1 + \frac{(1+K)\Xi(y,y')}{1-\Xi(y,y')} \le 1 + \frac{2(1+K)e^{-\delta(R-y)}}{1-e^{-\delta(R-y)}}.$$

We thus have the following upper and lower bound on $I_2(y)$

$$I_2(y) \leq \frac{1}{2}\left(1+\phi_n^+\right)\left(1+\frac{2(1+K)e^{-\delta(R-y)}}{1-e^{-\delta(R-y)}}\right)\int_0^{(1-\delta)(R-y)} \Phi(y,y')\alpha e^{-\alpha y'}\,\mathrm{d}y',$$

and

$$I_2(y) \geq \frac{1}{2}\left(1+\phi_n^-\right)\left(1-2(1+\sqrt{2})e^{-(R-y)}\right)\int_0^{(1-\delta)(R-y)} \Phi(y,y')\alpha e^{-\alpha y'}\,\mathrm{d}y'.$$

From this (56) follows since, $\lim_{n\to\infty}\varphi_n^{\pm}=0$ and by Lemma 6.8,

$$\lim_{n\to\infty}\sup_{0<y\leq(1-\varepsilon)R_n} \mu\left(\mathcal{B}_\infty\left(y\right)\right)^{-1}\frac{2\nu\alpha}{\pi}\int_0^{(1-\delta)(R-y)} \Phi(y,y')\alpha e^{-\alpha y'}\,\mathrm{d}y' = 1.$$

□

7 From G_{box} to G_∞ (Proving Proposition 5.5)

In this section we shall relate the clustering in the finite box model G_{box} to that of the infinite model. The main goal is to prove Proposition 5.5 which states that

$$\left|\mathbb{E}\left[c^*(k_n;G_{\text{box}})\right]-\gamma(k_n)\right| = o\left(s(k_n)\right).$$

Recall that G_{box} is obtained by restricting the Poisson Point Process $\mathcal{P}_{\alpha,\nu}$ to the box $\mathcal{R} = (-I_n, I_n]\times(0,R_n]$, with $I_n = \frac{\pi}{2}e^{R_n/2}$ and connecting two points $p_1,p_2\in\mathcal{R}$ if $|x_1-x_2|_{\pi e^{R_n/2}} \leq e^{(y_1+y_2)/2}$. We also recall that by definition of the norm $|.|_{\pi e^{R_n/2}}$ the left and right boundaries of $\mathcal{R}$ are identified. See Section 2.2 for more details. Due to this identification of the boundaries some triples of nodes that form triangles in the finite box model do not form a triangle in the infinite model. Therefore, to establish the required result we need to compute the asymptotic difference between triangle counts in both models.

For any $p\in\mathbb{R}\times\mathbb{R}_+$ we define for the finite box model,

$$T_{\text{box}}(p) = \sum_{p_1,p_2\in\mathcal{P}_n\setminus p}^{\neq} T_{\text{box}}(p,p_1,p_2)$$

where the sum is over all distinct pairs in $\mathcal{P}_n\setminus p$ and

$$T_{\text{box}}(p,p_1,p_2) = \mathbb{1}_{\{p_1\in\mathcal{B}_{\text{box}}(p)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\text{box}}(p)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\text{box}}(p_1)\}}.$$

Similarly, for the infinite model we define

$$T_\infty(y) = \sum_{p_1,p_2\in\mathcal{P}\setminus\{0,y\}}^{\neq} T_\infty(y,p_1,p_2),$$

where

$$T_\infty(y,p_1,p_2) = \mathbb{1}_{\{p_1\in\mathcal{B}_\infty(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_\infty(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_\infty(p_1)\}}.$$

We will first relate $\gamma(k_n)$ and $\mathbb{E}\left[c^*(k_n;G_{\text{box}})\right]$ using $T_\infty(y)$ and $T_{\text{box}}(y)$. Recall the definition of $\mathcal{K}_C(k_n)$

$$\mathcal{K}_C(k_n) = \left\{y\in\mathbb{R}_+ : \frac{k_n - C\sqrt{k_n\log(k_n)}}{\xi} \leq e^{\frac{y}{2}} \leq \frac{k_n + C\sqrt{k_n\log(k_n)}}{\xi}\right\},$$

Lemma 7.1. *Let $\gamma(k_n)$ be defined as in (13). Then as $n \to \infty$*

$$\gamma(k_n) = (1 + o(1)) \frac{1}{k_n^2 p_{k_n}} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_\infty(y)\right] \rho(y,k) \alpha e^{-\alpha y} \, \mathrm{d}y. \tag{57}$$

Moreover,

$$\mathbb{E}\left[c^*(k_n; G_{box})\right] = (1 + o(1)) \frac{1}{k_n^2 p_{k_n}} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_{box}(y)\right] \rho(y,k_n) \alpha e^{-\alpha y} \, \mathrm{d}y \tag{58}$$

as $n \to \infty$,

Proof. Recall that

$$P(y) = \mathbb{E}\left[\mathbb{1}_{\{u_1 \in \mathcal{B}_\infty(u_2)\}}\right],$$

where u_1 and u_2 are independent and distributed according to the probability density $\mu\left(\mathcal{B}_\infty\left(y\right)\right)^{-1} \mathbb{1}_{\{u_i \in \mathcal{B}_\infty(y)\}} f(x_i, y_i)$. It then follows from the Campbell-Mecke formula that

$$\begin{aligned}
\mathbb{E}\left[T_\infty(y)\right] &= \int \mathbb{1}_{\{p_1 \in \mathcal{B}_\infty(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_\infty(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_\infty(p_1)\}} f(x_1, y_1) f(x_2, y_2) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}y_1 \, \mathrm{d}y_2 \\
&= \mu\left(\mathcal{B}_\infty\left(y\right)\right)^2 P(y).
\end{aligned}$$

It then follows that,

$$\begin{aligned}
\gamma(k_n) &= \frac{1}{p_{k_n}} \cdot \int_0^\infty P(y) \rho(y,k) \alpha e^{-\alpha y} \, \mathrm{d}y \\
&= \frac{1}{p_{k_n}} \int_0^\infty \mathbb{E}\left[T_\infty(y)\right] \mu\left(\mathcal{B}_\infty\left(y\right)\right)^{-2} \rho(y,k) \alpha e^{-\alpha y_0} \, \mathrm{d}y_0 \\
&= (1 + o(1)) \frac{1}{k_n^2 p_{k_n}} \int_0^\infty \mathbb{E}\left[T_\infty(y)\right] \rho(y,k) \alpha e^{-\alpha y} \, \mathrm{d}y_0,
\end{aligned}$$

where the last line is due to a concentration of heights argument.

For (58) we recall that

$$c^*(k_n; G_{\text{box}}) = \frac{1}{\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \sum_{p \in \mathcal{P}} c_{\text{box}}(p) \mathbb{1}_{\{\deg_{\text{box}}(p) = k_n\}},$$

where $c_{\text{box}}(p)$ can be expressed as

$$c_{\text{box}}(p) = \frac{1}{\binom{\deg_{\text{box}}(p)}{2}} \sum_{p_1, p_2 \in \mathcal{P} \setminus p}^{\neq} T_{\text{box}}(p, p_1, p_2) = \frac{T_{\text{box}}(p)}{\binom{\deg_{\text{box}}(p)}{2}}.$$

By the Campbell-Mecje formula

$$\begin{aligned}
\mathbb{E}\left[c^*(k_n; G_{\text{box}})\right] &= \frac{1}{\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \int_{\mathcal{R}} \mathbb{E}\left[c_{\text{box}}(p) \mathbb{1}_{\{\deg_{\text{box}}(p) = k_n\}}\right] f(x,y) \, \mathrm{d}x \, \mathrm{d}y \\
&= \frac{1}{\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \int_{\mathcal{R}} \mathbb{E}\left[c_{\text{box}}(p) | \deg_{\text{box}}(p) = k_n\right] \rho_{\text{box}}(p, k_n) f(x,y) \, \mathrm{d}x \, \mathrm{d}y \\
&= (1 + o(1)) \frac{n}{\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[c_{\text{box}}(y) | \deg_{\text{box}}(y) = k_n\right] \rho(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y,
\end{aligned}$$

where the last line follows from a concentration of heights argument, for which we used that $\mathbb{E}\left[c_{\text{box}}(y) | \deg_{\text{box}}(y) = k_n\right] \leq \binom{k_n}{2}$. To analyze the conditional expectation we observe that, similar to the analysis of $\gamma(k_n)$, conditioned on their being k_n points in $\mathcal{B}_{\text{box}}\left(y\right)$, each points $u_i = (x_i, y_i)$ is independently distributed according to the probability density $\mu\left(\mathcal{B}_{\text{box}}(y)\right)^{-1} \mathbb{1}_{\{u_i \in \mathcal{B}_{\text{box}}(y)\}} f(x_i, y_i)$. Therefore,

$$
\begin{aligned}
\mathbb{E}\left[c_{\text{box}}(y) | \deg_{\text{box}}(y) = k_n\right] &= \binom{k_n}{2}^{-1} \mathbb{E}\left[\sum_{1 \le i < j \le k_n} \mathbb{1}_{\{u_i \in \mathcal{B}_{\text{box}}(u_j)\}}\right] \\
&= \mathbb{E}\left[u_1 \in \mathcal{B}_{\text{box}}(u_2)\right] \\
&= \mu\left(\mathcal{B}_{\text{box}}(y)\right)^{-2} \iint T_{\text{box}}(y, p_1, p_2) f(x_1, y_1) f(x_2, y_2) \, \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}x_2 \, \mathrm{d}y_2 \\
&= \mu\left(\mathcal{B}_{\text{box}}(y)\right)^{-2} \mathbb{E}\left[T_{\text{box}}(y)\right].
\end{aligned}
$$

and thus, by applying a concentration of heights argument on $\mu\left(\mathcal{B}_{\text{box}}(y)\right)^{-2}$,

$$
\mathbb{E}\left[c^*(k_n; G_{\text{box}})\right] = (1 + o(1)) \frac{n\mu\left(\mathcal{B}_{\text{box}}(2\log(k_n/\xi))\right)^{-2}}{\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_{\text{box}}(y)\right] \rho(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y.
$$

To finish the argument, we first note that $\mu\left(\mathcal{B}_{\text{box}}(2\log(k_n/\xi))\right)^{-2} = (1 + o(1)) k_n^2$, while

$$
\mathbb{E}\left[N_{\text{box}}(k_n)\right] = \int_{\mathcal{R}} \rho_{\text{box}}(y, k_n) f(x, y) \, \mathrm{d}x \, \mathrm{d}y,
$$

so that by a concentration of heights argument,

$$
\mathbb{E}\left[N_{\text{box}}(k_n)\right] = (1 + o(1)) n \int_0^\infty \rho(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y = (1 + o(1)) n p_{k_n}.
$$

We therefore conclude that

$$
\mathbb{E}\left[c^*(k_n; G_{\text{box}})\right] = (1 + o(1)) \frac{1}{k_n^2 p_{k_n}} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_{\text{box}}(y)\right] \rho(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y.
$$

□

Comparing (57) and (58), we conclude that to prove Proposition 5.5 it is enough to show that

$$
\left| \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_{\text{box}}(y) - T_\infty(y)\right] \rho(y, k) \alpha e^{-\alpha y} \, dy \right| = o\left(s(k_n) p_{k_n} k_n^2\right), \tag{59}
$$

which means we have to compute the expected difference in triangles between both models.

7.1 Comparing triangles between G_∞ and G_{box}

To analyze $T_{\text{box}}(y_0) - T_\infty(y_0)$ we first reiterate that the difference between the indicator $\mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p)\}}$ in the finite box model and $\mathbb{1}_{\{p_1 \in \mathcal{B}_\infty(p)\}}$ is that in G_{box} we identified the boundaries of the interval $[-\frac{\pi}{2} e^{R_n/2}, \frac{\pi}{2} e^{R_n/2}]$ and we stop at height $y = R_n$. This induces a difference in triangle counts between both models. To see this, note that for any $p = (x, y)$ with $0 \le y \le R_n$ we have that $\mathcal{B}_{\text{box}}(p) = \mathcal{B}_\infty(p) \cap \mathcal{R}$. This means that if $p', p_2 \in \mathcal{B}_{\text{box}}(p)$ and $p_2 \in \mathcal{B}_\infty(p') \cap \mathcal{R}$ then $p_2 \in \mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$ and hence (p, p', p_2) form a triangle both in G_{box} and G_∞. However, it could happen that there are points in the intersection $\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$ that are not in $\mathcal{B}_\infty(p) \cap \mathcal{B}_\infty(p')$. Let us denote this region by $\mathcal{T}(p\Delta p')$, see Figure 6 for an example of this region. Then, any $p_2 \in \mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p')$ creates a triangle with p and p' in G_{box} that is not present in G_∞. Finally, any point $p_2 \in \mathcal{B}_\infty(p) \cap \mathcal{B}_\infty(p6\prime)$ with height $y_2 > R$ creates a triangle with p, p' in G_∞ but not in G_{box}.

Let us now define the following triangle count function

$$
\widetilde{T}_{\text{box}}(p_0) = \sum_{(p_1, p_2) \in \mathcal{P} \setminus p_0}^{\neq} \widetilde{T}_{\text{box}}(p_0, p_1, p_2).
$$

where

$$\widetilde{T}_{\text{box}}(p_0, p_1, p_2) = \mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\text{box}}(p)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_\infty(p_1) \cap \mathcal{R}\}}.$$

Then $\widetilde{T}_{\text{box}}(p_0)$ only counts those triangles attached to p_0 that exist in both G_{box} and G_∞ and thus, by definition of the region $\mathcal{T}(p_0, p_1)$,

$$T_{\text{box}}(p_0) - \widetilde{T}_{\text{box}}(p_0) = \sum_{p_1, p_2 \in \mathcal{P}}^{\neq} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p_0)\}} \mathbb{1}_{\{p_2 \in \mathcal{T}(p_0, p_1)\}}.$$

The next result, which is crucial for the proof of Proposition 5.5, computes the expected measure of $\mathcal{T}_{\mathcal{P} \Delta \mathcal{P}_n}(p, p')$ with respect to p'.

Lemma 7.2. *Let $p_0 = (0, y)$ with $y \in \mathcal{K}_C(k_n)$. Then as $n \to \infty$,*

$$\mathbb{E}\left[\left|T_{box}(p_0) - \widetilde{T}_{box}(p_0)\right|\right] = y\, O\left(n^{-(2\alpha-1)}\right) + e^y\, O\left(n^{-(4\alpha-2)}\right).$$

The proof of the lemma is not difficult but cumbersome, since it involves computing many different integrals. We postpone this proof till the end of this section and proceed with the main goal, proving Proposition 5.5.

First we state a small lemma about the scaling of $s(k_n)$ that will be very useful.

Lemma 7.3. *Let $s(k_n)$ be as defined in (36). Then for any $k_n = o\left(n^{\frac{1}{2\alpha+1}}\right)$, as $n \to \infty$,*

$$n^{-(2\alpha-1)} = o\left(s(k_n)\right).$$

Proof. First let $\frac{1}{2} < \alpha < \frac{3}{4}$. Then

$$n^{-(2\alpha-1)} s(k_n)^{-1} = n^{-(2\alpha-1)} k_n^{4\alpha-2} = o\left(n^{-(2\alpha-1) + \frac{4\alpha-2}{2\alpha+1}}\right) = o\left(n^{-\frac{4\alpha^2-4\alpha+1}{2\alpha+1}}\right) = o(1),$$

since $4\alpha^2 - 4\alpha + 1 > 0$ for all $\alpha > \frac{1}{2}$. Similarly, for $\alpha \geq \frac{3}{4}$ we have that $4\alpha^2 > 2$ and hence,

$$n^{-(2\alpha-1)} s_\alpha(k_n) = o\left(n^{-(2\alpha-1)} k_n\right) = o\left(n^{-\frac{4\alpha^2-2}{2\alpha+1}}\right) = o(1).$$

□

We now proceed with proving the main result of this section.

Proof of Proposition 5.5. Let us write $\mathcal{R}' := \mathbb{R} \times \mathbb{R}_+ \setminus \mathcal{R}$ and let $p_0 = (0, y)$ denote the typical point. Next we recall that it is enough to show (59), so that in particular we have that $y \in \mathcal{K}_C(k_n)$.

Now

$$|T_{\text{box}}(p_0) - T_\infty(p_0)| = \left|T_{\text{box}}(p_0) - \widetilde{T}_{\text{box}}(p_0)\right| + \sum_{p_1, p_2 \in \mathcal{P} \cap \mathcal{R}'}^{\neq} T_\infty(p_0, p_1, p_2),$$

so that by the Campbell-Mecke formula

$$\begin{aligned} |\mathbb{E}\left[T_{\text{box}}(p_0) - T_\infty(p_0)\right]| \leq\ & \mathbb{E}\left[\left|T_{\text{box}}(p_0) - \widetilde{T}_{\text{box}}(p_0)\right|\right] \\ & + \int_{\mathcal{R}'} \int_{\mathcal{R}'} T_\infty(p_0, p_1, p_2) f(x_1, y_1) f(x_2, y_2)\, \mathrm{d}x_2\, \mathrm{d}y_2\, \mathrm{d}x_1\, \mathrm{d}y_1. \end{aligned}$$

Tobias: in the 1st integral we could have restricted to the ball of p, and possibly saved a lot. Does that not help? **Pim:** the first integral is taken care of by another Lemma so it does not matter I think. The part is taken care of by Lemma 7.2. For the other integral we have

$$\iint_{\mathcal{R}'} T_\infty(p_0, p_1, p_2) f(x_1, y_1) f(x_2, y_2)\, \mathrm{d}x_2\, \mathrm{d}y_2\, \mathrm{d}x_1\, \mathrm{d}y_1$$

$$\leq \left(\int_{\mathcal{R}'} \mathbb{1}_{\{p_1 \in \mathcal{B}_\infty(p_0)\}} f(x_1, y_1) \, \mathrm{d}x_1 \, \mathrm{d}y_1\right)^2 = O\left(\left(e^{y/2} \int_{R_n}^\infty e^{-(\alpha - \frac{1}{2})y_1} \, \mathrm{d}y_1\right)^2\right)$$
$$= O\left(e^y e^{-(2\alpha-1)R_n}\right) = O\left(e^y n^{-(4\alpha-2)}\right).$$

We thus we conclude, using Lemma 7.2, that,

$$|\mathbb{E}\left[T_{\text{box}}(p_0) - T_\infty(p_0)\right]| = O\left(yn^{-(2\alpha-1)} + n^{-(4\alpha-2)}e^y\right). \tag{60}$$

Therefore, since $y = O(\log(k_n))$ on $\mathcal{K}_C(k_n)$,

$$\int_{\mathcal{K}_C(k_n)} \rho(y, k_n) \left|\mathbb{E}\left[T_{\text{box}}(p_0) - T_\infty(p_0)\right]\right| e^{-\alpha y_0} \, \mathrm{d}y_0$$
$$= O(1) \left(\log(k_n) n^{-(2\alpha-1)} + k_n^2 n^{-(4\alpha-2)}\right) \int_0^\infty \rho(y_0, k_n) e^{-\alpha y_0} \, \mathrm{d}y_0$$
$$= O(1) \left(\log(k_n) n^{-(2\alpha-1)} + k_n^2 n^{-(4\alpha-2)}\right) p_{k_n} = o\left(s(k_n) p_{k_n} k_n^2\right),$$

where the last part follows from Lemma 7.3 and the fact that $s(k_n)^2 = o(s(k_n))$. This establishes (59) and hence finishes the proof. □

From the proof of Proposition 5.5 we obtain the following useful corollary, which will be used in Section 8.

Corollary 7.4. *Let $p_0 = (0, y)$. Then, as $n \to \infty$,*

$$\int_{-I_n}^{I_n} \int_{\mathcal{K}_C(k_n)} \rho_{box}(y, k_n) \mathbb{E}\left[\widetilde{T}_{box}(p_0)\right] f(x, y) \, \mathrm{d}x \, \mathrm{d}y = (1 + o(1)) n k_n^2 \int_0^\infty P(y) \rho(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y.$$

In particular,

$$\int_{\mathcal{K}_C(k_n)} \rho_{box}(y, k_n) \mathbb{E}\left[\widetilde{T}_{box}(p_0)\right] f(x, y) \, \mathrm{d}x \, \mathrm{d}y = \Theta\left(n k_n^{-(2\alpha-1)} s(k_n)\right).$$

Proof. We first write

$$\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(y) - T_\infty(y)\right|\right] \leq \mathbb{E}\left[\left|T_{\text{box}}(y) - \widetilde{T}_{\text{box}}(y)\right|\right] + \mathbb{E}\left[\left|T_{\text{box}}(y) - T_\infty(y)\right|\right].$$

Therefore, Lemma 7.2 and equation (60) imply that, uniformly for $y \in \mathcal{K}_C(k_n)$,

$$\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(y) - T_\infty(y)\right|\right] = O\left(\log(k_n) n^{-(2\alpha-1)} + k_n^2 n^{-(4\alpha-2)}\right) = o\left(s(k_n) k_n^2\right),$$

where the last part is due to Lemma 7.3. Next, since $\mathbb{E}\left[T_\infty(y)\right] = \mu\left(\mathcal{B}_\infty(y)\right)^2 P(y)$, we get

$$\mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\right] = \mathbb{E}\left[T_\infty(y)\right] + \mathbb{E}\left[\widetilde{T}_{\text{box}}(y) - T_\infty(y)\right] = k_n^2 P(y) + o\left(s(k_n) k_n^2\right),$$

uniformly on $\mathcal{K}_C(k_n)$. Therefore, we can apply a concentration of height argument to replace $\rho_{\text{box}}(y, k_n)$ with $\rho(y, k_n)$ and thus obtain

$$\int_{\mathcal{K}_C(k_n)} \rho_{\text{box}}(y, k_n) \mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\right] f(x, y) \, \mathrm{d}x \, \mathrm{d}y$$
$$= n k_n^2 \int_{\mathcal{K}_C(k_n)} \rho(y, k_n) \left(P(y) + o\left(s(k_n)\right)\right) \alpha e^{-\alpha y} \, \mathrm{d}y$$
$$= (1 + o(1)) n k_n^2 \int_{\mathcal{K}_C(k_n)} P(y) \rho(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y$$

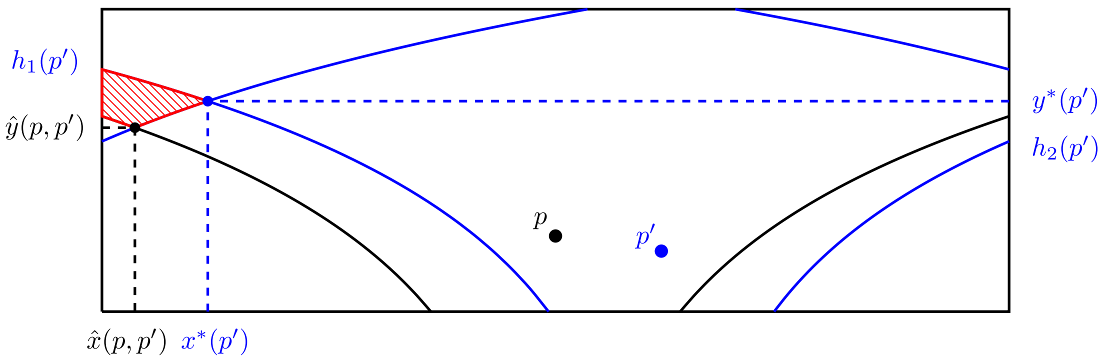

Figure 6: Example configuration of two points p and p' for which $\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$ is not a subset of $\mathcal{B}_{\infty}(p) \cap \mathcal{B}_{\infty}(p')$. The red region indicates the area belonging to $\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$ but not to $\mathcal{B}_{\infty}(p) \cap \mathcal{B}_{\infty}(p')$.

$$= (1 + o(1)) n k_n^2 \int_0^\infty P(y)\rho(y, k_n)\alpha e^{-\alpha y} \, \mathrm{d}y,$$

where we used that on $\mathcal{K}_C(k_n)$, $P(y) = \Theta(s(k_n))$. This proves the first statement. The second statement follows by observing that

$$\int_0^\infty P(y)\rho(y, k_n)\alpha e^{-\alpha y} \, \mathrm{d}y = p_{k_n}\gamma(k_n) = \Theta\left(k_n^{-(2\alpha+1)} s(k_n)\right).$$

□

7.2 Counting missing triangles

We now come back to computing the expected number of triangles attached to node at height y in G_{box} that are not present in G_{∞}.

Recall that $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p')$ denotes the region of points which form triangles with p and p' in G_{box} but not in G_{∞}. Figure 6 shows an example of a configuration where $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p') \neq \emptyset$. We observe that $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p') \neq \emptyset$ because the right boundary of the ball $\mathcal{B}_{\text{box}}(p')$ exists the right boundary of the box $\mathcal{R}$ and then, since we identified the boundaries, continues from the left so that $\mathcal{B}_{\text{box}}(p')$ covers part of the ball $\mathcal{B}_{\text{box}}(p)$ which would not be covered in the infinite limit model.

To further analyze this, let us introduce some notation. For any $p = (x, y) \in \mathcal{R}$ we will define the left and right boundary functions as, respectively,

$$b_p^-(z) = \begin{cases} 2\log(x - z) - y & \text{if } -\frac{\pi}{2}e^{R_n/2} \leq z \leq x - e^{y/2} \\ 2\log\left(\pi e^{R_n/2} + x - z\right) - y & \text{if } x - e^{(y+R_n)/2} + \pi e^{R_n/2} \leq z \leq \frac{\pi}{2}e^{R_n/2} \\ 0 & \text{else} \end{cases} \tag{61}$$

$$b_p^+(z) = \begin{cases} 2\log(z - x) - y & \text{if } x + e^{y/2} \leq z \leq \frac{\pi}{2}e^{R_n/2} \\ 2\log\left(\pi e^{R_n/2} + z - x\right) - y & \text{if } -\frac{\pi}{2}e^{R_n/2} \leq z \leq x + e^{(y+R_n)/2} - \pi e^{R_n/2} \\ 0 & \text{else} \end{cases} \tag{62}$$

Note that these functions describe the boundaries of the ball $\mathcal{B}_{\text{box}}(p)$. In particular, $p' = (x', y') \in \mathcal{B}_{\text{box}}(p)$ if and only if $y' \geq \min\left\{b_p^-(x'), b_p^+(x')\right\}$.

Since we have identified the left and right boundary of $\mathcal{R}$ we can assume, without loss of generality that $x = 0$. Due to symmetry it is then enough to restrict the analysis to the case where $x' > 0$. **Tobias:** does this not depend on p? **Pim:** I do not think so, because we can always take $p = (0, y)$ due to the invariance in the x-direction. I have updated the text to better reflect this. For this case there are two important points in the box $\mathcal{R}$. These are the intersection between the left boundary of p' and the right boundary of p', as it continues from the left side of the box, and the left boundary of p. We denote by $(x^*(p'), y^*(p'))$ the intersection between the left and right boundary of p' and by $(\hat{x}(p,p'), \hat{y}(p,p'))$ the intersection between the left boundary of p and the right boundary of p', see Figure 6.

Let us derive the expressions for the coordinates of these two points, starting with $(x^*(p'), y^*(p'))$. The x-coordinate $x^*(p')$ is the solution to the equation $b^+_{p'}(z) = b^-_{p'}(z)$ for $-\frac{\pi}{2}e^{R_n/2} \leq z \leq x + e^{(y+R_n)/2} - \pi e^{R_n/2}$. This equation becomes

$$2\log\left(\pi e^{R_n/2} + z - x'\right) - y' = 2\log\left(x' - z\right) - y',$$

whose solution is $x^*(p') := x' - \frac{\pi}{2}e^{R_n/2}$. Plugging this into either the left or right hand side of the above equation yields the y-coordinate $y^*(p') = 2\log\left(\frac{\pi}{2}e^{R_n/2}\right) - y'$. In a similar way, the x-coordinate $\hat{x}(p,p')$ is the solution to the equation $b^+_{p'}(z) = b^-_p(z)$ for $-\frac{\pi}{2}e^{R_n/2} \leq z \leq x + e^{(y+R_n)/2} - \pi e^{R_n/2}$, i.e.

$$2\log\left(\pi e^{R_n/2} + z - x'\right) - y' = 2\log\left(x - z\right) - y.$$

This solution is $\frac{x' - \pi e^{R_n/2}}{1 + e^{(y'-y)/2}}$ and again $\hat{y}(p,p')$ is obtained by plugging the solution into either the left or right hand side of the equation, yielding $\hat{y}(p,p') = 2\log\left(\frac{\pi e^{R_n/2} - x'}{e^{y/2} + e^{y'/2}}\right)$.

To summarize we have:

$$\begin{aligned}
x^*(p') &= x' - \frac{\pi}{2}e^{R_n/2} \\
y^*(p') &= 2\log\left(\frac{\pi}{2}e^{R_n/2}\right) - y' \\
\hat{x}(p,p') &= \frac{x' - \pi e^{R_n/2}}{1 + e^{(y'-y)/2}} \\
\hat{y}(p,p') &= 2\log\left(\frac{\pi e^{R_n/2} - x'}{e^{y/2} + e^{y'/2}}\right)
\end{aligned}$$

The crucial observation is that $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n} = \emptyset$ as long as the point $(x^*(p'), y^*(p'))$ is above the left boundary of p. This happens exactly when $y^*(p') > b^-_p(x^*(p'))$. Therefore the boundary of this event is given by the equation $y^*(p') = b^-_p(x^*(p'))$ which reads

$$2\log\left(\frac{\pi}{2}e^{R_n/2}\right) - y' = 2\log\left(\frac{\pi}{2}e^{R_n/2} - x'\right) - y.$$

Solving this equation gives us the function

$$b^*_p(z) = y - 2\log\left(1 - \frac{z}{\frac{\pi}{2}e^{R_n/2}}\right), \tag{63}$$

which is displayed by the red curve in Figure 7. It holds that $y^*(p') > b^-_p(x^*(p'))$ if and only if $y' < b^*_p(x')$ and hence we have that $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n} = \emptyset$ for all $p' \in \mathcal{R}$ for which $y' \geq b^*_p(x')$. We also note that when $y' = b^*_p(x')$ the two points $(x^*(p'), y^*(p'))$ and $(\hat{x}(p,p'), \hat{y}(p,p'))$ coincide.

