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.

May 28, 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, the parameter 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 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 $\gamma(k)$ which we give 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 show that $\gamma(k)$ scales differently, as k grows, for different ranges of α. More precisely, there exists constants $Q_{\alpha,\nu}$ depending on α and ν, such that as $k \to \infty$, $\gamma(k) \sim Q_{\alpha,\nu} k^{2-4\alpha}$ if $\frac{1}{2} < \alpha < \frac{3}{4}$, $\gamma(k) \sim Q_{\alpha,\nu} \log(k) k^{-1}$ if $\alpha = \frac{3}{4}$ and $\gamma = Q_{\alpha,\nu} k^{-1}$ when $\alpha > \frac{3}{4}$. These results contradict a prediction of Krioukov et al., which stated that the limiting values $\gamma(k)$ should always scale with k^{-1} as we let k grow.

*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

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 [24] 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.

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), such as 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 [33]. 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 $u \in \mathbb{H}$ is represented as (r, θ), where r is the hyperbolic distance between u and the origin O and θ as the angle between the line segment Ou 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 $u_i = (r_i, \theta_i)$, 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. [24], and proved rigorously by Gugelmann et al. [20], the degree sequence of the KPKVB model follows a power-law with exponent $2\alpha + 1$. Gugelmann et al. [20] 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 $\mathbb{P}(E_n) \to 1$ as $n \to \infty$. Here the events E_n might be defined in different probability spaces, in which case $\mathbb{P}$ will the corresponding probability measure.

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 [2] it is shown 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 [21] and [18] a.a.s. polylogarithmic upper and lower bounds

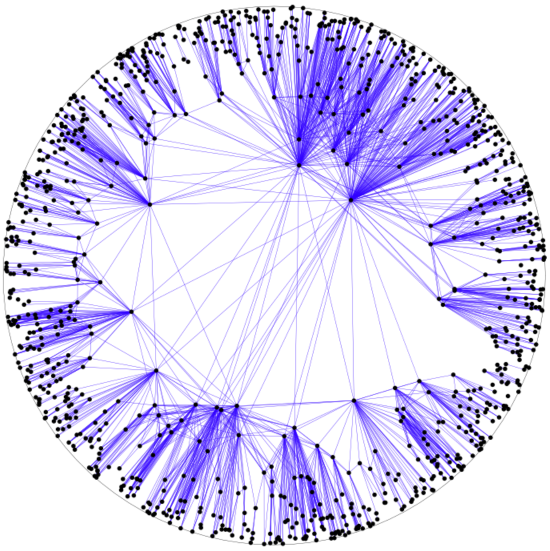

Figure 1: Simulation $G(n;\alpha,\nu)$ with $\alpha = 0.9$, $\nu = 0.2$ and $n = 5000$.

on the graph diameter of the largest component are shown, and in [29], 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 [7] and of the first and third authors [16] has established the "threshold for a giant component": if $\alpha < 1$ then there is a unique component of size linear in n no matter how small ν (i.e. the average degree); if $\alpha > 1$ all components are sublinear no matter the value of ν; and if $\alpha = 1$ 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 linearly sized component (all of these statements holding a.a.s.). Whether or not there is a giant component if $\alpha = 1$ and $\nu = \nu_c$ remains an open problem. In [21] and [22], Kiwi and Mitsche considered the size of the second largest component and showed that for $\alpha \in (1/2, 1)$, 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 [8] 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 [5] studied the size of the largest clique as well as the number of cliques of a given size. Boguña et al. [9] and Bläsius et al. [6] considered fitting the KPKVB model to data using maximum likelihood estimation. Kiwi and Mitsche [23] studied the spectral gap and related properties, and Bläsius et al.[4] considered the tree-width and related parameters of the KPKVB model. Recently Owada and Yogeshwaran [30] 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 this work we study the clustering coefficient in the KPKVB model. 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 of 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 [10] and for the evolution of graphs on more general spaces with negative curvature by the first author in [15].

We will study the other notion of clustering, the one which is also considered by Krioukov et al. [24] and Gugelmann et al. [20]. 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:

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

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 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. [20], 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 non-negative integer k the value

$$c(k;G) := \begin{cases} \dfrac{1}{N(k)} \displaystyle\sum_{\substack{v \in V(G), \\ \deg(v) = k}} c(v), & \text{if } N(k) \geq 1, \\ 0, & \text{else,} \end{cases} \tag{2}$$

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 [24, 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 [24]) in experiments on the infrastructure of the Internet obtained in [12]. 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, nor how $c(k; G(n;\alpha,\nu))$ 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 $k \ll n^{1/(2\alpha+1)}$. Proposition 1.4 clarifies the asymptotic behavior 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. [24].

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 $B^-(x;a,b) := \int_0^x u^{a-1}(1-u)^{b-1}\mathrm{d}u$ for 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 [14, p.255 Equation (2)], and let $G_{p,q}^{m,\ell}\left(z\middle|\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right)$ denote Meijer's G-Function [27], 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

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

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

A plot of γ can be found in Figure 2.

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.

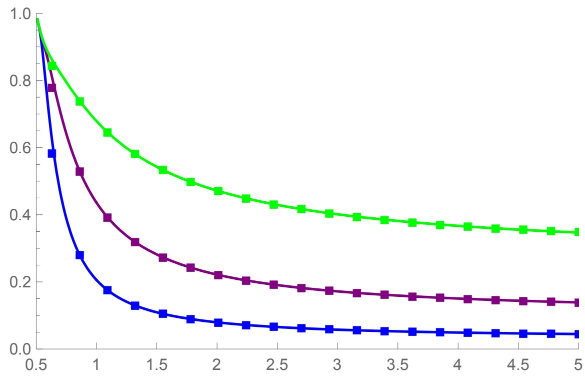

Figure 2: Plot of γ for α varying from 0.5 to 5 on the horizontal axis and for $\nu = \frac{1}{2}$ (blue), $\nu = 1$ (purple), $\nu = 2$ (green); simulations (squares in corresponding colour) with $n = 10000$ and 100 repetitions.

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

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

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 $(k_n)_{n\geq 1}$.

Theorem 1.3. *Let $\alpha > \frac{1}{2}, \nu > 0$ be fixed and let k_n be a sequence of non-negative integers 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 $\gamma(\cdot)$ is as in Theorem 1.2. In particular

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\frac{c(k_n; G_n)}{\gamma(k_n)} - 1\right|\right] = 0.$$

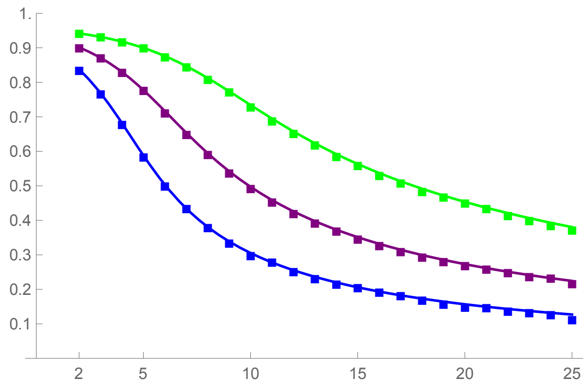

Figure 3: Plot $\gamma(k)$ for k varying from 2 to 25 on the horizontal axis, for $\alpha = 0.8$ and $\nu = \frac{1}{2}$ (blue), $\nu = 1$ (purple), $\nu = 2$ (green); simulations (squares in corresponding colour) with $n = 10000$ and 100 repetitions.

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

Note that Theorem 1.3 implies that the clustering function of the KPKVB model scales as $\gamma(k)$ as k grows, whose scaling is given in the above result. In particular, this contradicts the scaling conjectured in [24] for $\alpha \leq \frac{3}{4}$, and confirms it only for $\alpha > \frac{3}{4}$.

We remark that simultaneously and independently Stegehuis, van der Hofstad and van Leeuwaarden [32] used a completely different technique to obtain a similar, though less detailed, result on the $k \to \infty$ scaling of the clustering function in the KPKVB model.

1.2 Observations

There are a few other observation to be made regarding our results.

Uniform convergence. Our results for the local clustering function in particular imply uniform convergence of $c(k; G_n)$ for all $2 \leq k \leq a_n$ where $a_n \ll n^{\frac{1}{2\alpha+1}}$. To see this let

$$b_n = \arg \max_{2 \leq k \leq a_n} \mathbb{E}\left[\left| \frac{c(k; G_n)}{\gamma(k)} - 1 \right| \right].$$

Then $b_n \leq a_n \ll n^{\frac{1}{2\alpha+1}}$ and therefore by Theorem 1.3

$$\lim_{n\to\infty} \max_{2 \leq k \leq a_n} \mathbb{E}\left[\left| \frac{c(k; G_n)}{\gamma(k)} - 1 \right| \right] = \lim_{n\to\infty} \mathbb{E}\left[\left| \frac{c(b_n; G_n)}{\gamma(b_n)} - 1 \right| \right] = 0.$$

Near maximum scaling for k_n. Our results for the clustering function in the KPKVB model are valid for any sequence of non-negative integers $k_n \to \infty$ such that $k_n \ll n^{\frac{1}{2\alpha+1}}$. Although one would like to have results for any sequence $k_n \leq n - 1$, it turns out that $n^{1/(2\alpha+1)}$ is the near maximum scaling for which Theorem 1.1 can be true. To see why this is the case note that by definition of the local clustering function (2) we have that $c(k_n; G_n) = 0$ if $N_n(k_n) = 0$, where $N_n(k_n)$ denotes the number of vertices in a KPKVB graph $G(n; \alpha, \nu)$ with degree k_n. It follows by Markov's inequality that for any positive function $f : \mathbb{R}_+ \to \mathbb{R}_+$

$$\begin{aligned}\mathbb{E}\left[\left|\frac{c(k_n; G_n)}{f(k_n)} - 1\right|\right] &\geq \mathbb{E}\left[\left|\frac{c(k_n; G_n)}{f(k_n)} - 1\right| \mathbb{1}_{\{N_n(k_n)=0\}}\right] \\ &\geq \mathbb{P}\left(N_n(k_n) = 0\right) \geq 1 - \mathbb{E}\left[N_n(k_n)\right].\end{aligned}$$

We shall later establish (see Lemma 9.5) that $\mathbb{E}[N_n(k_n)]$ scales as $nk_n^{-(2\alpha+1)}$. Therefore if k_n is such that $k_n^{-(2\alpha+1)}n$ tends to zero as $n \to \infty$ we have that

$$\lim_{n\to\infty} \mathbb{E}\left[N_n(k_n)\right] = 0$$

and hence

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\frac{c(k_n; G_n)}{f(k_n)} - 1\right|\right] \geq 1 - \lim_{n\to\infty} \mathbb{E}\left[N_n(k_n)\right] = 1 \neq 0,$$

for any positive function f. This implies that we cannot expect a result like that of Theorem 1.3 to hold as soon as $k_n \gg n^{\frac{1}{2\alpha+1}}$.

Transition in scaling at $\alpha = 3/4$. It follows from Proposition 1.4 that there is a transition in the scaling of the local clustering function at $\alpha = 3/4$. This corresponds to an exponent $5/2$ for the probability mass function of the degree distribution. This transition is different from those often observed for networks with scale-free degree distributions, where transitions occur at integer values of the exponent. At this point, it is unclear what the underlying reason is for the appearance of the transition at this half exponent. Interestingly, a similar transition point has also been observed for both majority vote models [11] and flocking dynamics [28] on networks with scale-free degree degree distributions.

1.3 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 [16] 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. The main reason for this is that we push the possible scaling of k_n to its maximum and hence a great deal of work is needed to properly control the arising error terms and make sure these are of smaller order than $\gamma(k_n)$.

Finally, the Appendix includes some technical results on Meijer's G-function, Chernoff bounds for Poisson and Binomial random variables and the code used for simulations.

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 [16]. 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

$$f(x,y) = \frac{\alpha\nu}{\pi} e^{-\alpha y} \cdot \mathbb{1}_{\{y>0\}}. \tag{3}$$

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 $p \in \mathbb{R} \times (0,\infty)$, we write $\mathcal{B}_\infty(p)$ to denote the *ball* around p, i.e.

$$\mathcal{B}_\infty(p) = \{p' \in \mathbb{R} \times (0,\infty) : |x - x'| \le e^{\frac{y+y'}{2}}\}. \tag{4}$$

With this notation we then have that $\mathcal{B}_\infty(p) \cap \mathcal{P}$ denotes the set of neighbours 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 $\mu(S) = \int_S f(x,y)\, \mathrm{d}x\, \mathrm{d}y$.

2.2 The finite box model G_{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* $G_{\mathrm{box}} := G_{\mathrm{box}}(n;\alpha,\nu)$ has vertex set $\mathcal{V}_{\mathrm{box}} := \mathcal{P} \cap \mathcal{R}$ and edge set such that

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

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 $\mathcal{V}_{\mathrm{box}}$, 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 $\mathcal{B}_{\mathrm{box}}(p)$ as

$$\mathcal{B}_{\mathrm{box}}(p) = \left\{p' \in \mathcal{R} : |x - x'|_{\pi e^{R/2}} \le e^{\frac{y+y'}{2}}\right\}. \tag{5}$$

2.3 The Poissonized KPKVB model G_{Po}

Imagine that we have an infinite supply of i.i.d. points $u_1, u_2, \ldots$ 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 $u_1,\ldots,u_n$ 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 $u_1,\ldots,u_N$ 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,

$$g(r,\theta) = \frac{\alpha \sinh(\alpha r)}{2\pi(\cosh(\alpha R) - 1)} \cdot 1_{\{0 \le r \le R, -\pi < \theta \le \pi\}}. \tag{6}$$

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 [16] 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 ([16, Lemma 27]). *Let $\mathcal{V}_{\mathrm{Po}}$ denote the vertex set of $G_{\mathrm{Po}}(n;\alpha,\nu)$ and $\mathcal{V}_{box}$ the vertex set of $G_{\mathrm{box}}(n;\alpha,\nu)$. Define the map $\Psi : [0,R]\times(-\pi,\pi] \to \mathcal{R}$ by*

$$\Psi(r,\theta) = \left(\theta\frac{e^{R/2}}{2}, R-r\right). \tag{7}$$

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

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 $\Psi^{-1}(p)$ for $p \in \mathcal{R}$, i.e.

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

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 $r+r' \leq R$ or $r+r' > R$ and

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

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

$$|x-x'|_{\pi e^{R/2}} \leq \Phi(y,y') := \frac{1}{2}e^{R/2}\arccos\left(\frac{\cosh(R-y)\cosh(R-y')-\cosh R}{\sinh(R-y)\sinh(R-y')}\right). \tag{8}$$

The following lemma provides useful bounds on the function $\Phi(r,r')$. Note that in [16] the function Φ is written in terms of where $r := R-y, r' := R-y'$.

Lemma 2.2 ([16, 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 $y+y' < R$ we have that*

$$e^{\frac{1}{2}(y+y')} - Ke^{\frac{3}{2}(y+y')-R} \leq \Phi(y,y') \leq e^{\frac{1}{2}(y+y')} + Ke^{\frac{3}{2}(y+y')-R}, \tag{9}$$

Moreover:

$$\Phi(y,y') \geq e^{\frac{1}{2}(y+y')} \quad \text{if} \quad y,y' > K. \tag{10}$$

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

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

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

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

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

Remark 2.1 (Notion convention for points). *We will often be working with the finite box graph G_{Po} or the infinite graph G_∞, whose nodes are points in $\mathbb{R}\times\mathbb{R}_+$. For any point $p \in \mathbb{R}\times\mathbb{R}_+$ we will always use $p = (x,y)$. When considering different points $p, p' \in \mathbb{R}\times\mathbb{R}_+$, we will use primed coordinates to refer to p', i.e. $p' = (x',y')$, and similar with subscripts, i.e. $p_i = (x_i, y_i)$.*

2.5 The Campbell-Mecke formula

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 a specific incarnation, which follows from the Palm theory of Poisson point processes on metric spaces, see [25]. For this consider a Poisson point process $\mathcal{P}$ on some metric space $\mathcal{M}$ with density μ and let $\mathcal{N}$ denote the set of all possible point configurations in $\mathcal{M}$, equipped with the sigma algebra of the process $\mathcal{P}$. Then, for any natural number k and measurable function $h : \mathbb{R}^k \times \mathcal{N} \to \mathbb{R}$,

$$\mathbb{E}\left[\sum_{p_1,\dots,p_k \in \mathcal{P}}^{\neq} h(p_1,\dots,p_k,\mathcal{P})\right] = \int_{\mathcal{M}} \cdots \int_{\mathcal{M}} \mathbb{E}\left[h(x_1,\dots,x_k,\mathcal{P}\right] \mathrm{d}\mu(x_1)\dots\mathrm{d}\mu(x_k), \tag{11}$$

where the sum is over all distinct point $p_1,\dots,p_k \in \mathcal{P}$.

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 additional 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 intuition 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, using $\alpha > \frac{1}{2}$,

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

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

$$\rho(p, k) := \mathbb{P}\left(\mathrm{Po}(\mu(p)) = k\right), \tag{12}$$

where $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with expectation 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, for $k \in \mathbb{N}_0$:

$$p_k := \mathbb{P}(D = k) = \int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\, dy. \tag{13}$$

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. [20] gave for the limiting degree distribution of $G(n;\alpha,\nu)$. Using Stirling's approximation to the gamma function, we find that

$$p_k \sim 2\alpha\xi^{2\alpha}k^{-(2\alpha+1)} \quad \text{as } k \to \infty. \tag{14}$$

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

$$\int_0^\infty e^{\beta y}\rho(y,k)\alpha e^{-\alpha y}\, dy \sim 2\alpha\xi^{2(\beta+\alpha)}k^{-2(\beta+\alpha)-1}. \tag{15}$$

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 now *define*

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

(Where we take the expectation over both the Poisson point process $\mathcal{P}$ and $y \overset{\mathrm{d}}{=} \exp(\alpha)$, independently 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 first picking $N \overset{\mathrm{d}}{=} \mathrm{Po}(\mu(y_0))$ and then taking 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 intensity function of the Poisson point process, but set to zero outside of $\mathcal{B}_\infty((0,y_0))$ and re-normalized 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 re-normalized probability density function. This shows that, for every $k \geq 2$:

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

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 will write

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

with u_1, u_2 as above.