This analysis allows us to compute the expected difference in the number of triangles for the finite box model and the infinite model, for a typical node with height y, i.e. prove Lemma 7.2.

Proof of Lemma 7.2. Due to symmetry it is enough to show that

$$\int_0^{R_n}\int_0^{I_n} \mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p,p_1)\right) f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 = O\left(yn^{-(2\alpha-1)} + n^{-(2\alpha-1)}e^y\right) \tag{64}$$

The proof goes in two stages. First we compute $\mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p,p_1)\right)$ by splitting it over three disjoint regimes with respect to p_1, with $x_1 \geq 0$. Then we do the integration with respect to p_1.

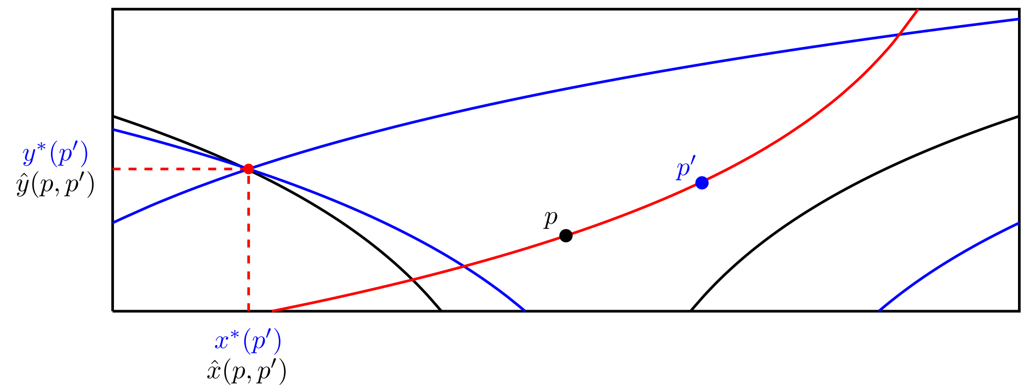

Figure 7: Example for a given p of the boundary function $x' \mapsto b_p^*(x')$, given by the red curve, which determines whether $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n} = \emptyset$. We see that when $y' = b_p^*(x')$ then $(\hat{x}(p,p'), \hat{y}(p,p')) = (x^*(p'), y^*(\prime))$.

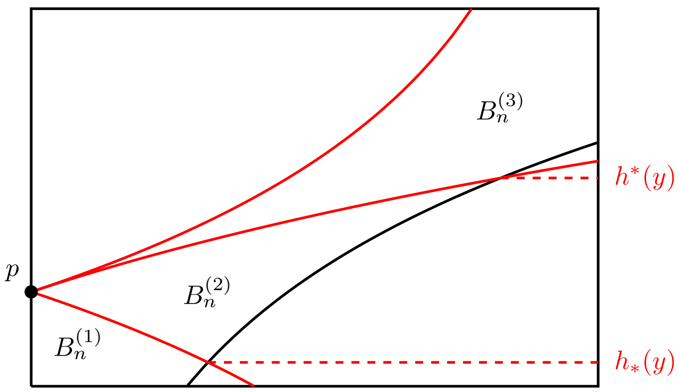

Figure 8: Three different areas $B_n^{(i)}$ used in the proof of Lemma 7.2.

Computing $\mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right)$

Recall that $I_n = \frac{\pi}{2} e^{R_n/2}$ and define the sets

$$A_n^{(1)} = \{p_1 \in \mathcal{R} \,:\, 0 \le y_1 \le y - 2\log(I_n/(I_n - x_1))\},$$
$$A_n^{(2)} = \left\{p_1 \in \mathcal{R} \,:\, y - 2\log(I_n/(I_n - x_1)) < y_1 \le y + 2\log\left(1 + \frac{x_1}{I_n}\right)\right\},$$
$$A_n^{(3)} = \left\{p_1 \in \mathcal{R} \,:\, y + 2\log\left(1 + \frac{x_1}{I_n}\right) < y_1 \le y + 2\log\left(\frac{I_n}{I_n - x_1}\right)\right\},$$

and let $B_n^{(i)} = \mathcal{B}_{\text{box}}(p) \cap A_n^{(i)}$, for $i = 1, 2, 3$, see Figure 8. Here the heights of the two intersections are given by

$$h_*(y) = y + 2\log\left(\frac{I_n}{I_n + e^y}\right) \tag{65}$$
$$h^*(y) = y + 2\log\left(\frac{I_n}{I_n - e^y}\right). \tag{66}$$

With these definitions we have that the union $B_n := \bigcup_{i=1}^n B_n^{(i)}$ denotes the area under the red curve in Figure 7 and hence, for all $p_1 \in \mathcal{R} \setminus B_n$ with $x_1 \ge 0$ we have that $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1) = \emptyset$. So we only need to consider $p_1 \in B_n$. We shall establish the following result:

$$\mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right) = \begin{cases} O\left(I_n^{-2\alpha} e^{\alpha y_1}\right) & \text{if } p_1 \in B_n^{(1)} \\ O\left(I_n^{-2\alpha} e^{\alpha y}\right) & \text{if } p_1 \in B_n^{(2)} \cup B_n^{(3)} \end{cases} \tag{67}$$

Depending on which regime p_1 belongs to, the set $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)$ has a different shape. We displayed these shapes in Figure 9 as a visual aid to follow the computations below.

Regime 1: $0 \le y_1 \le y - 2\log(I_n/(I_n - x_1))$ In this case the integral over p_2 splits into two parts

$$\mathcal{I}_n^{(1)}(p_1) := \int_{h_2(p_1)}^{y^*(p_1)} \int_{-I_n}^{x_1 + e^{(y_1+y_2)/2} - 2I_n} e^{-\alpha y_2} \, dx_2 \, dy_2$$
$$\mathcal{I}_n^{(2)}(p_1) := \int_{y^*(p_1)}^{h_1(p_1)} \int_{x^*(p_1)}^{x_1 - e^{(y_1+y_2)/2}} e^{-\alpha y_2} \, dx_2 \, dy_2.$$

We first compute $\mathcal{I}_n^{(1)}$.

$$\begin{aligned}
\mathcal{I}_n^{(1)}(p_1) &= \int_{h_2(p_1)}^{y^*(p_1)} \left(x_1 + e^{(y_1+y_2)/2} - I_n\right) e^{-\alpha y_2} \, dx_2 \, dy_2 \\
&\le e^{y_1/2} \int_{h_2(p_1)}^{y^*(p_1)} e^{-(\alpha - \frac{1}{2})y_2} \, dy_2 \\
&= \frac{2e^{y_1/2}}{2\alpha - 1} \left(e^{-(\alpha-\frac{1}{2})h_2(p_1)} - e^{-(\alpha-\frac{1}{2})y^*(p_1)}\right) \\
&= \frac{2e^{\alpha y_1}}{2\alpha - 1} I_n^{-(2\alpha-1)} \left(\left(1 - \frac{x_1}{I_n}\right)^{-(2\alpha-1)} - 1\right) \\
&= O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1}\right),
\end{aligned}$$

where we used that $x' \le e^{(y+y_1)/2} = o(I_n)$ for all $y_1 \le y$ and $y \in \mathcal{K}_C(k_n)$ so that

$$\left(\left(1 - \frac{x_1}{I_n}\right)^{-(2\alpha-1)} - 1\right) = O\left(\frac{x'}{I_n}\right) \quad \text{as } n \to \infty.$$

For $\mathcal{I}_n^{(2)}(p_1)$ we have

$$\begin{aligned}
\mathcal{I}_n^{(2)}(p_1) &= \int_{y^*(p_1)}^{h_1(p_1)} \left(I_n + x_1 - e^{(y_1+y_2)}\right) e^{-\alpha y_2} \, dx_2 \, dy_2 \\
&\le 2I_n \int_{y^*(p_1)}^{h_1(p_1)} e^{-\alpha y_2} \, dx_2 \, dy_2 \\
&= \frac{2}{\alpha} I_n \left(I_n^{-2\alpha} e^{\alpha y_1} - (I_n + x_1)^{-2\alpha} e^{-\alpha y_1}\right) \\
&= O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1}\right) = O\left(I_n^{-(2\alpha-1)} e^{\alpha y_1}\right).
\end{aligned}$$

We conclude that for $p_1 \in B_n^{(1)}$:

$$\mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right) = O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1}\right),$$

which establishes the first part of (67).

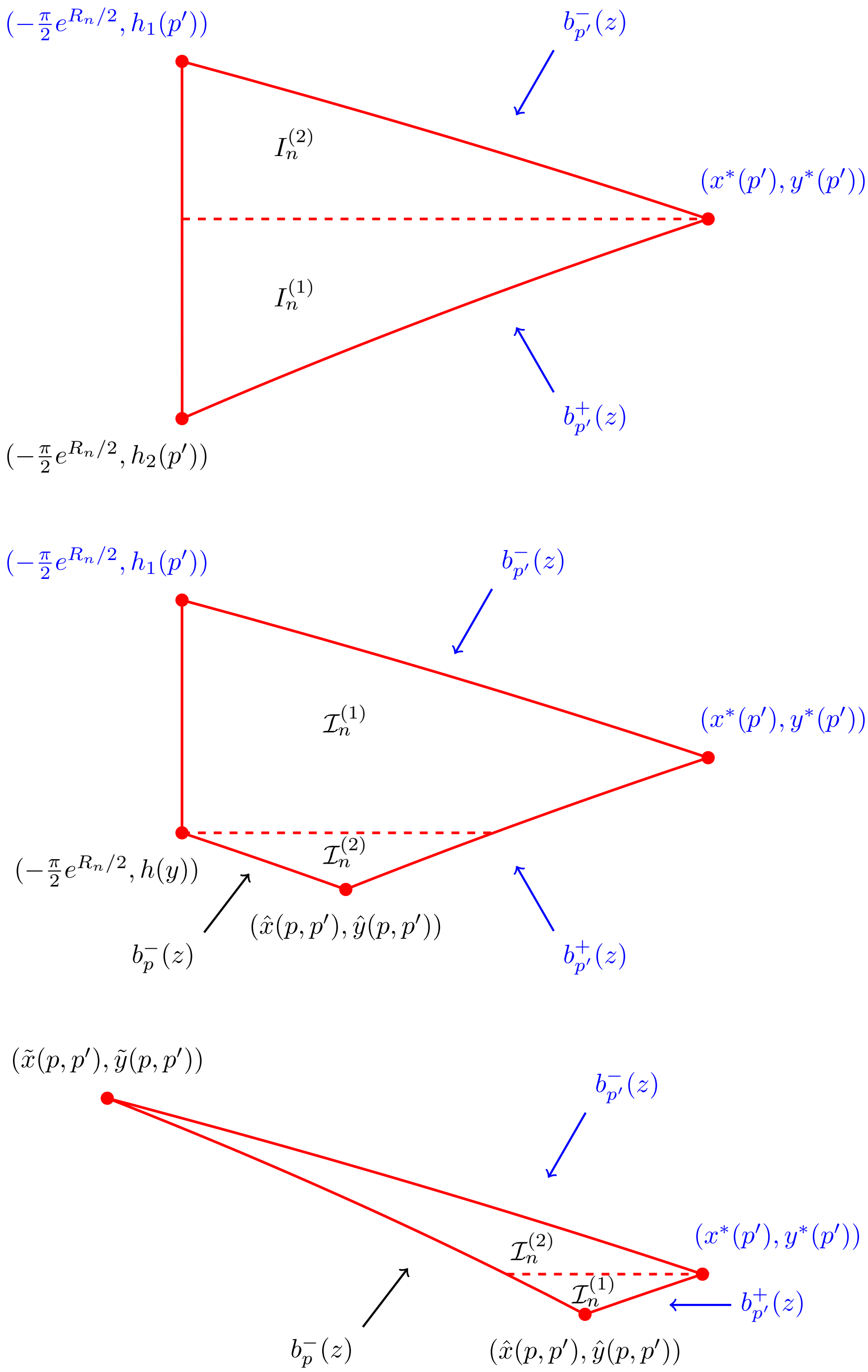

Figure 9: The different shapes of $\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)$ depending on the regime to which p_1 belongs. The top figure is for $p_1 \in B_n^{(1)}$, the middle for $p_1 \in B_n^{(2)}$ and the bottom one for $p_1 \in B_n^{(3)}$.

Regime 2: $y - 2\log(I_n/(I_n - x_1)) < y_1 \le y + 2\log\left(1 + \frac{x_1}{I_n}\right)$ Here we split the integration into two parts (see Figure 9). Recall that $x^*(p, p_1) = x_1 - I_n$. Then, for the first part we have

$$
\begin{aligned}
\mathcal{I}_n^{(1)}(p, p_1) &\le \int_{h(y)}^{h_1(p_1)} \int_{-I_n}^{x^*(p,p_1)} f(x_2, y_2) \, \mathrm{d}x_2 \, \mathrm{d}y_2 \\
&= O\left(x_1 \left(e^{-\alpha h(y)} - e^{-\alpha h_1(p_1)}\right)\right) \\
&= O\left(x_1 I_n^{-2\alpha} \left(e^{\alpha y} - e^{\alpha y_1} \left(1 + \frac{x_1}{I_n}\right)^{-2\alpha}\right)\right) \\
&= O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1} \left(\left(1 - \frac{x_1}{I_n}\right)^{-2\alpha} - \left(1 + \frac{x_1}{I_n}\right)^{-2\alpha}\right)\right) \\
&= O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1}\right) = O\left(I_n^{-(2\alpha-1)} e^{\alpha y}\right),
\end{aligned}
$$

were we used that $y \le y_1 + 2\log(I_n/(I_n - x_1))$ for $p_1 \in B_n^{(2)}$ for the third line and

$$
\left(1 - \frac{x_1}{I_n}\right)^{-2\alpha} - \left(1 + \frac{x_1}{I_n}\right)^{-2\alpha} = O\left(\frac{x_1}{I_n}\right) = O(1),
$$

for the last line.

For the second part we first compute that

$$
\begin{aligned}
x_1 + e^{(y_1+y_2)/2} - 2I_n + e^{(y+y_2)/2} &\le \left(e^{y/2} + e^{y_1/2}\right) e^{y_2/2} \\
&\le e^{y/2} \left(1 + \frac{I_n}{I_n - e^y}\right) e^{y_2/2} = O\left(e^{(y+y_2)/2}\right),
\end{aligned}
$$

since $y \in \mathcal{K}_C(k_n)$ and $k_n = o(\sqrt{n})$, so that $e^y = o(n) = o(I_n)$. Then we have

$$
\begin{aligned}
\mathcal{I}_n^{(2)} &= \int_{\hat{y}(p,p_1)}^{h(y)} \int_{-e^{(y+y_2)/2}}^{x_1 + e^{(y+y_1)/2} - 2I_n} f(x_2, y_2) \, \mathrm{d}x_2 \, \mathrm{d}y_2 \\
&= O\left(e^{y/2} \int_{\hat{y}(p,p_1)}^{h(y)} e^{-(\alpha - \frac{1}{2})y_2} \, \mathrm{d}y_2\right) \\
&= O\left(e^{y/2} \left(e^{-(\alpha-\frac{1}{2})\hat{y}(p,p_1)} - e^{-(\alpha-\frac{1}{2})h(y)}\right)\right) \\
&= O\left(e^{y/2} \left(\left(\frac{2I_n - x_1}{e^{y/2} + e^{y_1/2}}\right)^{-(2\alpha-1)} - I_n^{-(2\alpha-1)} e^{(\alpha-\frac{1}{2})y}\right)\right) \\
&= O\left(I_n^{-(2\alpha-1)} e^{\alpha y}\right),
\end{aligned}
$$

where for the last line we first used that $(2I_n - x_1)^{-(2\alpha-1)} \le I_n^{-(2\alpha-1)}$ and then

$$
\left(\left(e^{y/2} + e^{y_1/2}\right)^{2\alpha-1} - e^{(\alpha-\frac{1}{2})y}\right) \le e^{(\alpha-\frac{1}{2})y} \left(\left(1 + \sqrt{1 + \frac{x_1}{I_n}}\right)^{2\alpha-1} - 1\right) = O\left(e^{(\alpha-\frac{1}{2})y}\right).
$$

It then follows that for $p_1 \in B_n^{(2)}$

$$
\mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right) = O\left(I_n^{-(2\alpha-1)} e^{\alpha y}\right).
$$

Regime III $p_1 \in B_n^{(3)}$:

$$\begin{aligned}
\mathcal{I}_n^{(1)} &= \int_{y^*}^{\tilde{y}} \int_{-e^{(y+y_2)/2}}^{x_1 - e^{(y_1+y_2)/2}} f(x_2, y_2)\, \mathrm{d}x_2\, \mathrm{d}y_2 \\
&= O\left(\int_{y^*}^{\tilde{y}} x_1 e^{-\alpha y_2} - \left(e^{y_1/2} - e^{y/2} \right) e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 \right) \\
&= O\left(x_1 \int_{y^*}^{\tilde{y}} e^{-\alpha y_2}\, \mathrm{d}y_2 \right).
\end{aligned}$$

Now

$$\begin{aligned}
\int_{y^*}^{\tilde{y}} e^{-\alpha y_2}\, \mathrm{d}y_2 &= \frac{1}{\alpha}\left(e^{-\alpha y^*} - e^{-\alpha \tilde{y}} \right) = \frac{1}{\alpha}\left(I_n^{-2\alpha} e^{\alpha y_1} - \left(\frac{x_1}{e^{y_1/2} - e^{y/2}} \right)^{-2\alpha} \right) \\
&= \frac{I_n^{-2\alpha} e^{\alpha y_1}}{\alpha} \left(1 - \left(1 - e^{(y - y_1)/2} \right)^{2\alpha} \left(\frac{x_1}{I_n} \right)^{-2\alpha} \right) = O\left(I_n^{-2\alpha} e^{\alpha y_1} \right),
\end{aligned}$$

and hence we have

$$\mathcal{I}_n^{(1)} = O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1} \right).$$

For the second integral we have

$$\begin{aligned}
\mathcal{I}_n^{(2)} &= \int_{\hat{y}}^{y^*} \int_{-e^{(y+y_2)/2}}^{e^{(y_1+y_2)/2} + x_1 - 2I_n} f(x_2, y_2)\, \mathrm{d}x_2\, \mathrm{d}y_2 \\
&= O\left(\int_{\hat{y}}^{y^*} \left(e^{y/2} + e^{y_1/2} \right) e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 \right) \\
&= O\left(e^{y_1/2} \int_{\hat{y}}^{y^*} e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 \right).
\end{aligned}$$

For the integral we have

$$\begin{aligned}
\int_{\hat{y}}^{y^*} e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 &= \frac{2}{2\alpha - 1}\left(e^{-(\alpha - \frac{1}{2})\hat{y}} - e^{-(\alpha - \frac{1}{2})y^*} \right) \\
&= \frac{2}{2\alpha - 1}\left(\left(\frac{2I_n - x_1}{e^{y/2} + e^{y_1/2}} \right)^{-(2\alpha - 1)} - I_n^{-(2\alpha - 1)} e^{-(\alpha - \frac{1}{2})y_1} \right) \\
&= O\left(I_n^{-2\alpha} x_1 e^{-(\alpha - \frac{1}{2})y_1} \right)
\end{aligned}$$

so that

$$\mathcal{I}_n^{(2)} = O\left(I_n^{-(2\alpha - 1)} e^{(1-\alpha)y_1} \right) = O\left(I_n^{-2\alpha} x_1 e^{\alpha y} \right)$$

and hence for $p_1 \in B_n^{(3)}$

$$\mu\left(\mathcal{T}_{\mathcal{P} \Delta \mathcal{P}_n}(p, p_1) \right) = O\left(I_n^{-2\alpha} x_1 e^{\alpha y} \right) = O\left(I_n^{-(2\alpha - 1)} e^{\alpha y} \right).$$

Integration over p_1

We now proceed with the second part of the computation leading to (64). Here we will integrate $\mu(\mathcal{T}_{\mathcal{P} \Delta \mathcal{P}_n})(p, p_1)$ over the region $B_n := B_n^{(1)} \cup B_n^{(2)} \cup B_n^{(3)}$, see Figure 8. Let us first identify the boundaries of these areas.

The area $B_n^{(1)}$ is bounded from above by the line given by the equation

$$y_1 = y - 2\log\left(\frac{I_n}{I_n - x_1}\right).$$

Solving this for x_1 yields $x_1 = I_n\left(1 - e^{(y_1-y)/2}\right)$ and hence the area $B_n^{(1)}$ is given by

$$B_n^{(1)} = \left\{(x_1, y_1) \,:\, 0 \le y_1 \le y, \quad 0 \le x_1 \le I_n\left(1 - e^{(y_1-y)/2}\right) \wedge e^{(y+y_1)/2}\right\}.$$

In a similar way we have that $B_n^{(2)}$ is bounded from above by line

$$y_1 = y + 2\log\left(\frac{I_n}{I_n + x_1}\right),$$

which yields $x_1 = I_n\left(e^{(y_1-y)/2} - 1\right)$. The lower red boundary is the upper boundary of $B_n^{(2)}$ and hence we have

$$B_n^{(2)} = \left\{(x_1, y_1) \,:\, h_*(y) \le y_1 \le h^*(y),\; I_n\left(1 - e^{(y_1-y)/2}\right) \vee I_n\left(e^{(y_1-y)/2} - 1\right) \le x_1 \le e^{(y+y_1)/2}\right\}.$$

We continue is the same way to obtain for $B_n^{(3)}$

$$B_n^{(3)} = \left\{(x_1, y_1) \,:\, y \le y_1 \le R_n,\; I_n\left(1 - e^{(y-y_1)/2}\right) \le x_1 \le I_n\left(e^{(y_1-y)/2} - 1\right) \wedge e^{(y+y_1)/2} \wedge I_n\right\}.$$

We these characterizations of the areas we now integrate $\mu(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n})(p, p_1)$ over B_n, splitting the computations over the three different areas.

$\boldsymbol{p_1 \in B_n^{(1)}}$: We use that $I_n\left(1 - e^{(y_1-y)/2}\right) \wedge e^{(y+y_1)/2} \le I_n\left(1 - e^{(y_1-y)/2}\right)$ so that

$$\begin{aligned}
&\int_{B_n^{(1)}} \mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right) f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 \\
&\quad\le \int_0^y \int_0^{I_n(1-e^{(y_1-y)/2})} \mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right) f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 \\
&\quad= O\left(I_n^{-2\alpha} \int_0^y \int_0^{e^{(y+y_1)/2}} x_1\, \mathrm{d}x_1\, \mathrm{d}y_1\right) \\
&\quad= O\left(I_n^{-(2\alpha-1)} \int_0^y \left(1 - e^{(y_1-y)/2}\right)^2 \mathrm{d}y_1\right) \\
&\quad= O\left(I_n^{-(2\alpha-1)} y\right) = O\left(y n^{-(2\alpha-1)}\right).
\end{aligned}$$

$\boldsymbol{p_1 \in B_n^{(2)}}$: We will show that

$$\mu(B_n^{(2)}) = O\left(I_n^{-1} e^{(2-\alpha)y}\right), \tag{68}$$

which together with (67) yields

$$\begin{aligned}
\int_{B_n^{(2)}} \mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p, p_1)\right) f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 &= O\left(\mu(B_n^{(2)}) I_n^{-(2\alpha-1)} e^{\alpha y}\right) \\
&= O\left(I_n^{-2\alpha} e^{2y}\right).
\end{aligned}$$

The integration is split into two parts determined by $I_n\left(1 - e^{(y_1-y)/2}\right) \vee I_n\left(e^{(y_1-y)/2} - 1\right)$:

$$\mu(B_n^{(3)}) = \int_{h_*(y)}^y \int_{I_n(1-e^{(y_1-y)/2})}^{e^{(y+y_1)/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1$$

$$+\int_{y}^{h^*(y)}\int_{I_n(e^{(y_1-y)/2}-1)}^{e^{(y+y_1)/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1.$$

For the first integral we use that $e^{(y+y_1)/2} - I_n(1-e^{(y_1-y)/2}) \le e^{y_1/2}\left(e^{y/2}+e^{-y/2}\right)$ to obtain

$$\begin{aligned}
\int_{h_*(y)}^{y}\int_{I_n(1-e^{(y_1-y)/2})}^{e^{(y+y_1)/2}} & f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 \\
&= O\left(e^{y/2}\int_{h_*(y)}^{y} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1\right) \\
&= O\left(e^{y/2}\left(e^{-(\alpha-\frac{1}{2})y} - e^{-(\alpha-\frac{1}{2})y}\left(\frac{I_n}{I_n+e^y}\right)^{-(2\alpha-1)}\right)\right) \\
&= O\left(I_n^{-1}e^{(2-\alpha)y}\right).
\end{aligned}$$

For the second integral note that $e^{(y+y_1)/2} - I_n(e^{(y_1-y)/2}-1) \le e^{(y+y_1)/2}$ and hence

$$\begin{aligned}
\int_{y}^{h^*(y)}\int_{I_n(e^{(y_1-y)/2}-1)}^{e^{(y+y_1)/2}} & f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 \\
&= O\left(e^{y/2}\int_{y}^{h^*(y)} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1\right) \\
&= O\left(e^{y/2}\left(e^{-(\alpha-\frac{1}{2})y} - e^{-(\alpha-\frac{1}{2})y}\left(\frac{I_n}{I_n-e^y}\right)^{-(2\alpha-1)}\right)\right) \\
&= O\left(I_n^{-1}e^{(2-\alpha)y}\right),
\end{aligned}$$

so that (68) follows.

$\boldsymbol{p_1 \in B_n^{(3)}}$: For this area we show that

$$\mu(B_n^{(3)}) = O\left(e^{(1-\alpha)y}\right) \tag{69}$$

so that

$$\begin{aligned}
\int_{B_n^{(3)}} \mu\left(\mathcal{T}_{\mathcal{P}\Delta\mathcal{P}_n}(p,p_1)\right) f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 &= O\left(\mu(B_n^{(2)})I_n^{-(2\alpha-1)}e^{\alpha y}\right) \\
&= O\left(I_n^{-(2\alpha-1)}e^{y}\right).
\end{aligned}$$

Here the integral is split into three parts:

$$\begin{aligned}
\mu(B_n^{(3)}) = & \int_{y}^{h^*(y)}\int_{I_n(1-e^{(y-y_1)/2})}^{I_n(e^{(y_1-y)/2}-1)} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 \\
&+ \int_{h^*(y)}^{h(y)}\int_{I_n(1-e^{(y-y_1)/2})}^{e^{(y+y_1)/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 \\
&+ \int_{h(y)}^{R_n}\int_{I_n(1-e^{(y-y_1)/2})}^{I_n} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1.
\end{aligned}$$

Let us first focus on the first integral. Since $I_n(e^{(y_1-y)/2}-1) - I_n(1-e^{(y-y_1)/2}) \le I_n e^{(y_1-y)/2}$ we get, using similar arguments as above

$$\int_{y}^{h^*(y)}\int_{I_n(1-e^{(y-y_1)/2})}^{I_n(e^{(y_1-y)/2}-1)} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 = O\left(I_n e^{-y/2}\int_{y}^{h^*(y)} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1\right)$$

$$= O\left(I_n e^{-\alpha y}\left(1 - \left(\frac{I_n}{I_n - e^y}\right)^{-(2\alpha-1)}\right)\right)$$
$$= O\left(e^{(1-\alpha)y}\right).$$

Proceeding to the second integral, we first note that $e^{(y+y_1)/2} - I_n(1-e^{(y-y_1)/2}) = O\left(I_n e^{(y_1-y)/2}\right)$ so that similar calculations as before yield

$$\int_{h^*(y)}^{h(y)} \int_{I_n(1-e^{(y-y_1)/2})}^{e^{(y+y_1)/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 = O\left(I_n e^{-y/2} \int_{h^*(y)}^{h(y)} e^{-(\alpha-\frac{1}{2})y_1}\, \mathrm{d}y_1\right) = O\left(e^{(1-\alpha)y}\right).$$

□

8 Concentration for $c(k; G_{\mathbf{box}})$ (Proving Proposition 5.4)

In this section we establish a concentration result for the local clustering function $c^*(k; G_{\text{box}})$ in the finite box model G_{box}. Similar to the previous section we will focus on typical points $p = (0, y)$ with $y \in \mathcal{K}_C(k_n)$.

8.1 The main contribution of triangles

First we write

$$c^*(k_n; G_{\text{box}}) = \frac{T_{\text{box}}(k_n)}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]},$$

where

$$T_{\text{box}}(k_n) = \sum_{p \in \mathcal{P}} \mathbb{1}_{\{\deg_{\text{box}}(p) = k_n\}} \sum_{p_1, p_2 \in \mathcal{P} \setminus p}^{\neq} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\text{box}}(p)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\text{box}}(p_1)\}}$$

In particular, the variance of $c^*(k_n; G_{\text{box}})$ is determined by the variance of $T_{\text{box}}(k_n)$.