We now observe that

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

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,

$$\gamma(k) = \frac{1}{p_k} \cdot \int_0^\infty P(y_0)\rho(y_0, k)\alpha e^{-\alpha y_0}\, dy_0. \tag{16}$$

This also gives

$$\begin{aligned} \gamma &= \mathbb{E}[C] = \sum_{k\geq 2} \mathbb{E}[C|\, D = k]\, \mathbb{P}(D = k) \\ &= \int_0^\infty P(y_0) \left(\sum_{k=2}^\infty \rho(y_0, k) \right) \alpha e^{-\alpha y_0}\, dy_0 \\ &= \int_0^\infty P(y_0)\, (1 - \rho(y_0, 0) - \rho(y_0, 1))\, \alpha e^{-\alpha y_0}\, dy_0. \end{aligned} \tag{17}$$

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 $f \cdot \mathbb{1}_{\mathcal{B}_\infty((0,y_0))}/\mu(y_0)$. In particular

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

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 the intervals $[-e^{\frac{1}{2}(y_0+y_1)}, e^{\frac{1}{2}(y_0+y_1)}]$ and $[-e^{\frac{1}{2}(y_0+y_2)}, e^{\frac{1}{2}(y_0+y_2)}]$, respectively. We have that

$$P(y_0) = (\alpha - 1/2)^2 \int_0^\infty \int_0^\infty P(y_0, y_1, y_2) e^{-(\alpha-1/2)(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}y_1. \tag{18}$$

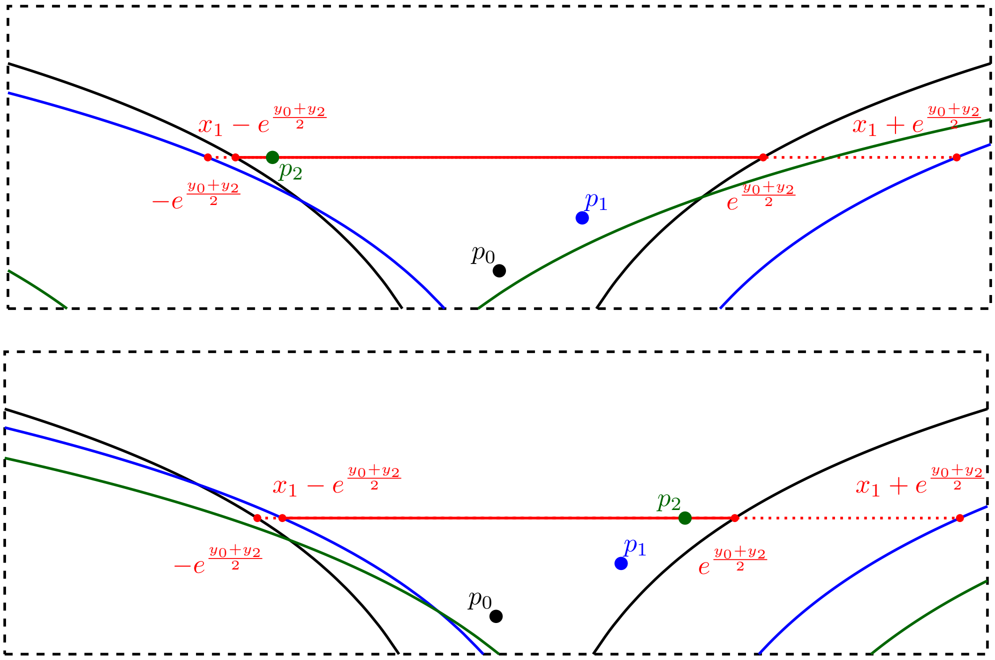

Figure 4: Situation for the intersections of the connection intervals considered in Lemma 3.3, with $y_0 < y_1 < y_2$ fixed and for different cases of $0 \leq x_1 \leq e^{(y_0+y_1)/2}$. The top figure shows the case where $0 \leq x_1 \leq e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2}$, while the bottom one shows 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 neighbourhoods are shown in, respectively, black, blue and green.

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

To compute the integral (18) 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 shows 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:

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

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 $e^{(y_1+y_2)/2} - e^{(y_0+y_2)/2} < x_1 \leq e^{(y_0+y_1)/2}$ then the probability that $|x_2 - x_1| \leq e^{(y_1+y_2)/2}$ 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

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

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

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

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

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

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

$$P(y_0, y_1, y_2) = (z_0/z_1) \, P(y_1, y_0, y_2) = (z_0/z_1) \, P(y_1, y_2, y_0). \tag{19}$$

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 (19). □

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

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

which is integrable over $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ (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 (18), 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:

$$I_1(y_0) := \int_{0<y_0<y_1<y_2} P(y_0, y_1, y_2) \cdot (\alpha - 1/2)^2 e^{-(\alpha-1/2)(y_1+y_2)} \, dy_2 \, dy_1$$
$$I_2(y_0) := \int_{0<y_1<y_0,y_2} P(y_0, y_1, y_2) \cdot (\alpha - 1/2)^2 e^{-(\alpha-1/2)(y_1+y_2)} \, dy_2 \, dy_1$$

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 (20), (25), (26) and (28).

Computing $I_1(y_0)$ Applying the change of variables $z_i := e^{-y_i/2}$, $i = 1, 2$, and Lemma 3.2 gives

$$\begin{aligned}
I_1(y_0) &= 4(\alpha - 1/2)^2 \cdot \int_{z_0>z_1>z_2>0} P(y_0, y_1(z), y_2(z)) z_1^{2\alpha-2} z_2^{2\alpha-2} \, dz_2 \, dz_1 \\
&= 4(\alpha - 1/2)^2 \cdot \Bigg(\int_{z_0>z_1>z_2>0} 1 \cdot z_1^{2\alpha-2} z_2^{2\alpha-2} \, dz_2 \, dz_1 \\
&\quad - \int_{\substack{z_0>z_1>z_2>0, \\ z_0<z_1+z_2}} G(z_0, z_1, z_2) \cdot z_1^{2\alpha-2} z_2^{2\alpha-2} \, dz_2 \, dz_1 \Bigg) \\
&=: 4(\alpha - 1/2)^2 (I_{11}(y_0) - I_{12}(y_0)).
\end{aligned}$$

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

$$\begin{aligned}
I_{11}(y_0) &= \int_0^{z_0} \int_0^{z_1} z_1^{2\alpha-2} z_2^{2\alpha-2} \, dz_2 \, dz_1 = \int_0^{z_0} z_1^{2\alpha-2} \left[\frac{z_2^{2\alpha-1}}{2\alpha - 1} \right]_0^{z_1} dz_1 \\
&= \frac{1}{2\alpha - 1} \cdot \int_0^{z_0} z_1^{4\alpha-3} \, dz_1 = \frac{1}{2(2\alpha - 1)^2} \cdot z_0^{4\alpha-2}.
\end{aligned} \tag{20}$$

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

$$J_{a,b,c}(z_0) := z_0^a \int_{\substack{z_0>z_1>z_2>0, \\ z_0<z_1+z_2}} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} \, dz_2 \, dz_1. \tag{21}$$

and note that

$$I_{1,2}(y_0) = \frac{1}{4}(J_{0,-1,1}(z_0) + J_{0,1,-1}(z_0) + J_{2,-1,-1}(z_0) + 2J_{0,0,0}(z_0) - 2J_{1,-1,0}(z_0) - 2J_{1,0,-1}(z_0)). \tag{22}$$

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

$$\begin{aligned}
J_{a,b,c}(z_0) &= z_0^a \int_{z_0/2}^{z_0} \int_{z_0-z_1}^{z_1} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} \, dz_2 \, dz_1 = z_0^a \int_{z_0/2}^{z_0} z_1^{b+2\alpha-2} \left[\frac{z_2^{c+2\alpha-1}}{c + 2\alpha - 1} \right]_{z_0-z_1}^{z_1} dz_1 \\
&= \frac{z_0^a}{c + 2\alpha - 1} \cdot \left(\int_{z_0/2}^{z_0} z_1^{b+c+4\alpha-3} \, dz_1 - \int_{z_0/2}^{z_0} z_1^{b+2\alpha-2} (z_0 - z_1)^{c+2\alpha-1} \, dz_1 \right) \\
&= \frac{z_0^{a+b+c+4\alpha-2}(1 - (1/2)^{b+c+4\alpha-2})}{(c + 2\alpha - 1)(b + c + 4\alpha - 2)} \\
&\quad - \frac{z_0^{a+b+c+4\alpha-3}}{c + 2\alpha - 1} \int_{z_0/2}^{z_0} (z_1/z_0)^{b+2\alpha-2} (1 - (z_1/z_0))^{c+2\alpha-1} \, dz_1
\end{aligned}$$

$$= \frac{z_0^{4\alpha-2}(1-(1/2)^{b+c+4\alpha-2})}{(c+2\alpha-1)(b+c+4\alpha-2)} - \frac{z_0^{4\alpha-2}}{c+2\alpha-1}\int_{1/2}^{1} u^{b+2\alpha-2}(1-u)^{c+2\alpha-1}\,\mathrm{d}u$$
$$= \frac{z_0^{4\alpha-2}(1-(1/2)^{b+c+4\alpha-2})}{(c+2\alpha-1)(b+c+4\alpha-2)} - \frac{z_0^{4\alpha-2}}{c+2\alpha-1}B^-(1/2; c+2\alpha, b+2\alpha-1),$$

where we have 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 (22) 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}\alpha^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}\alpha^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)\\
&\quad - (2\alpha-1)(\alpha-1)B^-(1/2; 2\alpha+1, 2\alpha-2)).
\end{aligned}$$

For the last step we use the identities

$$B^-(z; a, b) - B^-(z; a, b+1) = B^-(z; a+1, b), \tag{23}$$
$$B^-(z; a, b) + B^-(z; a, b+2) - 2B^-(z; a, b+1) = B^-(z; a+2, b). \tag{24}$$

to obtain

$$\begin{aligned}
I_{1,2}(y_0) &= \frac{\left(\frac{3}{64} - \frac{3}{16}2^{-4\alpha} + \alpha(-\frac{41}{128} + \frac{13}{16}2^{-4\alpha}) + \alpha^2(\frac{5}{8} - \frac{3}{4}2^{-4\alpha}) - \frac{15}{32}\alpha^3 + \frac{1}{8}\alpha^4\right) z_0^{4\alpha-2}}{4(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)\alpha}\\
&\quad - \frac{z_0^{4\alpha-2}B^-(1/2; 2\alpha+1, 2\alpha-2)}{8(\alpha-1)\alpha(2\alpha-1)}
\end{aligned} \tag{25}$$

Computing $I_2(y_0)$ We will follow a similar strategy as for $I_1(y_0)$. First, using the change of variables $z_i := e^{-y_i/2}$, $i = 1, 2$, we get

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

We start with the easy integral:

$$\begin{aligned}
I_{21}(y_0) &= z_0 \int_{1>z_1>\max(z_2,z_0); z_0, z_2>0} z_1^{2\alpha-3} z_2^{2\alpha-2} \, \mathrm{d}z_2 \, \mathrm{d}z_1 = z_0 \int_{z_0}^{1} \int_0^{z_1} z_1^{2\alpha-3} z_2^{2\alpha-2} \, \mathrm{d}z_2 \, \mathrm{d}z_1 \\
&= z_0 \int_{z_0}^{1} \left[\frac{z_2^{2\alpha-1}}{2\alpha-1} \right]_0^{z_1} z_1^{2\alpha-3} \, \mathrm{d}z_1 = \frac{z_0}{2\alpha-1} \int_{z_0}^{1} z_1^{4\alpha-4} \, \mathrm{d}z_1 = \frac{z_0 - z_0^{4\alpha-2}}{(4\alpha-3)(2\alpha-1)}.
\end{aligned} \tag{26}$$

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 function

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

and write

$$\begin{aligned}
I_{2,2}(y_0) &= \frac{1}{4} \left(J'_{0,-1,1}(z_0) + J'_{2,-1,-1}(z_0) + J'_{0,1,-1}(z_0) \right) \\
&\quad + \frac{1}{2} \left(J'_{1,-1,0}(z_0) - J'_{0,0,0}(z_0) - J'_{1,0,-1}(z_0) \right).
\end{aligned} \tag{27}$$

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

$$\begin{aligned}
J'_{a,b,c}(z_0) &= z_0^a \int_{z_0}^{1} \int_{z_1-z_0}^{z_1} z_1^{b+2\alpha-2} z_2^{c+2\alpha-2} \, \mathrm{d}z_2 \, \mathrm{d}z_1 \\
&= z_0^a \int_{z_0}^{1} \frac{1}{c+2\alpha-1} z_1^{b+2\alpha-2} (z_1^{c+2\alpha-1} - (z_1 - z_0)^{c+2\alpha-1}) \, \mathrm{d}z_1 \\
&= z_0^a \int_{z_0}^{1} \frac{1}{c+2\alpha-1} z_1^{b+c+4\alpha-3} \mathrm{d}z_1 - z_0^a \int_{z_0}^{1} \frac{1}{c+2\alpha-1} z_1^{b+2\alpha-2} (z_1 - z_0)^{c+2\alpha-1} \, \mathrm{d}z_1 \\
&= z_0^a \frac{1}{(c+2\alpha-1)(b+c+4\alpha-2)} (1 - z_0^{b+c+4\alpha-2}) \\
&\quad - \frac{z_0^a}{c+2\alpha-1} z_0^{b+c+4\alpha-2} B^-(1 - z_0; c+2\alpha, -b-c-4\alpha+2) \\
&= \frac{z_0^a - z_0^{4\alpha-2}}{(c+2\alpha-1)(b+c+4\alpha-2)} - \frac{z_0^{4\alpha-2} B^-(1 - z_0; c+2\alpha, -b-c-4\alpha+2)}{c+2\alpha-1}.
\end{aligned}$$

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

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

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

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 (27) 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 (23) and (24) this can be reduced to

$$
\begin{aligned}
I_{2,2}(y_0) = &\frac{1}{64\alpha(\alpha-1/2)^2(\alpha-1)} - \frac{(1-z_0)^{2\alpha}}{64\alpha(\alpha-1/2)^2(\alpha-1)} - \frac{z_0}{8(\alpha-1/2)(\alpha-1)(4\alpha-3)} \\
&+ \frac{z_0^2}{32(\alpha-1)^2} + \frac{(-6+25\alpha-48\alpha^2+44\alpha^3-16\alpha^4) z_0^{4\alpha-2}}{512\alpha(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)} \\
&+ \frac{z_0^{4\alpha-2} B^-(1-z_0; 2\alpha, 3-4\alpha)}{32(\alpha-1)(\alpha-1/2)^2}.
\end{aligned} \tag{28}
$$

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 $z_0^u B^-(1-z_0; a, b)$:

$$
\begin{aligned}
P(y_0) =& 2(I_1 + I_2) = 8(\alpha-1/2)^2 (I_{1,1} - I_{1,2} + I_{2,1} - I_{2,2}) \\
=& 8(\alpha-1/2)^2 \Bigg(\frac{1}{2(2\alpha-1)^2} z_0^{4\alpha-2} \\
&- \frac{\left(\frac{3}{64} - \frac{3}{16} 2^{-4\alpha} + \alpha(-\frac{41}{128} + \frac{13}{16} 2^{-4\alpha}) + \alpha^2(\frac{5}{8} - \frac{3}{4} 2^{-4\alpha}) - \frac{15}{32}\alpha^3 + \frac{1}{8} a^4\right) z_0^{4\alpha-2}}{4(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)\alpha} \\
&+ \frac{z_0^{4\alpha-2} B^-(1/2; 2\alpha+1, 2\alpha-2)}{8(\alpha-1)\alpha(2\alpha-1)} + \frac{z_0 - z_0^{4\alpha-2}}{(4\alpha-3)(2\alpha-1)} \\
&- \frac{1}{64\alpha(\alpha-1/2)^2(\alpha-1)} + \frac{(1-z_0)^{2\alpha}}{64\alpha(\alpha-1/2)^2(\alpha-1)} + \frac{z_0}{8(\alpha-1/2)(\alpha-1)(4\alpha-3)} \\
&- \frac{z_0^2}{32(\alpha-1)^2} - \frac{(-6+25\alpha-48\alpha^2+44\alpha^3-16\alpha^4) z_0^{4\alpha-2}}{512\alpha(\alpha-1/2)^2(\alpha-1)^2(\alpha-3/4)} \\
&- \frac{z_0^{4\alpha-2} B^-(1-z_0; 2\alpha, 3-4\alpha)}{32(\alpha-1)(\alpha-1/2)^2} \Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{8(\alpha-1)\alpha} + \frac{(\alpha-1/2)z_0}{\alpha-1} - \frac{(\alpha-1/2)^2 z_0^2}{4(\alpha-1)^2} \\
&\quad + z_0^{-2+4\alpha}\left(\frac{2^{-4\alpha-1}(3\alpha-1)}{\alpha(\alpha-1)^2} + \frac{(\alpha-1/2)B^-(1/2;1+2\alpha,-2+2\alpha)}{2(\alpha-1)\alpha}\right) \\
&\quad + \frac{(1-z_0)^{2\alpha}}{8(\alpha-1)\alpha} - \frac{z_0^{4\alpha-2}B^-(1-z_0;2\alpha,3-4\alpha)}{4(\alpha-1)}
\end{aligned}
$$

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 $\alpha = 3/4$

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 $B^-(1-z_0;2\alpha,3-4\alpha) = \int_0^{1-z_0} t^{2\alpha-1}(1-t)^{2-4\alpha}dt$). 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 (17) and (16), 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 $z^u B^-(1-z;v,w)$.

To compute J we observe that, by integration by parts,

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

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 $0 \le a < b \le 1$

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

In particular

$$\int_0^1 z^{u+2\alpha-k-1} e^{-\xi z^{-1}} \, \mathrm{d}z = \xi^{u+2\alpha-k} \Gamma^+(k-2\alpha-u, \xi). \tag{29}$$

where Γ^+ denotes the (upper) incomplete gamma function, and we have used the substitution $t = \xi/z$ which gives $\mathrm{d}z = -\xi t^{-2}\, \mathrm{d}t$. (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 $a, b, z \in \mathbb{C}$, $b \notin \mathbb{Z}_{\le 0}$, $Re(a), Re(z) > 0$, see [14, 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

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

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

$$\int_0^1 z_0^{2\alpha-k-1} e^{-\xi/z_0} (1-z_0)^{2\alpha} dz_0 = \Gamma(2\alpha+1) e^{-\xi} U(2\alpha+1, 1+k-2\alpha, \xi). \tag{30}$$

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:

$$\begin{aligned}\int_0^1 z^a e^{-\xi/z} B^-(1-z; b, c) \, \mathrm{d}z &= \int_0^1 z^a e^{-\xi/z} \int_0^{1-z} t^{b-1} (1-t)^{c-1} \, \mathrm{d}t \, \mathrm{d}z \\ &= \int_0^1 z^a e^{-\xi/z} \int_z^1 s^{c-1} (1-s)^{b-1} \, \mathrm{d}s \, \mathrm{d}z.\end{aligned}$$

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

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

$$
\begin{aligned}
&= \int_0^1 \int_0^s z^a e^{-\xi/z}\, \mathrm{d}z\, s^{c-1}(1-s)^{b-1}\, \mathrm{d}s \\
&= \int_0^1 \int_{\xi/s}^{\infty} \xi^{a+1} u^{-a-2} e^{-u}\, \mathrm{d}u\, s^{c-1}(1-s)^{b-1}\, \mathrm{d}s \\
&= \xi^{a+1} \int_0^1 \Gamma^+(-a-1, \xi/s) s^{c-1}(1-s)^{b-1}\, \mathrm{d}s. \qquad (31)
\end{aligned}
$$

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)

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

which holds for any $a \in \mathbb{R}$ and $s > 0$ (that for a fixed second argument, the upper incomplete gamma function is entire in the first argument, see [19, pp. 899, 1032ff.]). We can now evaluate the integral in (31) using several identities for Meijer's G-function. First, inserting the expression for the incomplete Gamma-function into (31) gives

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

Next we apply the inversion identity for Meijer's G-function (see [14, p. 209, 5.3.1.(9))]) to get

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

This expression is actually the Euler transform of Meijer's G-function (see [14, 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

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

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

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

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

$$
\int_0^1 z^a e^{-\xi/z} B^-(1-z; b, c)\, \mathrm{d}z = \xi^{6a-k-2}\Gamma(2\alpha) G^{3,0}_{2,3}\left(\xi\middle|\begin{matrix} 1, 3-2\alpha \\ 3-4\alpha, -6\alpha+k+2, 0 \end{matrix}\right). \qquad (32)
$$

Using equation (29), (30) and (32) we get

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

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

which is the expression in Theorem 1.1.