Next, recall the adjusted triangle count function

$$\widetilde{T}_{\text{box}}(p_0) = \sum_{(p_1, p_2) \in \mathcal{P} \setminus p_0}^{\neq} \widetilde{T}_{\text{box}}(p_0, p_1, p_2).$$

where

$$\widetilde{T}_{\text{box}}(p_0, p_1, p_2) = \mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p_0)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\text{box}}(p_0)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\infty}(p_1) \cap \mathcal{R}\}},$$

and recall the definition of $\mathcal{K}_{C,n}(k_n)$

$$\mathcal{K}_{C,n}(k_n) = \left\{p \in \mathcal{R} : \frac{k_n - C\sqrt{k_n \log(k_n)}}{\xi} \vee 1 \leq e^{\frac{y}{2}} \leq \frac{k_n + C\sqrt{k_n \log(k_n)}}{\xi}\right\}.$$

Slightly abusing notation, we will define the corresponding triangle degree function

$$\widetilde{T}_{\text{box}}(k, C) = \sum_{p \in \mathcal{P} \cap \mathcal{K}_{C,n}(k_n)} \mathbb{1}_{\{\deg_{\text{box}}(p) = k\}} \widetilde{T}_{\text{box}}(p). \tag{70}$$

The idea is that the main contribution of triangles of degree k_n to $T_{\text{box}}(k_n)$ is given by $\widetilde{T}_{\text{box}}(k_n, C)$. Therefore, in order to prove Proposition 5.4 it suffices to show that $\widetilde{T}_{\text{box}}(k_n, C)$ is sufficiently concentrated around its mean. This is done in the following proposition.

Proposition 8.1 (Concentration $\widetilde{T}_{\text{box}}(k_n, C)$)**.** *Let $\alpha > \frac{1}{2}$, $\nu > 0$ and let $(k_n)_{n\geq 1}$ be any positive sequence satisfying $k_n = o\left(n^{\frac{1}{2\alpha+1}}\right)$. Then for any $C > 0$, as $n \to \infty$,*

$$\mathbb{E}\left[\widetilde{T}_{box}(k_n, C)^2\right] = (1 + o(1))\, \mathbb{E}\left[\widetilde{T}_{box}(k_n, C)\right]^2.$$

We postpone the proof of this proposition till Section 8.3 and first use it to prove Proposition 5.4.

Proof of Proposition 5.4. We bound the expectation as follows,

$$\begin{aligned}\mathbb{E}\left[\left|c^*(k_n; G_{\text{box}}) - \mathbb{E}\left[c^*(k_n; G_{\text{box}})\right]\right|\right] \leq & \frac{\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(k_n, C) - \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right|\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \\ & + 2\mathbb{E}\left[\left|c^*(k_n; G_{\text{box}}) - \frac{\widetilde{T}_{\text{box}}(k_n, C)}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]}\right|\right].\end{aligned}$$

We will show that both terms are $o\left(s(k_n)\right)$.

First we note that $\mathbb{1}_{\{p_2 \in \mathcal{B}_\infty(p_1) \cap \mathcal{R}\}} \leq \mathbb{1}_{\{p_2 \in \mathcal{B}_{\text{box}}(p_1)\}}$ and hence $\widetilde{T}_{\text{box}}(p) \leq T_{\text{box}}(p)$. This implies that

$$\frac{\widetilde{T}_{\text{box}}(k_n, C)}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \leq c^*(k_n; G_{\text{box}}).$$

and therefore

$$\mathbb{E}\left[\left|c^*(k_n; G_{\text{box}}) - \frac{\widetilde{T}_{\text{box}}(k_n, C)}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]}\right|\right] = \mathbb{E}\left[c^*(k_n; G_{\text{box}})\right] - \frac{\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]}.$$

For the expectation of $\widetilde{T}_{\text{box}}(k_n, C)$ we use that

$$\mathbb{E}\left[\widetilde{T}_{\text{box}}(p) \middle| \deg_{\text{box}}(p) = k_n\right] = \binom{k_n}{2} \mu\left(\mathcal{B}_{\text{box}}(y)\right)^{-2} \mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right],$$

to get

$$\begin{aligned}\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] &= \int_{\mathcal{K}_{C,n}(k_n)} \mathbb{E}\left[\widetilde{T}_{\text{box}}(p) \middle| \deg_{\text{box}}(p) = k_n\right] \rho_{\text{box}}(y, k_n) f(x, y) \, \mathrm{d}x \, \mathrm{d}y \\ &= (1 + o(1)) \binom{k_n}{2} \int_{\mathcal{K}_{C,n}(k_n)} \mu\left(\mathcal{B}_{\text{box}}(y)\right)^{-2} \mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\right] \rho_{\text{box}}(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y \\ &= (1 + o(1)) \frac{1}{2} \int_{\mathcal{K}_{C,n}(k_n)} \mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\right] \rho_{\text{box}}(y, k_n) \alpha e^{-\alpha y} \, \mathrm{d}y \\ &= (1 + o(1)) n \binom{k_n}{2} \int_0^\infty P(y) \alpha e^{-\alpha y} \, \mathrm{d}y,\end{aligned}$$

where the last line is due to Corollary 7.4. Since $\mathbb{E}\left[N_{\text{box}}(k_n)\right] = (1 + o\left(1\right) n p_{k_n}$ it follows that

$$\frac{\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} = (1 + o(1)) \frac{\int_0^\infty P(y) \alpha e^{-\alpha y} \, \mathrm{d}y}{p_{k_n}} = (1 + o(1)) \gamma(k_n).$$

On the other hand, Proposition 5.5 implies that $\mathbb{E}\left[c^*(k_n; G_{\text{box}})\right] = (1 + o(1)) \gamma(k_n)$ and thus we conclude that

$$2\mathbb{E}\left[\left|c^*(k_n; G_{\text{box}}) - \frac{\widetilde{T}_{\text{box}}(k_n, C)}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]}\right|\right] = o\left(s(k_n)\right).$$

For the remaining term we use Proposition 8.1 to obtain

$$\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(k_n,C)-\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]\right|\right] \le \left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)^2\right]-\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]^2\right)^{\frac{1}{2}}$$
$$= o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]\right).$$

This implies

$$\frac{\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(k_n,C)-\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]\right|\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} = o\left(\frac{\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]}\right) = o\left(s(k_n)\right),$$

which finishes the proof. □

8.2 Joint neighborhoods and degrees in G_{box}

To prove Proposition 8.1 we need to understand the joint degree distribution in G_{box} and subsequently the joint neighborhoods of two points $p, p' \in \mathcal{R}$. We perform the analysis in this section. The two main results, which will be the crucial technical ingredients for the proof of Proposition 8.1, are Lemma 8.6 and Lemma 8.7.

Neighborhoods

We start with analyzing joint neighborhoods in G_{box}. Let $p, p' \in \mathcal{R}$. Then we denote by $\mathcal{N}_{\text{box}}(p\Delta p')$ the number of disjoint neighbors of p and p' in G_{box}, i.e. those points that belong to either $\mathcal{B}_{\text{box}}(p)$ or $\mathcal{B}_{\text{box}}(p')$ but not to their intersection. In addition we denote by $\mathcal{N}_{\text{box}}(p, p')$ the number of joint neighbors of p and p'. We shall establish lower bounds on the expected number of points in the disjoint neighborhood $\mathbb{E}\left[\mathcal{N}_{\text{box}}(p\Delta p')\right]$ as well as an asymptotic expression for the expected number points in the joint neighborhood. For this we will distinguish between the cases where the distance between the x-coordinates of p and p' is small or large. Figure 10-12 show the different situations that occur.

We start by analyzing the shape of the neighborhoods. Due to symmetry and the fact that we have identified the left and right boundaries of the box $\mathcal{R}$, we can, without loss of generality, assume that $p = (0, y)$ and $p' = (x', y')$ with $x' > 0$ and $y' \le y$. To understand the computation it is helpful to have a picture of the different situations. Figure 10 and Figure 11 show two different situations for small distance in the x-coordinates, in which case the number of disjoint neighbors is small. The case where this distance is large, and hence the number of disjoint neighbors is expected to be large, is shown in Figure 12. There are several different quantities that are important. The first are the heights $h_1(p')$ and $h_2(p')$ where, respectively, the left and right boundaries of the ball $\mathcal{B}_{\text{box}}(p')$ go outside the box $\mathcal{R}$. Note that when $x = 0$ then these heights are the same and we denote this by $h(y)$. We also need to know the coordinates $\hat{y}(p, p')$ and $\hat{x}(p, p')$ of the intersection of the right boundary of the neighborhood of p with the left boundary of the neighborhood of p'. Finally we will denote by $d(p, p')$ the distance between the lower right boundary of $\mathcal{B}_{\text{box}}(p)$ and the lower left of $\mathcal{B}_{\text{box}}(p')$, which is positive only when the bottom parts of both neighborhoods do not intersect, compare Figures 10 and 12. The derivation of the expression for these functions is similar to those in Section 7.2 and we omit the details here. The full expressions of all these functions are given below for further reference.

$$h(y) = R_n - y + 2\log\left(\frac{\pi}{2}\right) \tag{71}$$

$$h_1(p') = 2\log\left(x' + \frac{\pi}{2}e^{\frac{R_n}{2}}\right) - y' \tag{72}$$

$$h_2(p') = 2\log\left(\frac{\pi}{2}e^{\frac{R_n}{2}} - x'\right) - y' \tag{73}$$

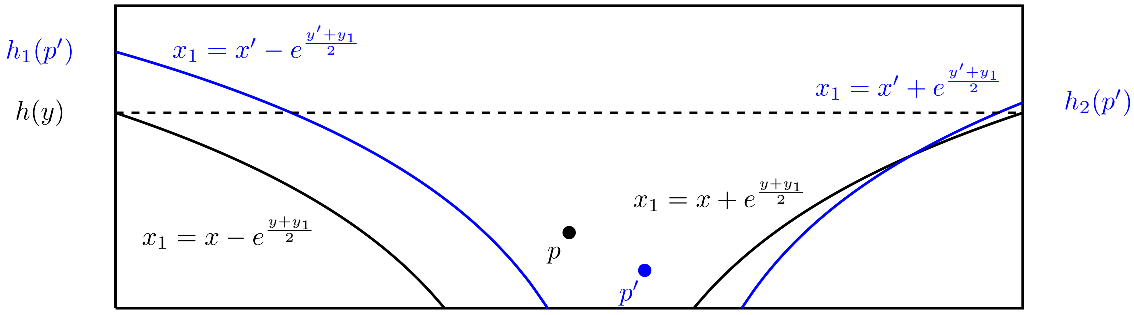

Figure 10: Schematic representation of the neighborhoods of p and p' in G_{box} when $|x-x'| \leq e^{\frac{y+y'}{2}}$ used for the proof of Lemma 8.2 and Lemma 8.4.

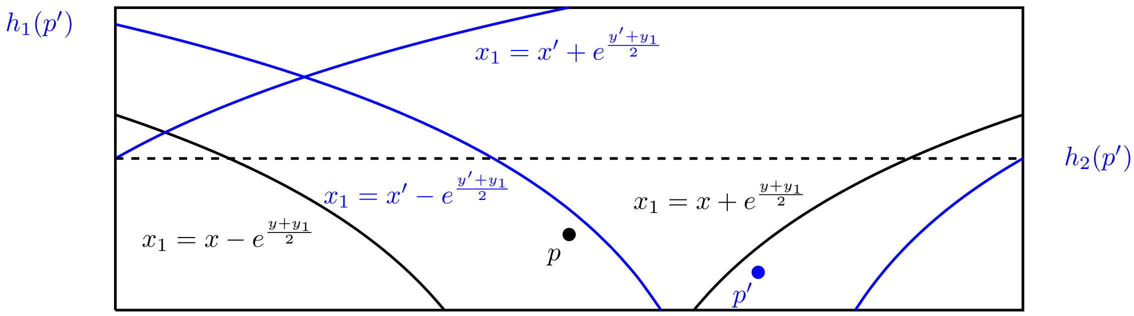

Figure 11: Schematic representation of the neighborhoods of p and p' in G_{box} when $e^{\frac{y+y'}{2}} < |x-x'| \leq e^{\frac{y}{2}} + e^{\frac{y'}{2}}$ used for the proof of Lemma 8.2 and Lemma 8.4.

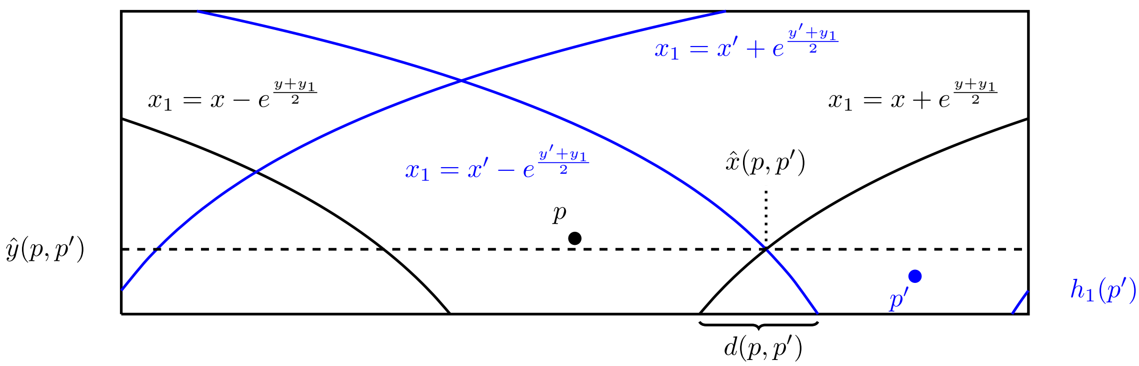

Figure 12: Schematic representation of the neighborhoods of p and p' in G_{box} when $|x-x'| > e^{\frac{y}{2}} + e^{\frac{y'}{2}}$ used for the proof of Lemma 8.3.

$$\hat{y}(p,p') = 2\log\left(\frac{|x-x'|}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}\right) \tag{74}$$

$$\hat{x}(p,p') = \frac{xe^{\frac{y'}{2}}+x'e^{\frac{y}{2}}}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}, \tag{75}$$

$$d(p,p') = |x-x'| - \left(e^{\frac{y}{2}}+e^{\frac{y'}{2}}\right). \tag{76}$$

We start with the result for points whose x-coordinates are close, which is when $d(p,p') < 0$.

Lemma 8.2. *Let $p, p' \in \mathcal{R}$ and denote $y^* = \min\{y, y'\}$. Then whenever $|x-x'| \leq e^{\frac{y}{2}}+e^{\frac{y'}{2}}$,*

$$\mathbb{E}\left[\mathcal{N}_{box}(p\Delta p')\right] \geq \frac{\nu}{\pi}|x'-x|\left(1-\left(\frac{2}{\pi}\right)^{2\alpha}e^{-\alpha(R_n-y^*)}\right)$$
$$+\xi\left|e^{\frac{y}{2}}-e^{\frac{y'}{2}}\right|\left(1-\left(\frac{2}{\pi}\right)^{2\alpha-1}e^{-(\alpha-\frac{1}{2})(R_n-y^*)}\right).$$

Proof. In order to proof the result we will consider the area in between the two left boundaries of the balls $\mathcal{B}_{\text{box}}(p)$ and $\mathcal{B}_{\text{box}}(p')$ up to the height $h^* := \min\{h(y), h_2(p')\}$, see Figure 10 and Figure 11 for reference. Note that here we do not need to consider different cases depending on whether $|x-x'| \leq e^{(y+y')/2}$ or $|x-x'| > e^{(y+y')/2}$.

By the Campbell-Mecke formula we get

$$\begin{aligned}
\mathbb{E}\left[\mathcal{N}_{\text{box}}(p\Delta p')\right] &= \mu\left(\mathcal{B}_{\text{box}}(p)\,\Delta B_{\mathcal{P},n}(p')\right) \\
&\geq \int_0^{h^*}\int_{x-e^{\frac{y+y_1}{2}}}^{x'-e^{\frac{y'+y_1}{2}}} f(x_1,y_1)\,dx_1\,dy_1 \\
&= \frac{\alpha\nu}{\pi}|x'-x|\int_0^{h^*} e^{-\alpha y_1}dy_1 + \frac{\alpha\nu}{\pi}\left|e^{\frac{y}{2}}-e^{\frac{y'}{2}}\right|\int_0^{h^*} e^{-(\alpha-\frac{1}{2})y_1}\,dy_1 \\
&= \frac{\nu}{\pi}|x'-x|\left(1-\left(\frac{2}{\pi}\right)^{2\alpha}e^{-\alpha(R_n-y^*)}\right) \\
&\quad+\xi\left|e^{\frac{y}{2}}-e^{\frac{y'}{2}}\right|\left(1-\left(\frac{2}{\pi}\right)^{2\alpha-1}e^{-(\alpha-\frac{1}{2})(R_n-y^*)}\right).
\end{aligned}$$

□

Now we will consider the case where $|x-x'| > e^{\frac{y}{2}}+e^{\frac{y'}{2}}$

Lemma 8.3. *Let $p, p' \in \mathcal{R}$. Then, whenever $|x-x'| > e^{\frac{y}{2}}+e^{\frac{y'}{2}}$,*

$$\mathbb{E}\left[\mathcal{N}_{box}(p\Delta p')\right] \geq \left(\mu(\mathcal{B}_{box}(p))+\mu(\mathcal{B}_{box}(p'))\right)\left(1-\phi_n(p,p')\right).$$

$$\phi_n(p,p') = \xi\left(\left(\frac{e^{y/2}+e^{y'/2}}{|x-x'|}\right)^{2\alpha-1} - e^{-(\alpha-\frac{1}{2})R_n}\right)$$

Proof. We will prove the results by using that

$$\mathbb{E}\left[\mathcal{N}_{\text{box}}(p\Delta p')\right] \geq \int_0^{\hat{y}(p,p')}\int_{-\frac{\pi}{2}e^{\frac{R_n}{2}}}^{\frac{\pi}{2}e^{\frac{R_n}{2}}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\text{box}}(p)\cup\mathcal{B}_{\text{box}}(p')\}}f(x_1,y_1)\,dx_1\,dy_1,$$

and computing the integral on the right. We refer to Figure 12 for further clarification.

Before we proceed we show that the neighborhoods of p and p' below $\hat{y}(p,p')$ are disjoint. This is clearly true when $\hat{y}(p,p') \leq h_2(p')$ so suppose that $\hat{y}(p,p') > h_2(p')$. Then, because we identified the right and left boundaries of the box $\mathcal{R}$ the right boundary of $\mathcal{B}_{\text{box}}(p')$ continues from the left boundary of the box and is described by the equation

$$x_1 = x' + e^{\frac{y'+y_1}{2}} - \pi e^{\frac{R_n}{2}}.$$

Now, let x'_{right} and x_{left} denote the x-coordinate of the intersection of the line $\hat{y}(p,p')$ with, respectively, the right boundary of $\mathcal{B}_{\text{box}}(p')$ and the left boundary of $\mathcal{B}_{\text{box}}(p)$. Then

$$\begin{aligned}
x'_{\text{right}} &= x' + e^{\frac{y'+\hat{y}(p,p')}{2}} - \pi e^{\frac{R_n}{2}} \\
&= x' + e^{\frac{\hat{y}(p,p')}{2}}\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right) - e^{\frac{y+\hat{y}(p,p')}{2}} - \pi e^{\frac{R_n}{2}} \\
&= x' + |x-x'| - e\frac{y+\hat{y}(p,p')}{2} - \pi e^{\frac{R_n}{2}} \\
&= x - e^{\frac{y+\hat{y}(p,p')}{2}} + 2|x-x'| - \pi e^{\frac{R_n}{2}} \\
&\leq x - e^{\frac{y+\hat{y}(p,p')}{2}} = x_{\text{left}},
\end{aligned}$$

and hence the neighborhoods of p and p' below $\hat{y}(p,p')$ are disjoint. It then follows that

$$\begin{aligned}
\mathbb{E}\left[\mathcal{N}_{\text{box}}(p,p')\right] &\geq \int_0^{\hat{y}(p,p')} \int_{-\frac{\pi}{2}e^{\frac{R_n}{2}}}^{\frac{\pi}{2}e^{\frac{R_n}{2}}} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p) \cup \mathcal{B}_{\text{box}}(p')\}} f(x_1,y_1)\, dx_1\, dy_1 \\
&= \int_0^{\hat{y}(p,p')} \int_{-\frac{\pi}{2}e^{\frac{R_n}{2}}}^{\frac{\pi}{2}e^{\frac{R_n}{2}}} \left(\mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p)\}} + \mathbb{1}_{\{p_1 \in \mathcal{B}_{\text{box}}(p')\}}\right) f(x_1,y_1)\, dx_1\, dy_1 \\
&= \left(\mu_{\alpha,\nu,n}(\mathcal{B}_{\text{box}}(p)) + \mu_{\alpha,\nu,n}(\mathcal{B}_{\text{box}}(p'))\right)\left(1 - \frac{2\alpha\nu}{\pi}\int_{\hat{y}(p,p')}^{R_n} e^{-(\alpha-\frac{1}{2})y_1}\, dy_1\right),
\end{aligned}$$

from which the result follows since

$$\begin{aligned}
\frac{2\alpha\nu}{\pi}\int_{\hat{y}(p,p')}^{R_n} e^{-(\alpha-\frac{1}{2})y_1}\, dy_1 &= \xi\left(e^{-(\alpha-\frac{1}{2})\hat{y}(p,p')} - e^{-(\alpha-\frac{1}{2})R_n}\right) \\
&= \xi\left(\left(\frac{e^{y/2}+e^{y'/2}}{|x-x'|}\right)^{2\alpha-1} - e^{-(\alpha-\frac{1}{2})R_n}\right).
\end{aligned}$$

□

Next we consider the number of common neighbors between two nodes p and p' in G_{box}, which we denote by $\mathcal{N}_{\mathcal{P}}(p,p')$.

Lemma 8.4. *Let $p, p' \in \mathcal{R}$. Then, whenever $|x-x'| > \left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)$,*

$$\mathbb{E}\left[\mathcal{N}_{box}(p,p')\right] = \left(\mu\left(\mathcal{B}_{box}(p)\right) + \mu\left(\mathcal{B}_{box}(p')\right)\right)\phi_n(p,p'),$$

where

$$\phi_n(p,p') = \frac{1+2\alpha}{4\alpha}\left(\frac{|x-x'|}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}\right)^{-(2\alpha-1)} + \frac{2\alpha-1}{4\alpha}|x'-x|e^{-\alpha R_n} - e^{-(\alpha-\frac{1}{2})R_n}.$$

Proof. Assume, without loss of generality, that $y \geq y'$ and $x \leq x'$ and consider the boundaries of the balls $\mathcal{B}_{\text{box}}(p)$ and $\mathcal{B}_{\text{box}}(p')$ as drawn in Figure 12. The left boundary of $\mathcal{B}_{\text{box}}(p)$ intersects the right boundary of $\mathcal{B}_{\text{box}}(p')$ if an only if $d(p,p') > 0$. We observe from the definition of $d(p,p')$, (76), that this is exactly the condition we imposed in the statement of the lemma. The y-coordinate of the intersection is then given by

$$h((p,p') := 2\log\left(\frac{|x-x'|}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}\right).$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[\mathcal{N}_{\mathrm{box}}(p,p')\right] &= \int_{h(p,p')}^{R_n}\int_{x'-e^{(y'+y_1)/2}}^{x+e^{(y+y_1)/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 \\ &= \left(e^{y/2}+e^{y'/2}\right)\frac{\alpha\nu}{\pi}\int_{h(p,p')}^{R_n} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1 \\ &\quad -(x'-x)\frac{\alpha\nu}{\pi}\int_{h(p,p')}^{R_n} e^{-\alpha y_1}\,\mathrm{d}y_1.\end{aligned}$$

For the first term we get

$$\begin{aligned}&\left(e^{y/2}+e^{y'/2}\right)\frac{\alpha\nu}{\pi}\int_{h(p,p')}^{R_n} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1 \\ &= \xi\left(e^{y/2}+e^{y'/2}\right)\left(e^{-(\alpha-\frac{1}{2})h(p,p')}-e^{-(\alpha-\frac{1}{2})R_n}\right) \\ &= \xi\left(e^{y/2}+e^{y'/2}\right)\left(\left(\frac{|x'-x|}{e^{y/2}+e^{y'/2}}\right)^{-(2\alpha-1)}-e^{-(\alpha-\frac{1}{2})R_n}\right).\end{aligned}$$

For the other term we compute

$$\begin{aligned}&(x'-x)\frac{\alpha\nu}{\pi}\int_{h(p,p')}^{R_n} e^{-\alpha y_1}\,\mathrm{d}y_1 \\ &= (x'-x)\frac{\nu}{\pi}\left(e^{-\alpha(p,p')}-e^{-\alpha R_n}\right) \\ &= \frac{\nu}{\pi}\left(e^{y/2}+e^{y'/2}\right)\left(\left(\frac{|x'-x|}{e^{y/2}+e^{y'/2}}\right)^{-(2\alpha-1)}-(x'-x)e^{-\alpha R_n}\right) \\ &= \xi\left(e^{y/2}+e^{y'/2}\right)\left(\frac{2\alpha-1}{4\alpha}\left(\frac{|x'-x|}{e^{y/2}+e^{y'/2}}\right)^{-(2\alpha-1)}-\frac{2\alpha-1}{4\alpha}|x'-x|e^{-\alpha R_n}\right).\end{aligned}$$

Combining these two results and noticing that $\xi e^{y/2} = \mu\left(\mathcal{B}_{\mathrm{box}}\left(p\right)\right)$ and similar for p' finishes the proof. □

Degrees

We now turn to the joint degree distribution of nodes in G_{box}. To ease notations we introduce the following short-hand notation for the joint degree distribution

$$\rho_{\mathrm{box}}(p,p',k,k') := \mathbb{P}\left(\mathrm{Po}\left(\mu\left(\mathcal{B}_{\mathrm{box}}\left(p\right)\right)\right)=k, \mathrm{Po}\left(\mu\left(\mathcal{B}_{\mathrm{box}}\left(p'\right)\right)\right)=k'\right).$$

We first establish an almost independence result for integrals over the joint degree distribution $\rho_{\mathrm{box}}(p,p',k,k')$ for the case where p and p' are sufficiently separated. For this we note that when $|x-x'| \gg e^{y/2}+e^{y'/2}$ it follows from Lemma 8.3

$$\mathbb{E}\left[\mathcal{N}_{\mathrm{box}}(p\Delta p')\right] = \left(\mu\left(\mathcal{B}_{\mathrm{box}}\left(p\right)\right)+\mu\left(\mathcal{B}_{\mathrm{box}}\left(p'\right)\right)\right)(1+o\left(1\right)),$$

for $p,p' \in \mathcal{K}_C(k_n)\times\mathcal{K}_C(k_n)$. This implies that the joint neighborhoods are almost independent and hence their degrees must be as well. To make this more precise, let $0<\varepsilon<\min\{(2\alpha-1)^{-1},1\}$ and define the following set

$$\mathcal{E}_\varepsilon(k_n) = \left\{(p,p')\in\mathcal{R}\times\mathcal{R} \,:\, y,y'\in\mathcal{K}_C(k_n) \text{ and } |x-x'|>k_n^{1+\varepsilon}\right\}. \tag{77}$$

The next result shows that for integrals over either $\mathcal{E}(k_n)$, when $k_n = \Theta(1)$, or $\mathcal{E}_\varepsilon(k_n)$, when $k_n \to \infty$, the joint degree distribution satisfies $\rho_n(p, p', k_n, k_n) = (1 + o(1))\rho(p, k_n)\rho(p', k_n)$.

Lemma 8.5. *Let $k_n \to \infty$ and $X = \mathrm{Po}(\lambda_1(n))$, $Y = \mathrm{Po}(\lambda_2(n))$, be two Poisson random variables where $\lambda_2(n) = O\left(k_n^{1-\varepsilon}\right)$, for some $0 < \varepsilon < 1$ and for some $C > 0$,*

$$k_n - C\sqrt{k_n \log(k_n)} \le \lambda_1(n) + \lambda_2(n) \le k_n + C\sqrt{k_n \log(k_n)}.$$

Then, if X_1 and X_2 are two independent copies of X,

$$\mathbb{P}\left(X_1 + Y = k_n, X_2 + Y = k_n\right) = (1 + o(1))\mathbb{P}\left(X_1 + Y = k_n\right)\mathbb{P}\left(X_2 + Y = k_n\right),$$

as $n \to \infty$.

Proof. First we write

$$\mathbb{P}\left(X_1 + Y = k_n, X_2 + Y = k_n\right) = \sum_{t=0}^{\infty} \mathbb{P}\left(X_1 = k_n - t\right)\mathbb{P}\left(X_2 = k_n - t\right)\mathbb{P}\left(Y = t\right).$$

Now fix a $C_1 > 0$ and define the set

$$A_n := \left\{t \in \mathbb{R}_+ \,:\, \lambda_2(n) - C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)} \le t \le \lambda_2(n) + C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right\}.$$

Then by a Chernoff bound (c.f. (126))

$$\begin{aligned}
&\sum_{t \in \mathbb{R}_+ \setminus A_n} \mathbb{P}\left(X_1 = k_n - t\right)\mathbb{P}\left(X_2 = k_n - t\right)\mathbb{P}\left(Y = t\right) \\
&\quad \le \mathbb{P}\left(Y > \lambda_2(n) + C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right) + \mathbb{P}\left(Y < \lambda_2(n) - C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right) \\
&\quad = \mathbb{P}\left(|\mathrm{Po}(\lambda_2(n)) - \lambda_2(n)| > C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right) = O\left(k_n^{-\frac{1+C_1^2}{2}}\right)
\end{aligned}$$

and hence

$$\mathbb{P}\left(X_1 + Y = k_n, X_2 + Y = k_n\right) = \sum_{t \in A_n} \mathbb{P}\left(X_1 = k_n - t\right)\mathbb{P}\left(X_2 = k_n - t\right)\mathbb{P}\left(Y = t\right) + O\left(k_n^{-(1+C_1^2)/2}\right).$$

Now take any $s \in A_n$ so that $|t - s| \le 2C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}$ and note that there exists a δ_n such that $|\delta_n| \le 2C\sqrt{k_n \log(k_n)}$ and $k_n - t = \lambda_1(n) + \delta_n$. It then follows that, uniformly in t, s and δ_n, as $n \to \infty$

$$\begin{aligned}
\frac{\mathbb{P}\left(X_2 = k_n - t\right)}{\mathbb{P}\left(X_2 = k_n - s\right)} &= \frac{\mathbb{P}\left(X_2 = k_n - t\right)}{\mathbb{P}\left(X_2 = k_n - t - (s-t)\right)} \\
&= \frac{(k_n - t - (s-t))!}{(k_n - t)!}\lambda_1(n)^{s-t} \\
&\sim (k_n - t - (s-t))^{-(s-t)}\lambda_1(n)^{s-t} \\
&= (\lambda_1(n) + \delta_n - (s-t))^{-(s-t)}\,\lambda_1(n)^{s-t} \\
&= \left(1 + \frac{\delta_n - (s-t)}{\lambda_1(n)}\right)^{s-t} \\
&\sim e^{\frac{(s-t)\delta_n}{\lambda_1(n)}} e^{-\frac{(s-t)^2}{\lambda_1(n)}} \sim 1,
\end{aligned}$$

where the last line follows since both $\frac{(s-t)\delta_n}{\lambda_1(n)} \to 0$ and $\frac{(s-t)^2}{\lambda_1(n)} \to 0$ as $n \to \infty$. In particular,

$$\mathbb{P}\left(X_2 = k_n - t\right) = (1 + o(1))\mathbb{P}\left(X_2 = k_n - s\right),$$

uniformly for all $t, s \in A_n$ and therefore, since

$$1 = \sum_{s=0}^{\infty} \mathbb{P}(Y = s) = (1 + o(1)) \sum_{s \in A_n} \mathbb{P}(Y = s),$$

we conclude that

$$\begin{aligned}
&\sum_{t \in A_n} \mathbb{P}(X_1 = k_n - t) \mathbb{P}(X_2 = k_n - t) \mathbb{P}(Y = t) \\
&\quad = (1 + o(1)) \sum_{t \in A_n} \mathbb{P}(X_1 = k_n - t) \mathbb{P}(X_2 = k_n - t) \mathbb{P}(Y = t) \sum_{s \in A_n} \mathbb{P}(Y = s) \\
&\quad = (1 + o(1)) \sum_{t \in A_n} \mathbb{P}(X_1 = k_n - t) \mathbb{P}(Y = t) \sum_{s \in A_n} \mathbb{P}(X_2 = k_n - s) \mathbb{P}(Y = s) \\
&\quad = (1 + o(1)) \mathbb{P}(X_1 + Y = k_n) \mathbb{P}(X_2 + Y = k_n),
\end{aligned}$$

from which the result follows. □

The above lemma makes it possible to factorize the joint degree distribution $\rho_{\text{box}}(p, p', k_n, k_n)$ in G_{box}, whenever $(p, p') \in \mathcal{E}_\varepsilon(k_n)$. In particular we have the following result.

Lemma 8.6. *Let $k_n \to \infty$ be such that $k_n = o\left(n^{\frac{1}{2\alpha+1}}\right)$ let $h : \mathbb{R}_+ \to \mathbb{R}$ be a uniformly bounded function. Then, as $n \to \infty$,*

$$\begin{aligned}
&\int_{\mathcal{E}_\varepsilon} \rho_{box}(p, p', k_n, k_n) h(y) h(y') f(x, y) f(x', y') \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}x' \, \mathrm{d}y' \\
&\qquad = (1 + o(1)) \int_{\mathcal{E}_\varepsilon} \rho(y, k_n) \rho(y', k_n) h(y) h(y') f(x, y) f(x', y') \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}x' \, \mathrm{d}y'.
\end{aligned}$$

Proof. Recall that $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with mean λ. Now define the random variables

$$\begin{aligned}
X_1(p, p') &:= \mathrm{Po}\left(\mu\left(\mathcal{B}_{\text{box}}(p) \setminus \mathcal{B}_{\text{box}}(p')\right)\right), \\
X_2(p, p') &:= \mathrm{Po}\left(\mu\left(\mathcal{B}_{\text{box}}(p') \setminus \mathcal{B}_{\text{box}}(p)\right)\right), \\
Y(p, p') &:= \mathrm{Po}\left(\mu\left(\mathcal{B}_{\text{box}}(p) \cup \mathcal{B}_{\text{box}}(p')\right)\right),
\end{aligned}$$

so that

$$\rho_{\text{box}}(p, p', k_n, k_n) = \mathbb{P}\left(X_1(p, p') + Y(p, p') = k_n, X_2(p, p') + Y(p, p') = k_n\right).$$

Next we note that due to symmetry $X_1(p, p')$ and $X_2(p, p')$ are independent copies of the same random variable $X(p, p')$. Moreover, on the set $\mathcal{E}_\varepsilon(k_n)$ the random variables $X(p, p')$ and $Y(p, p')$ satisfy the conditions of Lemma 8.5. Hence,

$$\rho_{\text{box}}(p, p', k_n, k_n) = (1 + o(1)) \rho_{\text{box}}(p, k_n) \rho_{\text{box}}(p', k_n) + O\left(k_n^{-\frac{(1+C)^2}{2}}\right).$$

where $C > 0$ can be selected arbitrarily large. Since the function h is uniformly bounded it follows that for some large enough $C > 0$,

$$\begin{aligned}
&\int_{\mathcal{E}_\varepsilon(k_n)} k_n^{-\frac{(1+C)^2}{2}} h(y) h(y') f(x, y) f(x', y') \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}x' \, \mathrm{d}y' \\
&= O(1) \, k_n^{-\frac{(1+C)^2}{2}} \left(\int_{\mathcal{R}} f(x, y) \, \mathrm{d}x \, \mathrm{d}y\right)^2 \\
&= o(1) \int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p, k_n) \rho_{\text{box}}(p', k_n) h(y) h(y') f(x, y) f(x', y') \, \mathrm{d}x \, \mathrm{d}y.
\end{aligned}$$

Therefore

$$\begin{aligned}&\int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p,p',k_n,k_n)h(y)h(y')f(x,y)f(x',y')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}x'\,\mathrm{d}y'\\
&= (1+o(1))\int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p,k_n)\rho_{\text{box}}(p',k_n)h(y)h(y')f(x,y)f(x',y')\,\mathrm{d}x\,\mathrm{d}y,\end{aligned}$$

and the result follows by applying a concentration of heights argument twice.

□

Next we show that when $k_n \to \infty$ and the expected number of disjoint neighbors of p and p' in $\mathcal{K}_C(k_n)$ grows only slightly with k_n then any fixed shift in the joint degree distribution does not effect its asymptotic behavior.

Pim: @All: The proof of this lemma follows almost the exact same steps as that of the previous one. Is there any way we could merge them?

Lemma 8.7. *Let $\alpha > \frac{1}{2}$, $\nu > 0$, $k_n \to \infty$ and fix $\varepsilon > 0$. Then for any fixed $i, j, i', j' \in \mathbb{Z}$ and $p, p' \in \mathcal{K}_C(k_n)$ such that $\mathbb{E}[\mathcal{N}_{box}(p\Delta p')] \ge k_n^\varepsilon$,*

$$\rho_n(p,p',k_n+i,k_n+i') = (1+o(1))\rho_n(p,p',k_n+j,k_n+j') \pm e^{-\Omega(k_n^\varepsilon)}.$$

Proof. Define

$$\begin{aligned}X_n &= \left|\mathcal{B}_{\text{box}}(p) \setminus \mathcal{B}_{\text{box}}(p')\right|,\\
Y_n &= \left|\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')\right|,\\
Z_n &= \left|\mathcal{B}_{\text{box}}(p') \setminus \mathcal{B}_{\text{box}}(p)\right|.\end{aligned}$$

Then it follows that X_n, Y_n and Z_n are independent Poisson random variables satisfying $\mathbb{E}[X_n] + \mathbb{E}[Y_n] = \mu(\mathcal{B}_{\text{box}}(p))$ and $\mathbb{E}[Z_n] + \mathbb{E}[Y_n] = \mu(B_{\mathcal{P},n}(p'))$ while

$$\begin{aligned}\rho_n(p,p',k_n+i,k_n+i') &= \mathbb{P}(X_n+Y_n = k+i, Z_n+Y_n = k+i')\\
&= \sum_{\ell=0}^{\infty} \mathbb{P}(Y_n=\ell)\,\mathbb{P}(X_n = k+i-\ell, Z_n = k+i'-\ell)\\
&= \sum_{\ell=0}^{\infty} \mathbb{P}(Y_n=\ell)\,\mathbb{P}(X_n = k+i-\ell)\,\mathbb{P}(Z_n = k+i'-\ell).\end{aligned}$$

Next define $\delta_n = k_n^{-\frac{1-\varepsilon}{2}}$, let n be large enough such that $0 < \delta_n < 1$ and note that by a Chernoff bound,

$$\mathbb{P}(|X_n - \mathbb{E}[X_n]| > \delta_n \mathbb{E}[X_n]) = O\left(e^{-\frac{\delta_n^2}{4(1+\delta_n)}\mathbb{E}[X_n]}\right),$$

and similar for Y_n and Z_n. Finally, we define

$$\begin{aligned}L_X(k_n) &= \{\ell : (1-\delta_n)\mathbb{E}[X_n] \le k+i-\ell \le (1+\delta_n)\mathbb{E}[X_n]\}\\
L_y(k_n) &= \{\ell : (1-\delta_n)\mathbb{E}[Y_n] \le \ell \le (1+\delta_n)\mathbb{E}[Y_n]\}\\
L_Z(k_n) &= \{\ell : (1-\delta_n)\mathbb{E}[Z_n] \le k+i'-\ell \le (1+\delta_n)\mathbb{E}[Z_n]\}\end{aligned}$$

We will now make distinguish between the cases $\mathbb{E}[Y_n] \le k_n/2$ and $\mathbb{E}[Y_n] > k_n/2$.

Let us first assume that $\mathbb{E}[Y_n] \le k_n/2$. Then, since $p \in \mathcal{K}_\varepsilon(k_n)$ and $\mu(\mathcal{B}_{\text{box}}(p)) = \Theta\left(e^{y/2}\right) = \Theta(k_n)$ it follows that $\mathbb{E}[X_n] = \Omega(k_n)$ and hence

$$\mathbb{P}(|X_n - \mathbb{E}[X_n]| > \delta_n \mathbb{E}[X_n]) = O\left(e^{-\frac{\delta_n^2}{4(1+\delta_n)}\mathbb{E}[X_n]}\right) = e^{-\Omega(k_n^{(1+\varepsilon)/2})} = e^{-\Omega(k_n^\varepsilon)}.$$

In particular, this implies

$$\sum_{\ell \notin L_X} \mathbb{P}(Y_n = \ell)\,\mathbb{P}(X_n = k+i-\ell)\,\mathbb{P}(Z_n = k+i'-\ell)$$
$$= O\left(\mathbb{P}\left(|X_n - \mathbb{E}[X_n]| > \delta_n \mathbb{E}[X_n]\right)\right) = e^{-\Omega(k_n^{(1+\varepsilon)/2})}.$$

Finally we note that, for $\ell \in L_X$, we have

$$\frac{\mathbb{P}(X_n = k+i-\ell)}{\mathbb{P}(X_n = k+j-\ell)} = \mathbb{E}[X_n]^{i-j}\,\frac{(k+j-\ell)!}{(k+i-\ell)!} \leq (1+\delta_n)^{2|i-j|}. \tag{78}$$

and observe that we have similar results for Z_n. Therefore,

$$\begin{aligned}
&\sum_{\ell \in L_X \cap L_Z} \mathbb{P}(Y_n = \ell)\,\mathbb{P}(X_n = k+i-\ell)\,\mathbb{P}(Z_n = k+i'-\ell) \\
&\quad\leq (1+\delta_n)^{2(j-i)+2(j'-i')} \sum_{\ell \in L_x \cap L_Z} \mathbb{P}(Y_n = \ell)\,\mathbb{P}(X_n = k+j-\ell)\,\mathbb{P}(Z_n = k+j'-\ell) \\
&\quad= (1+o(1))(1+\delta_n)^{2|i-j|+2|i'-j'|}\,\mathbb{P}(D_p = k_n + j, D_{p'} = k_n + j') \\
&\quad= (1+o(1))\mathbb{P}(D_p = k_n + j, D_{p'} = k_n + j')
\end{aligned}$$

and hence

$$\rho_n(p, p', k+i, k'+i') = (1+o(1))\rho_n(p, p', k+j, k'+j') + e^{-\Omega(k_n^{\varepsilon})}.$$

Now assume that $\mathbb{E}[Y_n] > k_n/2$. Then, since $\mathbb{E}[|\mathcal{N}_{\text{box}}(p\Delta p')|] \geq k_n^{\varepsilon}$ it follows that $\mathbb{E}[X_n] = \Omega(k_n^{\varepsilon})$ or $\mathbb{E}[Z_n] = \Omega(k_n^{\varepsilon})$. Without loss of generality we assume that $\mathbb{E}[X_n] = \Omega(k_n^{\varepsilon})$. Similar to (78) we have for Y_n

$$\frac{\mathbb{P}(Y_n = \ell)}{\mathbb{P}(Y_n = \ell + j' - i')} \leq (1+\delta_n)^{2(|i'-j'|)}.$$

Using similar computations as above we then have

$$\begin{aligned}
&\rho_n(p, p', k+i, k'+i') \\
&\quad= \sum_{\ell \in L_X \cap L_Y} \mathbb{P}(Y_n = \ell)\,\mathbb{P}(X_n = k+i-\ell)\,\mathbb{P}(Z_n = k+i'-\ell) + e^{-\Omega(k_n^{(1+\varepsilon)/2})} \\
&\quad= (1+o(1))\rho_n(p, p', k+j, k'+j') + e^{-\Omega(k_n^{(1+\varepsilon)/2})}.
\end{aligned}$$

□

8.3 Concentration for main triangle contribution

We now turn to Proposition 8.1. Before we dive into the proof let us first give a high level overview of the strategy and the flow of the arguments.

Recall (see (70)) that for any $C > 0$

$$\widetilde{T}_{\text{box}}(k, C) = \sum_{p \in \mathcal{P}_n \cap \mathcal{K}_{C,n}(k_n)} \mathbb{1}_{\{\deg_{\text{box}}(p) = k\}} \widetilde{T}_{\text{box}}(p)$$

Then we have

$$\widetilde{T}_{\text{box}}(k, C)^2 = \sum_{p, p' \in \mathcal{P}_n \cap \mathcal{K}_C(k)} \mathbb{1}_{\{D_{\mathcal{P},n}(p),\, D_{\mathcal{P},n}(p') = k\}} \sum_{(p_1,p_2),(p_1',p_2') \in \mathcal{P}_n \setminus (0,y)}^{\neq} \widetilde{T}_{\mathcal{P}}(p, p_1, p_2)\widetilde{T}_{\mathcal{P}}(p', p_1', p_2'),$$

This expression can be written as the sums of several terms, depending on how $\{p, p_1, p_2\}$ and $\{p', p_1', p_2'\}$ intersect. To this end we define, for $a \in \{0,1\}$ and $b \in \{0,1,2\}$,

$$I_{a,b} = \sum_{\substack{p,p' \in \mathcal{P}_n \cap \mathcal{K}_C(k) \\ |\{p\} \cap \{p'\}| = a}} \mathbb{1}_{\{D_{\mathcal{P},n}(p),\, D_{\mathcal{P},n}(p') = k\}} J_b(p,p'),$$

where

$$J_b(p,p') = \sum_{\substack{p_1,p_2,p_1',p_2' \in \mathcal{P}_n \\ |\{p_1,p_2\} \cap \{p_1',p_2'\}| = b}}^{\neq} T_{\mathcal{P},n}(p,p_1,p_2) T_{\mathcal{P},n}(p',p_1',p_2'),$$

with the summation being over all two distinct pairs (p_1,p_2) and (p_1',p_2'). Then we have

$$\widetilde{T}_{\text{box}}(k,C)^2 = \sum_{a=0}^{1} \sum_{b=0}^{2} I_{a,b}.$$

To prove Proposition 8.1 we will deal with each of the $I_{a,b}$ separately, showing that

$$\mathbb{E}\left[I_{0,0}\right] = (1+o(1))\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n)\right]^2 \tag{79}$$

and for all other combinations

$$\mathbb{E}\left[I_{a,b}\right] = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n)\right]^2\right). \tag{80}$$

Note that $J_b(p,p') \le J_0(p,p')$ and, since $I_{1,2} = \widetilde{T}_{\text{box}}(k_n, C)$, (80) holds for $I_{1,2}$.

To prove (**??**) and (**??**) we have to consider the cases $k_n = \Theta(1)$ and $k_n \to \infty$ separately. However, each case follows a similar strategy. First we define a set $\mathcal{E} \subseteq \mathcal{K}_C(k_n) \times \mathcal{K}_C(k_n)$ such that outside this set the contributions of $\mathbb{E}\left[I_{a,b}\right]$ are negligible while on this set the joint degree distribution factorizes. These will be the sets defined by (**??**) and (77). In particular, for these sets we have that $a = 1$ is not possible and hence we only need to consider the three different cases for b. For $a = 0 = b$, the factorization of the degree distributions will enable us to prove (**??**). For $b = 1, 2$ we derive sufficient bounds for $\mathbb{E}\left[J_b(p,p')\right]$ in terms of k_n for (p,p') in the set $\mathcal{E}$.

Proof of Proposition 8.1. Throughout this proof we set $i = |\{p', p_1, p_2, p_1', p_2'\} \cap \mathcal{B}_{\text{box}}(p)|$, $j = |\{p'\} \cap \mathcal{B}_{\text{box}}(p)|$ and define i', j' in a similar way by interchanging the primed and non-primed variables. In addition, we write $D_{\mathcal{P},n}(p,p',k,\ell)$ to denote the indicator that $|\mathcal{B}_{\text{box}}(p) \cap (\mathcal{P}_n \setminus \{p,p',p_1,p_2,p_1',p_2'\})| = k$ and $|\mathcal{B}_{\text{box}}(p') \cap (\mathcal{P}_n \setminus \{p,p',p_1,p_2,p_1',p_2'\})| = \ell$. Then,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{\{D_{\mathcal{P},n}(p)=k_n, D_{\mathcal{P},n}(p')=k_n\}} J_b(p,p')\right] \\
&= \mathbb{E}\left[\sum_{\substack{p_1,p_2,p_1',p_2' \in \mathcal{P}_n \\ |\{p_1,p_2\} \cap \{p_1',p_2'\}| = b}}^{\neq} D_{\mathcal{P},n}(p,p',k_n - i, k_n - i')\, \widetilde{T}_{\text{box}}(p,p_1,p_2) \widetilde{T}_{\text{box}}(p',p_1',p_2')\right],
\end{aligned}$$

where the sum is over all distinct pairs (p_1,p_2) and (p_1',p_2'). We also that

$$\mathbb{E}\left[T_{\mathcal{P}}(k_n)\right] = \Theta\left(n k_n^{-(2\alpha-1)} s_\alpha(k_n)\right).$$

Recall the definition of $\mathcal{E}_\varepsilon(k_n)$

$$\mathcal{E}_\varepsilon(k_n) = \left\{(p,p') \in \mathcal{K}_C(k_n) \times \mathcal{K}_C(k_n) \,:\, \mathbb{E}\left[|\mathcal{N}^c_{\text{box}}(p,p')|\right] \ge k_n^\varepsilon \text{ and } |x - x'| > k_n^{1+\varepsilon}\right\}$$

let $\mathcal{E}_\varepsilon(k_n)^c$ denote its complement and let $I^*_{a,b}$ denote the the part of $I_{a,b}$ where $p, p' \in \mathcal{P}_n \cap \mathcal{E}_\varepsilon(k_n)$. We first show that

$$\mathbb{E}\left[I_{a,b} - I^*_{a,b}\right] = o\left(\mathbb{E}\left[T_{\mathcal{P},n}(k_n)\right]^2\right), \tag{81}$$

so that for the remainder of the proof we only need to consider $p, p' \in \mathcal{E}_\varepsilon(k_n)$ and hence, we can apply Lemma 8.6. For this we note that by Lemma 8.2 and Lemma 8.3 we have for $p, p' \in \mathcal{K}_\varepsilon(k_n)$ that $\mathbb{E}\left[|\mathcal{N}_{\text{box}}(p\Delta p')|\right] \le k_n^\varepsilon$ implies that eventually and $|x-x'| \le k_n^{1+\varepsilon}$. In particular $|x-x'| \le k_n^{1+\varepsilon}$ for all $(p, p') \in \mathcal{E}_\varepsilon(k_n)^c$. Therefore, we have

$$
\begin{aligned}
&\mathbb{E}\left[I_{a,b} - I^*_{a,b}\right] \\
&\le \int_{\mathcal{K}_C(k_n)^2 \setminus \mathcal{E}_\varepsilon(k_n)} \rho(p, p', k-i, k-i')\mathbb{E}\left[J_b(p,p')\right] f(x,y) f(x',y') \, dx' \, dx \, dy' \, dy \\
&\le \int_{\mathcal{K}_C(k_n)^2 \setminus \mathcal{E}_\varepsilon(k_n)} \rho(p, p', k-i, k-i')\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] \mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\right] f(x,y) f(x',y') \, dx' \, dx \, dy' \, dy \\
&= O\left(\int_{\mathcal{K}_C(k_n)^2} \mathbb{1}_{\{|x-x'| \le k_n^{1+\varepsilon}\}} \rho_y(k)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] \mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\right] f(x,y) f(x',y') \, dx' \, dx \, dy' \, dy\right) \\
&= O\left(k_n^{1+\varepsilon} \binom{k_n}{2} \left(\int_{a_n^-}^{a_n^+} \Delta_{\mathcal{P}}(y') e^{-\alpha y'} \, dy'\right) \mathbb{E}\left[T_{\mathcal{P},n}(k_n)\right]\right) \\
&= O\left(k_n^{3+\varepsilon-2\alpha} s_\alpha(k_n) \mathbb{E}\left[T_{\mathcal{P},n}(k_n)\right]\right) \\
&= o\left(n k_n^{-(2\alpha-1)} s_\alpha(k_n) \mathbb{E}\left[T_{\mathcal{P},n}(k_n)\right]\right) = o\left(\mathbb{E}\left[T_{\mathcal{P},n}(k_n)\right]^2\right),
\end{aligned}
$$

which proves (81). Here we used that $k_n^{2+\varepsilon} = o(n)$ and $\mathbb{E}\left[T_{\mathcal{P},n}(k_n)\right] = \Theta\left(n k_n^{-(2\alpha+1)} s_\alpha(k_n)\right)$ for the last line.

We will now proceed to establish (79) and (80). We start with $I^*_{0,0}$

By Lemma 8.7

$$
\rho_(p, p', k_n - i, k_n - i') = (1+o(1))\rho(p, p', k_n - j, k_n - j') = (1+o(1))\rho_{\text{box}}(p, p', k_n, k_n)
$$

which now no longer depends on the other four points p_1, p_2, p'_1, p'_2. Hence, using the Campbell-Mecke formula, we get

$$
\mathbb{E}\left[I^*_{0,0}\right] = (1+o(1)) \int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p, p', k_n, k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] \mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\right] f(x,y) f(x',y') \, dx' \, dx \, dy' \, dy,
$$

Next, by Corollary 7.4 we have for $y \in K_C(k_n)$,

$$
\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] = (1+o(1))\mathbb{E}\left[\widetilde{T}_{\mathcal{P}}(p)\right] = (1+o(1))\binom{k_n}{2}\Delta_{\mathcal{P}}(y)
$$

and similar for p'. Hence, by Lemma 8.6

$$
\begin{aligned}
&(1+o(1)) \int_{\mathcal{E}_\varepsilon(k_n)} \rho_n(p, p', k_n, k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] \mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\right] f(x,y) f(x',y') \, dx' \, dx \, dy' \, dy \\
&\quad = (1+o(1))\binom{k_n}{2}^2 \int_{\mathcal{E}_\varepsilon(k_n)} \rho(y, k_n)\rho(y', k_n)\Delta_{\mathcal{P}}(y)\Delta_{\mathcal{P}}(y') f(x,y) f(x',y') \, dx' \, dx \, dy' \, dy \\
&\quad = (1+o(1))\left(\binom{k_n}{2} \int_{a_n^-}^{a_n^+} \int_{I_n} \rho(y, k_n)\Delta_{\mathcal{P}}(y) f(x,y) \, dx \, dy\right)^2 \\
&\quad = (1+o(1))\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right)^2,
\end{aligned}
$$

which proves (79).

Next we consider $I^*_{0,1}$. The proofs for the other two cases $I^*_{1,1}$ and $I^*_{0,2}$ follow using similar arguments and hence we omit them here.

Without loss of generality we will assume that $p_1 = p_1'$. Notice that now $i = |\{p', p_1, p_2, p_2'\} \cap \mathcal{B}_{\mathrm{box}}(p)|$ and $j = |\{p'\} \cap \mathcal{B}_{\mathrm{box}}(p)|$ and i', j' are defined, similarly, by interchanging the primed and non-primed variables. Then, if we define

$$T_{\mathcal{P},n}^{(0,1)}(p,p') = \sum_{(p_1,p_2)\in 2^{\mathcal{P}_n}} \sum_{p_2' \in \mathcal{P}_n} \widetilde{T}_{\mathrm{box}}(p,p_1,p_2)\widetilde{T}_{\mathrm{box}}(p',p_1,p_2'),$$

we have

$$\mathbb{E}\left[I_{0,1}^*\right] = (1+o(1)) \int_{\mathcal{E}_\varepsilon(k_n)} \rho_n(p,p',k_n,k_n)\mathbb{E}\left[T_{\mathcal{P},n}^{(0,1)}(p,p')\right] f(x,y)f(x',y')\, dx'\, dx\, dy'\, dy,$$

where we again used Lemma 8.7. We will show that

$$\mathbb{E}\left[T_{\mathcal{P},n}^{(0,1)}(p,p')\right] = o\left(k_n^4 s_\alpha(k_n)^2\right),$$

from which (80) follows since $k_n^4 s_\alpha(k_n)^2 = O\left(\mathbb{E}\left[\widetilde{T}_{\mathrm{box}}(p)\right]\mathbb{E}\left[\widetilde{T}_{\mathrm{box}}(p')\right]\right)$ on $\mathcal{K}_C(k_n) \times \mathcal{K}_C(k_n)$.

First we consider the contribution coming from $y_1 > 4\log(k_n)$. Since the integration of $T_{\mathcal{P}}(p,p_1,p_2)T_{\mathcal{P}}(p',p_1,p_2')$ over x_1, x_2 and x_2' is bounded by $O\left(e^y e^{\frac{y'}{2}} e^{\frac{y_1+y_2+y_2'}{2}}\right)$ it follows that contribution to $\mathbb{E}\left[T_{\mathcal{P},n}^{(0,1)}(p,p')\right]$ is bounded by

$$\begin{aligned} O\left(e^y e^{\frac{y'}{2}} \int_{4\log(k_n)}^{a_n^+} e^{-(\alpha-\frac{1}{2})y_1}\, dy_1\right) &= O\left(k_n^3 \int_{4\log(k_n)}^{a_n^+} e^{-(\alpha-\frac{1}{2})y_1}\, dy_1\right) \\ &= O\left(k_n^{3-(4\alpha-2)}\right) = o\left(k_n^4 s_\alpha(k_n)^2\right). \end{aligned}$$

To deal with the case where $y_1 \le 4\log(k_n)$ we define $b_n = 2\varepsilon\log(k_n))$ and will consider different cases for $\mathbb{E}\left[T_{\mathcal{P},n}^*(p,p')\right]$, depending on whether $y_2 \le b_n$ or $y_2 > b_n$ and similar for y_2'.

When $y_1 \le 4\log(k_n)$ and $y_2 > b_n$, the contribution to $\mathbb{E}\left[T_{\mathcal{P},n}^{(0,1)}(p,p')\right]$ is bounded by

$$\mathbb{E}\left[\widetilde{T}_{\mathrm{box}}(p)\right] O\left(e^{\frac{y'}{2}} \int_{b_n}^{a_n^+} e^{-(\alpha-\frac{1}{2})y_2}\, dy_2\right) = O\left(k_n^{1-\varepsilon}\right)\mathbb{E}\left[\widetilde{T}_{\mathrm{box}}(p)\right] = o\left(k_n^4 s_\alpha(k_n)^2\right).$$

Due to the symmetry in p_2 and p_2' the same results holds for the cases where $y_2 > b_n$.

Finally, when $y_1 \le 4\log(k_n)$ and both $y_2, y_2' \le b_n$ we have that

$$|x_2 - x_2'| \le |x_1 - x_2| + |x_1 - x_2'| \le e^{\frac{y_1}{2}}\left(e^{\frac{y_2}{2}} + e^{\frac{y_2'}{2}}\right) \le 2k_n^{2+\varepsilon}$$

whenever $T_{\mathcal{P}}(p,p_1,p_2)T_{\mathcal{P}}(p',p_1,p_2') > 0$ while both $|x - x_2|, |x' - x_2'| = O\left(k_n^{1+\varepsilon}\right)$. Hence it follows that

$$|x - x'| \le |x - x_2| + |x_2 - x_2'| + |x_2' - x'| = O\left(k_n^{2+\varepsilon}\right).$$

Next, by integrating only over x_2' and y_2' we get the contribution to $\mathbb{E}\left[T_{\mathcal{P},n}^{(0,1)}(p,p')\right]$ for this regime is bounded by

$$O\left(e^{\frac{y'}{2}}\mathbb{E}\left[\widetilde{T}_{\mathrm{box}}(p)\right]\right) = O\left(k_n\mathbb{E}\left[\widetilde{T}_{\mathrm{box}}(p)\right]\right) = o\left(k_n^4 s_\alpha(k_n)^2\right).$$

□

9 Equivalence for local clustering in $G_{\mathbf{Po}}$ and $G_{\mathbf{box}}$

In this section we establish the equivalence between $c^*(k; G_n)$ and $c^*(k; G_{\text{box}})$ as expressed in Proposition 5.3, using the coupling procedure explained in Section 2.4.