Similarly, we get

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

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

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 alternative, more explicit expression. This is what we will do in the current section. We will prove

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

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

and

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

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:

$$
\begin{aligned}
&P_\alpha(y_0(z_0))\\
&=\frac{1}{\alpha-1}\left(s_1(\alpha,z_0)+s_2(\alpha,z_0)+\frac{1}{\alpha-1}(s_3(\alpha,z_0)+s_4(\alpha,z_0))+s_5(\alpha,z_0)+s_6(\alpha,z_0)+s_7(\alpha,z_0)\right)
\end{aligned}
$$

where

$$
\begin{aligned}
s_1(\alpha,z_0) &= -\frac{1}{8\alpha}\\
s_2(\alpha,z_0) &= (\alpha-1/2)z_0
\end{aligned}
$$

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

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

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

In order to continue, we compute:

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

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

$$
\sum_{i \in \{1,2,5,6,7\}} s_i(1) + s_3'(1) + s_4'(1)
$$

$$
\begin{aligned}
&= -\frac{1}{8} + \frac{1}{2}z_0 - \frac{1}{4}z_0^2 + \frac{z_0^2}{32} + \frac{z_0^2}{4}\ln(\frac{z_0}{2}) + \frac{z_0^2}{4}B^-(1/2;3,0) + \frac{(1-z_0)^2}{8} - \frac{z_0^2}{4}B^-(1-z_0;2,-1) \\
&= -\frac{1}{8} + \frac{1}{2}z_0 - \frac{1}{4}z_0^2 + \frac{z_0^2}{32} + \left(\frac{z_0^2}{4}\ln(z_0) - \frac{z_0^2}{4}\ln 2\right) + \left(-\frac{5z_0^2}{32} + \frac{z_0^2}{4}\ln 2\right) \\
&\qquad + \left(\frac{1}{8} - \frac{z_0}{4} + \frac{z_0^2}{8}\right) + \left(\frac{z_0^2}{4} - \frac{z_0}{4} - \frac{z_0^2}{4}\ln z_0\right) &= 0,
\end{aligned}
$$

using that

$$
\begin{aligned}
B^-(\frac{1}{2};3,0) &= \int_0^{\frac{1}{2}} t^2(1-t)^-1dt = \int_{\frac{1}{2}}^1 (1-s)^2 s^{-1} ds \\
&= \int_{\frac{1}{2}}^1 s^{-1} - 2 + sds = -2 + \frac{1}{2} - \ln\frac{1}{2} + 1 - \frac{1}{8} = -\frac{5}{8} + \ln 2
\end{aligned}
$$

and

$$
\begin{aligned}
B^-(1-z_0;2,-1) &= \int_0^{1-z_0} t(1-t)^{-2}dt = \int_{z_0}^1 (1-s)s^{-2}ds \\
&= \int_{z_0}^1 s^{-2} - s^{-1}ds = -1 + z_0^{-1} + \ln z_0.
\end{aligned}
$$

Finally, it follows that as $\alpha \to 1$,

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

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

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

$$
\begin{aligned}
&- z_0 \ln z_0 - z_0^2(\ln z_0)^2 - z_0^2 \int_0^{1-z_0} \ln\left(\frac{\sqrt{t}}{1-t}\right) t(1-t)^{-2}\,\mathrm{d}t \\
=& \frac{5}{4}z_0 - \frac{9}{16}z_0^2 + \frac{z_0^2}{2}(\ln\frac{z_0}{2})^2 + \frac{z_0^2}{8}\ln\frac{z_0}{2} + \frac{(1-z_0)^2}{4}\ln(1-z_0) \\
&+ \frac{3}{8}z_0^2\ln(z_0) + z_0^2\ln(z_0)\ln 2 + \frac{z_0^2\ln 2}{4} + z_0^2/2(11/8 - 1/4\ln 2 - 3/2\ln(2)^2 - \mathrm{Li}_2(1/2)) \\
&- z_0\ln z_0 - z_0^2(\ln z_0)^2 + z_0(1 + \frac{1}{2}(2-z_0)\ln(z_0) + \frac{1}{2}z_0\ln(z_0)^2 - \frac{1}{2}(1-z_0)\ln(1-z_0) \\
&+ \frac{1}{2}z_0\,\mathrm{Li}_2(z_0)) - z_0^2 - \frac{1}{2}z_0^2\,\mathrm{Li}_2(1) \\
=& \frac{9}{4}z_0 - \frac{25}{16}z_0^2 + \frac{z_0^2}{2}(\ln\frac{z_0}{2})^2 + \frac{z_0^2}{8}\ln\frac{z_0}{2} + \frac{(1-z_0)^2}{4}\ln(1-z_0) \\
&- \frac{1}{8}z_0^2\ln(z_0) + z_0^2\ln(z_0)\ln 2 + \frac{z_0^2\ln 2}{4} + z_0^2/2(11/8 - 1/4\ln 2 - 3/2\ln(2)^2 \\
&- \mathrm{Li}_2(1/2) - \mathrm{Li}_2(1) + \mathrm{Li}_2(z_0)) - \frac{1}{2}z_0^2(\ln z_0)^2 - \frac{1}{2}z_0(1-z_0)\ln(1-z_0)
\end{aligned}
$$

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

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

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

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

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

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 (17) and (16) 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 (14). 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 (14) and (15) to obtain the following scaling result

$$
\frac{\int_0^\infty e^{-\beta y}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}{\int_0^\infty \rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y}\sim \xi^{2\beta}k^{-2\beta}. \tag{33}
$$

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 (evaluating the expressions for $\alpha = 3/4$)

$$\frac{2}{y}e^{\frac{y}{2}}P(y) = \frac{2}{y}B^-(1-e^{-\frac{y}{2}};3/2,0) - \frac{4}{y}\frac{e^{\frac{y}{2}}\left((1-e^{-\frac{y}{2}})^{3/2}-1\right)}{3} - \frac{1}{y} - \frac{e^{-\frac{y}{2}}}{4y} + \frac{2}{y}\left(\frac{5}{3} - \frac{2B^-(\frac{1}{2};5/2,-1/2)}{3}\right)$$

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

$$e^{\frac{y}{2}}\left(\left(1-e^{-\frac{y}{2}}\right)^{3/2}-1\right) \sim -\frac{3}{2}, \tag{34}$$

which implies that

$$\lim_{y\to\infty}\frac{4}{y}\frac{e^{\frac{y}{2}}\left((1-e^{-\frac{y}{2}})^{3/2}-1\right)}{3} = 0.$$

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

$$\frac{2}{y}B^-(1-e^{-\frac{y}{2}};3/2,0) = -\frac{B^-(1-z;3/2,0)}{\log(z)}.$$

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

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

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

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

$$+ e^{-(4\alpha-3)\frac{y}{2}} Q_\alpha + \frac{(\alpha-\frac{1}{2})^2}{4(\alpha-1)^2} e^{-\frac{y}{2}}.$$

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 (34)

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

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

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

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

$$\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y = (1 + o(1)) \int_{a^-(k)}^{a^+(k)} P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y. \tag{35}$$

To establish (35) 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

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

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 $\mu(y) > k$. 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 $P(y) \le 1$,

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

$$= O(1)\left(\rho(a^-(k),k) + \rho(a^+(k),k)\right),$$

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

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

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, for sufficiently large k,

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

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 $P(y) = h(y)(1+o(1))$, uniformly in y, as $y \to \infty$, then

$$\int_0^\infty P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \sim \int_0^\infty h(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y, \tag{36}$$

as $y \to \infty$.

We now proceed with the proof of Proposition 1.4, which is split over the different cases for α.

Proof when $1/2 < \alpha < 3/4$ By Proposition 3.7 and (36) 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 (33) with $\beta = 2\alpha - 1$.

Proof when $\alpha = 3/4$ Similar to the previous case Proposition 3.7 and (36) 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 (33) because of the additional logarithmic term. To prove the result we first show that

$$\int_{a^-(k)}^{a^+(k)} P(y)\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y \sim \int_{a^-(k)}^{a^+(k)} \frac{y}{2}e^{-y/2}\rho(y,k)\alpha e^{-\alpha y}\,\mathrm{d}y. \tag{37}$$

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

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

and similarly, a lower bound

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

Now observe that as $k \to \infty$,

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

and therefore it follows that

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

and

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

This proves (37).

Next we note that by (33) with $\beta = 1/2$ we have

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

Therefore, since by (35),

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

it follows from (37) that as $k \to \infty$,

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

when $\alpha = 3/4$.

Proof when $\alpha > 3/4$ Again, by Proposition 3.7, equation (36) and (33) with $\beta = 1/2$, it follows that as $k \to \infty$,

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

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 Clustering function, proving Theorem 1.2

In this subsection, we want to show that the clustering function of the KPKVB model $c_n(k) \xrightarrow{\mathbb{P}} \gamma(k)$ for a fixed k. The key idea is that for the poissonized KPKVB model the coupling with the box model is guaranteed to be exact (in the sense that it also preserves edges) for all vertices up to height $H = \frac{R}{4}$ (which is all we need to show the convergence of the clustering function for constant k).

The overall proof proceeds in four steps: firstly poissonizing the KPKVB model, secondly truncating the poissonized KPKVB model at height $H = H_n \to \infty$, thirdly using that the coupling between the poissonized KPKVB and box model is exact when truncated at this height H and finally showing that the clustering function in the truncated box model agrees with the integral for $\gamma(k)$.

We will encouter the following models: we will start with the KPKVB model $G_n = G(n; \alpha, \nu)$, then the Poissonized KPKVB model $G_{\mathrm{Po}} = G_{Po}(n; \alpha, \nu)$, moving to the truncated Poissonized KPKVB model $G_{Po,H}$ which is the subgraph of G_{Po} induced on all vertices with height at most H resp. radial coordinate at least $R - H$, then the coupling will bring us to the truncated box model $G_{box,H}$ which is the subgraph of the box model G_{box} induced by all vertices with height at most H.

4.1.1 Poissonization

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. Then*

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

Proof. We consider the coupling where we have an infinite supply of i.i.d. points $u_1, u_2, \ldots$ chosen according to the (α, R)-quasi uniform distribution, the vertices of $G(n; \alpha, \nu)$ are $u_1, \ldots, u_n$ and the vertices of $G_{\mathrm{Po}}(n; \alpha, \nu)$ are $u_1, \ldots, u_N$ with $N \overset{\mathrm{d}}{=} \mathrm{Po}(n)$ independently of $u_1, u_2, \ldots$.

Due to the coupling, the KPKVB graph or its poissonization is contained in the other, i.e. we can use new graph labels $G_1 \subset G_2$; e.g. if $N \geq n$, then $G_1 \overset{d}{=} G(n; \alpha, \nu)$ and $G_2 \overset{d}{=} G_{\mathrm{Po}}(n; \alpha, \nu)$. We write $V_{1,k}$ for the set of vertices of G_1 which have k neighbours in G_1 and $V_{2,k}$ for the set of vertices of G_2 which have k neighbours in G_2. Write $N_{1,k} = |V_{1,k}|$ and $N_{2,k} = |V_{2,k}|$.

Write W_1 for the set of all vertices of G_1 which have k neighbours in G_1 but no neighbour in $G_2 \backslash G_1$. Write W_2 for the set of all vertices of G_1 which have k neighbours in G_1 and at least one neighbour in $G_2 \backslash G_1$. Write W_3 for the set of all vertices of G_2 which have strictly less than k neighbours in G_1, but k neighbours in G_2.

With this notation, we can rewrite:

$$\begin{aligned}
|c(k; G_n) - c(k; G_{\mathrm{Po}})| &= \left| \sum_{v \in V_{1,k}} \frac{c_{G_1}(v)}{N_{1,k}} - \sum_{v \in V_{2,k}} \frac{c_{G_2}(v)}{N_{2,k}} \right| \\
&= \left| \sum_{v \in W_1} \frac{c_{G_1}(v)}{N_{1,k}} + \sum_{v \in W_2} \frac{c_{G_1}(v)}{N_{1,k}} - \sum_{v \in W_1} \frac{c_{G_2}(v)}{N_{2,k}} - \sum_{v \in W_3} \frac{c_{G_2}(v)}{N_{2,k}} \right|.
\end{aligned}$$

For every vertex of W_1, the clustering coefficient in G_1 and in G_2 have the same value. Hence, we can simplify:

$$|c(k; G_n) - c(k; G_{\mathrm{Po}})| = \left| \sum_{v \in W_1} c_{G_1}(v) \frac{N_{2,k} - N_{1,k}}{N_{1,k} N_{2,k}} + \sum_{v \in W_2} \frac{c_{G_1}(v)}{N_{1,k}} - \sum_{v \in W_3} \frac{c_{G_2}(v)}{N_{2,k}} \right|.$$

Using the triangle inequality for the absolute value and the fact that $|c_G(v)| \leq 1$,

$$|c(k; G_n) - c(k; G_{\mathrm{Po}})| \leq \frac{|W_1||N_{2,k} - N_{1,k}|}{N_{1,k} N_{2,k}} + \frac{|W_2|}{N_{1,k}} + \frac{|W_3|}{N_{2,k}}.$$

Using that $|W_1| \leq N_{1,k}$,

$$|c(k;G_n) - c(k;G_{\mathrm{Po}})| \leq \frac{|N_{2,k} - N_{1,k}|}{N_{2,k}} + \frac{|W_2|}{N_{1,k}} + \frac{|W_3|}{N_{2,k}}.$$

From the Chernoff bound for Poisson random variables, it follows that for any $c > 0$, $N \in [n - c\sqrt{n \log n}, n + c\sqrt{n \log n}]$ a.a.s.. This implies that a.a.s. the vertex sets of G_1 and G_2 differ by at most $c\sqrt{n \log n}$ many vertices. In particular also $|N_{1,k} - N_{2,k}|, |W_2|, |W_3| = O(\sqrt{n \log n})$ a.a.s.

By the paper by Gugelmann et al. [20] it holds that $N_{1,k}$ or $N_{2,k}$ is $\Theta(n)$ a.a.s. for fixed degree k and as $|N_{1,k} - N_{2,k}| = O(\sqrt{n \log n})$ a.a.s., it follows that both $N_{1,k}, N_{2,k} = \Theta(n)$ a.a.s. Putting everything together yields the claim:

$$|c(k;G_n) - c(k;G_{\mathrm{Po}})| \leq \frac{|N_{2,k} - N_{1,k}|}{N_{2,k}} + \frac{|W_2|}{N_{1,k}} + \frac{|W_3|}{N_{2,k}} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

□

4.1.2 Truncation

Consider the truncated Poissonized KPKVB model $G_{Po,H}$ with $H = H_n = \frac{R}{4}$.

Lemma 4.2. *Let $N_1(H)$ denote the number of vertices with radial coordinate at most $R - H \in [0, R]$ in the Poissonized KPKVB model G_{Po}. Then $\mathbb{E}[N_1(H)] = o(n)$.*

Proof. By the Poisson distribution of the number of points inside the disk with radius $R - H$, by the asymptotic behaviour of cosh and using that all $R, R - H, H \to \infty$,

$$\mathbb{E}[N_1(H)] = n\frac{\cosh(\alpha(R - H)) - 1}{\cosh(\alpha R) - 1} = (1 + o(1))ne^{-\alpha H} = o(n)$$

□

Lemma 4.3. *Let $N_2(H)$ denote the number of vertices with at least one neighbour with radial coordinate at most $R-H \in [0, R]$ in the Poissonized KPKVB model. If $H \to \infty$, then $\mathbb{E}[N_2(H)] = o(n)$.*

Proof. Among the vertices with at least one neighbour with radial coordinate at most $R - H$, we distinguish between those that themselves have radial coordinate at most $R - H_0$ and those that themselves have radial coordinate at least $R - H_0$ where $H_0 = \log H$. We will show that the expectation of both numbers is sublinear in n.

Since $H \to \infty$ we also have that $H_0 = \log H \to \infty$. Hence by Lemma 4.2, the expected number of vertices with radial coordinate at most $R - H_0$ (in particular those vertices among them with at least one neighbour with radial coordinate at most $R - H$), $\mathbb{E}[N_1(H_0) = o(n)]$ is sublinear. For a vertex with radial coordinate at least $R - H_0$, the probability that it has a neighbour with radial coordinate at most $R - H$, tends to zero. Therefore, by the Campbell-Mecke formula the number of such vertices is sublinear. □

Lemma 4.4. *Let $\alpha > \frac{1}{2}$, $\nu > 0$ and $k \geq 2$ be fixed. Let G_{Po} denote the Poissonized KPKVB model and $G_{Po,H}$ denote the truncated Poissonized KPKVB model. Then*

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

Proof. Let $V_n(k)$ denote the set of all vertices with degree k in the Poissonized KPKVB model G_{Po} and $V_{n,H}(k)$ the set of all vertices with degree k in the truncated Poissonized KPKVB model $G_{Po,H}$. Note that a vertex may have a smaller degree and different clustering coefficient in G_{Po} and $G_{Po,H}$ (even if it is a vertex of both graphs). We write $c_n(v)$ for the clustering coefficient of a vertex v in G_{Po} and $c_{n,H}(v)$ for the clustering coefficient in $G_{Po,H}$. Let $N_n(k) = |V_n(k)|$ denote the number of degree k vertices in G_{Po}. Then as $N_n(k) = \Theta(nk^{-(2\alpha+1)})$ a.a.s., there is a constant $a_1(k) > 0$ such that a.a.s. $N_n(k) \geq a_1(k)n$. Let $N_{n,H}(k) = |V_{n,H}(k)|$ be the number of degree k vertices in $G_{Po,H}$.

Using a telescoping sum, we obtain that

$$\begin{aligned}|c(k; G_{\mathrm{Po}}) - c(k; G_{Po,H})| \leq &\left| c(k; G_{\mathrm{Po}}) - \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_n(v) \right| \\ &+ \left| \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_n(v) - \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_{n,H}(v) \right| \\ &+ \left| \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_{n,H}(v) - c(k; G_{Po,H}) \right|.\end{aligned}$$

Now we use $X_1 := |V_n(k) \backslash V_{n,H}(k)|$, $X_2 := \sum_{v \in V_{n,H}(k)} |c_n(k) - c_{n,H}(k)|$ and $X_3 := |N_{n,H}(k) - N_n(k)|$ to bound the three differences. In the first difference, after factorizing $N_n(k)$, all terms in $V_{n,H}(k)$ cancel and the remaining terms are all bounded by one:

$$\left| c(k; G_{\mathrm{Po}}) - \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_n(v) \right| = \left| \frac{1}{N_n(k)} \sum_{v \in V_n \backslash V_{n,H}(k)} c_n(v) \right| \leq \frac{X_1}{N_n(k)}.$$

In the second difference, the sums can be aligned and then the triangle inequality for the absolute value applied:

$$\left| \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_n(v) - \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_{n,H}(v) \right| \leq \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} |c_n(v) - c_{n,H}(v)| = \frac{X_2}{N_n(k)}.$$

In the third difference, we can simplify:

$$\begin{aligned}\left| \frac{1}{N_n(k)} \sum_{v \in V_{n,H}(k)} c_{n,H}(v) - c(k; G_{Po,H}) \right| &= \left| \sum_{v \in V_{n,H}(k)} c_{n,H} \left(\frac{1}{N_n(k)} - \frac{1}{N_{n,H}(k)} \right) \right| \\ &= \frac{|N_{n,H}(k) - N_n(k)|}{N_n(k) N_{n,H}(k)} \sum_{v \in V_{n,H}(k)} c_{n,H}(k).\end{aligned}$$

Thus, we end up with the upper bound, using that $c_{n,H}(k) \leq 1$ and $N_{n,H}(k) = |V_{n,H}(k)|$,

$$|c(k; G_{\mathrm{Po}}) - c(k; G_{Po,H})| \leq \frac{X_1}{N_n(k)} + \frac{X_2}{N_n(k)} + \frac{X_3}{N_n(k)} \frac{\sum_{v \in V_{n,H}(k)} c_{n,\leq H}(k)}{N_{n,H}(k)} \leq \frac{X_1 + X_2 + X_3}{N_n(k)}.$$

Now using that a.a.s. $N_n(k) \geq a_1(k)n$ gives that a.a.s.

$$|c(k; G_{\mathrm{Po}}) - c(k; G_{Po,H})| \leq \frac{X_1 + X_2 + X_3}{a_1(k)n}.$$

By Lemma 4.2, we have that $\mathbb{E}[X_1] = \mathbb{E}[X_3] = o(n)$ and by Lemma 4.3, we have that $\mathbb{E}[X_2] = o(n)$. We conclude that $\mathbb{E}|c(k; G_{\mathrm{Po}}) - c(k; G_{Po,H})| = o(1)$. From this, convergence in probability follows by Markov's inequality: for $\epsilon > 0$,

$$\mathbb{P}(|c(k; G_{\mathrm{Po}}) - c(k; G_{Po,H})| \geq \epsilon) \leq \frac{\mathbb{E}|c(k; G_{\mathrm{Po}}) - c(k; G_{Po,H})|}{\epsilon} = o(1).$$

□

4.1.3 Coupling

Lemma 4.5. *The clustering function in the truncated Poissonized KPKVB model $G_{Po,H}$ and in the truncated box model $G_{box,H}$ agree, i.e. $c(k; G_{Po,H}) = c(k; G_{box,H})$.*

Proof. As all vertices involved have height $\leq H$ resp. radial coordinate $\geq R - H$ and as $H \leq \frac{R}{4}$, all vertices involved (in the disk) will have radial coordinate $\geq \frac{3}{4}R$, therefore by Lemma 2.3, edges exist between such vertices in the disk if and only if they exist in the box. So the two graphs are isomorphic and hence the clustering functions agree. □

4.1.4 Deriving the limit integral expression

Lemma 4.6. *Let $\alpha > \frac{1}{2}$, $\nu > 0$ and $k \geq 2$ be fixed. Then*

$$c(k; G_{box,H}) \xrightarrow[n\to\infty]{\mathbb{P}} \gamma(k).$$

Let $\mathcal{R}_H = \mathcal{R}([0, H])$ denote the truncated box. For a vertex v of the vertex set $V = V(G_{box,H})$, writing $\Gamma(v)$ for the set of all neighbours of vertex v, let

$$h(V, v) = \binom{k}{2}^{-1} \sum_{uw \in \binom{\Gamma(v)}{2}} \mathbb{1}_{\{uw \in E(G_{box,H})\}} \mathbb{1}_{\{D_{G_{box,H}}(v) = k\}}.$$

Let $X = \frac{1}{n} \sum_{v \in V(G_{box,H})} h(V(G_{box,H}), v)$. Let $Z = \frac{N_{box,H}(k)}{n} = \frac{1}{n} \sum_{v \in G_{box,H}} \mathbb{1}_{\{D_{G_{box,H}}(v) = k\}}$. Then we have that $c(k; G_{box,H}) = \frac{X}{Z}$. We will show separately that numerator X and denominator Z converge in probability (where the limit of Z is a positive constant). From this, the claim of the lemma will follow as convergence in probability is closed under taking quotients.