Recall that $\mathcal{P}$ denotes a Poisson process on $\mathbb{R} \times \mathbb{R}_+$, with intensity $f(x,y)$, $I_n = \frac{\pi}{2} e^{R_n/2}$, $\mathcal{R} = (-I_n, I_n] \times (0, R_n]$ and $\mathcal{V}_n = \mathcal{P} \cap \mathcal{R}$. In addition we define for any interval $I \subseteq \mathbb{R}_+$, $\mathcal{R}(I) := (-I_n, I_n] \times I$ and denote by $\mathcal{B}_{\text{box}}(p)$ the *ball*

$$\mathcal{B}_{\text{box}}(p) = \left\{ p' \in \mathcal{V}_n : |x - x'|_{\pi e^{R_n/2}} < e^{\frac{y+y'}{2}} \right\}.$$

Note that when $p \in \mathcal{V}_n$ then $\mathcal{B}_{\text{box}}(p)$ denotes its neighborhood in the graph G_{box}. Note that the above definition implies that for all $y \in [0, R_n]$ we have

$$\mathcal{R}([R_n - y - 2\ln(\pi/2), R_n]) \subseteq \mathcal{B}_{\text{box}}((0,y)) \tag{82}$$

- this is a fact which we are going to use several times in our analysis. **Pim:** @All: Maybe we should either add figure to illustrate this or we could refer to a previous figure.

For any Borel-measurable subset $S \subseteq \mathbb{R} \times \mathbb{R}_+$, we let

$$\mu(S) = \int_S f(x,y)\, dx\, dy = \frac{\nu\alpha}{\pi} \int_S e^{-\alpha y} dy.$$

Thus, the number of points of $\mathcal{P}_{\alpha,\nu}$ inside S is distributed as $\text{Po}(\mu_{\alpha,\nu,}(S))$.

Finally, we recall the map Ψ from (5)

$$\Psi(r,\theta) = \left(\theta \frac{e^{R_n/2}}{2}, R_n - r\right),$$

and remind the reader that $\mathcal{B}(p)$ denotes the image under Ψ of the ball of hyperbolic radius R_n around the point $\Psi^{-1}(p)$ and that under the coupling between the hyperbolic random graph and the finite box model, described in Section 2.4, two points p and p' are connected if and only if

$$|x - x'|_{\pi e^{r_n/2}} \le \Phi(R_n - y, R_n - y'),$$

where the function Φ can be approximated, for $y + y' < R_n$, using Lemma 2.2 by

$$e^{\frac{1}{2}(y+y')} - K e^{\frac{3}{2}(y+y') - R_n} \le \Phi(R_n - y, R_n - y') \le e^{\frac{1}{2}(y+y')} + K e^{\frac{3}{2}(y+y') - R_n}.$$

9.1 Some results on the hyperbolic geometric graph

We start with some basic results for the hyperbolic random geometric graph. Observe that (8) from Lemma 2.2 implies the following.

Corollary 9.1.

$$\mathcal{B}_\infty(p) \cap \mathcal{R}([K, R_n]) \subseteq \mathcal{B}(p) \cap \mathcal{R}(K, R_n).$$

Furthermore, Lemma 2.2 enables us to determine the measure of a ball around a given point $p = (0, y)$ - this is will be fairly useful in our subsequent analysis.

Let $p \in \mathcal{R}$. Then we can see that the curve $x' = e^{\frac{1}{2}(y+y')}$ with $x' \ge 0$ meets the right boundary of $\mathcal{R}_N$, that is, the line $x' = \frac{\pi}{2} e^{R_n/2}$ at $y' = R_n - y + 2\ln\frac{\pi}{2}$. Hence, any point $p' \in \mathcal{R}([R_n - y + 2\ln\frac{\pi}{2}, R_n])$ is included in $\mathcal{B}_\infty(p)$. In other words,

$$\mathcal{B}_\infty(p) \cap \mathcal{R}([R_n - y + 2\ln\frac{\pi}{2}, R_n]) = \mathcal{R}([R_n - y + 2\ln\frac{\pi}{2}, R_n]).$$

This together with the fact that for any $u' = (r', \theta')$,

$$r' < y = R_n - r \Rightarrow d_{\mathbb{H}}(\Psi^{-1}(p), u') \le R_n$$

implies that

$$(\mathcal{B}(p) \triangle \mathcal{B}_\infty(p)) \cap \mathcal{R}([R_n - y + 2\ln\frac{\pi}{2}, R_n]) = \emptyset, \tag{83}$$

where $A \triangle B$ denotes the symmetric difference. We can now compute the expected number of points in $\mathcal{B}(p) \triangle \mathcal{B}_\infty(p)$, i.e. those that belong are a neighbor of p in only one of the two models.

Lemma 9.2. *Let $0 \le y_n < R_n$ be such that $R_n - y_n \to \infty$ and write $p_n = (x_n, y_n)$. Then we have, as $n \to \infty$,*

$$\mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)) = \Theta(1) \cdot \begin{cases} e^{(1/2-\alpha)R_n + \alpha y_n}, & \text{if } \alpha < 3/2 \\ (R_n - y_n)e^{3y/2 - R_n}, & \text{if } \alpha = 3/2 \\ e^{3y_n/2 - R_n}, & \text{if } \alpha > 3/2 \end{cases}.$$

Proof. Let $r_n := R_n - y$. Lemma 2.2 implies that for such a p_n, if a point p belongs to $\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n) \cap \mathcal{R}([0, r_n])$ then

$$|x_n - x| = \Theta(1) \cdot e^{\frac{3}{2}(y_n + y) - R_n}.$$

Now, if $p \in [r_n, r_n + 2\ln\frac{\pi}{2})]$ and also $p \in \mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)$, then

$$|x_n - x| = \frac{\pi}{2} e^{R_n/2} - e^{\frac{1}{2}(y_n + y)}.$$

Finally, (83) implies that no point in $\mathcal{R}([r_n + 2\ln\frac{\pi}{2}, R_n])$ belongs to $\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)$. We first compute the expected number of points in $p \in \mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)$ that have $R_n - y \le r_n$. The result depends on the value of α, yielding the following three cases

$$\begin{aligned} \mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n) \cap \mathcal{R}([0, r_n])) &= \Theta(1) \cdot e^{3y_n/2 - R_n} \int_0^{r_n} e^{(3/2-\alpha)y} \, dy \\ &= \Theta(1) \cdot \begin{cases} e^{(1/2-\alpha)R_n + \alpha y_n}, & \text{if } \alpha < 3/2 \\ (R_n - y_n)e^{3y_n/2 - R_n}, & \text{if } \alpha = 3/2 \\ e^{3y_n/2 - R_n}, & \text{if } \alpha > 3/2 \end{cases}. \end{aligned}$$

Next we compute the number of remaining points in $\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)$,

$$\begin{aligned} \mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n) \cap \mathcal{R}([r_n, R_n])) &= \frac{\nu\alpha}{\pi} \int_{r_n}^{r_n + 2\ln\frac{\pi}{2}} \left(\frac{\pi}{2} e^{R_n/2} - e^{\frac{1}{2}(y_n + y)}\right) e^{-\alpha y} \, dy \\ &= O(1) \cdot e^{R_n/2} \int_{r_n}^{r_n + 2\ln\frac{\pi}{2}} e^{-\alpha y} \, dy = O(1) \cdot e^{R_n/2} e^{-\alpha r_n} \\ &= O(1) \cdot e^{(1/2-\alpha)R_n + \alpha y_n}. \end{aligned}$$

Now note that for any $\alpha > 3/2$, we have

$$((1/2 - \alpha)R_n + \alpha y_n) - (3y_n/2 - R_n) = (3/2 - \alpha)(R_n - y_n) \to -\infty,$$

by our assumption on y_n. For $\alpha = 3/2$, these two quantities are equal. From these observations, we deduce that

$$\mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)) = \Theta(1) \cdot \begin{cases} e^{(1/2-\alpha)R_n + \alpha y_n}, & \text{if } \alpha < 3/2 \\ r_n e^{3y_n/2 - R_n}, & \text{if } \alpha = 3/2 \\ e^{3y_n/2 - R_n}, & \text{if } \alpha > 3/2 \end{cases}.$$

$\square$

9.2 Equivalence clustering $G_{\mathbf{Po}}$ and $G_{\mathbf{box}}$

Here we prove Proposition 5.3. We first establish a few results regarding the number of nodes of degree k_n in both the Poissonized KPKVB graph G_{Po} and the finite box model G_{box}.

Lemma 9.3. *Let $\alpha > 1/2$, $\nu > 0$ and $\{k_n\}_{n\geq 1}$ be a sequence such that $k_n = O\left(n^{1/(2\alpha+1)}\right)$. Then*

$$\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] = \Theta\left(1\right) n k_n^{-(2\alpha+1)}, \tag{84}$$

and

$$\mathbb{E}\left[N_{box}(k_n)\right] = \Theta\left(1\right) n k_n^{-(2\alpha+1)}. \tag{85}$$

Moreover,

$$\lim_{n\to\infty}\left|\frac{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]}{\mathbb{E}\left[N_{box}(k_n)\right]} - 1\right| = 0. \tag{86}$$

Proof. Recall that

$$\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] = \int_{\mathcal{R}} \rho_{\mathrm{Po}}(y, k_n) f(x,y)\, \mathrm{d}x\, \mathrm{d}y.$$

Then by Lemma 6.6 and a concentration argument

$$\begin{aligned}\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] &= (1 + o\left(1\right)) \int_{\mathcal{R}} \rho(y, k_n) f(x,y)\, \mathrm{d}x\, \mathrm{d}y \\ &= (1 + o\left(1\right)) n \int_0^{R_n} \rho(y, k_n) f(x,y)\, \mathrm{d}x\, \mathrm{d}y = \Theta\left(1\right) n k_n^{-(2\alpha+1)}.\end{aligned}$$

Similarly,

$$\mathbb{E}\left[N_{\mathrm{box}}(k_n)\right] = (1 + o\left(1\right)) \int_{\mathcal{R}} \rho(y, k_n) f(x,y)\, \mathrm{d}x\, \mathrm{d}y$$

From which both (85) and (86) follow. □

Recall that Proposition 5.3 states

$$\lim_{n\to\infty} s(k_n)^{-1}\, \mathbb{E}\left[\left|c^*(k_n; G_{\mathrm{Po}}) - c^*_{\mathcal{P},n}(k_n)\right|\right] = 0.$$

Since for $\alpha > 3/4$, $s_{3/4}(k_n) = \log(k_n)^{-1} s_\alpha(k_n) = o\left(s_\alpha(k_n)\right)$ it suffices to prove the following two cases:

1. if $1/2 < \alpha \leq 3/4$, then

$$\lim_{n\to\infty} k_n^{4\alpha-2} \cdot \mathbb{E}\left[\left|c^*(k_n; G_{\mathrm{Po}}) - c^*(k_n; G_\infty)\right|\right] = 0,$$

2. if $3/4 < \alpha$, then

$$\lim_{n\to\infty} k_n \cdot \mathbb{E}\left[\left|c^*(k_n; G_{\mathrm{Po}}) - c^*(k_n; G_\infty)\right|\right] = 0.$$

Recall the definition of $\mathcal{K}_C(k_n)$

$$\mathcal{K}_C(k_n) = \left\{y \in \mathbb{R}_+ : \frac{k_n - C\kappa_n}{\xi_{\alpha,\nu}} \vee 0 \leq e^{\frac{y}{2}} \leq \frac{k_n + C\kappa_n}{\xi_{\alpha,\nu}} \wedge e^{R_n/2}\right\},$$

with $C > 0$ and

$$\kappa_n = \sqrt{k_n \log(k_n)}.$$

The following lemma will be frequently used in the proof of Proposition 5.3.

Lemma 9.4. *Let $t, r \in \mathbb{R}$ be fixed and let $\hat{\rho}(y,k)$ be any of the three probability functions $\rho_{\mathrm{Po}}(y,k), \rho_{box}(y,k)$ or $\rho(y,k)$. Then for any sequence k_n of positive integers with $k_n = O\left(n^{\frac{1}{2\alpha+1}}\right)$ and $C > 0$ large enough,*

$$\int_{\mathcal{K}_C} e^{ty} \hat{\rho}_n(y, k_n - r) e^{-\alpha y} \, \mathrm{d}y = O(1) \, k_n^{-2\alpha - 1 + 2t}$$

as $n \to \infty$.

Proof. Note that on $\mathcal{K}_C(k_n)$ we have that $e^{ty} = \Theta\left(k_n^{2t}\right)$. Hence, by a concentration argument

$$\begin{aligned}\int_{\mathcal{K}_C} e^{ty} \hat{\rho}_n(y, k_n - r) e^{-\alpha y} \, \mathrm{d}y &= \Theta\left(k_n^{2t}\right) \int_{\mathcal{K}_C} \hat{\rho}_n(y, k_n - r) e^{-\alpha y} \, \mathrm{d}y \\ &= O\left(k_n^{2t}\right) n \mathbb{E}\left[N_\infty(k_n)\right] = O(1) \, n k_n^{-2\alpha-1+2t}.\end{aligned}$$

□

Proof of Proposition 5.3. To keep notations concise we abbreviate $\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]$ and $\mathbb{E}\left[N_\infty(k_n)\right]$ by $\overline{n}_{\widetilde{\mathbb{H}}}(k_n)$ and $\overline{n}_{\mathcal{P}}(k_n)$, respectively. We will also suppress the subscripts n in most expression regarding the graphs G_{Po} and G_{box}. Finally we will write $P_{\mathrm{Po}}(p)$ to denote the probability that two random neighbors of a point p in G_{Po} for a triangle. Then we have

$$\begin{aligned}\mathbb{E}\left[\left|c^*(k_n; G_{\mathrm{Po}}) - c^*(k_n; G_\infty)\right|\right] &= \binom{k_n}{2}^{-1} \mathbb{E}\left[\left|\sum_{p \in \mathcal{P}} \frac{\mathbb{1}_{\{\deg_{\mathrm{Po}}(p) = k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} P_{\mathrm{Po}}(p) - \frac{\mathbb{1}_{\{\deg_\infty(p) = k_n\}}}{\overline{n}_{\mathcal{P}}(k_n)} P(p)\right|\right] \\ \leq \binom{k_n}{2}^{-1} &\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \mathbb{E}\left[\left|\sum_{p \in \mathcal{P}} \mathbb{1}_{\{\deg_{\mathrm{Po}}(y) = k_n\}} P_{\mathrm{Po}}(p) - \mathbb{1}_{\{\deg_\infty(p) = k_n\}} P(p)\right|\right] \\ &+ \binom{k_n}{2}^{-1} \left|\frac{1}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} - \frac{1}{\overline{n}_{\mathcal{P}}(k_n)}\right| \mathbb{E}\left[\sum_{p \in \mathcal{P}} \mathbb{1}_{\{\deg_\infty(p) = k_n\}} P(p)\right]\end{aligned}$$

The last term can be rewritten as

$$\left|1 - \frac{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}{\overline{n}_{\mathcal{P}}(k_n)}\right| \mathbb{E}\left[c^*(k_n; G_\infty)\right] = \left|1 - \frac{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}{\overline{n}_{\mathcal{P}}(k_n)}\right| \gamma(k_n)(1 + o(1)),$$

where we used Proposition 5.5 (See Section 7). The first term in this product converges to zero by Lemma 9.3 while the second term scales as $s(k_n)$. Hence

$$\left|1 - \frac{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}{\overline{n}_{\mathcal{P}}(k_n)}\right| \mathbb{E}\left[c^*(k_n; G_\infty)\right] = o\left(s(k_n)\right),$$

and therefore we are left to analyze the other term. By the Campbell-Mecke formula we have that

$$\begin{aligned}&\mathbb{E}\left[\left|\sum_{p \in \mathcal{P}} \frac{\mathbb{1}_{\{\deg_n(y) = k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} P_{\mathrm{Po}}(p) - \frac{\mathbb{1}_{\{\deg_\infty(p) = k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} P(p)\right|\right] \\ &= \int_{\mathcal{R}} \mathbb{E}\left[\left|\frac{\mathbb{1}_{\{\deg_n(y) = k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} P_{\mathrm{Po}}(y) - \frac{\mathbb{1}_{\{\deg_\infty(y) = k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} P(y)\right|\right] f(x,y) \, \mathrm{d}y \, \mathrm{d}x.\end{aligned}$$

Since

$$\begin{aligned}\mathbb{E}\left[\frac{\mathbb{1}_{\{\deg_n(y) = k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} P_{\mathrm{Po}}(y)\right] &\leq \binom{k_n}{2} \rho_{\widetilde{\mathbb{H}}}(y, k_n) \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \\ &= \binom{k_n}{2} \rho_{\widetilde{\mathbb{H}}}(y, k_n) \Theta\left(\overline{n}_{\mathcal{P}}(k_n)^{-1}\right)\end{aligned}$$

$$= \Theta\left(n^{-1}k_n^{2\alpha+3}\right)\rho_{\widetilde{\mathbb{H}},n}(y,k_n)$$

and similar for the other term, it follows that

$$\mathbb{E}\left[\left|\frac{\mathbb{1}_{\{\deg_n(y)=k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}P_{\mathrm{Po}}(y) - \frac{\mathbb{1}_{\{\deg_\infty(y)=k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}P(y)\right|\right] \\ \leq \Theta\left(n^{-1}k_n^{2\alpha+3}\right)\left(\rho_{\widetilde{\mathbb{H}},n}(k,n)+\rho_n(y,k_n)\right).$$

Therefore, by a concentration argument (c.f. Corollary **??**), it is enough to consider the integral

$$\int_{\mathcal{K}_C(k_n)}\mathbb{E}\left[\left|\frac{\mathbb{1}_{\{\deg_n(y)=k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}P_{\mathrm{Po}}(y) - \frac{\mathbb{1}_{\{\deg_\infty(y)=k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}P(y)\right|\right]e^{-\alpha y}\,\mathrm{d}y\,\mathrm{d}x, \tag{87}$$

where we also used that $f(x,y)$ is simply a constant multiple of the function $e^{-\alpha y}$. We shall proceed by expanding the integrand and analyzing the individual terms. With a slight abuse of notation we shall write y instead of $(0,y)$ in expression such as $\mathcal{B}(y)$. In addition we write $D_{\mathbb{H}}(y,k_n;\mathcal{P})$ for the indicator which is equal to 1 if and only if $\mathcal{B}((0,y))$ contains k_n points from $\mathcal{P}\setminus\{(0,y)\}$. We define $D_{\mathcal{P}}(y,k_n;\mathcal{P})$ analogously for the ball $\mathcal{B}_\infty((0,y))$.

We need to split the integrand over several terms and then analyze each of these separately. Applying the Campbel-Mecke formula yields

$$\mathbb{E}\left[\left|\frac{\mathbb{1}_{\{\deg_n(y)=k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}P_{\mathrm{Po}}(y) - \frac{\mathbb{1}_{\{\deg_\infty(y)=k_n\}}}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}P(y)\right|\right] \leq \\ \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{P}\setminus\{(0,y)\}}^{\neq}\left|\frac{D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}\Delta_{\mathbb{H}}(y,p_1,p_2) \right.\right. \\ \left.\left. - \frac{D_{\mathcal{P}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}\Delta_{\mathcal{P}}(y,p_1,p_2)\right|\right],$$

where the sum ranges over all distinct pairs of points in $\mathcal{P}\setminus\{(0,y)\}$. In what follows, we will set $\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p') = \mathcal{B}(p')\triangle\mathcal{B}_\infty(p')$ and $\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}(p') = \mathcal{B}(p')\cap\mathcal{B}_\infty(p')$. We will now bound the sum that is inside the expectation. Note that each summand is the absolute value of the difference between two quantities that are either equal to 0 or of order $\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}$ $(\overline{n}_{\mathcal{P}}(k_n)^{-1})$. We will split these summands into 5 classes which are all combinations of $p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}$ for which only one of the two terms of this difference is non-zero.

1. both p_1 and p_2 have $y_1,y_2<(1-\varepsilon)R_n\wedge(R_n-y)$ and
   (a) p_1 is in $\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}(y)$ but $p_2\in\mathcal{B}(p_1)\setminus\mathcal{B}_\infty(p_1)$ and $\mathcal{B}(y)$ contains exactly k_n-2 or k_n-1 other points (depending on whether $p_2\in\mathcal{B}(y)$ or not).
   (b) p_1 is in $\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}(y)$ but $p_2\in\mathcal{B}_\infty(p_1)\setminus\mathcal{B}(p_1)$ and $\mathcal{B}_\infty(y)$ contains exactly k_n-2 or k_n-1 other points (depending on whether $p_2\in\mathcal{B}_\infty(y)$ or not).
2. the above cases but with $y_1\geq(1-\varepsilon)R_n\wedge(R_n-y)$.
3. $y_1\geq K$ and $p_1\in\mathcal{B}(y)\setminus\mathcal{B}_\infty(y)$ and $p_2\in\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}(y)$ - here we use Corollary 9.1 which implies that if $p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)$ and $y_1\geq K$, then in fact $p_1\in\mathcal{B}(y)\setminus\mathcal{B}_\infty(y)$.
4. $y_1<K$ and $p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)$ and $p_2\in\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}(y)$.

We bound this sum by the following expression:

$$\sum_{(p_1,p_2)\in\mathcal{P}\setminus\{(0,y)\}}^{\neq}\left|\frac{D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)}\Delta_{\mathbb{H}}(y,p_1,p_2)\right.$$

$$- \frac{D_{\mathcal{P}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\})}{\overline{n}_{\widetilde{\mathbb{H}}}(k_n)} \Delta_{\mathcal{P}}(y, p_1, p_2) \Bigg|$$

$$\leq \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y_1, y_2 < (1-\varepsilon) R_n \wedge (R_n - y)}} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathbb{H} \cap \mathcal{P}}((0,y))\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1)\}} D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \tag{88}$$

$$\leq \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y_1, y_2 < (1-\varepsilon) R_n \wedge (R_n - y)}} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathbb{H} \cap \mathcal{P}}((0,y))\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1)\}} D_{\mathcal{P}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \tag{89}$$

$$+ \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y_1 \geq (1-\varepsilon) R_n \wedge (R_n - y)}} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathbb{H} \cap \mathcal{P}}(y)\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1) \cap \mathcal{B}(y)\}} D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \tag{90}$$

$$+ \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y_1 \geq (1-\varepsilon) R_n \wedge (R_n - y)}} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathbb{H} \cap \mathcal{P}}(y)\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1) \cap \mathcal{B}(y)\}} D_{\mathcal{P}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \tag{91}$$

$$+ 2\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y(p_1) \geq K}} \mathbb{1}_{\{p_1 \in \mathcal{B}(y) \setminus \mathcal{B}_{\infty}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y) \cap \mathcal{B}_{\infty}(y)\}} D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \tag{92}$$

$$+ (\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} + \overline{n}_{\mathcal{P}}(k_n)^{-1}) \sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y(p_1) < K}} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y) \cap \mathcal{B}_{\infty}(y)\}} \tag{93}$$

In the following sections we will give upper bounds on the expected values of each one of these partial sums.

The sums (88) **and** (89) We will analyze (88). The analysis of the other sum (89) is similar. Note first that for any two points p_1, p_2 the following holds: $p_1 \in \mathcal{B}(y)$ and $p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1) \cap \mathcal{B}(y)$, then $p_2 \in \mathcal{B}(y)$ and $p_1 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_2) \cap \mathcal{B}(y)$. Using this symmetry, it suffices to consider distinct pairs $(p_1, p_2) \in \mathcal{P} \setminus \{(0, y)\}$ with $0 \leq y_2 \leq y_1 \leq R - y$. Let $\mathcal{D}$ denote the set of these pairs.

We are going to consider several sub-cases and, thereby, split the domain $\mathcal{D}$ into the corresponding sub-domains. Let $\omega = \omega(n) \to \infty$ as $n \to \infty$ be a slowly growing function and set $y_\omega := y + \omega$. We let

$$\begin{aligned}
\mathcal{D}_1 &= \{(p_1, p_2) \in \mathcal{D} \ : \ y \leq y_1 \leq R_n/2, \ y_\omega \leq y_2 \leq y_1\}, \\
\mathcal{D}_2 &= \{(p_1, p_2) \in \mathcal{D} \ : \ y_1 \leq R_n/2, \ y_2 \leq y_\omega\} \text{ and} \\
\mathcal{D}_3 &= \{(p_1, p_2) \in \mathcal{D} \ : R_n/2 < y_1 \leq R_n - y, \ y_2 \leq y_1\}.
\end{aligned}$$

Note that $\mathcal{D} \subseteq \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$. Hence, we can write

$$\begin{aligned}
&\mathbb{E}\left[\sum_{\substack{p_1, p_2 \in \mathcal{P} \setminus \{(0,y)\} \\ y_1, y_2 \leq (1-\varepsilon) R_n \wedge (R_n - y)}} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1) \cap \mathcal{B}(y)\}} D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \right] \\
&= \sum_{i=1}^{3} \mathbb{E}\left[\sum_{(p_1, p_2) \in \mathcal{D}_i} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1) \cap \mathcal{B}(y)\}} \cdot D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \right].
\end{aligned} \tag{94}$$

We bound each one of the above three summands as follows:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{(p_1, p_2) \in \mathcal{D}_1} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H} \triangle \mathcal{P}}(p_1) \cap \mathcal{B}(y)\}} D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \right] \\
&\leq \mathbb{E}\left[\sum_{(p_1, p_2) \in \mathcal{D}_1} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\}} D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\}) \right] := \mathcal{I}_n^{(1)}(y),
\end{aligned} \tag{95}$$

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_2} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\
&\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_2} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\}}\, D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] := \mathcal{I}_n^{(2)}(y)
\end{aligned}
\tag{96}
$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_3} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\
&\leq \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_3} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot \mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}\, D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] := \mathcal{I}_n^{(3)}.
\end{aligned}
\tag{97}
$$

We will bound each term using the Campbell-Mecke formula and show for $i = 1, 2, 3$ that for $1/2 < \alpha < 3/4$

$$
\lim_{n\to\infty} k_n^{4\alpha-2}\binom{k_n}{2}^{-1} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(i)} e^{-\alpha}\,\mathrm{d}y = 0, \tag{98}
$$

and for $\alpha \geq 3/4$

$$
\lim_{n\to\infty} k_n\binom{k_n}{2}^{-1} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(i)} e^{-\alpha}\,\mathrm{d}y = 0. \tag{99}
$$

For the first term (95), we note that

$$
\mathbb{P}\left(D_{\mathbb{H}}(y) = k_n - 2; \mathcal{P}\setminus\{p_1,p_2\}\right) = \rho_{\widetilde{\mathbb{H}}}(y, k_n-2).
$$

and hence $\mathcal{I}_n^{(1)}(y)$ becomes

$$
\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\int_{-I_n}^{I_n}\int_{y}^{R_n/2}\int_{-I_n}^{I_n}\int_{y_\omega}^{y_1} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}(y)\}}\,\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1. \tag{100}
$$

Next, Lemma 2.2 implies that for $y' \leq R_n - y$, we have that if $(x', y') \in \mathcal{B}(y)$, then $|x'| < (1+K)e^{y/2+y'/2}$, where $K > 0$ is as in Lemma 2.2. Using these observations, we obtain:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{p_1,p_2\in\mathcal{D}_1} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\cap\mathcal{P}}((0,y))\}}\cdot \mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}\cdot D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\
&\qquad = \rho_{\widetilde{\mathbb{H}}}(y,k_n-2)e^{y}\int_{y}^{R_n/2} e^{y_1/2}\int_{y_\omega}^{y_1} e^{y_2/2}e^{-\alpha y_2}\cdot e^{-\alpha y_1}dy_2dy_1.
\end{aligned}
$$

Now, the double integral becomes

$$
\begin{aligned}
&\int_{y}^{R_n/2} e^{y_1/2}\int_{y_\omega}^{y_1} e^{y_2/2}e^{-\alpha y_2}\cdot e^{-\alpha y_1}dy_2dy_1 = \\
&O(1)\cdot\int_{y}^{R_n/2} e^{y_1/2-\alpha y_1}\cdot e^{(1/2-\alpha)y_\omega}dy_1 \\
&= O(1)\cdot e^{(1/2-\alpha)y_\omega}\cdot\int_{y}^{R_n/2} e^{y_1/2-\alpha y_1}dy_1 \\
&= O(1)\cdot e^{(1/2-\alpha)y_\omega+(1/2-\alpha)y} \\
&\ll e^{(1-2\alpha)y},
\end{aligned}
\tag{101}
$$

since $y_\omega = y + \omega$ and $\omega \to \infty$. We then deduce that

$$\mathbb{E}\left[\sum_{p_1,p_2 \in \mathcal{D}_1} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathbb{H}\cap\mathcal{P}}((0,y))\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\}} \cdot D_{\mathbb{H}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\})\right] \ll \rho_{\widetilde{\mathbb{H}}}(y, k_n - 2) e^{(1-2\alpha)y}. \tag{102}$$

We now integrate this with respect to y and determine its contribution to (87) is

$$\begin{aligned}&\binom{k_n}{2}^{-1} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \rho_{\widetilde{\mathbb{H}}}(y, k_n - 2) e^{(1-2\alpha)y} e^{-\alpha y} \, dy \, dx \\ &= O\left(k^{2\alpha-1} k_n^{-6\alpha+1}\right) = O\left(k_n^{-4\alpha}\right),\end{aligned}$$

where we used Lemma 9.4 with $s = 1 - 2\alpha$.

Since $k_n^{-4\alpha} = o\left(\min\{k_n^{-4\alpha+2}, k_n^{-1}\}\right)$ for all $\alpha > 1/2$ we deduce that for $1/2 < \alpha < 3/4$

$$\lim_{n\to\infty} k_n^{4\alpha-2} \binom{k_n}{2}^{-1} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y} \, dy = 0,$$

while for $\alpha \geq 3/4$

$$\lim_{n\to\infty} k_n \binom{k_n}{2}^{-1} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y} \, dy = 0.$$

We will now bound the term in (96). Using similar observations as for the previous term we get that $\mathcal{I}_n^{(2)}(y)$ equals

$$\rho_{\widetilde{\mathbb{H}}}(y, k_n - 2) \int_{-I_n}^{I_n} \int_0^{R_n/2} \int_{-I_n}^{I_n} \int_0^{y_\omega} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}((0,y))\}} e^{-\alpha(y_1+y_2)} \, dy_2 d dx_2 \, dy_1 \, dx_1.$$

Now, Lemma 2.2 implies that for $y_2 \leq R_n - y_1$, we have that if $(x_2, y_2) \in \mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}((x_1, y_1))$, then x_2 lies in an interval of length $K e^{3y/2+3y'/2-R_n}$, where $K > 0$ is again the constant in Lemma 2.2. Using these observations we obtain:

$$\mathcal{I}_n^{(2)}(y) = \rho_{\widetilde{\mathbb{H}}}(y, k_n - 2) e^{y/2} \int_0^{R_n/2} e^{y_1/2+3y_1/2} \int_0^{y_\omega} e^{3y_2/2-R_n} e^{-\alpha y_2} \cdot e^{-\alpha y_1} \, dy_2 \, dy_1. \tag{103}$$

Now the latter integral is

$$\begin{aligned}&e^{-R_n}\left(\int_0^{R_n/2} e^{(2-\alpha)y_1} \, dy_1\right)\left(\int_0^{y} e^{(3/2-\alpha)y_2} \, dy_2\right) \\ &= O(1)\, e^{-R_n}\left(\begin{cases} e^{(1-\alpha/2)R_n} & \text{if } \frac{1}{2} < \alpha < 2 \\ R_n & \text{if } \alpha \geq 2 \end{cases}\right)\left(\begin{cases} e^{(3/2-\alpha)y} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ y & \text{if } \alpha \geq \frac{3}{2} \end{cases}\right) \\ &= O(1) \begin{cases} e^{-\frac{\alpha}{2}R_n} e^{(3/2-\alpha)y} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ y e^{\frac{\alpha}{2}R_n} & \text{if } \frac{3}{2} \leq \alpha < 2 \\ y R_n e^{-R_n} & \text{if } \alpha \geq 2. \end{cases}\end{aligned}$$

Since $y \leq R_n = O(\log(n))$ we conclude that

$$\mathcal{I}_n^{(2)}(y) = O(1)\, \rho_{\widetilde{\mathbb{H}}}(y, k_n - 2) \begin{cases} n^{-\alpha} e^{(3/2-\alpha)y} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ n^{-\alpha} \log(n) & \text{if } \frac{3}{2} \leq \alpha < 2 \\ n^{-2} \log(n)^2 & \text{if } \alpha \geq 2. \end{cases}$$

We proceed with the integration of this with respect to y and determine its contribution to (87) is

$$O(1)\binom{k_n}{2}^{-2}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)e^{-\alpha y}\,\mathrm{d}y\cdot\begin{cases}n^{-\alpha}k_n^{3-2\alpha} & \text{if } \frac{1}{2}<\alpha<\frac{3}{2}\\ n^{-\alpha}\log(n) & \text{if } \frac{3}{2}\leq\alpha<2\\ n^{-2}\log(n)^2 & \text{if } \alpha\geq 2,\end{cases}$$

$$=O(1)\cdot\begin{cases}n^{-\alpha}k_n^{1-2\alpha} & \text{if } \frac{1}{2}<\alpha<\frac{3}{2}\\ n^{-\alpha}\log(n)k_n^{-2} & \text{if } \frac{3}{2}\leq\alpha<2\\ n^{-2}\log(n)^2k_n^{-2} & \text{if } \alpha\geq 2.\end{cases}$$

Now for $1/2<\alpha<3/4$ it holds that $4\alpha^2-\alpha+1>0$. Hence since $k_n=O\left(n^{\frac{1}{2\alpha+1}}\right)$, we have

$$k_n^{4\alpha-2}n^{-\alpha}k_n^{1-2\alpha}=n^{-\alpha}k_n^{2\alpha-1}=O\left(n^{-\alpha+\frac{2\alpha-1}{2\alpha+1}}\right)=O\left(k_n^{-\frac{4\alpha^2-\alpha+1}{2\alpha+1}}\right)=o(1),$$

from which we deduce that

$$\lim_{n\to\infty}k_n^{4\alpha-2}\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(2)}(y)e^{-\alpha y}\,\mathrm{d}y=0.$$

For $\alpha\geq 3/4$ we have that both $n^{-\alpha}\log(n)k_n^{-1}$ and $n^{-2}\log(n)^2k_n^{-1}$ converge to zero as $n\to\infty$ and hence in this case

$$\lim_{n\to\infty}k_n\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(2)}(y)e^{-\alpha y}\,\mathrm{d}y=0.$$

We will now consider the term in (97). Recall that $\mathcal{D}_3$ consists of all pairs $(p_1,p_2)\in\mathcal{D}$ such that $R_n/2<y_1\leq(1-\varepsilon)R_n\wedge(R_n-y)$ and $y_2\leq y_\omega$ with the property that $p_1\in\mathcal{B}(y)$ and $p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)$. So, in particular, $p_2\in(\mathcal{B}(p_1)\cup\mathcal{B}_\infty(p_1))\cap\mathcal{B}(y)$.

We will consider this intersection more closely. We use Lemma 2.2 to define a ball around p_1 that contains both $\mathcal{B}(p_1)$ and $\mathcal{B}_\infty(p_1)$. For $K>0$, we define, for any point $p_1=(x_1,y_1)\in\mathbb{R}\times\mathbb{R}_+$,

$$\check{\mathcal{B}}_{\mathbb{H},n}(p_1):=\{(x',y')\ :\ y'<R_n-y_1,\ |x_1-x'|<(1+K)e^{\frac{1}{2}(y_1+y')}\}.\tag{104}$$

It is an implication of Lemma 2.2 that

$$(\mathcal{B}(p_1)\cup\mathcal{B}_\infty(p_1))\cap\mathcal{R}([0,R_n-y_1])\subseteq\check{\mathcal{B}}_{\mathbb{H},n}(p_1)$$

Therefore, any point $p_2=(x_2,y_2)\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)$ with $y_2\leq R-y_1$ must belong to $\check{\mathcal{B}}_{\mathbb{H},n}(p_1)\cap\check{\mathcal{B}}_{\mathbb{H},n}((0,y))$.

We will use this in order to derive a lower bound on y_2 as a function of x_1,y_1. Let us suppose without loss of generality that $x_1<0$. The left boundary of $\check{\mathcal{B}}_{\mathbb{H},n}((0,y))$ is given by the equation $x'=(1+K)e^{\frac{1}{2}(y+y')}$ whereas the right boundary of $\check{\mathcal{B}}_{\mathbb{H},n}(p_1)$ is given by the curve having equation $x'=x_1+(1+K)e^{\frac{1}{2}(y_1+y')}$. The equation that determines the intersection point $(\hat{x},\hat{y})$ of these curves is

$$x_1+(1+K)e^{(y_1+\hat{y})/2}=(1+K)e^{(y+\hat{y})/2}.$$

We can solve the above for $\hat{y}$

$$|x_1|=(1+K)e^{\hat{y}/2}\left(e^{y_1/2}+e^{y/2}\right).$$

But $y_1>R_n/2$ and $y<(1+\varepsilon)R_n/(2\alpha+1)$. So if ε is small enough depending on α, we have

$$|x_1|=(1+K)e^{\hat{y}/2}\left(e^{y_1/2}+e^{y/2}\right)=(1+K+o(1))e^{\hat{y}/2+y_1/2}.$$

Let c_K denote the multiplicative term $1+K+o(1)$, which appears in the above. The above yields

$$\hat{y} = \Big(2\log(|x_1|e^{-y_1/2}) - \log c_K\Big) \vee 0 := c(x_1,y_1). \tag{105}$$

In particular, note that $\hat{y}=0$ if and only if $|x_1| \le c_K e^{y_1/2}$. Moreover, since $p_1 \in \mathcal{B}(y)$ and $x_1 \le R_n - y$, we also have that $|x_1| \le e^{(y+y_1)/2}(1+o(1))$. This upper bound on $|x_1|$ together with (112), imply that for n sufficiently large, we have $\hat{y} \le y$. This observation will be used below, where we integrate over y_2, thus ensuring that the integrals are non-zero.

We conclude that

$$p' \in \check{\mathcal{B}}_{\mathbb{H},n}(y) \cap \check{\mathcal{B}}_{\mathbb{H},n}((x_1,y_1)) \Rightarrow y' \ge c(x_1,y_1),$$

Therefore we have

$$\mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}} \le \mathbb{1}_{\left\{y_2 \ge c(x_1,y_1),\, p_2 \in \check{\mathcal{B}}_{\mathbb{H},n}((0,y))\right\}}. \tag{106}$$

If we integrate this over x_2, y_2 we get

$$\begin{aligned}
&\int_{-I_n}^{I_n}\int_0^{y_\omega} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha y_2} dy_2 dx_2 \\
&\le \int_{-I_n}^{I_n}\int_0^{y_\omega} \mathbb{1}_{\left\{y_2 \ge c(x_1,y_1),\, p_2 \in \check{\mathcal{B}}_{\mathbb{H},n}(y)\right\}} e^{-\alpha y_2} dy_2 dx_2 \le (1+K)\cdot e^{y/2}\int_{c(x_1,y_1)}^{y_\omega} e^{y_2/2-\alpha y_2} dy_2 \\
&= O(1)\cdot e^{y/2+(1/2-\alpha)c(x_1,y_1)}.
\end{aligned}$$

Note also that

$$\mathbb{E}\left[D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] = \rho_{\widetilde{\mathbb{H}}}(y,k_n-2),$$

uniformly over all $(p_1,p_2) \in \mathcal{D}_3$.

So the Campbell-Mecke formula yields that $\mathcal{I}_n^{(3)}(y)$ equals:

$$\begin{aligned}
&O(1)\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\, e^{y/2}\int_{-I_n}^{I_n}\int_{R_n/2}^{(R_n-y)\wedge(1-\varepsilon)R_n} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} e^{(1/2-\alpha)c(x_1,y_1)-\alpha y_1} dy_1 dx_1 \\
&= O(1)\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\, e^{y/2}\int_{-I_n}^{I_n}\int_{R_n/2}^{(R_n-y)\wedge(1-\varepsilon)R_n} \mathbb{1}_{\left\{p_1\in\check{\mathcal{B}}_{\mathbb{H},n}(y)\right\}} e^{(1/2-\alpha)c(x_1,y_1)-\alpha y_1} dy_1 dx_1.
\end{aligned} \tag{107}$$

Due to the symmetry of $\check{\mathcal{B}}_{\mathbb{H},n}(y)$, the integration over x_1 is:

$$O(1)\cdot e^{y/2}\cdot \int_0^{(1+K)e^{y/2+y_1/2}} e^{c(x_1,y_1)(1/2-\alpha)} dx_1$$

We will split this integral into two parts according to the value of $c(x_1,y_1)$:

$$\int_0^{(1+K)e^{y/2+y_1/2}} e^{c(x_1,y_1)(1/2-\alpha)} dx_1 = \int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} e^{c(x_1,y_1)(1/2-\alpha)} dx_1 + \int_0^{c_K e^{y_1/2}} dx_1.$$

The first integral becomes:

$$\begin{aligned}
&\int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} e^{c(x_1,y_1)(1/2-\alpha)} dx_1 = \int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} e^{c(x_1,y_1)/2(1-2\alpha)} dx_1 \\
&= O(1)\cdot \int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} x_1^{1-2\alpha} e^{-\frac{y_1}{2}(1-2\alpha)} dx_1 \\
&= O(1)\cdot e^{-y_1/2+\alpha y_1}\cdot e^{\frac{(y+y_1)}{2}2(1-\alpha)} \\
&= O(1)\cdot e^{y_1/2+y(1-\alpha)}.
\end{aligned}$$

The second integral trivially gives:

$$\int_0^{c_K e^{y_1/2}} dx_1 = O(1) \cdot e^{y_1/2} = O(1) \cdot e^{y_1/2 + y(1-\alpha)}.$$

We conclude that

$$e^{y/2} \cdot \int_0^{(1+K)e^{y/2+y_1/2}} e^{c(x_1,y_1)(1/2-\alpha)} dx_1 = O(1) \cdot e^{y_1/2+y(3/2-\alpha)}.$$

Now, we integrate this with respect to y_1 and get

$$e^{y(3/2-\alpha)} \int_{R_n/2}^{R_n-y} e^{(1/2-\alpha)y_1} dy_1 = O(1) \cdot e^{y(3/2-\alpha)} e^{(1/2-\alpha)R_n/2} = O(1) \cdot n^{1/2-\alpha} \cdot e^{y(3/2-\alpha)},$$

from which we deduce

$$\mathcal{I}_n^{(3)}(y) = O(1) \cdot n^{1/2-\alpha} e^{y(3/2-\alpha)} \rho_{\widetilde{\mathbb{H}}}(y, k_n - 2). \tag{108}$$

We now apply Lemma 9.4 and get

$$\begin{aligned}
&\binom{k_n}{2}^{-2} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y} \, \mathrm{d}y \\
&= O\left(1\right) k_n^{2\alpha-1} n^{-1} \int_{\mathcal{K}_C(k_n)} e^{(3/2-\alpha)y} \rho_{\widetilde{\mathbb{H}}}(y, k_n - 2) e^{-\alpha y} \, \mathrm{d}y \\
&= O\left(1\right) k_n^{1-2\alpha} n^{-(\alpha-1/2)}.
\end{aligned}$$

Since for $\alpha > 1/2$, $k_n = O\left(n^{\frac{1}{2\alpha+1}}\right) = o\left(n^{1/2}\right)$ we have that $k_n^{4\alpha-2} k_n^{1-2\alpha} n^{-(\alpha-1/2)} = o\left(1\right)$ and hence for $1/2 < \alpha < 3/4$.

$$\lim_{n\to\infty} k_n^{4\alpha-2} \binom{k_n}{2}^{-2} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y} \, \mathrm{d}x \, \mathrm{d}y = 0,$$

For $\alpha \geq 3/4$ we observe that $2\alpha^2 + 2\alpha - 2 < 0$. Hence, since

$$k_n n^{-(\alpha-1/2)} k_n^{1-2\alpha} = O\left(n^{-(\alpha-1/2)} n^{\frac{2-2\alpha}{2\alpha+1}}\right) = O\left(n^{-\frac{2\alpha^2+2\alpha-2}{2\alpha+1}}\right) = o\left(1\right).$$

we get for $\alpha \geq 3/4$

$$\lim_{n\to\infty} k_n \binom{k_n}{2}^{-2} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y} \, \mathrm{d}x \, \mathrm{d}y = 0.$$

The sums (90) and (91) Again, we will only consider (90) since the analysis for the other term is similar. Recall that in this case, we consider pairs (p_1, p_2), with $p_1 = (x_1, y_1)$ satisfying $y_1 \geq (R_n - y) \wedge (1-\varepsilon)R_n$, and $p_1 \in \mathcal{B}(y)$, $p_2 \in \mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1) \cap \mathcal{B}(y)$. We split this into three sub-domains: i. $y_2 \geq R_n - y$; ii. $R_n - y_1 \leq y_2 \leq R_n - y$ and iii. $y_2 < R_n - y_1$. Similar to the analysis above we define

$$\begin{aligned}
\mathcal{D}_1 &:= \{(p_1, p_2) \ : \ p_1, p_2 \ \in \mathcal{P} \setminus \{(0, y)\}, \ y_1 \geq (1-\varepsilon)R_n \wedge (R_n - y), \ R_n - y \leq y_2 \leq R_n\} \\
\mathcal{D}_2 &:= \{(p_1, p_2) \ : \ p_1, p_2 \ \in \mathcal{P} \setminus \{(0, y)\}, \ y_1 \geq (1-\varepsilon)R_n \wedge (R_n - y), \ R_n - y_1 \leq y_2 \leq R_n - y\} \\
\mathcal{D}_3 &:= \{(p_1, p_2) \ : \ p_1, p_2 \ \in \mathcal{P} \setminus \{(0, y)\}, \ y_1 \geq (1-\varepsilon)R_n \wedge (R_n - y), \ y_2 \leq R_n - y_1\}
\end{aligned}$$

and write, for $i = 1, 2, 3$,

$$\mathcal{I}_n^{(i)}(y) := \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_i} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \cdot \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}} \cdot D_{\mathbb{H}}(y, k_n-2; \mathcal{P}\setminus\{p_1,p_2\})\right].$$

In the first case, note that $y_1 + y_2 \geq 2(R_n - y) > R_n$, since $2y < 2(1+\varepsilon)\frac{R_n}{2\alpha+1} < R_n$. Thus, $p_2 \in \mathcal{B}(p_1)$. Furthermore, $y_2 > R_n - y_1 + 2\ln(\pi/2)$, which implies that $p_2 \in \mathcal{B}_\infty(p_1)$ too. Hence, the contribution from these pairs is zero.

The Campbell-Mecke formula yields that:

$$\begin{aligned}\mathcal{I}_n^{(1)}(y) = O(1)&\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\times\\ &\int_{-I_n}^{I_n}\int_{R_n-y}^{R_n} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\cdot e^{-\alpha(y_2+y_1)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1.\end{aligned}$$

We proceed to bound the integral:

$$\begin{aligned}&\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\int_{-I_n}^{I_n}\int_{R_n-y}^{R_n} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1\\ &\leq \int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\int_{-I_n}^{I_n}\int_{R_n-y}^{R_n} e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1\\ &= \left(\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n} e^{-\alpha y_1}\,\mathrm{d}y_1\,\mathrm{d}x_1\right)\left(\int_{-I_n}^{I_n}\int_{R_n-y}^{R_n} e^{-\alpha y_2}\,\mathrm{d}y_2\,\mathrm{d}x_2\right).\end{aligned}$$

We evaluate

$$\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n} e^{-\alpha y_1}dy_1dx_1 = O(1)\cdot n\cdot e^{-\alpha R_n + ((\varepsilon R_n)\vee y))\alpha} = O(1)\cdot n\cdot e^{-\alpha R_n+\alpha y+\alpha\varepsilon R_n}$$

and

$$\int_{-I_n}^{I_n}\int_{R_n-y}^{R_n} e^{-\alpha y_2}dy_2dx_2 = O(1)\cdot n\cdot e^{-\alpha R_n+\alpha y}.$$

Also, $n\cdot e^{-\alpha R_n} = O(1)\cdot e^{(1/2-\alpha)R_n}$, whereby we deduce that

$$\begin{aligned}&\int_{\mathcal{D}_1} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}dy_2dx_2dy_1dx_1\\ &\quad = O(1)\cdot e^{(1-2\alpha)R_n+2\alpha y+\alpha\varepsilon R_n} = O(1)\cdot n^{2(1-2\alpha)+2\alpha\varepsilon}\cdot e^{2\alpha y}.\end{aligned}$$

With these computations we obtain

$$\begin{aligned}&\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(1)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y\\ &= O(1)n^{(2(1-2\alpha)+2\alpha\varepsilon}\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)} e^{2\alpha y}\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)e^{-\alpha y}\,\mathrm{d}y\,\mathrm{d}x\\ &= O(1)n^{(2(1-2\alpha)+2\alpha\varepsilon}k_n^{-2}\,\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\,n\,k_n^{2\alpha-1} = O\left(1\right)n^{(2(1-2\alpha)+2\alpha\varepsilon}k_n^{4\alpha-2}.\end{aligned}$$

Thus, for $1/2 < \alpha < 3/4$, we have

$$k_n^{4\alpha-2}\,n^{2(1-2\alpha)+2\alpha\varepsilon}\,k_n^{4\alpha-2} = n^{2\alpha\varepsilon}\cdot\left(\frac{k_n^2}{n}\right)^{2(2\alpha-1)} = o(1),$$

provided that $\varepsilon = \varepsilon(\alpha) > 0$ is small enough, and hence for such ε

$$\lim_{n\to\infty} k_n^{4\alpha-2}\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(1)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.$$

When $\alpha > 3/4$ we have $2\alpha - 1 > 1/2$ and we get

$$k_n \cdot n^{2(1-2\alpha)+2\alpha\varepsilon}\cdot k_n^{4\alpha-2} \ll k_n \cdot n^{-1/2+2\alpha\varepsilon}\cdot k_n^{4\alpha-2}\cdot n^{1-2\alpha} = o(1),$$

provided that ε is small enough, depending on α, so that

$$\lim_{n\to\infty} k_n\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(1)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.$$

We now consider the second sub-domain $\mathcal{D}_2$. The Campbell-Mecke formula yields that:

$$\begin{aligned}
\mathcal{I}_n^{(2)}(y) &= \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_2}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1\})\right]\\
&= O(1)\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\cdot\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\times\\
&\quad\int_{-I_n}^{I_n}\int_{R_n-y_1}^{R_n-y}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1.
\end{aligned}$$

We bound the integral as follows:

$$\begin{aligned}
&\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\int_{-I_n}^{I_n}\int_{R_n-y_1}^{R_n-y}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1\\
&\le\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\int_{-I_n}^{I_n}\int_{R_n-y_1}^{R_n-y}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1.
\end{aligned}$$

Now,

$$\begin{aligned}
&\int_{-I_n}^{I_n}\int_{R_n-y_1}^{R_n-y}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}\cdot e^{-\alpha y_2}dy_2dx_2 = O(1)\cdot e^{y/2}\int_{R_n-y_1}^{R_n-y}e^{(1/2-\alpha)y_2}dy_2\\
&= O(1)\cdot e^{y/2+(1/2-\alpha)(R_n-y_1)}.
\end{aligned}$$

We then integrate with respect to y_1:

$$\begin{aligned}
O(1)\cdot e^{y/2}\cdot&\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}e^{(1/2-\alpha)(R_n-y_1)}e^{-\alpha y_1}dy_1dx_1\\
&\le O(1)\cdot e^{y/2+(1/2-\alpha)R_n}\cdot\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}e^{(\alpha-1/2)y_1}e^{-\alpha y_1}dy_1dx_1\\
&= O(1)\cdot e^{y/2+(1-\alpha)R_n-((1-\varepsilon)R_n\wedge(R_n-y))/2}\\
&= O(1)\cdot e^{y/2+(1/2-\alpha)R_n+((\varepsilon R_n)\vee y)/2}\\
&= O(1)\cdot e^{y+(1/2-\alpha)R_n+\varepsilon R_n} = O(1)\cdot n^{1-2\alpha+\varepsilon}\cdot e^{y}.
\end{aligned}$$

Therefore, the contribution of this term to (87) is:

$$\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(2)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y$$

$$
\begin{aligned}
&= O\left(n^{1-2\alpha+\varepsilon}\right)\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)e^{y}e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y\\
&= O\left(1\right)n^{1-2\alpha+\varepsilon},
\end{aligned}
$$

where we used Lemma 9.4 with $s=1$.

For $1/2<\alpha<3/4$, we have

$$
k_n^{4\alpha-2}\cdot n^{1-2\alpha+\varepsilon}=n^{\varepsilon}\left(\frac{k_n^2}{n}\right)^{2\alpha-1}=o(1),
$$

provided that $\varepsilon=\varepsilon(\alpha)>0$ is small enough yielding

$$
\lim_{n\to\infty}k_n^{4\alpha-2}\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(2)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y=0.
$$

Similarly, for $\alpha>3/4$ we have $2\alpha-1>1/2$ and we get

$$
k_n\cdot n^{1-2\alpha+\varepsilon}\ll n^{-1/2+\varepsilon}\cdot k_n=o(1),
$$

provided that ε is small enough, so that

$$
\lim_{n\to\infty}k_n\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(2)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y=0.
$$

For the third sub-domain $\mathcal{D}_3$ we shall use (106) which states that if $p_2=(x_2,y_2)\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}\left(p_1\right)\cap\mathcal{B}\left(y\right)$ and $y_2\leq R_n-y_1$, then $y_2\geq c(x_1,y_1)$, where $c(x_1,y_1)=\left(2\log(|x_1|e^{-y_1/2})-\log c_K\right)\vee 0$ (cf. (112)). Moreover, $p_2\in\check{\mathcal{B}}_{\mathbb{H},n}(p_1)$.

Again, we will use the Campbell-Mecke formula:

$$
\begin{aligned}
\mathcal{I}_n^{(3)}(y)&=\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_3}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}\cdot D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right]\\
&=O(1)\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\times\\
&\qquad\int_{-I_n}^{I_n}\int_0^{R_n-y_1}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}dy_2dx_2dy_1dx_1
\end{aligned}
$$

The inner integral with respect to $p_2:=(x_2,y_2)$ is

$$
\begin{aligned}
&\int_{-I_n}^{I_n}\int_0^{R_n-y_1}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}e^{-\alpha y_2}dy_2dx_2\\
&\qquad\leq\int_{-I_n}^{I_n}\int_0^{R_n-y_1}\mathbb{1}_{\left\{y_2\geq c(x_1,y_1),p_2\in\check{\mathcal{B}}_{\mathbb{H},n}((0,y))\right\}}e^{-\alpha y_2}dy_2dx_2\\
&\qquad=O(1)e^{y/2}\int_{c(x_1,y_1)}^{R_n-y_1}e^{y_2/2-\alpha y_2}dy_2\\
&\qquad=O(1)e^{y/2+(1/2-\alpha)c(x_1,y_1)}.
\end{aligned}
$$

Thus, we get

$$
\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\int_{-I_n}^{I_n}\int_0^{R_n-y_1}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p_1)\cap\mathcal{B}(y)\}}\times
$$

$$e^{-\alpha(y_1+y_2)}dy_2dx_2dy_1dx_1$$
$$\leq O(1)\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n} e^{y/2+(1/2-\alpha)c(x_1,y_1)}e^{-\alpha y_1}dy_1dx_1.$$

Due to symmetry, to bound the integral it is enough to integrate this with respect to x_1 from 0 to I_n. We will split this integral into two parts according to the value of $c(x_1, y_1)$:

$$\int_0^{I_n} e^{c(x_1,y_1)(1/2-\alpha)}dx_1 = \int_{c_K e^{y_1/2}}^{I_n} e^{c(x_1,y_1)(1/2-\alpha)}dx_1 + \int_0^{c_K e^{y_1/2}} dx_1.$$

The first integral becomes:

$$\int_{c_K e^{y_1/2}}^{I_n} e^{c(x_1,y_1)(1/2-\alpha)}dx_1 = O(1)\cdot\int_{c_K e^{y_1/2}}^{I_n} x_1^{1-2\alpha}e^{-\frac{y_1}{2}(1-2\alpha)}dx_1$$
$$= \begin{cases} O(R_n)\cdot e^{-y_1/2+\alpha y_1}\cdot e^{\frac{R_n}{2}2(1-\alpha)} & \text{if } \alpha\leq 1\\ O(1)\cdot e^{-y_1/2+\alpha y_1+2(1-\alpha)y_1/2} & \text{if } \alpha>1\end{cases}$$
$$= \begin{cases} O(R_n)\cdot e^{(\alpha-1/2)y_1}\cdot n^{2(1-\alpha)} & \text{if } \alpha\leq 1\\ O(1)\cdot e^{y_1/2} & \text{if } \alpha>1\end{cases}.$$

The second integral trivially gives:

$$\int_0^{c_K e^{y_1/2}} dx_1 = O(1)\cdot e^{y_1/2}.$$

Putting these two together we conclude that

$$e^{y/2}\cdot\int_0^{I_n} e^{c(x_1,y_1)(1/2-\alpha)}dx_1 = O(1)\cdot e^{y_1/2+y(3/2-\alpha)}.$$

Now, we integrate these with respect to y_1:

$$n^{2(1-\alpha)}\cdot\int_{(1-\varepsilon)R_n\wedge(R_n-y)}^{R_n} e^{(\alpha-1/2)y_1-\alpha y_1}dy_1 = O(1)\cdot n^{2(1-\alpha)}\cdot e^{-R_n/2+\varepsilon R_n/2+y/2} \tag{109}$$
$$= O(1)\cdot n^{1-2\alpha+\varepsilon}\cdot e^{y/2}. \tag{110}$$

Therefore, we conclude that

$$\mathcal{I}_n^{(3)}(y) = O\left(R_n\right)n^{1-2\alpha+\varepsilon(2\alpha-1)}\,e^{y/2}\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)$$

and hence, using again Lemma 9.4,

$$\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(3)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y$$
$$= O\left(R_n\right)n^{1-2\alpha+\varepsilon(2\alpha-1)}k_n^{-2}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)} e^{y/2}\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y$$
$$= O\left(R_n\right)n^{1-2\alpha+\varepsilon(2\alpha-1)}.$$

It follows that for $\varepsilon=\varepsilon(\alpha)$ small enough

$$k_n^{4\alpha-2}R_n n^{1-2\alpha+\varepsilon(2\alpha-1)} = R_n n^{\varepsilon(2\alpha-1)}\left(\frac{k_n^2}{n}\right)^{2\alpha-1} = o\left(1\right)$$

and hence for $\alpha > 1/2$,

$$\lim_{n\to\infty} k_n^{4\alpha-2}\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(3)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.$$

Since $4\alpha - 2 \geq 1$ when $\alpha \geq 3/4$ it immediately follows that

$$\lim_{n\to\infty} k_n\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(3)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.$$

The sum of (92) We will give an upper bound for

$$\mathbb{E}\left[\sum_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y(p_1)\geq K}}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\setminus\mathcal{B}_\infty(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\,D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right].$$

Let us set $p=(0,y)$. Recall that $\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}\,(y)\cap\mathcal{R}([R_n-y+2\log\left(\frac{\pi}{2}\right),R_n])=\emptyset$. Thus, the summand in the above sum is equal to 0, when $y_1 > R_n - y + 2\log(\pi/2)$.