By the Campbell-Mecke formula, we have

$$\mathbb{E}[X] = \frac{1}{n} \int_{\mathcal{R}_H} \mathbb{E} h(V, v) \mu(dv) = \int_0^H \mathbb{E} h(V, (0, y)) \alpha e^{-\alpha y} dy.$$

Then, we condition on the additional vertex $(0, y) \in \mathcal{R}_H$ having degree k,

$$\mathbb{E}[X] = \int_0^H \mathbb{E}[h(V, (0, y)) | D_{G_{box,H}}(0, y) = k] \mathbb{P}(D_{G_{box,H}}(0, y) = k) \alpha e^{-\alpha y} dy.$$

We observe that for $H \to \infty$, the integrand converges point-wise for $y \in (0, \infty)$ to $P(y)\rho(y, k)\alpha e^{-\alpha y}$ and is always bounded by the integrable function $\alpha e^{-\alpha y}$. Therefore, by the theorem of dominated convergence, as $H \to \infty$,

$$\mathbb{E}[X] \to \int_0^\infty P(y)\rho(y, k)\alpha e^{-\alpha y} dy.$$

Note that for X^2, the summation can be split into off-diagonal (with $v_1 \neq v_2$) and diagonal (with $v_1 = v_2$) terms, i.e. write $X^2 = \frac{1}{n^2}(Y_1 + Y_2)$ where

$$Y_1 = \sum_{v_1 \neq v_2 \in V(G_{box,H})} h(V(G_{box,H}), v_1) h(V(G_{box,H}), v_2),$$

and $Y_2 = \sum_{v_1 \in V(G_{box,H})} h(V(G_{box,H}), v_1)^2$. As $h \leq 1$, $Y_2 \leq nX \leq n$ and so, $\mathbb{E}Y_2 \leq n$, we have by the Campbell-Mecke formula,

$$\mathbb{E}[Y_1] = \int_{\mathcal{R}_H} \int_{\mathcal{R}_H} \mathbb{E}[h(V(G_{box,H}), v_1) h(V(G_{box,H}), v_2)] \mu(dv_1) \mu(dv_2).$$

If $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$ have horizontal distance $|x_1 - x_2| > 2e^H$, their neighbourhood balls are disjoint w.r.t. the connection rule of the truncated box model $G_{box,H}$. To see this, assume that $v_3 = (x_3, y_3)$ is adjacent to both v_1 and v_2. Then $|x_1 - x_3| \leq e^{\frac{1}{2}(y_1+y_3)}$ and $|x_2 - x_3| \leq e^{\frac{1}{2}(y_2+y_3)}$. As $y_1, y_2, y_3 \leq H$, it follows that $|x_1 - x_3| \leq e^H$ and $|x_2 - x_3| \leq e^H$, which implies that $|x_1 - x_2| \leq |x_1 - x_3| + |x_2 - x_3| \leq 2e^H$, contradiction.

Due to the disjointedness of the neighbourhood balls and the properties of the Poisson process $h(V, v_1)$ and $h(V, v_2)$ are independent. The horizontal distance is uniform in $[0, \frac{\pi n}{2\nu}]$. Therefore, the probability that two vertices have horizontal distance $\leq 2e^H$ is $\leq \frac{4\nu e^H}{\pi n} = o(1)$. Writing d_H for the horizontal distance, we can continue the calculation of the second moment,

$$\begin{aligned}
\mathbb{E}[Y_1] &= \iint_{\substack{\mathcal{R}_H \times \mathcal{R}_H: \\ |x_1 - x_2| > 2e^H}} \mathbb{E}[h(V, v_1)]\mathbb{E}[h(V, v_2)]\mu(dv_1)\mu(dv_2) \\
&\quad + \iint_{\substack{\mathcal{R}_H \times \mathcal{R}_H: \\ |x_1 - x_2| \leq 2e^H}} \mathbb{E}[h(V, v_1)]\mathbb{E}[h(V, v_2)]\mu(dv_1)\mu(dv_2) \\
&= (1 + o(1)) \iint_{\mathcal{R}_H \times \mathcal{R}_H} \mathbb{E}[h(V, v_1)]\mathbb{E}[h(V, v_2)]\mu(dv_1)\mu(dv_2) \\
&= (1 + o(1))(n\mathbb{E}X)^2.
\end{aligned}$$

It follows that $\mathbb{E}\left[X^2\right] = (1 + o(1))\mathbb{E}\left[X\right]^2$, which implies $Var(X) \to 0$ using that $\mathbb{E}[X]$ tends to a constant (see above). We conclude

$$X \xrightarrow[n \to \infty]{\mathbb{P}} \int_0^\infty P(y)\rho(y, k)\alpha e^{-\alpha y} dy.$$

In a similar way, by the Campbell-Mecke formula,

$$\mathbb{E}[Z] = \frac{1}{n} \int_{\mathcal{R}_H} \mathbb{P}(D_{G_{box,H}}(v) = k)\mu(dv) = \int_0^H \mathbb{P}(D_{G_{box,H}}(0, y) = k)\alpha e^{-\alpha y} dy.$$

By the theorem of dominated convergence, it follows that

$$\mathbb{E}[Z] \to \int_0^\infty \rho(y, k)\alpha e^{-\alpha y} dy.$$

The statement that $Var(Z) \to 0$ follows by the same steps as for X, just using $h(V, v) = \mathbb{1}_{\{D_{G_{box,H}}(v) = k\}}$ instead. Hence,

$$\mathbb{E}[Z] \xrightarrow[n \to \infty]{\mathbb{P}} \int_0^\infty \rho(y, k)\alpha e^{-\alpha y} dy.$$

Finally note that

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

4.1.5 Putting everything together

Proof of Theorem 1.2: The overall statement follows from the lemmas together with the fact that convergence in probability is closed under addition. Write

$$c(k; G_n) - \gamma(k) = c(k; G_n) - c(k; G_{\text{Po}}) + c(k; G_{\text{Po}}) - c(k; G_{Po,H}) + c(k; G_{Po,H}) - \gamma(k).$$

The first difference converges in probability to zero due to the Poissonization Lemma 4.1, the second difference due to the truncation Lemma 4.4, then for the third difference we use $c(k; G_{Po,H}) = c(k; G_{box,H})$ from Lemma 4.5 and then Lemma 4.6 on the derivation of the limit integral. ■

4.2 Overall clustering coefficient, proving Theorem 1.1

In this subsection, we want to infer that the clustering coefficient of the KPKVB random graph $c(G_n) \xrightarrow{\mathbb{P}} \gamma$ by using that $c(k; G_n) \xrightarrow{\mathbb{P}} \gamma(k)$ for fixed k.

Proof of Theorem 1.1: Let $\epsilon > 2\delta > 0$. As the p_k form the probability mass function of an $\mathbb{N}_0$-valued random variable, we have that $\sum_{k=0}^{\infty} p_k = 1$, which implies that there is $K_\delta > 0$ such that $1 - \delta < \sum_{k=0}^{K_\delta} p_k \leq 1$. From Equations (13), (16) and (17), we know that $\gamma = \sum_{k=2}^{\infty} \gamma(k) p_k$. From this, it follows that $(1-\delta)\gamma < \sum_{k=2}^{K_\delta} \gamma(k) p_k \leq 1$ and hence that $|\gamma - \sum_{k=2}^{K_\delta} \gamma(k) p_k| \leq \delta$. Write $p(k; G_n) = \frac{N_n(k)}{n}$ for the fraction of degree k vertices in the KPKVB model. From the definition of the clustering coefficient and function, it follows immediately that $c(G_n) = \sum_{k=2}^{\infty} c(k; G_n) p(k; G_n)$. By the initial observation, it follows again that $(1-\delta) c(k; G_n) < \sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n) \leq 1$, and hence, $|c(G_n) - \sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n)| \leq \delta$. Putting these observations together:

$$\begin{aligned}
|c(G_n) - \gamma| &\leq \left| c(G_n) - \sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n) \right| + \left| \sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n) - \sum_{k=2}^{K_\delta} \gamma(k) p_k \right| + \left| \sum_{k=2}^{K_\delta} \gamma(k) p_k - \gamma \right| \\
&\leq \delta + \left| \sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n) - \sum_{k=2}^{K_\delta} \gamma(k) p_k \right| + \delta.
\end{aligned}$$

Therefore,

$$\mathbb{P}\left(|c(G_n) - \gamma| \geq \epsilon \right) \leq \mathbb{P}\left(\left| \sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n) - \sum_{k=2}^{K_\delta} \gamma(k) p_k \right| \geq \epsilon - 2\delta \right) \to 0,$$

because $c(k; G_n) \xrightarrow{\mathbb{P}} \gamma(k)$ for $k = 2, \ldots, K_\delta$ and $p(k; G_n) \xrightarrow{\mathbb{P}} p_k$ together imply that

$$\sum_{k=2}^{K_\delta} c(k; G_n) p(k; G_n) \xrightarrow{\mathbb{P}} \sum_{k=2}^{K_\delta} \gamma(k) p_k.$$

$\blacksquare$

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

The proof of Theorem 1.3 follows the same strategy as outlined in Section 2 and executed in Section 4. However, the fact that $k = k_n \to \infty$ as $n \to \infty$, introduces significant technical challenges, especially for k_n close the the maximum scale $n^{\frac{1}{2\alpha+1}}$. For example, the coupling between G_{Po} and G_{box} 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. Moreover, 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 in the strategy outlined in Section 2 only introduces error terms that are of smaller order, i.e. that are $o(\gamma(k_n))$. This will turn out to require a great deal of care in bounding all error terms we encounter.

In this section we explain the challenges associated 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. To this end we define the scaling function

$$s(k) = \begin{cases} k^{-(4\alpha-2)} & \text{if } \frac{1}{2} < \alpha < \frac{3}{4}, \\ \log(k) k^{-1} & \text{if } \alpha = \frac{3}{4}, \\ k^{-1} & \text{if } \alpha > \frac{3}{4}, \end{cases} \tag{38}$$

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, unless stated otherwise, $\{k_n\}_{n\geq 1}$ will always denote a sequence of non-negative 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,

$$c^*(k; G) = \frac{1}{\mathbb{E}[N_k]} \sum_{\substack{v \in V(G) \\ \deg(v)=k}} c(v). \tag{39}$$

Notice that the only difference between $c(k; G)$ and $c^*(k; G)$ is that we replace N_k by its expectation $\mathbb{E}[N_k]$. The advantage is that now, the only randomness is in the formation of triangles. In addition, note that since $\mathbb{E}[N_k] > 0$ 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. Below briefly describe each of the intermediate steps leading up to the proof of Theorem 1.3. We start with an observation on the dependence between the degree of a point $p = (x, y)$ and its height y.

5.1 Restricting the height for vertices with degree k_n

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

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

The reason for this was that the integrand included the function $\rho(y, k) = \mathbb{P}(\mathrm{Po}(\mu(y)) = k)$, where $\mathrm{Po}(\lambda)$ 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 $\mathbb{P}(\mathrm{Po}(\mu_n(y)) = k_n) := \hat{\rho}_n(y, k_n)$, for some $\mu_n(y)$ which is asymptoticly equivalent to $\mu(y)$. In these cases we also want to be able to restrict our integration around those heights y for which $y \approx 2\log(k_n/\xi)$. Such results are established in Section 6. Moreover, we prove that if μ_n corresponds to either $\mu(\mathcal{B}(y))$ or $\mu(\mathcal{B}_{\mathrm{box}}(y))$, then for a certain classs of functions h

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

i.e. we may replace $\hat{\rho}_n(y, k_n)$ in the integrand with $\rho(y, k_n)$.

5.2 Adjusted clustering and the Poissonized KPKVB model

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

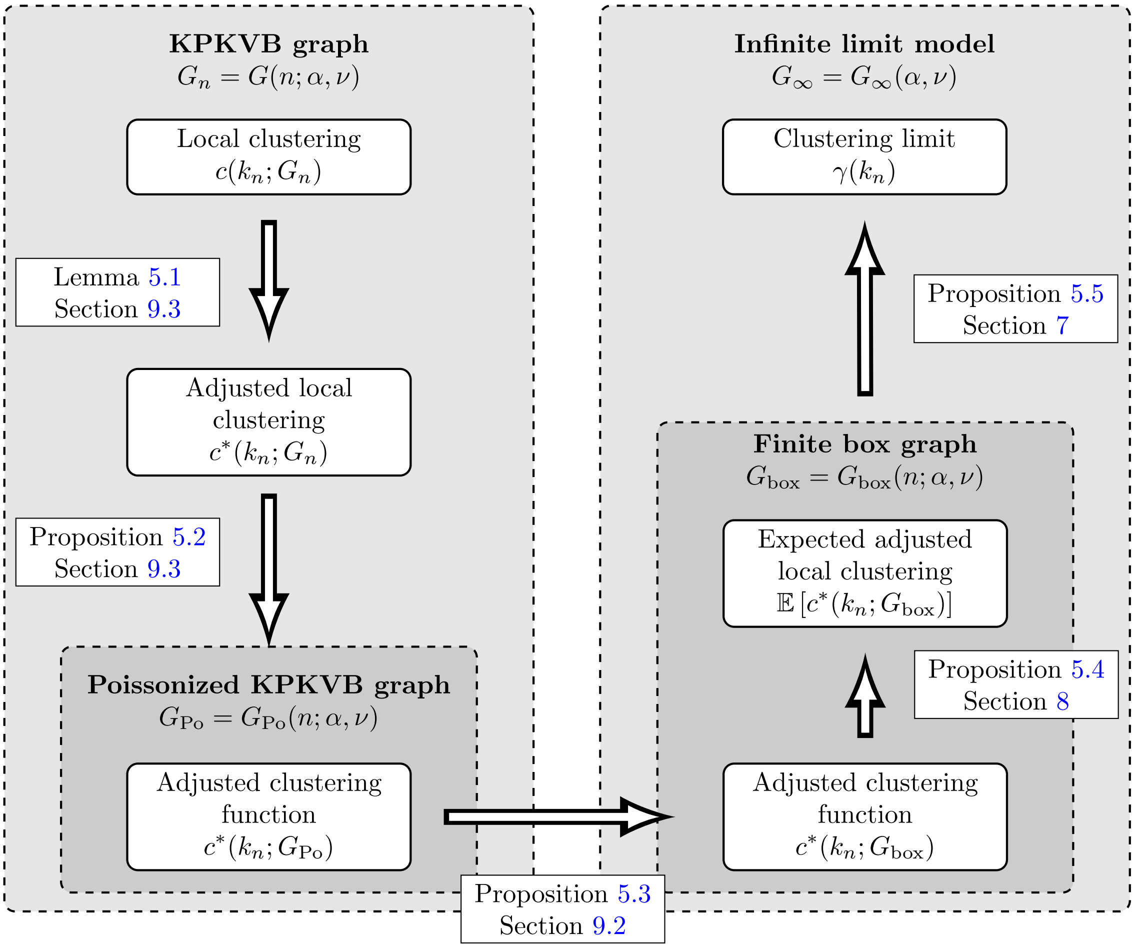

Figure 5: Overview of the proof strategy for Theorem 1.2. The left column denote the models in which the true hyperbolic balls are used while the right column contains the models that use an approximation of these. The most important part is the transition between these to setting which is accomplished by Proposition 5.3.

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 for KPKVB graphs G_n behaves similarly to that in the G_{Po}. The proof, found in Section 9.3, is based on a standard coupling between a Binomial Point Process and Poisson Point Process.

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

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

5.3 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 step is one of the key technical challenges we face in proving Theorem 1.3.

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 $y < R/4$ for large enough n. In addition, the main contribution of triangles would also come from nodes with heights $y' < R/4$. 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 $R/4 = \frac{1}{2}\log(n) - \frac{1}{2}\log(\nu)$ we get that $R/4 = o(y_n)$ for all $\alpha > (3-4\varepsilon)/2$ and hence $y_n > R/4$ 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 adjusted clustering function). *As $n \to \infty$,*

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

Tobias: Maybe we could replace these three with a statement on $|c(k_n; G_n) - c^*(k_n; G_{\mathrm{box}})|$, 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.

Together, the three results described so far imply that the difference between the clustering function for a KPKVB graph and the adjusted clustering function for the finite box graph G_{box} converges to zero faster than the proposed scaling $\gamma(k_n)$ in Theorem 1.3. Hence, to prove this theorem it is enough to prove it for $c^*(k; G_{\mathrm{box}})$.

5.4 From the finite box to the infinite model

To compute the limit of the adjusted clustering function $c^*(k; G_{\mathrm{box}})$ we first prove in Section 8 that it is concentrated around its mean $\mathbb{E}\left[c^*(k_n; G_{\mathrm{box}})\right]$.

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 adjusted clustering function in G_{box}). *As $n \to \infty$,*

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

This result represents another technical challenge 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 neighbourhoods 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 to compute expectation $\mathbb{E}\left[c^*(k_n; G_{\mathrm{box}})\right]$ and show that it is asymptotically equivalent to $\gamma(k_n)$ as $n \to \infty$. To accomplish this we move to the infinite limit model G_∞ and show that the difference between the expected value of $c^*(k; G_{\mathrm{box}})$ 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 $n \to \infty$,*

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

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 proof of Proposition 5.5 therefore relies on analyzing the number of triangles coming from these additional edges and showing that their contribution to the 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 $N_{\mathrm{Po}}(k)$ denotes number of nodes with degree k in G_{Po} and $\rho_{box}(y,k) = \mathbb{P}\left(\mathrm{Po}(\mu(\mathcal{B}_\infty(y))) = k\right)$, i.e. the degree distribution in G_{box} for a point $p = (x, y)$.*

5.5 Proof of the main results

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

Proof of Theorem 1.3. Note that the second of the theorem follows immediately from the first.

To prove the first statement, we rewrite $c(k_n; G_n)$ as

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

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

$$\begin{aligned} \mathbb{E}\left[|c(k_n; G_n) - \gamma(k_n)|\right] &\le \mathbb{E}\left[|c(k_n; G_n) - c^*(k_n; G_n)|\right] + \mathbb{E}\left[|c^*(k_n; G_n) - c^*(k_n; G_{\mathrm{Po}})|\right] \\ &\quad + \mathbb{E}\left[|c^*(k_n; G_{\mathrm{Po}}) - c^*(k_n; G_{\mathrm{box}})|\right] + \mathbb{E}\left[|c^*(k_n; G_{\mathrm{box}}) - \mathbb{E}c^*(k_n; G_{\mathrm{box}})|\right] \\ &\quad + \left|\mathbb{E}\left[c^*(k_n; G_{\mathrm{box}})\right] - \gamma(k_n)\right| \end{aligned}$$

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 second term, Proposition 5.3 for the coupling between the Poissonized KPKVB and 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:

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

which establishes the first statement of the theorem and finishes the proof. □

6 Concentration of heights for vertices with degree k

Here we show that if we integrate with respect to the function $\hat{\rho}_n(y, k_n) = \mathbb{P}\left(\mathrm{Po}(\mu_n(y)) = k_n\right)$ then we may restrict integration with respect to the *height* y to an interval on which $\mu_n(y) = \Theta\left(k_n\right)$. We will refer to such a result as a *concentration of heights* result. In addition, if $\mu_n(y)$ is equivalent to $\mu(y)$ on this interval, then we may replace $\hat{\rho}_n(y, k_n)$ in the integral with $\rho(y, k_n) = \mathbb{P}\left(\mathrm{Po}(\mu(y)) = k_n\right)$ (the degree distribution of the typical point in G_∞).