Recall the definition of the extended ball $\check{\mathcal{B}}_{\mathbb{H},n}(p)$ around p (104) that contains both $\mathcal{B}\,(p)$ and $\mathcal{B}_\infty\,(p)$

$$\check{\mathcal{B}}_{\mathbb{H},n}(p) := \{p' : y' < R_n - y,\ |x'| < (1+K)e^{\frac{1}{2}(y+y')}\},$$

and that we have $\mathbb{E}\left[D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] = \rho_{\widetilde{\mathbb{H}}}(y,k_n-2)$.

Observe that,

$$\mathcal{B}\,(p)\cap\mathcal{R}([0,r(p)))\subseteq\check{\mathcal{B}}_{\mathbb{H},n}(p)$$

and

$$\mathcal{B}\,(p)\cap\mathcal{R}([r(p),R_n]) = \mathcal{R}([r(p),R_n]).$$

We thus conclude that

$$\mathcal{B}\,(p)\subseteq\check{\mathcal{B}}_{\mathbb{H},n}(p)\cup\mathcal{R}([r(p),R_n]). \tag{111}$$

Hence, if we set

$$h_y(p_1,\mathcal{P}) := \mathbb{1}_{\{p_1\in\mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\cdot\left(\mu\left(\check{\mathcal{B}}_{\mathbb{H},n}(p_1)\cap\check{\mathcal{B}}_{\mathbb{H},n}(p)\right)+\mu\left(\mathcal{R}([R_n-y,R_n])\right)\right),$$

then

$$\begin{aligned}
&\mathbb{1}_{\{p_1\in\mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\cdot\mathbb{E}\left[\left(\sum_{p_2\in\mathcal{P}\setminus\{p,p_1\}}\mathbb{1}_{\{p_2\in\mathcal{B}(p)\cap\mathcal{B}_\infty(p_1)\}}\right)\cdot D_{\mathbb{H}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right]\\
&= O(1)\cdot\mathbb{1}_{\{p_1\in\mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\cdot\mu(\mathcal{B}\,(p)\cap\mathcal{B}\,(p_1))\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)\\
&\leq O(1)\cdot h_y(p_1,\mathcal{P})\rho_{\widetilde{\mathbb{H}}}(y,k_n-2).
\end{aligned}$$

To calculate the expectation of the above function we need to approximate the intersection of the two balls $\check{\mathcal{B}}_{\mathbb{H},n}(p)$ and $\check{\mathcal{B}}_{\mathbb{H},n}(p_1)$, where $p_1=(x_1,y_1)$. Let us assume without loss of generality that $x_1>0$. The right boundary of $\check{\mathcal{B}}_{\mathbb{H},n}(p)$ is given by the equation $x = x(y_1) = (1+K)e^{\frac{1}{2}(y+y_1)}$ whereas the left boundary of $\check{\mathcal{B}}_{\mathbb{H},n}(p_1)$ is given by the curve $x = x(y_1) = x_1-(1+K)e^{\frac{1}{2}(y+y_1)}$.

The equation that determines the intersecting point of the two curves is

$$x_1-(1+K)e^{(\hat{y}+y_1)/2} = (1+K)e^{(\hat{y}+y)/2},$$

where $\hat{y}$ is the y-coordinate of the intersecting point. We can solve the above for $\hat{y}$

$$x_1 = (1+K)e^{\hat{y}/2}\left(e^{y/2}+e^{y_1/2}\right).$$

But since $p_1 = (x_1, y_1) \in \mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p)$, we also have $x_1 > e^{\frac{y+y_1}{2}}$. Therefore,

$$e^{\hat{y}/2} > \frac{1}{1+K}\,\frac{e^{\frac{y+y_1}{2}}}{e^{y/2}+e^{y_1/2}} \geq \frac{1}{2(1+K)}\,\frac{e^{\frac{y_1+y}{2}}}{e^{\max\{y,y_1\}/2}} > \frac{1}{2(1+K)}\,e^{\min\{y,y_1\}/2}.$$

The above yields

$$\hat{y} > \min\{y, y_1\} - 2\log(2(1+K)) := c(y_1, y). \tag{112}$$

which, in turn, implies the following

$$p \in \check{\mathcal{B}}_{\mathbb{H},n}((0,y)) \cap \check{\mathcal{B}}_{\mathbb{H},n}(p_1) \Rightarrow y(p) \geq c(y_1, y). \tag{113}$$

We thus conclude that

$$\mathcal{B}(p_1) \cap \mathcal{B}(p) \subseteq \left(\check{\mathcal{B}}_{\mathbb{H},n}(p) \cap \mathcal{R}([c(y_1,y), R_n])\right) \cup \mathcal{R}([R_n - y, R_n]),$$

which in turn implies that

$$\mu\left(\check{\mathcal{B}}_{\mathbb{H},n}(p_1) \cap \mathcal{B}(p)\right) \leq \mu\left(\check{\mathcal{B}}_{\mathbb{H},n}(p) \cap \mathcal{R}([c(y_1,y), R_n]\right) + \mu(\mathcal{R}([R_n - y, R_n])).$$

Therefore,

$$\begin{aligned} h_y(p_1, \mathcal{P}) \leq\ & \mathbb{1}_{\{p_1 \in \mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\mu\left(\check{\mathcal{B}}_{\mathbb{H},n}(p) \cap \mathcal{R}([c(y_1,y), R_n])\right) \\ & + \mathbb{1}_{\{p_1 \in \mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\mu\left(\mathcal{R}([R_n - y, R_n])\right). \end{aligned}$$

Now, the Campbell-Mecke formula gives

$$\begin{aligned} & \mathbb{E}\left[\sum_{\substack{p_1,p_2 \in \mathcal{P}\setminus\{(0,y)\} \\ y(p_1)\geq K}} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\setminus\mathcal{B}_\infty(y)\}}\mathbb{1}_{\{p_2 \in \mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\, D_{\mathbb{H}}(y, k_n - 2; \mathcal{P}\setminus\{p_1,p_2\})\right] \\ & \leq \mathbb{E}\left[\left(\sum_{p_1\in\mathcal{P}} h_y(p_1, \mathcal{P}\setminus\{p_1\})\right)\right] \\ & = \frac{\nu\alpha}{\pi}\int_{\mathcal{R}} \mathbb{E}\left[h_y(p_1, \mathcal{P}\setminus\{p_1\})\right] e^{-\alpha y_1}\, \mathrm{d}x_1\, \mathrm{d}y_1 \\ & \leq \frac{\nu\alpha}{\pi}\int_{\mathcal{R}} \mathbb{1}_{\{p_1 \in \mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\mu\left(\check{\mathcal{B}}_{\mathbb{H},n}(p) \cap \mathcal{R}([c(y_1,y), R_n])\right) e^{-\alpha y_1}\, \mathrm{d}x_1\, \mathrm{d}y_1 && (114) \\ & \quad + \frac{\nu\alpha}{\pi}\int_{\mathcal{R}} \mathbb{1}_{\{p_1 \in \mathcal{B}(p)\setminus\mathcal{B}_\infty(p)\}}\mu\left(\mathcal{R}([R_n - y, R_n])\right) e^{-\alpha y_1}\, \mathrm{d}x_1\, \mathrm{d}y_1. && (115) \end{aligned}$$

Recall that $(\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}((0,y))) \cap \mathcal{R}([R_n - y + 2\log\left(\frac{\pi}{2}\right), R_n]) = \emptyset$. We will first calculate the measures μ appearing in (114) and (115). The first one is:

$$\begin{aligned} \mu\left(\check{\mathcal{B}}_{\mathbb{H},n}(y) \cap \mathcal{R}([c(y_1,y), R_n])\right) & \leq (1+K)\frac{\nu\alpha}{\pi}\cdot e^{y/2}\int_{c(y_1,y)}^{R_n} e^{-(\alpha-\frac{1}{2})y'}\, dy' \\ & = O\left(e^{\frac{y}{2}-(\alpha-\frac{1}{2})\min\{y,y_1\}}\right). \end{aligned}$$

The second term is:

$$\mu\left(\mathcal{R}([R_n - y, R_n])\right) = \frac{\nu\alpha}{\pi}\int_{R_n - y}^{R_n} \pi e^{\frac{R_n}{2}} e^{-\alpha y'}\, dy' = O\left(e^{\frac{R_n}{2}} e^{-\alpha(R_n - y)}\right) = O\left(e^{\alpha y - (\alpha - \frac{1}{2})R_n}\right).$$

Using these, we get

$$\int_{\mathcal{R}([0, R_n - y_n + 2\ln\frac{\pi}{2}])} \mathbb{E}\left[h_y(p_1, \mathcal{P}\setminus\{p_1\})\right] e^{-\alpha y_1}\, dx_1\, dy_1$$

$$= O(1) \int_{\mathcal{R}([0,R_n-y+2\ln\frac{\pi}{2}])} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(p)\}} e^{\frac{y}{2}-(\alpha-\frac{1}{2})\min\{y,y_1\}-\alpha y_1}\, dx_1\, dy_1 \tag{116}$$

$$+ O(1) \int_{\mathcal{R}([0,R_n-y+2\ln\frac{\pi}{2}])} \mathbb{1}_{\{p_1\in\mathcal{B}((0,y))\}} e^{\alpha y-(\alpha-\frac{1}{2})R_n-\alpha y_1}\, dx_1\, dy_1. \tag{117}$$

Now, Lemma 2.2 implies that for any $y \in [0, R_n - y_n + 2\ln\frac{\pi}{2}]$, we have

$$\int_{-I_n}^{I_n} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}\, dx_1 \le 2Ke^{\frac{3}{2}(y_1+y)-R_n}.$$

Therefore, (116) is

$$\begin{aligned}
&O(1)\cdot e^{2y-R_n}\int_0^{R_n-y+2\ln\frac{\pi}{2}} e^{\frac{3y_1}{2}-(\alpha-\frac{1}{2})\min\{y_1,y\}-\alpha y_1}\, dy_1\\
&\quad= O(1)\cdot e^{2y-R_n}\left(\int_0^y e^{\frac{3y_1}{2}-(2\alpha-\frac{1}{2})y_1}\, dy_1 + e^{-(\alpha-\frac{1}{2})y}\int_y^{R_n-y+2\ln\frac{\pi}{2}} e^{(\frac{3}{2}-\alpha)y_1}\, dy_1\right)\\
&\quad= O(1)\left(\begin{cases} e^{(4-2\alpha)y-R_n}, & \text{if } \alpha<2\\ R_n\cdot e^{2y-R_n}, & \text{if } \alpha\ge 2\end{cases} + \begin{cases} e^{-(\alpha-\frac{1}{2})R_n+y}, & \text{if } \alpha<3/2\\ R_n\cdot e^{2(2-\alpha)y-R_n}, & \text{if } \alpha\ge 3/2\end{cases}\right).
\end{aligned}$$

Similarly, for (117) we have

$$\begin{aligned}
&\int_{\mathcal{R}([0,R_n-y+2\ln\frac{\pi}{2}])} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}((0,y))\}} e^{\alpha y-(\alpha-\frac{1}{2})R_n-\alpha y_1}\, dx_1\, dy_1\\
&\quad= e^{\frac{3y}{2}-R_n+\alpha y-(\alpha-\frac{1}{2})R_n}\cdot\int_0^{R_n-y+2\ln\frac{\pi}{2}} e^{\frac{3y_1}{2}-\alpha y_1}\, dy_1\\
&\quad= O(1)\cdot\begin{cases} e^{\frac{3y}{2}-R_n+\alpha y-(\alpha-\frac{1}{2})R_n+(\frac{3}{2}-\alpha)(R_n-y)}, & \text{if } \alpha<3/2\\ R_n\cdot e^{(\frac{3}{2}+\alpha)y-(\alpha+\frac{1}{2})R_n}, & \text{if } \alpha\ge 3/2\end{cases}\\
&\quad= O(1)\cdot\begin{cases} e^{-(2\alpha-1)R_n+2\alpha y}, & \text{if } \alpha<3/2\\ R_n\cdot e^{(\frac{3}{2}+\alpha)y-(\alpha+\frac{1}{2})R_n}, & \text{if } \alpha\ge 3/2\end{cases}.
\end{aligned}$$

We thus conclude, using $2(2-\alpha)y \le y$ for $\alpha > 3/2$, that

$$\mathbb{E}\left[\left(\sum_{p_1\in\mathcal{P}\setminus\{p\}} h_y(p_1,\mathcal{P}\setminus\{p_1\})\right)\right] \le O(1)\cdot\left(\mathcal{I}_n^{(1)}(y)+\mathcal{I}_n^{(2)}(y)+\mathcal{I}_n^{(3)}(y)\right), \tag{118}$$

where

$$\begin{aligned}
\mathcal{I}_n^{(1)}(y) &= \begin{cases} e^{(4-2\alpha)y-R_n}, & \text{if } \alpha<2\\ R_n\cdot e^{2y-R_n}, & \text{if } \alpha\ge 2\end{cases},\\
\mathcal{I}_n^{(2)}(y) &= \begin{cases} e^{-(\alpha-\frac{1}{2})R_n+y}, & \text{if } \alpha<3/2\\ R_n\cdot e^{y-R_n}, & \text{if } \alpha\ge 3/2\end{cases}\\
\mathcal{I}_n^{(3)}(y) &= \begin{cases} e^{-(2\alpha-1)R_n+2\alpha y}, & \text{if } \alpha<3/2\\ R_n\cdot e^{(\frac{3}{2}+\alpha)y-(\alpha+\frac{1}{2})R_n}, & \text{if } \alpha\ge 3/2\end{cases}.
\end{aligned}$$

We now need to calculate:

$$\binom{k_n}{2}^{-1}\cdot\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\int_{\mathcal{K}_C(k_n)}\mathbb{E}\left[\left(\sum_{p_1\in\mathcal{P}} h_y(p_1,\mathcal{P}\setminus\{p_1\})\right)\right]\cdot\rho_{\widetilde{\mathbb{H}}}(y,k_n-2)e^{-\alpha y}\,\mathrm{d}y\,\mathrm{d}x.$$

Firstly, note that as $\overline{n}_{\widetilde{\mathbb{H}}}(k_n) = \Theta(1) \cdot n \cdot k_n^{-(2\alpha+1)}$, we have

$$\binom{k_n}{2}^{-1} \cdot \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} = O(1) \cdot \frac{k_n^{2\alpha-1}}{n}.$$

Also, $\mathbb{E}\left[\left(\sum_{p_1 \in \mathcal{P}} h_y(p_1, \mathcal{P} \setminus \{p_1\})\right)\right]$ is given as the sum of $\mathcal{I}_n^{(1)}(y), \mathcal{I}_n^{(2}(y)$ and $\mathcal{I}_n^{(3)}(y)$ (cf. (118)). Setting

$$J_3 = \frac{k_n^{2\alpha-1}}{n} \cdot (M_1 + M_2 + M_3)$$

with

$$M_i = \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(i)}(y) \rho_{\widetilde{\mathbb{H}}}(y, k_n - 1) e^{-\alpha y} \, \mathrm{d}y \, \mathrm{d}x$$

it follows that

$$\binom{k_n}{2}^{-1} \overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left(\sum_{p_1 \in \mathcal{P} \setminus \{(0,y)\}} h_y(p_1, \mathcal{P} \setminus \{p_1\})\right)\right] \rho_{\widetilde{\mathbb{H}}}(y, k_n - 1) e^{-\alpha y} \, \mathrm{d}y \, \mathrm{d}x$$
$$= O(1) \cdot J_3$$

Computing each of the integral separately we obtain, using Lemma 9.4 and the fact that $n = \nu e^{R_n/2}$,

$$M_1 := \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) \rho_{\widetilde{\mathbb{H}}}(y, k_n - 1) e^{-\alpha y} \, \mathrm{d}y \, \mathrm{d}x = O(1) \cdot \begin{cases} \frac{k_n^{7-6\alpha}}{n}, & \text{if } \alpha < 2 \\ R_n^2 \frac{k_n^{3-2\alpha}}{n}, & \text{if } \alpha \geq 2 \end{cases}.$$

$$M_2 := \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) \rho_{\widetilde{\mathbb{H}}}(y, k_n - 1) e^{-\alpha y} dy = O(1) \cdot \begin{cases} e^{-(\alpha-1)R_n} k_n^{-2\alpha+1}, & \text{if } \alpha < 3/2 \\ R_n \frac{k_n^{7-6\alpha}}{n}, & \text{if } \alpha \geq 3/2 \end{cases}$$
$$= \begin{cases} \frac{k_n^{1-2\alpha}}{n^{2(\alpha-1)}}, & \text{if } \alpha < 3/2 \\ R_n \frac{k_n^{7-6\alpha}}{n}, & \text{if } \alpha \geq 3/2 \end{cases}$$

and finally

$$M_3 := \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) \rho_{\widetilde{\mathbb{H}}}(y, k_n - 1) e^{-\alpha y} dy = O(1) \cdot \begin{cases} e^{-(2\alpha-3/2)R_n} k_n^{2\alpha-1}, & \text{if } \alpha < 3/2 \\ R_n e^{-\alpha R_n} k_n^2, & \text{if } \alpha \geq 3/2 \end{cases}$$
$$= O(1) \cdot \begin{cases} \frac{k_n^{2\alpha-1}}{n^{4\alpha-3}}, & \text{if } \alpha < 3/2 \\ R_n \cdot \frac{k_n^2}{n^{2\alpha}}, & \text{if } \alpha \geq 3/2 \end{cases}.$$

Now, we will consider the two cases according to the value of α. Assume first that $1/2 < \alpha < 3/4$. In this case, we want to show that

$$\lim_{n \to \infty} k_n^{4\alpha-2} \cdot J_3 = 0. \tag{119}$$

Using the above expression for J_3, we have

$$k_n^{4\alpha-2} \cdot J_3 = O(1) \cdot \frac{k_n^{6\alpha-3}}{n} \cdot \left(\frac{k_n^{7-6\alpha}}{n} + \frac{k_n^{2(1-\alpha)}}{n^{2(\alpha-1)}} + \frac{k_n^{2\alpha-1}}{n^{4\alpha-3}}.\right)$$

We wish to show that each one of the above three terms is $o(1)$ for $k_n = O(n^{\frac{1}{2\alpha+1}})$. For the first one we have

$$\frac{k_n^{6\alpha-3}}{n} \cdot \frac{k_n^{7-6\alpha}}{n} = \frac{k_n^4}{n^2} = O(1) \cdot n^{\frac{4}{2\alpha+1}-2} = o(1),$$

since $4 < 4\alpha + 2$ when $1/2 < \alpha$. The second one yields:

$$\frac{k_n^{6\alpha-3}}{n} \cdot \frac{k_n^{-2\alpha+1}}{n^{2(\alpha-1)}} = \frac{k_n^{4\alpha-2}}{n^{2\alpha-1}} = O(1)\frac{n^{\frac{4\alpha-2}{2\alpha+1}}}{n^{2\alpha-1}}.$$

We need to show that $\frac{4\alpha-2}{2\alpha+1} < 2\alpha - 1$. Indeed, rearranging this yields, $4\alpha - 2 < 4\alpha^2 - 1$, which is equivalent to $0 < 4\alpha^2 - 4\alpha + 1 = (2\alpha - 1)^2$. This holds for all $\alpha > 1/2$.

Finally, the third one yields:

$$\frac{k_n^{6\alpha-3}}{n} \cdot \frac{k_n^{2\alpha-1}}{n^{4\alpha-3}} = \frac{k_n^{8\alpha-4}}{n^{2(2\alpha-1)}} = \frac{k_n^{4(2\alpha-1)}}{n^{2(2\alpha-1)}}.$$

But $k_n^4 \leq O(1) \cdot n^{\frac{4}{2\alpha+1}} = o(n^2)$, as $2\alpha + 1 > 2$.

For $\alpha \geq 3/4$, we would like to show that

$$\lim_{n\to\infty} k_n \cdot J_3 = 0. \tag{120}$$

Firstly, if $3/4 < \alpha < 3/2$ we have,

$$k_n \cdot J_3 = O(1) \cdot \frac{k_n^{2\alpha}}{n} \cdot \left(\frac{k_n^{7-6\alpha}}{n} + \frac{k_n^{-2\alpha+1}}{n^{2(\alpha-1)}} + \frac{k_n^{2\alpha-1}}{n^{4\alpha-3}}.\right)$$

As above we will deal with the three term of this. For the first one we have

$$\frac{k_n^{2\alpha}}{n}\frac{k_n^{7-6\alpha}}{n} = \frac{k_n^{7-4\alpha}}{n^2} \leq \frac{k_n^4}{n^2} = o(1).$$

The second one yields:

$$\frac{k_n^{2\alpha}}{n}\frac{k_n^{-2\alpha+1}}{n^{2(\alpha-1)}} = \frac{k_n}{n^{2\alpha-1}} \leq \frac{k_n}{n^{1/2}} = O(1)\frac{n^{\frac{1}{2\alpha+1}}}{n^{1/2}} = o(1).$$

Finally, the third one yields:

$$\frac{k_n^{2\alpha}}{n} \cdot \frac{k_n^{2\alpha-1}}{n^{4\alpha-3}} = \frac{k_n^{4\alpha-1}}{n^{2(2\alpha-1)}} = O(1)\frac{n^{\frac{4\alpha-1}{2\alpha+1}}}{n^{2(2\alpha-1)}}.$$

We need to show that $\frac{4\alpha-1}{2\alpha+1} < 2(2\alpha - 1)$, which is equivalent to $8\alpha^2 - 4\alpha - 1 > 0$; this is indeed the case for any $\alpha \geq 3/4$.

For $3/2 \leq \alpha < 2$, it is only M_2 and M_3 that change values. In particular, for any $\alpha \geq 3/2$ we have

$$\frac{k_n}{n} \cdot M_2 = O(1) \cdot R_n \cdot \frac{k_n^{2\alpha}}{n} \cdot \frac{k_n^{7-6\alpha}}{n} = o(1),$$

as above. Also,

$$\frac{k_n}{n} \cdot M_3 = O(1) \cdot R_n \cdot \frac{k_n}{n} \cdot \frac{k_n^2}{n^{2\alpha}} = R_n \cdot \frac{k_n^3}{n^{2\alpha+1}} = o(1),$$

since $k_n = o(n^{1/2})$ (and, therefore, $k_n^3 = o(n^{3/2})$) but $2\alpha + 1 > 2$.

If $\alpha \geq 2$ too, then M_1 changes value and we have

$$\frac{k_n}{n} \cdot M_1 = O(1) \cdot R_n^2 \cdot \frac{k_n}{n} \cdot \frac{k_n^{3-2\alpha}}{n} = \frac{k_n^{4-2\alpha}}{n^2} = o(1),$$

since $\alpha \geq 2$.

The sum of (93) Now, we will first give an upper bound on the term

$$\mathbb{E}\left[\sum_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y_1<K}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right].$$

Using the Campbell-Mecke formula, we write

$$\begin{aligned}
&\mathbb{E}\left[\sum_{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\},\ y_1<K} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right] = \\
&\le \int_{-I_n}^{I_n}\int_0^K\int_{-I_n}^{I_n}\int_0^{R_n} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}e^{-\alpha y_2}e^{-\alpha y_1}dx_2dy_2dx_1dy_1\\
&\le \mu(\mathcal{B}(y))\cdot\int_{-I_n}^{I_n}\int_0^K \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}e^{-\alpha y_1}dx_1dy_1.
\end{aligned}$$

By Lemma 6.9 and a concentration argument, the first factor is

$$\mu(\mathcal{B}(y)) = O(1)e^{y/2}.$$

We bound the second factor using Lemma 2.2. In particular, (7) implies that if $(x_1,y_1)\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}((0,y))$, then because $y_1<K$

$$|x_1-e^{(y+y_1)/2}| \le e^{(y+y_1)/2}\cdot Ke^{y+y_1-R_n} = O(1)e^{(y+y_1)/2}\cdot e^{y-R_n}.$$

Therefore,

$$\int_{-I_n}^{I_n}\int_0^K \mathbb{1}_{\{(x_1,y_1)\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}((0,y))\}}e^{-\alpha y_1}dx_1dy_1 = O(1)\cdot e^{y-R_n}\cdot\int_0^K e^{(y+y_1)/2}e^{-\alpha y_1}dy_1 = O(1)\cdot e^{3y/2-R_n},$$

and hence

$$\mathbb{E}\left[\sum_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y_1<K}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right] = O(1)\cdot e^{2y-R_n}.$$

Now, we integrate this over y:

$$\begin{aligned}
e^{-R_n}\int_{\mathcal{K}_C(k_n)} e^{2y-\alpha y}dydx &= O(1)ne^{-R_n}\int_{I_\varepsilon(k_n)} e^{2y-\alpha y}dy\\
&= O(1)\cdot n^{-1}\int_{I_\varepsilon(k_n)} e^{2y-\alpha y}dy\\
&= O(1)\cdot n^{-1}\cdot\begin{cases} k_n^{2(2-\alpha)(1+\varepsilon)}, & \text{if } \alpha<2\\ \log k_n, & \text{if } \alpha=2\\ 1, & \text{if } \alpha>2\end{cases}.
\end{aligned}$$

We deduce that

$$\binom{k_n}{2}^{-1}\overline{n}_{\widetilde{\mathbb{H}}}(k_n)^{-1}\mathbb{E}\left[\sum_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y_1<K}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathbb{H}\triangle\mathcal{P}}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right] = O(1)\cdot n^{-1}\cdot k_n^{2\alpha-1}.$$

To finish the argument assume first that $1/2<\alpha\le 3/4$. In this case, we will consider

$$k_n^{4\alpha-2}\cdot n^{-2}\cdot k_n^{2\alpha-1+4-2\alpha} = n^{-2}\cdot k_n^{4\alpha+1}.$$

But, $\alpha \leq 3/4$, we have $4\alpha + 1 \leq 4$ and $k_n = o(n^{1/2})$, whereby $k_n^{4\alpha+1} = o(n^2)$.

Now, suppose that $3/4 < \alpha < 2$. Here, we will consider

$$k_n \cdot n^{-2} \cdot k_n^{2\alpha-1+4-2\alpha} = \frac{k_n^2}{n^2} = o(1).$$

When $\alpha \geq 2$, we will bound $\log k_n$ and 1 by k_n and we will consider

$$k_n \cdot n^{-2} \cdot k_n^{2\alpha-1+1} = n^{-2} k_n^{2\alpha+1}.$$

But $k_n = O(1) \cdot n^{\frac{1}{2\alpha+1}}$, whereby $k_n^{2\alpha+1} = O(n)$ and the above term is therefore $o(1)$.

□

9.3 Coupling G_n to G_{Po}

Now that we have established the equivalence of the local clustering function between the Poisson hyperbolic graph G_{Po} and G_{box} the final step is to relate local clustering in G_{Po} to the original hyperbolic random graph G_n. As mentioned in Section 2.4, this is done by moving from $c(k; G_n)$ to the adjusted local clustering $c^*(k; G_n)$ (Lemma 5.1) and then to $c^*(k; G_{\text{Po}})$ (Proposition 5.2). We prove Proposition 5.2 first and will end this section with the proof of Lemma 5.1.

To achieve the results we consider the standard coupling between the binomial and Poisson process. That is, we take a sequence of i.i.d. random elements $z_1, z_2, \dots$ uniformly on the hyperbolic disk of radius R_n, i.e. according to the distribution (1). Then the original hyperbolic random graph consists of the first n points and the poissonized version of the first $N \stackrel{d}{=} \text{Po}(n)$ many points (N is a Poisson random variable with mean n). Under this coupling $N_n(k) = \sum_{j=1}^n \mathbb{1}_{\{\deg_n(z_j)=k\}}$ denotes the (random) number of degree k vertices in the KPKVB hyperbolic random graph model with n vertices and $N_{\text{Po}}(k) = \sum_{j=1}^N \mathbb{1}_{\{\deg_{\text{Po}}(z_j)=k\}}$ denotes the (random) number of degree k vertices in the poissonized KPKVB model.

We start with a result that relates the number of nodes with degree k_n in both models.

Lemma 9.5. *Let $\{k_n\}_{n\geq 1}$ be sequence of natural numbers with $0 \leq k_n \leq n-1$ and $k_n = o(n^{\frac{1}{2\alpha+1}})$. Then*

$$\mathbb{E}\left[|N_n(k_n) - N_{\text{Po}}(k_n)|\right] = o\left(\mathbb{E}\left[N_{\text{Po}}(k_n)\right]\right) = o\left(nk^{-(2\alpha+1)}\right),$$

and furthermore,

$$\mathbb{E}\left[N_{\text{Po}}(k_n)\right] = \Theta\left(k_n^{-(2\alpha+1)} n\right).$$

Tobias: I thought it might be a good idea to show somewhere in the paper that the result of Gugelmann et al. that $N_k = (1 + o(1))p_k n$ a.a.s., extends for all $k \ll n^{1/(2\alpha+1)}$ – where p_k the expression in terms on Γ^+, ξ etc. Gugelmann et al. had show it only until some small power of n. Adding this just means we have to compute the leading constants for the expectation $\mathbb{E}N_k$, and bound the variance, in addition to what we do here. Somehow I had expected we would include that. In fact Markus had prepared a short write up of the variance part of the argument. Should not be too hard to add the rest, no?
We already have the degree sequence for the infinite model somewhere in the paper. **Pim:** Will happen but I honestly had no time to put it in.

Proof. The second result follows directly from the first and Lemma 9.3. Therefore we will only need to prove the first statement of the lemma.

We use the Chernoff concentration result for a Poisson random (126), which states that with probability $n^{-C^2/2}$ the Poisson random variable N with expectation n is contained in the interval $[n - C\sqrt{n \log n}, n + C\sqrt{n \log n}]$. We proceed by bounding the effect on the number of degree k_n vertices by adding or removing $C\sqrt{n \log n}$ many vertices to $G_n(\alpha, \nu$ and from G_{Po}, respectively.

Define the events

$$A_n^{\pm} := \{N \in [n, n \pm C\sqrt{n \log n}]\},$$

Tobias: weird notation **Pim:** Will be adjusted later and let $A_n = A_n^- \cup A_n^+$. Then,

$$\begin{aligned}\mathbb{E}\left[\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right|\right] &\leq \mathbb{E}\left[\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right| \middle| A_n\right] + O\left(n^{1-C^2/2}\right)\\ &= \mathbb{E}\left[\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right| \middle| A_n\right] + o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right)\end{aligned}$$

by choosing C large enough, e.g. $C > \sqrt{2}$. What is left to show is that for any $C > 0$

$$\mathbb{E}\left[\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right| \middle| A_n\right] = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right).$$

Let $V_{\mathbb{H},n}(k_n)$ be the set of degree k_n vertices in the KPKVB graph G_n and $V_{\widetilde{\mathbb{H}},n}(k_n)$ be the set of degree k_n vertices in the Poisson graph G_{Po}. Then

$$\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right| = \left|V_{\mathbb{H},n}(k_n) \Delta V_{\widetilde{\mathbb{H}},n}(k_n)\right|,$$

where $A \Delta B$ denotes the symmetric difference between two sets A and B.

We first consider the case $N \in [n, n + C\sqrt{n \log n}]$, i.e. the event A_n^+. For $z \in V_{\widetilde{\mathbb{H}},n}(k_n) \setminus V_{\mathbb{H},n}(k_n)$, z has degree k_n in the Poisson graph, but not in the binomial graph; **Tobias:** I dont particularly like the name "binomial graph". Maybe just use "standard KPKVB graph" or "fixed number of node graph" **Pim:** I replaced it. so as $N \geq n$, either z or one of its k_n neighbors must have been removed during the transition from the Poisson graph to the binomial graph. On the event A_n, at most $C\sqrt{n \log n}$ many vertices are removed. The probability of hitting a degree k_n vertex or one of its neighbors is at most $\frac{k_n+1}{N} \leq \frac{k_n+1}{n}$. Therefore, by the union bound, the probability that a particular degree k_n vertex of the Poisson graph is removed is upper bounded by $C\sqrt{n \log n}\frac{k_n+1}{n}$. Hence, the expected number of degree k_n vertices that disappear in the transition from the Poisson graph to the KPKVB graph is bounded by

$$\mathbb{E}\left[\left|V_{\widetilde{\mathbb{H}},n}(k_n) \setminus V_{\mathbb{H},n}(k_n)\right| \middle| A_n^{(1)}\right] \leq \mathbb{E}[N_{\mathrm{Po}}(k_n)] C\sqrt{n \log n}\frac{k_n + 1}{n} = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right),$$

Tobias: absolute bar missing above **Pim:** Fixed. where the last line follows since for $\alpha > 1/2$,

$$k_n\sqrt{\frac{\log(n)}{n}} = o\left(n^{\frac{1}{2\alpha+1}}\sqrt{\frac{\log(n)}{n}}\right) = o\left(n^{-\frac{2\alpha-1}{4\alpha+2}}\sqrt{\log(n)}\right) = o(1).$$

For $z \in V_{\mathbb{H},n}(k_n) \setminus V_{\widetilde{\mathbb{H}},n}(k_n)$, z is a degree k_n vertex in the binomial graph, but must have degree $k_n + \ell$ in the Poisson graph (where $1 \leq \ell \leq c\sqrt{n \log n}$). By linearity of expectation the expected number of degree $k_n + \ell$ vertices of the Poisson graph which turn into degree k_n vertices of the binomial graph is equal to the expected number of degree $k_n + \ell$ vertices in the Poisson graph times the probability that a degree $k_n + \ell$ vertex turns into a degree k_n vertex in the transition back, from the Poisson graph to the binomial graph. The probability of choosing uniformly a set of ℓ neighbors of a degree $k_n + \ell$ vertex of the Poisson graph is given by $\frac{k_n+\ell}{N} \cdots \frac{k_n+1}{N-\ell+1}$. **Tobias:** what exactly is the chance experiment here? You are doing ordered, which is not what I would have expected. Maybe that can spelled out. **Pim:** Will ask Markus to add more explanation (it is his proof). Now, using $k_n = o\left(n^{\frac{1}{2\alpha+1}}\right) = o\left(C\sqrt{n \log n}\right)$ for $\alpha > \frac{1}{2}$, $\ell \leq C\sqrt{n \log n}$ and $N - \ell + 1 \geq n$, this probability is bounded from above by $(C+1)^\ell(\frac{\sqrt{n \log n}}{n})^\ell = ((C+1)\sqrt{\frac{\log n}{n}})^\ell$ which is upper bounded by $(\frac{1}{2})^\ell$ for n large enough, i.e. $n \geq n_0$. Therefore, using the geometric series, we conclude

$$\begin{aligned}\mathbb{E}\left[\left|V_{\mathbb{H},n}(k_n) \setminus V_{\widetilde{\mathbb{H}},n}(k_n)\right| \middle| A_n^+\right] &\leq \sum_{\ell=1}^{\sqrt{n \log n}} \mathbb{E}\left[N_{\widetilde{\mathbb{H}},n}(k_n + \ell)\right]\left((c+1)\sqrt{\frac{\log n}{n}}\right)^\ell\\ &\leq \sum_{\ell=1}^{\sqrt{n \log n}} \Theta\left(n(k_n+\ell)^{-2\alpha-1}\right)\left((c+1)\sqrt{\frac{\log n}{n}}\right)^\ell\end{aligned}$$

$$= O\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right) \sum_{\ell=1}^{\sqrt{n\log n}} \left((c+1)\sqrt{\frac{\log n}{n}}\right)^{\ell} = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right),$$

and hence

$$\mathbb{E}\left[\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right| \middle| A_n^+\right] = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right).$$

The case $N \in [n - C\sqrt{n\log n}, n)$ (event A_n^-) follows by similar arguments. As $N < n$, a vertex $z \in V_{\widetilde{\mathbb{H}},n}(k_n) \setminus V_{\mathbb{H},n}(k_n)$ with degree k_n in the Poisson graph must have a strictly larger degree in the binomial graph, i.e. in the transition from the Poisson graph to the binomial graph, a vertex must have been dropped in the neighborhood of z. By the union bound, this can be upper bounded by the number of additional vertices (of the binomial graph) times the probability that a random point falls into the neighborhood of a degree k_n vertex. We obtain

$$\mathbb{E}\left[\left|V_{\widetilde{\mathbb{H}},n}(k_n) \setminus V_{\mathbb{H},n}(k_n)\right| \middle| A_n^{(2)}\right] = O\left(\sqrt{n\log n}\frac{k_n}{n}\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right) = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right)$$

Tobias: Not so fast! You seem assume that dropping into the ball around a **degree** k vertex has order k/n. That has to be argued. It is of course only true if the height of the particular vertex is appropriate. Please correct / provide the relevant details. I guess we may have to appeal to "concentration of heights" here. **Pim:** You are right. I overlooked this when proof reading Markus proof. I did not have time to fix it now, but the conclusion is still true. Will fix it at a later stage. A vertex $z \in V_{\mathbb{H},n}(k_n) \setminus V_{\widetilde{\mathbb{H}},n}(k_n)$ could be one of the additional vertices in the binomial graph or it is a degree $k_n - \ell$ vertex of the Poisson graph which receives exactly ℓ new vertices in its neighborhood in the transition from the Poisson graph to the binomial graph. The probability that one of the additional vertices of the binomial graph (compared to the smaller Poisson graph) has degree k_n is asymptotically of the order $k_n^{-(2\alpha+1)}$ (as can be seen by considering the alternative coupling between the binomial and the Poisson process, where instead of taking $z_1, \dots, z_N$ for the Poisson process, we take the points $z_n, z_{n-1}, \dots, z_{n-N+1}$ (resp. points with index larger than n after we hit z_1): **Tobias:** I don't quite follow the construction here. **Pim:** I will ask Markus to add more explanation. for this graph, we have that the expected number of degree k_n vertices is $\Theta\left(nk^{-2\alpha-1}\right)$, so the probability that a vertex chosen uniformly from the Poisson graph has degree k is $\Theta(k^{-2\alpha-1})$). Therefore, the expected number of additional points with degree k_n is $O\left(\sqrt{n\log n}k_n^{-2\alpha-1}\right) = o\left(nk_n^{-2\alpha-1}\right) = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right)$. The expected number of degree $k_n - \ell$ vertices of the Poisson graph which receive exactly ℓ new vertices can be bounded in a sum resp. series similarly as done for $z \in V_{\mathbb{H},n}(k_n) \setminus V_{\widetilde{\mathbb{H}},n}(k_n)$ in the case $N \geq n$. We therefore conclude that

$$\mathbb{E}\left[\left|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)\right| \middle| A_n^-\right] = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right),$$

which finishes the proof. □

We note that this result together with Lemma 9.3 implies Lemma **??**. Next we prove Proposition 5.2, which states

$$\lim_{n\to\infty} s(k_n)\mathbb{E}\left[\left|c^*(k_n; G_n) - c^*(k_n; G_{\mathrm{Po}})\right|\right] = 0.$$

Proof of Proposition 5.2. First we note that Proposition 5.3, Proposition 5.4, Proposition 5.5 together imply that

$$\mathbb{E}\left[c^*(k_n; G_{\mathrm{Po}})\right] = (1 + o(1))s(k_n)$$

Therefore it suffices to show that

$$\mathbb{E}\left[\left|c^*(k_n; G_n) - c^*(k_n; G_{\mathrm{Po}})\right|\right] = o\left(\mathbb{E}\left[c^*(k_n; G_{\mathrm{Po}})\right]\right).$$

For this we observe that we are looking at the modified clustering coefficient, where we divide by the expected number of degree k_n vertices. As the expected numbers of degree k_n vertices in G_{Po} and G_n are asymptotically equivalent (see Lemma 9.5), it is therefore sufficient to consider the sum of the clustering coefficients of all vertices of degree k_n. Given again the standard coupling between the binomial and Poisson process (as used in the proof of Lemma 9.5), we denote by $V_{\mathbb{H},n}(k_n)$ the set of degree k_n vertices in G_n and by $V_{\widetilde{\mathbb{H}},n}(k_n)$ the set of degree k_n vertices in the graph G_{Po}. If a vertex is contained in both sets, it must have the same degree in both the Poisson and binomial graph, and given the nature of the coupling, the neighbourhoods are therefore the same and hence also their clustering coefficients agree.

The difference of the sum of the clustering coefficients therefore comes from all the clustering coefficients of the symmetric difference $V_{\mathbb{H},n}(k_n)\Delta V_{\widetilde{\mathbb{H}},n}(k_n)$. This symmetric difference is again a Poisson process, whose expected number of points is $\mathbb{E}\left[|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)|\right] = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right)$ by Lemma 9.5. Therefore we have that

$$\mathbb{E}\left[|c^*(k_n;G_n) - c^*(k_n;G_{\mathrm{Po}})|\right] \leq \frac{\mathbb{E}\left[|N_{\mathrm{Po}}(k_n) - N_{\mathrm{Po}}(k_n)|\right]}{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]}\mathbb{E}\left[c^*(k_n;G_{\mathrm{Po}})\right] = o\left(1\right)\mathbb{E}\left[c^*(k_n;G_{\mathrm{Po}})\right],$$

which finishes the proof. □

We end this section with the proof of Lemma 5.1, whose statement is

$$\mathbb{E}\left[|c^*(k_n;G_n) - c(k_n;G_n)|\right] = o\left(s(k_n)\right).$$

Proof of Lemma 5.1. Let $0 < \delta < 1$ and define the event

$$A_n = \left\{|N_{\mathrm{Po}}(k_n) - \mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]| \leq \mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]^{\frac{1+\delta}{2}}\right\}.$$

Since $N_{\mathrm{Po}}(k_n) = \sum_{i=1}^n \mathbb{1}_{\{\deg_{\mathrm{Po}}(i)=k_n\}}$ it follows from Lemma D.1, with $c = \mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]^{-\frac{1-\delta}{2}}$, that

$$\mathbb{P}\left(A_n\right) \geq 1 - O\left(e^{-\frac{\mathbb{E}[N_{\mathrm{Po}}(k_n)]^\delta}{2}}\right) = 1 - O\left(e^{-\frac{n^\delta k_n^{-\delta(2\alpha+1)}}{2}}\right), \tag{121}$$

where the last part is due to Lemma 9.5.

On the event A_n

$$\left|\frac{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]}{N_{\mathrm{Po}}(k_n)} - 1\right| \leq \frac{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]^{\frac{1+\delta}{2}}}{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] + \mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]^{\frac{1+\delta}{2}}} \leq \mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]^{-\frac{1-\delta}{2}}.$$

Therefore we have

$$\begin{aligned}\mathbb{E}\left[|c^*(k_n;G_n) - c(k_n;G_n)|\right] &\leq \mathbb{E}\left[|c^*(k_n;G_n) - c(k_n;G_n)|\,\mathbb{1}_{\{A_n\}}\right] + O\left(1 - \mathbb{P}\left(A_n\right)\right)\\ &= \mathbb{E}\left[c^*(k_n;G_n)\left|\frac{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]}{N_{\mathrm{Po}}(k_n)} - 1\right|\mathbb{1}_{\{A_n\}}\right] + O\left(e^{-\frac{n^\delta k_n^{-\delta(2\alpha+1)}}{2}}\right)\\ &\leq \mathbb{E}\left[c^*(k_n;G_n)\right]\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]^{-\frac{1-\delta}{2}} + O\left(e^{-\frac{n^\delta k_n^{-\delta(2\alpha+1)}}{2}}\right).\end{aligned}$$

The second term is clearly $o\left(s(k_n)\right)$. The first term is clearly $o\left(\mathbb{E}\left[c^*(k_n;G_n)\right]\right)$ which is $o\left(s(k_n)\right)$ by Proposition 5.2. □

Tobias: Maybe a "conclusion and further work" section here, before the references and appendix?
Also, appendix usually comes after references in my experience.

Acknowledgments We are grateful to Dmitri Krioukov for pointing us to the problem of local clustering and helpful discussions.

References

[1] Mohammed Amin Abdullah, Nikolaos Fountoulakis, and Michel Bode. Typical distances in a geometric model for complex networks. *Internet Mathematics*, aug 2017.

[2] Thomas Bläsius, Tobias Friedrich, and Anton Krohmer. Cliques in hyperbolic random graphs. *Algorithmica*, 80(8):2324–2344, 2018.

[3] Michel Bode, Nikolaos Fountoulakis, and Tobias Müller. On the largest component of a hyperbolic model of complex networks. *The Electronic Journal of Combinatorics*, 22(3):3–24, 2015.

[4] Elisabetta Candellero and Nikolaos Fountoulakis. Clustering and the hyperbolic geometry of complex networks. *Internet Mathematics*, 12(1-2):2–53, 2016.

[5] Kimberly Claffy, Young Hyun, Ken Keys, Marina Fomenkov, and Dmitri Krioukov. Internet mapping: from art to science. In *Conference For Homeland Security, 2009. CATCH'09. Cybersecurity Applications & Technology*, pages 205–211. IEEE, 2009.

[6] Brian Davies. *Integral transforms and their applications*, volume 41. Springer Science & Business Media, 2012.

[7] Arthur Erdélyi, Wilhelm Magnus, Fritz Oberhettinger, Francesco G Tricomi, et al. Higher transcendental functions, vol. 1, 1953.

[8] Nikolaos Fountoulakis and Tobias Müller. Law of large numbers for the largest component in a hyperbolic model of complex networks. *The Annals of Applied Probability*, 28(1):607–650, 2018.

[9] David Freedman. Another note on the borel–cantelli lemma and the strong law, with the poisson approximation as a by-product. *The Annals of Probability*, pages 910–925, 1973.

[10] Izrail Solomonovich Gradshteyn and Iosif Moiseevich Ryzhik. *Table of Integrals, Series, and Products*. Elsevier, 2015.

[11] Luca Gugelmann, Konstantinos Panagiotou, and Ueli Peter. Random hyperbolic graphs: degree sequence and clustering. In *International Colloquium on Automata, Languages, and Programming*, pages 573–585. Springer, 2012.

[12] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010.

[13] Yudell L Luke. *Mathematical functions and their approximations*. Academic Press, 2014.

[14] CS Meijer. On the G-function. I-VIII. In *Nederl. Akad. Wetensch., Proc.*, volume 49, 1946.

[15] Mathew Penrose. *Random geometric graphs*. Number 5. Oxford university press, 2003.

[16] Clara Stegehuis, Remco van der Hofstad, and Johan SH van Leeuwaarden. Scale-free network clustering in hyperbolic and other random graphs. *arXiv*, 2018.

A Meijer's G-function

Recall that $\Gamma(z)$ denotes the Gamma function. Let p, q, m, ℓ be four integers satisfying $0 \leq m \leq q$ and $0 \leq \ell \leq p$ and consider two sequences $\mathbf{a}_p = \{a_1, \dots, a_p\}$ and $\mathbf{b}_q = \{b_1, \dots, b_q\}$ of reals such that $a_i - b_j$ is not a positive integer for all $1 \leq i \leq p$ and $1 \leq j \leq q$ and $a_i - a_j$ is not an integer for all distinct indices $1 \leq i, j \leq p$. Then, with ι denoting the complex unit, Meijer's G-Function [14] is defined as

$$G^{m,\ell}_{p,q}\left(z\middle|\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right) = \frac{1}{2\pi\iota}\int_L \frac{\prod_{j=1}^m \Gamma(b_j - t)\prod_{j=1}^{\ell}\Gamma(1-a_j+t)}{\prod_{j=m+1}^q \Gamma(1-b_j+t)\prod_{j=\ell+1}^p \Gamma_n(a_j - t)} z^t \,\mathrm{d}t, \tag{122}$$

where the path L is an upward oriented loop contour which separates the poles of the function $\prod_{j=1}^m \Gamma(b_j - t)$ from those of $\prod_{j=1}^n \Gamma(1-a_j+t)$ and begins and ends at $+\infty$ or $-\infty$.

The Meijer's G-Function is of a very general nature and has relation to many known special functions such as the Gamma function and the generalized hypergeometric function. For more details, such as many identities for $G^{m,\ell}_{p,q}\left(z\middle|\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right)$ see [10, 13].

For our purpose we need the following identity which follows from an Mellin transform operation.

Lemma A.1. *For any $a \in \mathbb{R}$ and $\xi, s > 0$,*

$$\Gamma^+(-a-1, \xi/s) = G^{2,0}_{1,2}\left(\frac{\xi}{s}\middle|\begin{matrix}1\\ -a-1, 0\end{matrix}\right)$$

Proof. Let $x > 0$ and $q \in \mathbb{R}$ and note that as the Γ-function is the Mellin transform of e^{-x}, by the inverse Mellin transform formula, we have $e^{-x} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(p)x^{-p}dp$ for $c > 0$ (see [6, p.196]). Applying the change of variable $p(r) = q - r$ yields $e^{-x} = \frac{1}{2\pi i}\int_{c+q-i\infty}^{c+q+i\infty}\Gamma(q-r)x^{r-q}dr$, then multiplying both sides with $-x^{q-1}$ gives $-x^{q-1}e^{-x} = -\frac{1}{2\pi i}\int_{c+q-i\infty}^{c+q+i\infty}\Gamma(q-r)x^{r-1}dr$. Now, integrating both sides gives $\int_x^\infty t^{q-1}e^{-t}dt = \frac{1}{2\pi i}\int_{c+q-i\infty}^{c+q+i\infty}\frac{\Gamma(q-r)}{-r}x^r dr$. On the left-hand side is the incomplete gamma function and on the right-hand side with using $-r = \frac{\Gamma(1-r)}{\Gamma(-r)}$ is the Meijer G-function, i.e. $\Gamma^+(q,x) = G^{2,0}_{1,2}\left(\begin{matrix}1\\ q, 0\end{matrix}\middle| x\right)$. The claim follows by plugging in $q = -a-1$ and $x = \frac{\xi}{s}$. $\square$

B Incomplete Beta function

Here we derive the asymptotic behavior for the function $B^-(1-z; 2\alpha, 3-4\alpha)$ as $z \to 0$, which is used to analyze the asymptotic behavior of $P(y)$, see Section 3.3.

Lemma B.1. *We have the following asymptotic results for $B^-(1-z; 2\alpha, 3-4\alpha)$*

1. *For $1/2 < \alpha < 3/4$*
$$\lim_{z\to 0} B^-(1-z, 2\alpha, 3-4\alpha) = B(2\alpha, 3-4\alpha).$$

2. *When $\alpha = 3/4$,*
$$\lim_{z\to 0}\frac{B^-(1-z, 2\alpha, 3-4\alpha)}{\log(z)} = -1.$$

3. *For $\alpha > 3/4$,*
$$\lim_{z\to 0} z^{4\alpha-3}B^-(1-z, 2\alpha, 3-4\alpha) = \frac{1}{4\alpha-3}.$$

Proof. We use the hypergeometric representation of the incomplete Beta function,

$$B^-(x, a, b) = \frac{x^a}{2a}F(a, 1-b, a+1, x),$$

where F denote the hypergeometric function [REF]. In particular we have that

$$B^-(1-z; 2\alpha, 3-4\alpha) = \frac{(1-z)^{2\alpha}}{2\alpha}F(2\alpha, 4\alpha-2, 2\alpha+1, 1-z).$$

The behavior of $F(a, b, c, 1-z)$ as $z \to 0$ depend on the real part of the sum of $c-a-b$ and whether $c = a+b$ [REF]. Since in our case a, b, c will be real it only depends on the sum of $c-a-b$. For $c-a-b > 0$ we have

$$\lim_{z\to 0} F(a, b, c, 1-z) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \tag{123}$$

if $c = a+b$ then

$$\lim_{z\to 0} \frac{F(a, b, c, 1-z)}{\log(z)} = -\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, \tag{124}$$

and finally, when $c-a-b < 0$

$$\lim_{z\to 0} \frac{F(a, b, c, 1-z)}{z^{c-a-b}} = \frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}. \tag{125}$$

In our case we have,

$$B^-(1-z; 2\alpha, 3-4\alpha) = \frac{(1-z)^{2\alpha}}{2\alpha} F(a, b, c, 1-z),$$

with $a := 2\alpha$, $b := 4\alpha - 2$ and $c := 2\alpha + 1$. Therefore,

$$c - a - b = 2\alpha + 1 - 2\alpha - (4\alpha - 2) = 3 - 4\alpha.$$

Now if $\alpha < 3/4$ then $c-a-b > 0$ and hence

$$\lim_{z\to 0} B^-(1-z; 2\alpha, 3-4\alpha) = \frac{1}{2\alpha}\frac{\Gamma(2\alpha+1)\Gamma(3-4\alpha)}{\Gamma(1)\Gamma(3-2\alpha)} = \frac{\Gamma(2\alpha)\Gamma(3-4\alpha)}{\Gamma(3-2\alpha)} = B(2\alpha, 3-4\alpha),$$

where we used that $\Gamma(2\alpha+1) = 2\alpha\Gamma(2\alpha)$.

When $\alpha = 3/4$ then $c-a-b = 0$ and therefore (124) implies that

$$\lim_{z\to 0} \frac{B^-(1-z; 2\alpha, 3-4\alpha)}{\log(z)} = -\frac{1}{2\alpha}\frac{\Gamma(6\alpha-2)}{\Gamma(2\alpha)\Gamma(4\alpha-2)} = -\frac{\Gamma(5/2)}{\frac{3}{2}\Gamma(3/2)} = -1.$$

Finally, when $\alpha > 3/4$, $c-a-b = 3-4\alpha < 0$ and using (125) we get

$$\lim_{z\to 0} z^{4\alpha-3} B^-(1-z, 2\alpha, 3-4\alpha) = \frac{1}{2\alpha}\frac{\Gamma(2\alpha+1)\Gamma(4\alpha-3)}{\Gamma(2\alpha)\Gamma(4\alpha-2)} = \frac{\Gamma(4\alpha-3)}{\Gamma(4\alpha-2)} = \frac{1}{4\alpha-3}.$$

□

C Some results on functions

Lemma C.1. *For any $0 < \lambda < 1$ there exists a $K > 0$, such that for all $0 < x \le (1-\lambda)2$*

$$\frac{1}{2}\arccos(1-x)\left(1 - \frac{(1+\sqrt{2})x}{1+x}\right) \le \frac{x}{\sqrt{1-(1-x)^2}} \le \frac{1}{2}\arccos(1-x)\left(1 + \frac{(K+1)x}{1-x}\right).$$

In particular, as $x \to 0$,

$$\frac{x}{\sqrt{1-(1-x)^2}} \sim \frac{1}{2}\arccos(1-x).$$

Proof. First we observe that for all $0 < x < 2$

$$0 < \sqrt{2x}\left(1 - \frac{x}{\sqrt{8}}\right) \le \arccos(1-x) \le \sqrt{2x}\left(1 + \frac{x}{\sqrt{8}}\right)$$

while for every $0 < \lambda < 1$, there exists a $K > 0$ such that for all $0 < x \leq (1-\lambda)2$,

$$0 < \frac{1}{\sqrt{2x}}\left(1 - \frac{x}{2}\right) \leq \frac{1}{\sqrt{1-(1-x)^2}} \leq \frac{1}{\sqrt{2x}}\left(1 + Kx\right).$$

It then follows that for all $0 < x \leq (1-\lambda)2$,

$$\begin{aligned}
\frac{x}{\sqrt{1-(1-x^2)}} &\leq \frac{1}{2}\sqrt{2x}\left(1 + K\frac{x}{\sqrt{2}}\right) \\
&\leq \frac{1}{2}\arccos(1-x)\frac{1+Kx}{1-\frac{x}{\sqrt{8}}} \\
&\leq \frac{1}{2}\arccos(1-x)\left(1 + \frac{(K+1)x}{1-x}\right),
\end{aligned}$$

and

$$\begin{aligned}
\frac{x}{\sqrt{1-(1-x^2)}} &\geq \frac{1}{2}\sqrt{2x}\left(1 - \frac{x}{2}\right) \\
&\geq \frac{1}{2}\arccos(1-x)\frac{1-\frac{x}{2}}{1+\frac{x}{\sqrt{8}}} \\
&\geq \frac{1}{2}\arccos(1-x)\left(1 - \frac{(1+\sqrt{2})x}{1+x}\right),
\end{aligned}$$

which finishes the proof. □

D Some results for random variables

We start with the following concentration result which follows from [9, Theorem 4], together with the note directly after it.

Lemma D.1. *Let X_n be a sum of n, possibly dependent, indicators and $c > 0$. Then*

$$\mathbb{P}\left(|X_n - \mathbb{E}\left[X_n\right]| > c\mathbb{E}\left[X_n\right]\right) \leq 2e^{-\frac{c^2\mathbb{E}[X_n]^2}{2}}.$$

Let $H(x) = x\log(x) - x + 1$. Then by a Chernoff bound, see for instance [15, Lemma 1.2],

$$\begin{aligned}
\mathbb{P}\left(\mathrm{Po}(\lambda) \geq k\right) &\leq e^{-\lambda H(k/\lambda)} \quad \text{for all } k \geq \lambda \\
\mathbb{P}\left(\mathrm{Po}(\lambda) \leq k\right) &\leq e^{-\lambda H(k/\lambda)} \quad \text{for all } k \leq \lambda.
\end{aligned}$$

Note that $H(x) \leq (x-1)^2/2$ for all $0 \leq x \leq 1$. Therefore,

$$\mathbb{P}\left(|\mathrm{Po}(\lambda) - \lambda| \geq x\right) \geq 1 - e^{-\lambda H(1-x/\lambda)} - e^{-\lambda H(1+x/\lambda)}$$

$\mathbb{P}(X > k) \leq e^{-\mu H(\frac{k}{\mu})}$ and for $k < \mu$, $\mathbb{P}(X < k) \leq e^{-\mu H(\frac{k}{\mu})}$, where $H(x) = x\ln x - x + 1$, [15]), it follows that $\mathbb{P}(N \in [n - c\sqrt{n\log n}, n + c\sqrt{n\log n}]) \geq 1 - e^{-nH(\frac{n-c\sqrt{n\log n}}{n})} - e^{-nH(\frac{n+c\sqrt{n\log n}}{n})} \geq 1 - 2e^{-n\frac{c^2 n\log n}{n^2}} = 1 - 2e^{-c^2\log n} = 1 - 2n^{-c^2}$ (where we have used that $H(x) = (x-1)^2$ for x close to 1)

By a Chernoff bound we have

$$\mathbb{P}\left(|\mathrm{Po}(\lambda) - \lambda| \geq x\right) \leq 2e^{-\frac{x^2}{2(\lambda+x)}}. \tag{126}$$

In particular, if $\lambda_n \to \infty$, then, for any $0 < \varepsilon < 1$,

$$\mathbb{P}\left(|\mathrm{Po}(\lambda_n) - \lambda_n| \geq \lambda_n^{\frac{1+\varepsilon}{2}}\right) \leq 2e^{-\frac{\lambda_n^{\varepsilon}}{2\left(1+\lambda_n^{-(1-\varepsilon)/2}\right)}} = O\left(e^{-\lambda_n^{\varepsilon}}\right).$$