We start with a concentration of heights result 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 present a generalization of this result (Lemma 6.1) and use this to establish concentration of heights results for the Poissonized KPKVB G_{Po} and finite box model G_{box}.

Finally we provide a general result that allow to substitute $\hat{\rho}_n(y, k_n)$ in the integrand with $\rho(y, k_n)$ and show that this holds in particular for the degree distributions in G_{Po} and G_{box}, given by, respectively $\rho_{\mathrm{Po}}(y, k_n) := \mathbb{P}\left(\mathrm{Po}(\mu\left(\mathcal{B}\left(y\right)\right)) = k_n\right)$ and $\rho_{\mathrm{box}}(y, k_n) := \mathbb{P}\left(\mathrm{Po}(\mu\left(\mathcal{B}_{\mathrm{box}}\left(y\right)\right)) = k_n\right)$.

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 $\beta < \alpha$,*
$$\int_{\mathbb{R}_+ \setminus (a_n^-, a_n^+)} \rho(y, k_n) h(y) \alpha e^{-\alpha y} \, \mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right), \tag{40}$$
as $n \to \infty$.

2. *If in addition $C > \sqrt{4\alpha+1}$ and $h(a_n) \sim h(b_n)$ whenever $a_n \sim b_n$, as $n \to \infty$. Then,*
$$\int_0^\infty h(y)\rho(y, k_n)\alpha e^{-\alpha y} \, \mathrm{d}y \sim 2\alpha\xi^{2\alpha} h(2\log(k_n/\xi)) k_n^{-(2\alpha+1)}, \tag{41}$$
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 $h(y)$,

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

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

$$\int_{\mathbb{R}_+ \setminus (a_n^-, a_n^+)} h(y)\rho(y, k_n)\alpha e^{-\alpha y} \, \mathrm{d}y = O(1)\left(\rho(a_n^-, k_n) + \rho(a_n^+, k_n)\right). \tag{42}$$

We shall now bound the terms $\rho(a_n^{\pm}, k_n)$. We explicitly show the bound for $\rho(a_n^+, k_n)$, the computation for $\rho(a_n^-, k_n)$ is similar. 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 as $x - \log(1+x) \sim x^2/2$ as $x \to 0$, we get

$$\begin{aligned}\rho(a_n^+, k_n) &\leq \sqrt{2\pi} k_n^{-1/2} e^{-k_n\left(\epsilon_n + C\frac{\kappa_n}{k_n} - \log\left(1+\epsilon_n+C\frac{\kappa_n}{k_n}\right)\right)} \\ &\sim (2\pi)^{-1/2} k_n^{-1/2} e^{-\frac{k_n(\epsilon_n + C\kappa_n/k_n)^2}{2}} \\ &= O\left(k_n^{-(1+C^2)/2}\right), \end{aligned} \tag{43}$$

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

$$\rho(a_n^-, k_n) \leq \Theta(1)\, k_n^{-1/2} e^{-\frac{k_n(\epsilon_n - C\kappa_n/k_n)^2}{2}} = O\left(k_n^{-(1+C^2)/2}\right). \tag{44}$$

Plugging (44) and (43) into (42) 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 $C > \sqrt{4\alpha+1}$, (40) 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 (14)),

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

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 $c_n \sim 2\log(k_n/\xi)$. 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 $n \to \infty$. □

Note that we can tune the error in (40) by selecting an appropriately large $C > 0$, i.e. by restricting 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 $y \to \infty$*

$$\int_0^\infty g(y)\rho(y, k_n)\alpha e^{-\alpha y} \, dy \sim 2\alpha\xi^{2\alpha} h(2\log(k_n/\xi))k_n^{-(2\alpha+1)}, \tag{45}$$

as $n \to \infty$.

Proof. By assumption, g satisfies the conditions of 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 $C > 0$ we define

$$\mathcal{K}_C(k_n) = \left\{ y \in \mathbb{R}_+ : \frac{k_n - C\sqrt{a_n \log(k_n)}}{\xi} \vee 1 \leq e^{\frac{y}{2}} \leq \frac{k_n + C\sqrt{k_n \log(k_n)}}{\xi} \right\}. \tag{46}$$

The following corollary allows us to bound integrals of functions $h_n(y)$ by considering their 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 $0 \leq y \leq R$. Then, as $n \to \infty$,*

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

In particular,

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

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

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

By assumption on $h_n(y)$,

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

and

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

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

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

The result then follows since

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

□

For functions $h_n(y) = k_n^s h(y)$ we obtain an asymptotic equivalent expression for the associated integral.

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 $h_n(y) = O(k_n^s)\, h(y)\rho(y, k_n)$, uniformly on $0 \le y \le R$. Then,*

$$\int_{\mathcal{R}} h_n(y)\rho(y,k_n) f(x,y)\, \mathrm{d}x\, \mathrm{d}y \sim 2\alpha\xi^{2\alpha}\, n\, h_n(2\log(k_n/\xi)) k_n^{-(2\alpha+1)}, \tag{47}$$

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. □

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 we can obtain some general crude bound $h_n(y) = O(k_n^s)\, e^{\alpha/2y}$ for some $s > 0$ for $y \in \mathbb{R}_+$. Then, if we can obtain a more precise bound $h_n(y) = (1 + o(1))k_n^2 e^{\alpha/2y}$ uniformly on $\mathcal{K}_C(k_n)$, 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 with*

$$h(y) := \begin{cases} e^{-\frac{y}{2}(4\alpha-2)} c_\alpha \xi^{4\alpha-2} & \text{if } \frac{1}{2} < \alpha < \frac{3}{4}, \\ \frac{y}{2} e^{-\frac{y}{2}} & \text{if } \alpha = \frac{3}{4}, \\ e^{-\frac{y}{2}} \frac{\alpha - \frac{1}{2}}{\alpha - \frac{3}{4}} & \text{if } \alpha > \frac{3}{4}. \end{cases}$$

Hence, Proposition 1.4 directly follows from Proposition 3.7 and a concentration of heights argument (Corollary 6.2).

6.2 A more general concentration of heights argument

Although powerful, the current versions of the concentration of heights arguments are only valid for the function $\rho(y, k_n) := \mathbb{P}(\mathrm{Po}(\mu(\mathcal{B}_\infty(y))) = k_n)$, which uses the neighbourhoods 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 for the degree distribution function in these models. To be more precise, let us define

$$\rho_{\mathrm{Po}}(y,k) = \mathbb{P}(\mathrm{Po}(\mu(\mathcal{B}(y))) = k)$$

and

$$\rho_{\mathrm{box}}(y,k) = \mathbb{P}(\mathrm{Po}(\mu(\mathcal{B}_{\mathrm{box}}(y))) = k)$$

Then we want when Lemma 6.1 to remain true if we replace $\rho(y, k_n)$ with either the function $\rho_{\mathrm{Po}}(y, k_n)$ or $\rho_{\mathrm{box}}(y, k_n)$. To establish these result we first prove a more general version of Lemma 6.1.

Lemma 6.5. *Let $\alpha > \frac{1}{2}, \nu > 0$ and $k_n \to \infty$ such that $k_n = o(n)$, $\ell_n = (1+\epsilon_n)k_n$, with $\epsilon_n \to 0$ and define*

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

In addition, let $\hat{\rho}_n(y,k) = \mathbb{P}(\mathrm{Po}(\hat{\mu}_n(y)) = k)$, where $\hat{\mu}_n(y)$ is a differentiable function that satisfies for some $0 < \varepsilon < 1$

$$\text{i) } \lim_{n\to\infty} \sup_{0\le y\le(1-\varepsilon)R} \frac{\hat{\mu}_n(y)}{\mu(\mathcal{B}_\infty(y))} = 1 \quad \text{and} \qquad \text{ii) } \lim_{n\to\infty} \sup_{0\le y\le(1-\varepsilon)R} \frac{\hat{\mu}_n'(y)}{\mu(\mathcal{B}_\infty(y))} = \frac{1}{2}.$$

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 $\beta < \alpha$,*
$$\int_{\mathbb{R}_+\setminus(a_n^-,a_n^+)} \hat{\rho}_n(y,k_n)h(y)\alpha e^{-\alpha y}\,\mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right), \tag{48}$$
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,*
$$\int_0^\infty h(y)\hat{\rho}_n(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \sim \int_0^\infty h(y)\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y, \tag{49}$$
as $n \to \infty$.

Proof.

For simplicity we write $\mu(y) := \mu(\mathcal{B}_\infty(y)) = \xi e^{y/2}$ throughout the proof. Observe that $\mu'(y) = \mu(y)/2$ and $\mu^{-1}(yz) = \mu^{-1(y)} + \mu^{-1}(z)$.

Proof of statement 1. Let $0 < \varepsilon < 1$ be such that conditions i) and ii) hold, take $\varepsilon' = \min\{\varepsilon/3, 1/3\} < 1/2$ and let $\delta > 0$ be such that $\hat{\mu}_n(\delta) \ge \xi$. We first show that we can restrict to integration over $(\delta, (1-\varepsilon')R)$ that the integration over $(1-\varepsilon')R \le y \le R$ is negligible.

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

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

By construction $\varepsilon' < \varepsilon$, and hence by condition i) we have that $\hat{\mu}_n((1-\varepsilon')R) = \Theta\left(\mu((1-\varepsilon')R\right) = \Theta\left(n^{2(1-\varepsilon')}\right)$. Therefore, since $k_n = o(n)$, it follows that $\hat{\mu}_n((1-\varepsilon')R)/k_n = \omega\left(n^{1-2\varepsilon'}\right) \to \infty$ as $n \to \infty$. We can now use Stirling's approximation to bound $\hat{\rho}_n((1-\varepsilon')R,k_n)$ as

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

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

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

In a similar fashion we get that

$$\int_0^\delta h(y)\hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y = O\left(1\right)\hat{\rho}_n(\delta,k_n)$$
$$= O\left(1\right)k_n^{-1/2}e^{-k_n(\log(k_n)-\hat{\mu}_n(\delta))} = O\left(k_n^{-(1+C^2)/2}\right).$$

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

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

and

$$\int_{a_n^+}^{(1-\varepsilon')R} \hat{\rho}_n(y,k_n)e^{(\beta-\alpha)y}\,\mathrm{d}y = O\left(k_n^{-(1+C^2)/2}\right). \tag{51}$$

To prove this we first establish a result that will also help with proving statement 2.

Let $(a,b) \subseteq (\delta,(1-\varepsilon')R)$ and consider the change of variable $z = \mu^{-1}(\hat{\mu}_n(y))$. Then, writing $\hat{a} = \hat{\mu}_n^{-1}(\mu(a))$ and similar for $\hat{b}$, we get

$$\int_a^b \rho(z,k_n)e^{-\alpha z}\,\mathrm{d}z = \int_a^b \mathbb{P}\left(\mathrm{Po}(\mu(z)) = k_n\right)e^{-\alpha z}\,\mathrm{d}z$$
$$= \int_{\hat{a}}^{\hat{b}} \mathbb{P}\left(\mathrm{Po}(\hat{\mu}_n(y)) = k_n\right)e^{-\alpha\mu^{-1}(\hat{\mu}_n(y))}\frac{\hat{\mu}_n'(y)}{\mu'(\mu^{-1}(\hat{\mu}_n(y)))}\,\mathrm{d}y,$$

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 $\hat{\mu}_n(y)$ satisfies conditions i) and ii). Since $\mu'(y) = \mu(y)/2$ it follows that, uniformly on (a,b),

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

Moreover, we have

$$e^{-\alpha\mu^{-1}(\hat{\mu}_n(y))} = e^{-\alpha(y+\mu^{-1}(1+o(1)))} = (1+o\left(1\right))e^{-\alpha y},$$

uniformly on (a,b). These results then imply

$$\int_a^b \rho(z,k_n)e^{-\alpha z}\,\mathrm{d}z = (1+o\left(1\right))\int_{\hat{a}}^{\hat{b}} \hat{\rho}_n(y,k_n)e^{-\alpha y}\,\mathrm{d}y. \tag{52}$$

Now, using (52) with $a = \mu^{-1}(\hat{\mu}_n(\delta))$ and $b = \mu^{-1}(\hat{\mu}_n(a_n^-)$ we get

$$\int_a^b \rho(y,k_n)e^{(\beta-\alpha)y}\,\mathrm{d}y = (1+o\left(1\right))\int_\delta^{a_n^-} \hat{\rho}_n(y,k_n)e^{(\beta-\alpha)y}\,\mathrm{d}y.$$

Since $\mu^{-1}(\hat{\mu}_n(a_n^-) = (1+o\left(1\right))a_n^-$ and $b \geq 0$ the left hand side is

$$O\left(1\right)\int_0^{a_n^-} \rho(y,k_n)e^{(\beta-\alpha)y}\,\mathrm{d}y,$$

and hence (50) follows from Lemma 6.1. The proof of (51) follows in a similar way.

Proof of statement 2. The proof of the second statement follows the same line of reasoning as above, using (52). First, the mean value theorem for definite integrals implies that

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

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

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

This however follows immediately from (52) by picking $a = \mu^{-1}(\hat{\mu}_n(a_n^-))$ and $b = \mu^{-1}(\hat{\mu}_n(a_n^+))$. □

To apply this lemma to $\rho_{\mathrm{Po}}(y,k_n)$ or $\rho_{\mathrm{box}}(y,k_n)$ we need to show that both these functions satisfies the conditions i) and ii) in Lemma 6.5. We will do this in the next two sections and establish that the results from Corollaries 6.2, 6.3 and 6.4 hold when we replace $\rho(y,k_n)$ with either $\rho_{\mathrm{box}}(y,k_n)$ or $\rho_{\mathrm{Po}}(y,k_n)$.

6.3 Concentration of heights for the finite box model

The following lemma immediately implies that $\mu(\mathcal{B}_{\mathrm{box}}(y))$ satisfies the conditions of Lemma 6.5.

Lemma 6.6. *For all $y > 2\log(\pi/2)$,*

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

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

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

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

$$\mu(\mathcal{B}_{box}(p)) = \mu(\mathcal{B}_{\infty}(p))\left(1-e^{-(\alpha-\frac{1}{2})R}\right).$$

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

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

Therefore, if $y \leq 2\log(\pi/2)$ this intersection occurs above the height R of the box $\mathcal{R}$ 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 $\mu(\mathcal{B}_{\infty}(p)) = \xi e^{\frac{y}{2}}$ where $\xi = \frac{4\alpha\nu}{(2\alpha-1)\pi}$. Then, after some simple algebra, we have that

$$\begin{aligned}
\mu(\mathcal{B}_{\mathrm{box}}(p)) &= \int_0^{h(y)}\int_{-\frac{\pi}{2}e^{\frac{R}{2}}}^{\frac{\pi}{2}e^{\frac{R}{2}}} \mathbb{1}_{\left\{|x'|\leq e^{\frac{y+y'}{2}}\right\}} f_{\alpha,\nu}(x',y')\,dx'\,dy' \\
&\quad + \int_{h(y)}^{R}\int_{-\frac{\pi}{2}e^{\frac{R}{2}}}^{\frac{\pi}{2}e^{\frac{R}{2}}} f_{\alpha,\nu}(x',y')\,dx'\,dy' \\
&= \frac{2\alpha\nu}{\pi}e^{\frac{y}{2}}\int_0^{h(y)} e^{-(\alpha-\frac{1}{2})y'}\,dy' + \alpha\nu e^{\frac{R}{2}}\int_{h(y)}^{R} e^{-\alpha y'}\,dy'
\end{aligned}$$

$$
\begin{aligned}
&= \xi e^{\frac{y}{2}} \left(1 - \left(\frac{\pi}{2}\right)^{-(2\alpha-1)} e^{-(\alpha-\frac{1}{2})(R-y)}\right) \\
&\quad + \nu e^{\frac{R}{2}} \left(\left(\frac{\pi}{2}\right)^{-2\alpha} e^{-\alpha(R-y)} - e^{-\alpha R}\right) \\
&= \mu(\mathcal{B}_\infty(p))\left(1 - \phi_n(y)\right).
\end{aligned}
$$

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

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

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

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

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

□

From the definition of $\phi_n(y)$ in Lemma 6.6 it is immediate that $\rho_{\text{box}}(y,k)$ satisfies the conditions for $\hat{\rho}_n(y,k)$ in Lemma 6.5. We thus have the following corollary.

Corollary 6.7. *The statements in Corollaries 6.2, 6.3 and 6.4 hold when we replace $\rho(y,k_n)$ with $\rho_{box}(y,k_n)$.*

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 $y + y' < R$, $p' \in \mathcal{B}(p)$ exactly when $|x - x'|_{\pi e^{R/2}} \leq \Phi(y, y')$. In this setting, the coupling lemma (Lemma 2.2) gives that

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

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 $\mathcal{R}$ and not of the hyperbolic disc $\mathcal{D}_R$, i.e. the balls $\mathcal{B}(p)$ "live" in the finite box and not on the hyperbolic disc. We start with the following preliminary result.

Lemma 6.8. *Let $\Phi(y,y')$ be defined as in (8). Then, for any $0 \leq \delta < 1$*

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

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

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

$$+ \frac{2\nu\alpha}{\pi} K e^{\frac{3y}{2} - R} \int_0^{(1-\delta)(R-y)} e^{(\frac{3}{2}-\alpha)y'} \, \mathrm{d}y'$$

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

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

Therefore we get

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

and hence

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

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

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

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

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

□

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* $0 < \varepsilon < 1$

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

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

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

For the first part we have that

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

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

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

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

$$
\begin{aligned}
&\mu\left(\mathcal{B}((0,y)) \cap \mathcal{R}([R-y,R])\right) \\
&\quad = \int_{R-y}^{R} \int_{I_n} f_{\alpha,\nu}(x',y') \, \mathrm{d}x' \, \mathrm{d}y' = \nu\alpha e^{R/2}\left(e^{-\alpha(R-y)} - e^{-\alpha R}\right) \\
&\quad = \mu\left(\mathcal{B}_\infty(0,y)\right) \frac{2\alpha-1}{4\pi}\left(e^{-(\alpha-\frac{1}{2})(R-y)} - e^{-(\alpha-\frac{1}{2})R-y/2}\right),
\end{aligned} \tag{53}
$$

from which we conclude that

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

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 6.5. To show that condition ii) is also satisfied we need to analyze

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

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

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

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 6.5.

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

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

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

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

where

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

with $\Phi(y,y')$ defined as in (8) and, see (53),

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

Taking the derivative of the last expression gives

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

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

$$\lim_{n\to\infty}\sup_{0<y\leq(1-\varepsilon)R}\mu\left(\mathcal{B}_\infty\left(y\right)\right)^{-1}\frac{2\alpha\nu}{\pi}\frac{\partial}{\partial y}\int_0^{R-y}\Phi(r,r')e^{-\alpha y'}\,\mathrm{d}y'=\frac{1}{2}. \tag{54}$$

We start with some preliminary computations. For convenience we define

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

so that

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

Next, following the same calculation as in the proof of [16, 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

$$2e^{-(R-y-y')}h_1(y)h_2(y)\leq\Xi(y,y')\leq 2e^{-(R-y-y')}. \tag{55}$$

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 $\Xi(y,y')$ 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,

$$\begin{aligned}\frac{\partial}{\partial y}\Phi(R-y,R-y')&=\frac{1}{2}e^{R/2}\frac{1}{\sqrt{1-\left(1-\Xi(y,y')\right)^2}}\frac{\partial}{\partial y}\Xi(y,y')\\&=\frac{\frac{1}{2}e^{R/2}\Xi(y,y')}{\sqrt{1-\left(1-\Xi(y,y')\right)^2}}\left(1+\varphi_n(y,y')\right).\end{aligned} \tag{56}$$

Applying the Leibniz's rule we then get

$$\frac{\partial}{\partial y}\int_0^{R-y}\Phi(y,y')\alpha e^{-\alpha y'}\,\mathrm{d}y'$$

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

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

$$
\lim_{n\to\infty} \sup_{0<y\le(1-\varepsilon)R} \mu\left(\mathcal{B}_\infty(y)\right)^{-1} I_t(y) = 0, \quad \text{for } t = 1,3 \tag{57}
$$

while

$$
\lim_{n\to\infty} \sup_{0\le y\le(1-\varepsilon)R} \mu\left(\mathcal{B}_\infty(y)\right)^{-1} \frac{2\nu\alpha}{\pi} I_2(y) = \frac{1}{2}. \tag{58}
$$

This then implies (54) and finishes the proof.

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

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

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

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

$$\leq 2e^{-\alpha(R-y)}\sqrt{1-e^{-\delta(R-y)}}$$
$$\leq 2e^{-\alpha(R-y)}.$$

We therefore conclude that

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

which implies

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

Finally, to show (58) we define

$$\varphi_n^- := \sup_{0<y\leq(1-\varepsilon)R}\inf_{0\leq y'\leq(1-\delta)(R-y)}\varphi_n(y,y')$$
$$\varphi_n^+ := \sup_{0<y\leq(1-\varepsilon)R}\sup_{0\leq y'\leq(1-\delta)(R-y)}\varphi_n(y,y')$$

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

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

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),

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

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

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

and

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

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

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

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

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

□

We now conclude that similar to Corollary 6.7, the concentration of heights results also hold for $\rho_{\mathrm{Po}}(y, k_n)$.

Corollary 6.11. *The statement in Corollaries 6.2, 6.3 and 6.4 hold when we replace $\rho(y, k_n)$ with $\rho_{\mathrm{Po}}(y, k_n)$.*

Remark 6.3 (Generalized concentration of heights arguments)**.** *Since the results from Corollaries 6.2, 6.3 and 6.4 hold for any of the three functions $\rho(y, k_n)$, $\rho_{box}(y, k_n)$ and $\rho_{\mathrm{Po}}(y, k_n)$, we will will refer only to one of these three when using a concentration of heights argument of for any of the three models G_{Po}, G_{box} and G_∞.*

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

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

Recall that G_{box} is obtained by restricting the Poisson point process $\mathcal{P}_{\alpha,\nu}$ to the box $\mathcal{R} = (-I_n, I_n] \times (0, R]$, with $I_n = \frac{\pi}{2} e^{R/2}$ and connecting two points $p_1, p_2 \in \mathcal{R}$ if and only if $|x_1 - x_2|_{\pi e^{R/2}} \le e^{(y_1+y_2)/2}$. We also recall that by definition of the norm $|.|_{\pi e^{R/2}}$ 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 a triangle 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. To keep notation concise we write $|\cdot|_n$ for the norm $|\cdot|_{\pi e^{R/2}}$.

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

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

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

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

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

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

Here we slightly abuse notation and write $\mathcal{B}_\infty(y)$ for $\mathcal{B}_\infty((0, y))$. We will adopt this notational convention throughout the remainder of this section, to keep notation concise. We further write $N_{\mathrm{box}}(k)$ to denote the number of vertices with degree k in G_{box}.

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

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

Lemma 7.1. *Let $\gamma(k_n)$ be defined as in (16). Then as $n \to \infty$*

$$\gamma(k_n) = (1 + o(1)) \frac{1}{k_n^2 p_{k_n}} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_\infty(y)\right] \rho(y, k) \alpha e^{-\alpha y} \, \mathrm{d}y. \tag{59}$$

Moreover,

$$\mathbb{E}\left[c^*(k_n;G_{\text{box}})\right] = (1+o(1))\frac{1}{k_n^2 p_{k_n}}\int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_{box}(y)\right]\rho(y,k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \tag{60}$$

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 (60) 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

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

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

$$\mathbb{E}\left[c_{\text{box}}(y)\middle|\deg_{\text{box}}(y)=k_n\right] = \binom{k_n}{2}^{-1}\mathbb{E}\left[\sum_{1\leq i<j\leq k_n}\mathbb{1}_{\{u_i\in\mathcal{B}_{\text{box}}(u_j)\}}\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[u_1 \in \mathcal{B}_{\text{box}}\left(u_2\right)\right] \\
&= \mu\left(\mathcal{B}_{\text{box}}\left(y\right)\right)^{-2} \iint T_{\text{box}}(y, p_1, p_2) f(x_1, y_1) f(x_2, y_2) \, \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}x_2 \, \mathrm{d}y_2 \\
&= \mu\left(\mathcal{B}_{\text{box}}\left(y\right)\right)^{-2} \mathbb{E}\left[T_{\text{box}}(y)\right].
\end{aligned}
$$

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

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

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

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

We therefore conclude that

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

$\square$

Comparing (59) and (60), we conclude that to prove Proposition 5.5 it is enough to show that

$$
\left| \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[T_{\text{box}}(y) - T_\infty(y)\right] \rho(y, k) \alpha e^{-\alpha y} \, dy \right| = o\left(s(k_n) p_{k_n} k_n^2\right), \tag{61}
$$

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 $[-\frac{\pi}{2}e^{R/2}, \frac{\pi}{2}e^{R/2}]$ and we stop at height $y = R$. This induces a difference in triangle counts between both models. To see this, note that for any $p = (x, y)$ with $0 \le y \le R$ we have that $\mathcal{B}_{\text{box}}\left(p\right) = \mathcal{B}_\infty\left(p\right) \cap \mathcal{R}$. This means that if $p', p_2 \in \mathcal{B}_{\text{box}}\left(p\right)$ and $p_2 \in \mathcal{B}_\infty\left(p'\right) \cap \mathcal{R}$ then $p_2 \in \mathcal{B}_{\text{box}}\left(p\right) \cap \mathcal{B}_{\text{box}}\left(p'\right)$ 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 $\mathcal{B}_{\text{box}}\left(p\right) \cap \mathcal{B}_{\text{box}}\left(p'\right)$ that are not in $\mathcal{B}_\infty\left(p\right) \cap \mathcal{B}_\infty\left(p'\right)$. Let us denote this region by $\mathcal{T}(p, p')$, see Figure 6 for an example of this region. Then, any $p_2 \in \mathcal{T}(p, p')$ creates a triangle with p and p' in G_{box} that is not present in G_∞. Finally, any point $p_2 \in \mathcal{B}_\infty\left(p\right) \cap \mathcal{B}_\infty\left(p'\right)$ 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 $\mathcal{T}(p_0, p_1)$,

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

The next result, which is crucial for the proof of Proposition 5.5, computes the expected measure of $\mathcal{T}(p, p')$ 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 (38). 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 (61), 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

$$\begin{aligned}|\mathbb{E}\left[T_{\text{box}}(p_0) - T_\infty(p_0)\right]| &\leq \mathbb{E}\left[\left|T_{\text{box}}(p_0) - \widetilde{T}_{\text{box}}(p_0)\right|\right] \\ &\quad + \int_{\mathcal{R}'} \int_{\mathcal{R}'} T_\infty(p_0, p_1, p_2) f(x_1, y_1) f(x_2, y_2) \, dx_2 \, dy_2 \, dx_1 \, dy_1.\end{aligned}$$

The first part is taken care of by Lemma 7.2. For the other integral we have

$$\begin{aligned}&\iint_{\mathcal{R}'} T_\infty(p_0, p_1, p_2) f(x_1, y_1) f(x_2, y_2) \, dx_2 \, dy_2 \, dx_1 \, dy_1 \\ &\quad \leq \left(\int_{\mathcal{R}'} \mathbb{1}_{\{p_1 \in \mathcal{B}_\infty(p_0)\}} f(x_1, y_1) \, dx_1 \, dy_1\right)^2 = O\left(\left(e^{y/2} \int_R^\infty e^{-(\alpha - \frac{1}{2}) y_1} \, dy_1\right)^2\right)\end{aligned}$$

$$= O\left(e^{y}e^{-(2\alpha-1)R}\right) = O\left(e^{y}n^{-(4\alpha-2)}\right).$$

Thus we conclude, using Lemma 7.2, that,

$$|\mathbb{E}\left[T_{\text{box}}(p_0) - T_{\infty}(p_0)\right]| = O\left(yn^{-(2\alpha-1)} + n^{-(4\alpha-2)}e^{y}\right). \tag{62}$$

Therefore, on $\mathcal{K}_C(k_n)$,

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

where the last part follows from Lemma 7.3 and the fact that $s(k_n)^2 = o\left(s(k_n)\right)$. This establishes (61) 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. Recall that

$$\rho_{\text{box}}(y,k_n) = \mathbb{P}\left(\text{Po}(\mu(\mathcal{B}_{\text{box}}(y))) = k_n\right),$$

denote the degree distribution of a point $p_0 = (0,y)$ in G_{box}.

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

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

In particular,

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

Proof. We first write

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

Therefore, Lemma 7.2 and equation (62) 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

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

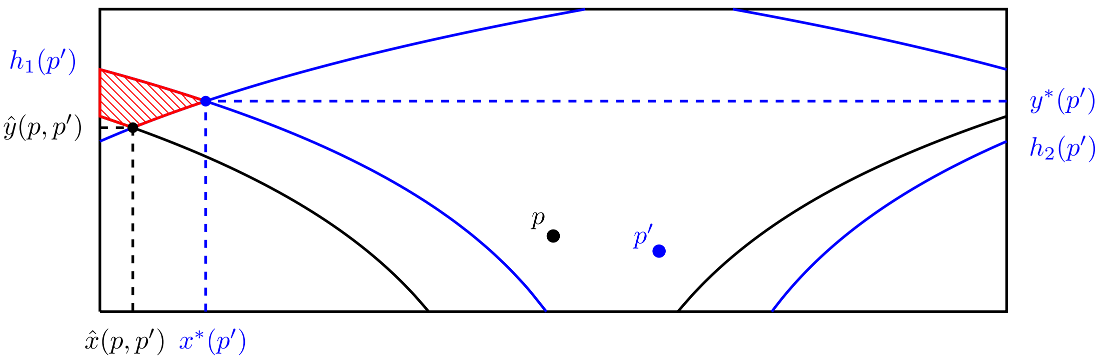

Figure 6: Example configuration of two points p and p' for which $\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$ is not a subset of $\mathcal{B}_{\infty}(p) \cap \mathcal{B}_{\infty}(p')$. The red region indicates the area belonging to $\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$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

$$
\begin{aligned}
&= (1 + o(1)) n k_n^2 \int_{\mathcal{K}_C(k_n)} P(y)\rho(y, k_n)\alpha e^{-\alpha y}\, dy \\
&= (1 + o(1)) n k_n^2 \int_0^{\infty} P(y)\rho(y, k_n)\alpha e^{-\alpha y}\, dy,
\end{aligned}
$$

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

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

□

7.2 Counting missing triangles

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

Recall that $\mathcal{T}(p,p')$ 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 $\mathcal{T}(p,p') \neq \emptyset$. We observe that $\mathcal{T}(p,p') \neq \emptyset$ because the right boundary of the ball $\mathcal{B}_{\text{box}}(p')$ exits 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,

$$
b_p^-(z) = \begin{cases} 2\log(x - z) - y & \text{if } -\frac{\pi}{2}e^{R/2} \leq z \leq x - e^{y/2} \\ 2\log\left(\pi e^{R/2} + x - z\right) - y & \text{if } x - e^{(y+R)/2} + \pi e^{R/2} \leq z \leq \frac{\pi}{2}e^{R/2} \\ 0 & \text{otherwise} \end{cases} \tag{63}
$$

$$
b_p^+(z) = \begin{cases} 2\log(z - x) - y & \text{if } x + e^{y/2} \leq z \leq \frac{\pi}{2}e^{R/2} \\ 2\log\left(\pi e^{R/2} + z - x\right) - y & \text{if } -\frac{\pi}{2}e^{R/2} \leq z \leq x + e^{(y+R)/2} - \pi e^{R/2} \\ 0 & \text{otherwise} \end{cases} \tag{64}
$$

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 $-\frac{\pi}{2}e^{R/2} \leq z \leq x + e^{(y+R)/2} - \pi e^{R/2}$. This equation becomes

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

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

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

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

To summarize we have:

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

The crucial observation is that $\mathcal{T}(p, p') = \emptyset$ 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

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

Solving this equation gives us the function

$$b_p^*(z) = y - 2\log\left(1 - \frac{z}{\frac{\pi}{2}e^{R/2}}\right), \tag{65}$$

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 $\mathcal{T}(p, p') = \emptyset$ 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

$$\int_0^R \int_0^{I_n} \mu\left(\mathcal{T}(p, p_1)\right) f(x_1, y_1) \, dx_1 \, dy_1 = O\left(yn^{-(2\alpha-1)} + n^{-(2\alpha-1)}e^y\right) \tag{66}$$

The proof goes in two stages. First we compute $\mu\left(\mathcal{T}(p, p_1)\right)$ 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.

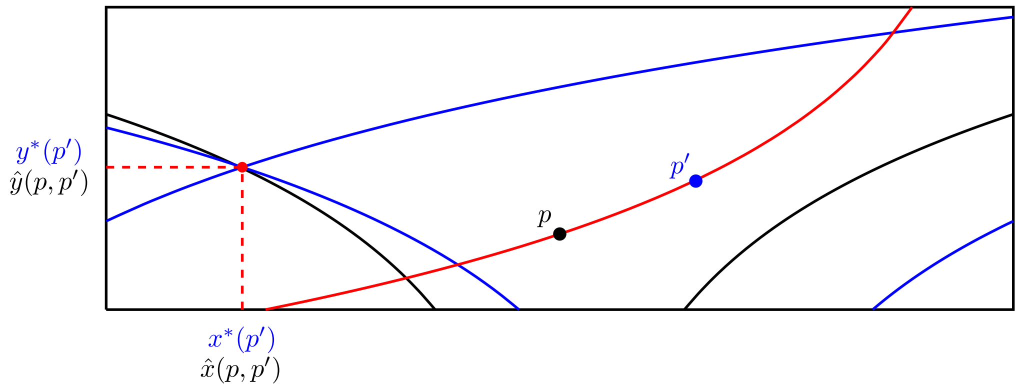

Figure 7: Example for a given p of the boundary function $x' \mapsto b_p^*(x')$, given by the red curve, which determines whether $\mathcal{T}(p,p') = \emptyset$. We see that when $y' = b_p^*(x')$ then $(\hat{x}(p,p'), \hat{y}(p,p')) = (x^*(p'), y^*(\prime))$.

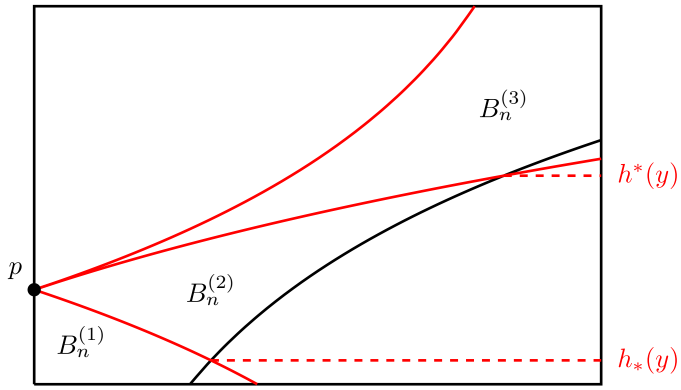

Figure 8: Three different areas $B_n^{(i)}$ used in the proof of Lemma 7.2.

Computing $\mu\left(\mathcal{T}(p,p_1)\right)$

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

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

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

$$
h_*(y) = y + 2\log\left(\frac{I_n}{I_n + e^y}\right) \tag{67}
$$

$$
h^*(y) = y + 2\log\left(\frac{I_n}{I_n - e^y}\right). \tag{68}
$$

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 $\mathcal{T}(p,p_1) = \emptyset$. So we only need to consider $p_1 \in B_n$. We shall establish the following result:

$$\mu\left(\mathcal{T}(p,p_1)\right) = \begin{cases} O\left(I_n^{-2\alpha} e^{\alpha y_1}\right) & \text{if } p_1 \in B_n^{(1)} \\ O\left(I_n^{-2\alpha} e^{\alpha y}\right) & \text{if } p_1 \in B_n^{(2)} \cup B_n^{(3)} \end{cases} \tag{69}$$

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

Case $p_1 \in B_n^{(1)}$: $0 \le y_1 \le y - 2\log(I_n/(I_n - x_1))$ In this case the integral over p_2 splits into two parts

$$\mathcal{I}_n^{(1)}(p_1) := \int_{h_2(p_1)}^{y^*(p_1)} \int_{-I_n}^{x_1 + e^{(y_1+y_2)/2} - 2I_n} e^{-\alpha y_2} \, \mathrm{d}x_2 \, \mathrm{d}y_2$$
$$\mathcal{I}_n^{(2)}(p_1) := \int_{y^*(p_1)}^{h_1(p_1)} \int_{x^*(p_1)}^{x_1 - e^{(y_1+y_2)/2}} e^{-\alpha y_2} \, \mathrm{d}x_2 \, \mathrm{d}y_2.$$

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

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

where we used that $x_1 \le e^{(y+y_1)/2} = o(I_n)$ for all $y_1 \le y$ and $y \in \mathcal{K}_C(k_n)$ so that

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

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

$$\begin{aligned}
\mathcal{I}_n^{(2)}(p_1) &= \int_{y^*(p_1)}^{h_1(p_1)} \left(I_n + x_1 - e^{(y_1+y_2)}\right) e^{-\alpha y_2} \, \mathrm{d}y_2 \\
&\le 2I_n \int_{y^*(p_1)}^{h_1(p_1)} e^{-\alpha y_2} \, \mathrm{d}x_2 \, \mathrm{d}y_2 \\
&= \frac{2}{\alpha} I_n \left(I_n^{-2\alpha} e^{\alpha y_1} - (I_n + x_1)^{-2\alpha} e^{-\alpha y_1}\right) \\
&= O\left(I_n^{-2\alpha} x_1 e^{\alpha y_1}\right) = O\left(I_n^{-(2\alpha-1)} e^{\alpha y_1}\right).
\end{aligned}$$

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

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

which establishes the first part of (69).

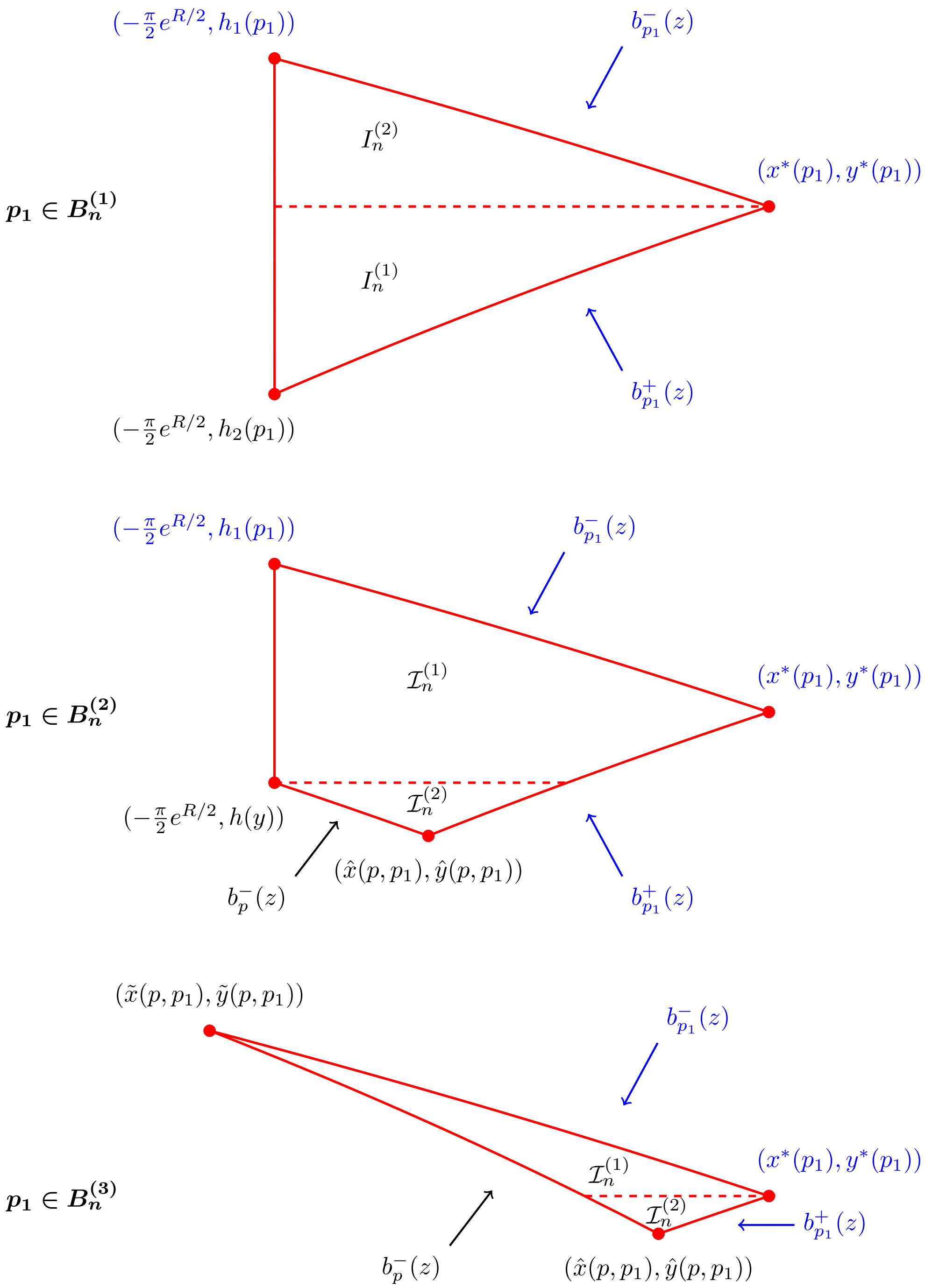

Figure 9: The different shapes of $\mathcal{T}(p, p_1)$ depending on the regime to which p_1 belongs. The top figure is for $p_1 \in B_n^{(1)}$, the middle for $p_1 \in B_n^{(2)}$ and the bottom one for $p_1 \in B_n^{(3)}$.

Case $p_1 \in B_n^{(2)}$: $y - 2\log(I_n/(I_n - x_1)) < y_1 \le y + 2\log\left(1 + \frac{x_1}{I_n}\right)$ 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 us the upper bound on y_1 to compute that

$$
\begin{aligned}
x_1 + e^{(y_1+y_2)/2} - 2I_n + e^{(y+y_2)/2} &\le \left(e^{y/2} + e^{y_1/2}\right) e^{y_2/2} \\
&\le e^{y/2}\left(2 + \frac{x_1}{I_n}\right) e^{y_2/2} = O\left(e^{(y+y_2)/2}\right),
\end{aligned}
$$

since $|x_1| \le I_n$. 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 $p_1 \in B_n^{(2)}$

$$
\mu\left(\mathcal{T}(p, p_1)\right) = O\left(I_n^{-(2\alpha-1)} e^{\alpha y}\right).
$$

Case $p_1 \in B_n^{(3)}$: $y + 2\log(1 + x_1/I_n) < y_1 \leq y + 2\log(I_n/(I_n - x_1))$

$$
\begin{aligned}
\mathcal{I}_n^{(1)} &= \int_{y^*}^{\tilde{y}} \int_{-e^{(y+y_2)/2}}^{x_1 - e^{(y_1+y_2)/2}} f(x_2, y_2)\, \mathrm{d}x_2\, \mathrm{d}y_2 \\
&= O\left(\int_{y^*}^{\tilde{y}} x_1 e^{-\alpha y_2} - \left(e^{y_1/2} - e^{y/2}\right) e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 \right) \\
&= O\left(x_1 \int_{y^*}^{\tilde{y}} e^{-\alpha y_2}\, \mathrm{d}y_2 \right).
\end{aligned}
$$

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, using that $y \leq y_1$ for $p_1 \in B_n^{(3)}$,

$$
\begin{aligned}
\mathcal{I}_n^{(2)} &= \int_{\hat{y}}^{y^*} \int_{-e^{(y+y_2)/2}}^{e^{(y_1+y_2)/2} + x_1 - 2I_n} f(x_2, y_2)\, \mathrm{d}x_2\, \mathrm{d}y_2 \\
&= O\left(\int_{\hat{y}}^{y^*} \left(e^{y/2} + e^{y_1/2}\right) e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 \right) \\
&= O\left(e^{y_1/2} \int_{\hat{y}}^{y^*} e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 \right).
\end{aligned}
$$

For the integral we have

$$
\begin{aligned}
&\int_{\hat{y}}^{y^*} e^{-(\alpha - \frac{1}{2})y_2}\, \mathrm{d}y_2 = \frac{2}{2\alpha - 1}\left(e^{-(\alpha - \frac{1}{2})\hat{y}} - e^{-(\alpha - \frac{1}{2})y^*} \right) \\
&\quad = \frac{2}{2\alpha - 1} \left(\left(\frac{2I_n - x_1}{e^{y/2} + e^{y_1/2}} \right)^{-(2\alpha - 1)} - I_n^{-(2\alpha-1)} e^{(\alpha - \frac{1}{2})y_1} \right) = O\left(I_n^{-(2\alpha-1)} e^{(\alpha - \frac{1}{2})y_1} \right),
\end{aligned}
$$

where we used the upper bound on y_1 and the fact that $2I_n - x_1 = \Theta(I_n)$ for all $x_1 \in [-I_n, I_n]$. We conclude that

$$
\mathcal{I}_n^{(2)} = O\left(I_n^{-(2\alpha-1)} x_1 e^{\alpha y} \right)
$$

and hence for $p_1 \in B_n^{(3)}$

$$
\mu\left(\mathcal{T}(p, p_1)\right) = O\left(I_n^{-2\alpha} x_1 e^{\alpha y}\right) = O\left(I_n^{-(2\alpha-1)} e^{\alpha y} \right).
$$

Integration $\mu(\mathcal{T}(p, p_1))$ with respect to p_1

We now proceed with the second part of the computation leading to (66). Here we will integrate $\mu(\mathcal{T}(p, p'))(p, p_1)$ 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 in the same way for $B_n^{(3)}$

$$B_n^{(3)} = \left\{(x_1, y_1) \,:\, y \le y_1 \le R,\ I_n\left(1 - e^{(y-y_1)/2}\right) \le x_1 \le I_n\left(e^{(y_1-y)/2} - 1\right) \wedge e^{(y+y_1)/2} \wedge I_n\right\}.$$

We these characterizations of the areas we now integrate $\mu(\mathcal{T}(p, p_1))$ over B_n, splitting the computations over the three different areas.

Integration over $B_n^{(1)}$: 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

$$\begin{aligned}
&\int_{B_n^{(1)}} \mu\left(\mathcal{T}(p, p_1)\right) f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 \\
&\quad \le \int_0^y \int_0^{I_n(1-e^{(y_1-y)/2})} \mu\left(\mathcal{T}(p, p_1)\right) f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 \\
&\quad = O\left(I_n^{-2\alpha} \int_0^y \int_0^{e^{(y+y_1)/2}} x_1\, \mathrm{d}x_1\, \mathrm{d}y_1\right) \\
&\quad = O\left(I_n^{-(2\alpha-1)} \int_0^y \left(1 - e^{(y_1-y)/2}\right)^2 \mathrm{d}y_1\right) \\
&\quad = O\left(I_n^{-(2\alpha-1)} y\right) = O\left(y n^{-(2\alpha-1)}\right).
\end{aligned}$$

Integration over $B_n^{(2)}$: We will show that

$$\mu(B_n^{(2)}) = O\left(I_n^{-1} e^{(2-\alpha)y}\right), \tag{70}$$

which together with (69) yields

$$\begin{aligned}
\int_{B_n^{(2)}} \mu\left(\mathcal{T}(p, p_1)\right) f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 &= O\left(\mu(B_n^{(2)}) I_n^{-(2\alpha-1)} e^{\alpha y}\right) \\
&= O\left(I_n^{-2\alpha} e^{2y}\right).
\end{aligned}$$

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 $e^{(y+y_1)/2} - I_n(1-e^{(y_1-y)/2}) \leq e^{y_1/2}\left(e^{y/2}+e^{-y/2}\right)$ to obtain

$$\begin{aligned}
&\int_{h_*(y)}^{y}\int_{I_n(1-e^{(y_1-y)/2})}^{e^{(y+y_1)/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1\\
&\qquad= O\left(e^{y/2}\int_{h_*(y)}^{y} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1\right)\\
&\qquad= O\left(e^{y/2}\left(e^{-(\alpha-\frac{1}{2})y} - e^{-(\alpha-\frac{1}{2})y}\left(\frac{I_n}{I_n+e^y}\right)^{-(2\alpha-1)}\right)\right)\\
&\qquad= O\left(I_n^{-1}e^{(2-\alpha)y}\right).
\end{aligned}$$

For the second integral note that $e^{(y+y_1)/2} - I_n(e^{(y_1-y)/2}-1) \leq e^{(y+y_1)/2}$ and hence

$$\begin{aligned}
&\int_{y}^{h^*(y)}\int_{I_n(e^{(y_1-y)/2}-1)}^{e^{(y+y_1)/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1\\
&\qquad= O\left(e^{y/2}\int_{y}^{h^*(y)} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1\right)\\
&\qquad= O\left(e^{y/2}\left(e^{-(\alpha-\frac{1}{2})y} - e^{-(\alpha-\frac{1}{2})y}\left(\frac{I_n}{I_n-e^y}\right)^{-(2\alpha-1)}\right)\right)\\
&\qquad= O\left(I_n^{-1}e^{(2-\alpha)y}\right),
\end{aligned}$$

so that (70) follows.

Integration over $B_n^{(3)}$: For this case we show that

$$\mu(B_n^{(3)}) = O\left(e^{(1-\alpha)y}\right) \tag{71}$$

so that

$$\begin{aligned}
\int_{B_n^{(3)}} \mu\left(\mathcal{T}(p,p_1)\right) f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 &= O\left(\mu(B_n^{(2)})I_n^{-(2\alpha-1)}e^{\alpha y}\right)\\
&= O\left(I_n^{-(2\alpha-1)}e^{y}\right).
\end{aligned}$$

Here the integral is split into three parts:

$$\begin{aligned}
\mu(B_n^{(3)}) = &\int_{y}^{h^*(y)}\int_{I_n(1-e^{(y-y_1)/2})}^{I_n(e^{(y_1-y)/2}-1)} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1\\
&+\int_{h^*(y)}^{h(y)}\int_{I_n(1-e^{(y-y_1)/2})}^{e^{(y+y_1)/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1\\
&+\int_{h(y)}^{R}\int_{I_n(1-e^{(y-y_1)/2})}^{I_n} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1.
\end{aligned}$$

Let us first focus on the first integral. Since $I_n(e^{(y_1-y)/2}-1) - I_n(1-e^{(y-y_1)/2}) \leq I_n e^{(y_1-y)/2}$ we get, using similar arguments as above

$$\int_{y}^{h^*(y)}\int_{I_n(1-e^{(y-y_1)/2})}^{I_n(e^{(y_1-y)/2}-1)} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 = O\left(I_n e^{-y/2}\int_{y}^{h^*(y)} e^{-(\alpha-\frac{1}{2})y_1}\,\mathrm{d}y_1\right)$$

$$
\begin{aligned}
&= O\left(I_n e^{-\alpha y}\left(1-\left(\frac{I_n}{I_n-e^y}\right)^{-(2\alpha-1)}\right)\right)\\
&= O\left(e^{(1-\alpha)y}\right).
\end{aligned}
$$

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 $c(k;G_{\mathrm{box}})$ (Proving Proposition 5.4)

In this section we establish a concentration result for the local clustering function $c^*(k;G_{\mathrm{box}})$ 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

Recall that $N_{\mathrm{box}}(k_n)$ denotes the number of vertices in G_{box} with degree k_n. We first write

$$
c^*(k_n;G_{\mathrm{box}}) = \frac{T_{\mathrm{box}}(k_n)}{\binom{k_n}{2}\mathbb{E}\left[N_{\mathrm{box}}(k_n)\right]},
$$

where

$$
T_{\mathrm{box}}(k_n) = \sum_{p\in\mathcal{P}}\mathbb{1}_{\{\deg_{\mathrm{box}}(p)=k_n\}}\sum_{(} p_1,p_2)\in\mathcal{P}\setminus\{p\}^{\neq}\mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{box}}(p)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{box}}(p)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{box}}(p_1)\}}
$$

In particular, the variance of $c^*(k_n;G_{\mathrm{box}})$ is determined by the variance of $T_{\mathrm{box}}(k_n)$.

Next, recall the adjusted triangle count function

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

where

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

as well as the definition of $\mathcal{K}_C(k_n)$

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

and write $\mathcal{R}(k_n,C)=[-I_n,I_n]\times\mathcal{K}_C(k_n)$ for the part of the box $\mathcal{R}$ with heights in $\mathcal{K}_C(k_n)$. Slightly abusing notation, we will define the corresponding triangle degree function

$$
\widetilde{T}_{\mathrm{box}}(k_n,C) = \sum_{p\in\mathcal{P}\cap\mathcal{R}(k_n,C)}\mathbb{1}_{\{\deg_{\mathrm{box}}(p)=k_n\}}\widetilde{T}_{\mathrm{box}}(p). \tag{72}
$$

and with that a different clustering function.

$$
\widetilde{c}_{\mathrm{box}}(k_n) = \frac{\widetilde{T}_{\mathrm{box}}(k_n,C)}{\binom{k_n}{2}\mathbb{E}\left[N_{\mathrm{box}}(k_n)\right]}. \tag{73}
$$

The idea is that the main contribution of triangles of degree k_n to the triangle count $T_{\text{box}}(k_n)$ 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 last part 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 first use this result to prove Proposition 5.4. The remainder of this section is devoted to the proof of Proposition 8.1. The final proof can be found in Section 8.3.

Proof of Proposition 5.4. We bound the expectation as follows,

$$\begin{aligned}\mathbb{E}\left[|c^*(k_n; G_{\text{box}}) - \mathbb{E}\left[c^*(k_n; G_{\text{box}})\right]|\right] \leq \;& \frac{\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(k_n, C) - \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right|\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \\ &+ 2\mathbb{E}\left[|c^*(k_n; G_{\text{box}}) - \widetilde{c}_{\text{box}}(k_n)|\right].\end{aligned}$$

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

$$\widetilde{c}_{\text{box}}(k_n) = \frac{\widetilde{T}_{\text{box}}(k_n, C)}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} \leq c^*(k_n; G_{\text{box}}).$$

and therefore

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

For the expectation of $\widetilde{T}_{\text{box}}(k_n, C)$ we use that

$$\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle|\deg_{\text{box}}(p) = k_n\right] = \binom{k_n}{2}\mu\left(\mathcal{B}_{\text{box}}\left(y\right)\right)^{-2}\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right],$$

to get

$$\begin{aligned}\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] &= \int_{\mathcal{R}(k_n,C)} \mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle|\deg_{\text{box}}(p) = k_n\right]\rho_{\text{box}}(y, k_n)f(x, y)\,\mathrm{d}x\,\mathrm{d}y \\ &= (1 + o(1))\binom{k_n}{2}\int_{\mathcal{R}(k_n,C)} \mu\left(\mathcal{B}_{\text{box}}\left(y\right)\right)^{-2}\mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\right]\rho_{\text{box}}(y, k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \\ &= (1 + o(1))\frac{1}{2}\int_{\mathcal{R}(k_n,C)} \mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\right]\rho_{\text{box}}(y, k_n)\alpha e^{-\alpha y}\,\mathrm{d}y \\ &= (1 + o(1))n\binom{k_n}{2}\int_0^\infty P(y)\rho(y, k_n)\alpha e^{-\alpha y}\,\mathrm{d}y,\end{aligned}$$

where the last line is due to Corollary 7.4. In particular, since the last integral is $\Theta\left(k_n^{-(2\alpha+1)}s(k_n)\right)$ we conclude that

$$\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] = \Theta\left(nk_n^{-(2\alpha-1)}s(k_n)\right). \tag{74}$$

Since $\mathbb{E}\left[N_{\text{box}}(k_n)\right] = (1 + o\left(1\right)np_{k_n}$ it follows that

$$\widetilde{c}_{\text{box}}(k_n) = \frac{\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]}{\binom{k_n}{2}\mathbb{E}\left[N_{\text{box}}(k_n)\right]} = (1 + o\left(1\right))\frac{\int_0^\infty P(y)\alpha e^{-\alpha y}\,\mathrm{d}y}{p_{k_n}} = (1 + o(1))\gamma(k_n).$$

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

$$2\mathbb{E}\left[|c^*(k_n; G_{\text{box}}) - \widetilde{c}_{\text{box}}(k_n)|\right] = o\left(\gamma(k_n)\right) = o\left(s(k_n)\right).$$

For the remaining term we use Hölder's inequality and Proposition 8.1 to obtain

$$\begin{aligned}\mathbb{E}\left[\left|\widetilde{T}_{\text{box}}(k_n, C) - \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right|\right] &\le \left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)^2\right] - \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]^2\right)^{\frac{1}{2}} \\ &= o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right).\end{aligned}$$

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. □

We note that the above proof establishes the following important result

Corollary 8.2. *Let $k_n \to \infty$. Then, as $n \to \infty$,*

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

8.2 Joint neighbourhoods and degrees in G_{box}

To prove Proposition 8.1 we need to understand the joint degree distribution in G_{box}. This subsequently requires us to analyse the joint neighbourhoods in G_{box} of two points $p, p' \in \mathcal{R}$. We start with a general result for near independent Poisson random variables.

Lemma 8.3. *Let $k_n \to \infty$ and $X_1 = \text{Po}(\lambda_1(n))$, $X_2 = \text{Po}(\lambda_2(n))$ and $Y = \text{Po}(\lambda_3(n))$, be three Poisson random variables where $\lambda_3(n) = O\left(k_n^{1-\varepsilon}\right)$, for some $0 < \varepsilon < 1$ and for some $C > 0$,*

$$k_n - C\sqrt{k_n \log(k_n)} \le \lambda_i(n) + \lambda_3(n) \le k_n + C\sqrt{k_n \log(k_n)},$$

for $i = 1, 2$. Then, as $n \to \infty$

$$\mathbb{P}\left(X_1 + Y = k_n, X_2 + Y = k_n\right) = (1+o(1))\mathbb{P}\left(X_1 + Y = k_n\right)\mathbb{P}\left(X_2 + Y = k_n\right).$$

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

$$A_n := \left\{t \in \mathbb{R}_+ \,:\, \lambda_3(n) - C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)} \le t \le \lambda_3(n) + C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right\}.$$

Then by a Chernoff bound (c.f. (132))

$$\begin{aligned}&\sum_{t \in \mathbb{R}_+ \setminus A_n} \mathbb{P}\left(X_1 = k_n - t\right)\mathbb{P}\left(X_2 = k_n - t\right)\mathbb{P}\left(Y = t\right) \\ &\quad\le \mathbb{P}\left(Y > \lambda_3(n) + C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right) + \mathbb{P}\left(Y < \lambda_3(n) - C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right) \\ &\quad= \mathbb{P}\left(|\text{Po}(\lambda_3(n)) - \lambda_3(n)| > C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}\right) \le 2k_n^{-\frac{C_1}{4}}.\end{aligned}$$

and hence

$$\mathbb{P}\left(X_1+Y=k_n, X_2+Y=k_n\right)=\sum_{t\in A_n}\mathbb{P}\left(X_1=k_n-t\right)\mathbb{P}\left(X_2=k_n-t\right)\mathbb{P}\left(Y=t\right)+O\left(k_n^{-(1+C_1^2)/2}\right).$$

Next, for $i=1,2$ we have by assumption on $\lambda_i(n)+\lambda_3(n)$ that

$$\begin{aligned}\mathbb{P}\left(X_i+Y=k_n\right)&\geq\frac{\left(k_n+C\sqrt{k_n\log(k_n)}\right)^{k_n}}{k_n!}e^{-\left(k_n+C\sqrt{k_n\log(k_n)}\right)}\\&\geq e^{-1}k_n^{-1/2}\left(1+C\sqrt{\frac{\log(k_n)}{k_n}}\right)^{k_n}e^{-C\sqrt{k_n\log(k_n)}}\\&\geq e^{-1}k_n^{-1/2}e^{k_n\log\left(1+C\sqrt{\frac{\log(k_n)}{k_n}}\right)-C\sqrt{k_n\log(k_n)}}\\&\geq e^{-1}k_n^{-1/2}e^{-\frac{C^2}{2}\log(k_n)}=e^{-1}k_n^{-\frac{1+C^2}{2}}.\end{aligned}$$

where we also used that $\log(1+x)\geq x-x^2/2$, for $0\leq x\leq 1$ and $k_n!\geq e\sqrt{k_n}k_n^{k_n}e^{-k_n}$. Hence, by taking $C_1>4(1+C^2)$ we get that $k_n^{-\frac{C_1}{4}}=o\left(\mathbb{P}\left(X_1+Y=k_n\right)\mathbb{P}\left(X_1+Y=k_n\right)\right)$. It remains to show that

$$\begin{aligned}&\sum_{t\in A_n}\mathbb{P}\left(X_1=k_n-t\right)\mathbb{P}\left(X_2=k_n-t\right)\mathbb{P}\left(Y=t\right)\\&\qquad=(1+o(1))\mathbb{P}\left(X_1+Y=k_n\right)\mathbb{P}\left(X_2+Y=k_n\right).\end{aligned}$$

For this take any $s\in A_n$ so that $|t-s|\leq 2C_1\sqrt{k_n^{1-\varepsilon}\log(k_n)}$ and note that there exists a δ_n satisfying $|\delta_n|\leq 2C\sqrt{k_n\log(k_n)}$, for n large enough, such that $k_n-t=\lambda_1(n)+\delta_n$. It then follows that, uniformly in t,s and δ_n, as $n\to\infty$

$$\begin{aligned}\frac{\mathbb{P}\left(X_2=k_n-t\right)}{\mathbb{P}\left(X_2=k_n-s\right)}&=\frac{\mathbb{P}\left(X_2=k_n-t\right)}{\mathbb{P}\left(X_2=k_n-t-(s-t)\right)}\\&=\frac{(k_n-t-(s-t))!}{(k_n-t)!}\lambda_1(n)^{s-t}\\&\sim(k_n-t-(s-t))^{-(s-t)}\lambda_1(n)^{s-t}\\&=\left(\lambda_1(n)+\delta_n-(s-t)\right)^{-(s-t)}\lambda_1(n)^{s-t}\\&=\left(1+\frac{\delta_n-(s-t)}{\lambda_1(n)}\right)^{s-t}\\&\sim e^{\frac{(s-t)\delta_n}{\lambda_1(n)}}e^{-\frac{(s-t)^2}{\lambda_1(n)}}\sim 1,\end{aligned}$$

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

$$1=\sum_{s=0}^{\infty}\mathbb{P}\left(Y=s\right)=(1+o(1))\sum_{s\in A_n}\mathbb{P}\left(Y=s\right),$$

we conclude that

$$\sum_{t\in A_n}\mathbb{P}\left(X_1=k_n-t\right)\mathbb{P}\left(X_2=k_n-t\right)\mathbb{P}\left(Y=t\right)$$

$$= (1 + o(1)) \sum_{t \in A_n} \mathbb{P}(X_1 = k_n - t) \mathbb{P}(X_2 = k_n - t) \mathbb{P}(Y = t) \sum_{s \in A_n} \mathbb{P}(Y = s)$$
$$= (1 + o(1)) \sum_{t \in A_n} \mathbb{P}(X_1 = k_n - t) \mathbb{P}(Y = t) \sum_{s \in A_n} \mathbb{P}(X_2 = k_n - s) \mathbb{P}(Y = s)$$
$$= (1 + o(1)) \mathbb{P}(X_1 + Y = k_n) \mathbb{P}(X_2 + Y = k_n),$$

from which the result follows. □

To see how this lemma can be applied to analyze the joint degree distribution in G_{box}, fix two points $p, p' \in \mathcal{R}$ and denote by

$$\rho_{\text{box}}(p, p', k, k') := \mathbb{P}(\text{Po}(\mu(\mathcal{B}_{\text{box}}(p))) = k, \text{Po}(\mu(\mathcal{B}_{\text{box}}(p'))) = k'). \tag{75}$$

the joint degree distribution. Then if we define,

$$X_1(p, p') := \text{Po}(\mu(\mathcal{B}_{\text{box}}(p) \setminus \mathcal{B}_{\text{box}}(p'))),$$
$$X_2(p, p') := \text{Po}(\mu(\mathcal{B}_{\text{box}}(p') \setminus \mathcal{B}_{\text{box}}(p))),$$
$$Y(p, p') := \text{Po}(\mu(\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')))$$

it follows that

$$\rho_{\text{box}}(p, p', k_n, k_n) = \mathbb{P}(X_1(p, p') + Y(p, p') = k_n, X_2(p, p') + Y(p, p') = k_n).$$

Now, if $y, y' \in \mathcal{K}_C(k_n)$ the three Poisson random variables defined above satisfy the condition of Lemma 8.3 regarding the sum $\lambda_i(n) + \lambda_3(n)$. Therefore, if in addition $\mu(\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')) = O\left(k_n^{1-\varepsilon}\right)$, for some $0 < \varepsilon < 1$, we have that

$$\rho_{\text{box}}(p, p', k_n, k_n) = (1 + o(1)) \rho_{\text{box}}(p, k_n) \rho_{\text{box}}(p', k_n).$$

To make this more precise, we define, for any $0 < \varepsilon < 1$, the following set

$$\mathcal{E}_\varepsilon(k_n) = \left\{(p, p') \in \mathcal{R} \times \mathcal{R} \,:\, y, y' \in \mathcal{K}_C(k_n) \text{ and } |x - x'|_n > k_n^{1+\varepsilon}\right\}, \tag{76}$$

where $|x|_n = \min\{|x|, \pi e^{R/2} - |x|\}$ denotes the norm on the finite box $\mathcal{R}$ where the left and right boundaries are identified. We will show (see Corollary 8.5) that for all $(p, p') \in \mathcal{E}_\varepsilon(k_n)$ it holds that $\mu(\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')) = O\left(k_n^{1-\varepsilon}\right)$ and hence the joint degree distribution factorizes on this set. We will use this set later in Section 8.3 to prove Proposition 8.1. The main idea behind the above result is that if p and p' are sufficiently separated in the x-direction, then the overlap of their neighbourhoods $\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')$ is of smaller order than $\mu(\mathcal{B}_{\text{box}}(p)) + \mu(\mathcal{B}_{\text{box}}(p'))$. We shall therefore proceed with analyzing the joint neighbourhoods in G_{box}.

Common neighbourhoods

Let $p, p' \in \mathcal{R}$ and denote by $\mathcal{N}_{\text{box}}(p, p')$ the number of common neighbours of p and p'. We shall establish an upper bound on the expected number of joint neighbours when p and p' are sufficiently separated. Observe that $\mathbb{E}[\mathcal{N}_{\text{box}}(p, p')] = \mu(\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p'))$.

We start by analyzing the shape of joint neighbourhood. 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 $x' > 0$. To understand the computation it is helpful to have a picture. Figure 10 shows such an example. There are several different quantities that are important. The first are the heights where the left and right boundaries of the ball $\mathcal{B}_{\text{box}}(p)$ hit the boundaries of the box $\mathcal{R}$. Since $x = 0$ these heights are the same and we denote their common value by $h(y)$. We also need to know the coordinates $\hat{y}_{\text{right}}(p, p')$ and $\hat{x}_{\text{right}}(p, p')$ of the intersection of the right boundary of the neighbourhood of p with the left boundary of the neighbourhood of p' and those for the intersection of the left boundary of the neighbourhood of p with the right boundary of the neighbourhood of p', which we denote by $\hat{y}_{\text{left}}(p, p')$ and $\hat{x}_{\text{left}}(p, p')$. Finally we

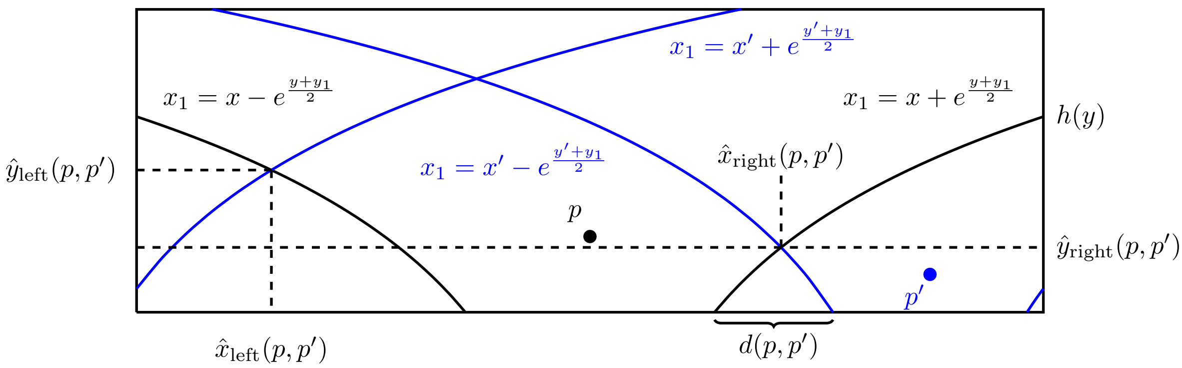

Figure 10: Schematic representation of the neighbourhoods of p and p' in G_{box} when $|x-x'| > e^{\frac{y}{2}} + e^{\frac{y'}{2}}$ used for the proof of Lemma 8.4. Note that although here $p' \notin \mathcal{B}_{\text{box}}(p)$, this is not true in general. This situation was merely chosen to improve readability of the figure.

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 neighbourhoods do not intersect, as is the case in Figure 10. The condition $d(p,p') > 0$ is exactly the right notion for p and p' being sufficiently separated.

Note that $\hat{y}_{\text{left}}(p,p')$ and $\hat{x}_{\text{left}}(p,p')$ correspond to, respectively, $\hat{y}(p,p')$ and $\hat{x}(p,p')$ considered in Section 7.2. The derivation of $\hat{y}_{\text{right}}(p,p')$ and $\hat{x}_{\text{right}}(p,p')$ is done in a similar manner and we omit the details here. The full expressions of all these functions are given below for further reference.

$$h(y) = R - y + 2\log\left(\frac{\pi}{2}\right) \tag{77}$$

$$h_1(p') = 2\log\left(x' + \frac{\pi}{2}e^{\frac{R}{2}}\right) - y' \tag{78}$$

$$h_2(p') = 2\log\left(\frac{\pi}{2}e^{\frac{R}{2}} - x'\right) - y' \tag{79}$$

$$\hat{y}_{\text{right}}(p,p') = 2\log\left(\frac{x'}{e^{\frac{y}{2}} + e^{\frac{y'}{2}}}\right) \tag{80}$$

$$\hat{x}_{\text{right}}(p,p') = \frac{x'}{1 + e^{\frac{y'-y}{2}}}, \tag{81}$$

$$\hat{y}_{\text{left}}(p,p') = 2\log\left(\frac{\pi e^{R/2} - x'}{e^{\frac{y}{2}} + e^{\frac{y'}{2}}}\right), \tag{82}$$

$$\hat{x}_{\text{left}}(p,p') = \frac{x' - \pi e^{R/2}}{1 + e^{\frac{y'-y}{2}}}, \tag{83}$$

$$d(p,p') = |x - x'|_n - \left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right). \tag{84}$$

The following result shows that if $d(p,p') > 0$, then the expected number of common neighbours is $o\left(\mu\left(\mathcal{B}_{\text{box}}(p)\right) + \mu\left(\mathcal{B}_{\text{box}}(p')\right)\right)$.

Lemma 8.4. *Let $p, p' \in \mathcal{R}$. Then, whenever $|x - x'|_n > \left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)$,*

$$\mathbb{E}\left[\mathcal{N}_{box}(p,p')\right] \leq \left(\mu\left(\mathcal{B}_{box}(p)\right) + \mu\left(\mathcal{B}_{box}(p')\right)\right)\phi_n(p,p'),$$

where

$$\phi_n(p,p') = 2\left(\frac{|x-x'|}{e^{\frac{y}{2}} + e^{\frac{y'}{2}}}\right)^{-(2\alpha-1)} + \frac{3\nu^{2\alpha+1}e^{-(\alpha-\frac{1}{2})R}e^{\alpha y}}{2\pi^{2\alpha}\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)} + \frac{\nu e^{-(\alpha-\frac{1}{2})R}}{e^{\frac{y}{2}} + e^{\frac{y'}{2}}}.$$

Proof. Again, without loss of generality we assume that $p = p_0 = (0, y)$ and $p' = (x', y')$ with $0 \leq x' \leq \frac{\pi}{2} e^{R/2}$. Consider the following two areas

$$A_1 := \left\{ p_1 = (x_1, y_1) \in \mathcal{R} \,:\, y_1 \geq \hat{y}_{\mathrm{left}}(p_0, p'),\, -\frac{\pi}{2} e^{R/2} \vee -e^{\frac{y+y_1}{2}} \leq x_1 \leq x' + e^{\frac{y'+y_2}{2}} - \pi e^{R/2} \right\}$$

$$A_2 := \left\{ p_1 = (x_1, y_1) \in \mathcal{R} \,:\, y_1 \geq \hat{y}_{\mathrm{right}}(p_0, p'),\, x' - e^{\frac{y'+y_1}{2}} \leq x_1 \leq \frac{\pi}{2} e^{R/2} \wedge e^{\frac{y+y_2}{2}} \right\}.$$

That is A_1 and A_2 describe all points $p_1 \in \mathcal{B}_{\mathrm{box}}(p_0) \cap \mathcal{B}_{\mathrm{box}}(p)$ with $y_1 \geq \hat{y}_{\mathrm{left}}(p_0, p')$ and $y_1 \geq \hat{y}_{\mathrm{right}}(p_0, p')$, respectively. It now follows that

$$\mathbb{E}\left[\mathcal{N}_{\mathrm{box}}(p, p')\right] \leq \mu(A_1) + \mu(A_2),$$

and we will proceed by computing the measures of both sets.

For A_1 we have (see also Figure 6)

$$\mu(A_1) = \int_{\hat{y}_{\mathrm{left}}(p_0,p')}^{h(y)} \int_{-e^{\frac{y+y_1}{2}}}^{x'+e^{\frac{y'+y_1}{2}} - \pi e^{R/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 + \int_{h(y)}^{R} \int_{-\frac{\pi}{2} e^{R/2}}^{x'+e^{\frac{y'+y_1}{2}} - \pi e^{R/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1.$$

For the first integral we get

$$\begin{aligned}
&\int_{\hat{y}_{\mathrm{left}}(p_0,p')}^{h(y)} \int_{-e^{\frac{y+y_1}{2}}}^{x'+e^{\frac{y'+y_1}{2}} - \pi e^{R/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 \\
&= \frac{\alpha\nu}{\pi}\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right) \int_{\hat{y}_{\mathrm{left}}(p_0,p')}^{h(y)} e^{-(\alpha - \frac{1}{2})y_1}\, \mathrm{d}y_1 - \frac{\alpha\nu}{\pi}\left(\pi e^{R/2} - x'\right) \int_{\hat{y}_{\mathrm{left}}(p_0,p')}^{h(y)} e^{-\alpha y_1}\, \mathrm{d}y_1 \\
&= \xi \left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)\left(e^{-(\alpha-\frac{1}{2})\hat{y}_{\mathrm{left}}(p_0,p')} - e^{-(\alpha-\frac{1}{2})h(y)}\right) - \frac{\nu}{\pi}\left(\pi e^{R/2} - x'\right)\left(e^{-\alpha \hat{y}_{\mathrm{left}}(p_0,p')} - e^{-\alpha h(y)}\right).
\end{aligned}$$

Similarly, the second integral equals

$$\begin{aligned}
&\int_{h(y)}^{R} \int_{-\frac{\pi}{2} e^{R/2}}^{x'+e^{\frac{y'+y_1}{2}} - \pi e^{R/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 \\
&\quad = \xi e^{\frac{y'}{2}}\left(e^{-(\alpha-\frac{1}{2})h(y)} - e^{-(\alpha-\frac{1}{2})R}\right) - \frac{\nu}{\pi}\left(\pi e^{R/2} - x'\right)\left(e^{-\alpha h(y)} - e^{-\alpha R}\right).
\end{aligned}$$

Adding these expressions yields

$$\begin{aligned}
\mu(A_1) &= \xi\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right) e^{-(\alpha-\frac{1}{2})\hat{y}_{\mathrm{left}}(p_0,p')} - \xi e^{\frac{y}{2}} e^{-(\alpha-\frac{1}{2})h(y)} - \xi e^{\frac{y'}{2}} e^{-(\alpha-\frac{1}{2})R} \\
&\quad - \frac{\nu}{\pi}\left(\pi e^{R/2} - x'\right) e^{-\alpha \hat{y}_{\mathrm{left}}(p_0,p')} + \nu e^{R/2} e^{-\alpha h(y)} + \frac{\nu}{\pi}\left(\pi e^{R/2} - x'\right) e^{-\alpha R} \\
&\leq \xi\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right) e^{-(\alpha-\frac{1}{2})\hat{y}_{\mathrm{left}}(p_0,p')} + \nu e^{R/2} e^{-\alpha h(y)} + \nu e^{-(\alpha-\frac{1}{2})R}.
\end{aligned}$$

We now proceed with A_2. Here we have

$$\mu(A_2) = \int_{\hat{y}_{\mathrm{right}}(p_0,p')}^{h(y)} \int_{x'-e^{\frac{y'+y_1}{2}}}^{e^{\frac{y+y_1}{2}}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1 + \int_{h(y)}^{R} \int_{x'-e^{\frac{y'+y_1}{2}}}^{\frac{\pi}{2} e^{R/2}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1.$$

Computing each integral separately we get

$$\int_{\hat{y}_{\mathrm{right}}(p_0,p')}^{h(y)} \int_{x'-e^{\frac{y'+y_1}{2}}}^{e^{\frac{y+y_1}{2}}} f(x_1, y_1)\, \mathrm{d}x_1\, \mathrm{d}y_1$$

$$= \xi \left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)\left(e^{-(\alpha-\frac{1}{2})\hat{y}_{\text{right}}(p_0,p')} - e^{-(\alpha-\frac{1}{2})h(y)}\right) - \frac{\nu}{\pi}x'\left(e^{\alpha\hat{y}_{\text{right}}(p_0,p')} - e^{-\alpha h(y)}\right),$$

and

$$\int_{h(y)}^{R}\int_{x'-e^{\frac{y'+y_1}{2}}}^{\frac{\pi}{2}e^{R/2}} f(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1$$
$$= \xi e^{\frac{y'}{2}}\left(e^{-(\alpha-\frac{1}{2})h(y)} - e^{-(\alpha-\frac{1}{2})R}\right) + \frac{\nu}{\pi}\left(\frac{\pi}{2}e^{R/2} - x'\right)\left(e^{-\alpha h(y)} - e^{-\alpha R}\right).$$

Combining these results we get

$$\begin{aligned}\mu(A_2) &= \xi\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)e^{-(\alpha-\frac{1}{2})\hat{y}_{\text{right}}(p_0,p')} - \xi e^{\frac{y}{2}}e^{-(\alpha-\frac{1}{2})h(y)} - \xi e^{\frac{y'}{2}}e^{-(\alpha-\frac{1}{2})R}\\ &\quad - \frac{\nu}{\pi}x'e^{-\alpha\hat{y}_{\text{right}}(p_0,p')} + \frac{\nu}{2}e^{R/2}e^{-\alpha h(y)} - \frac{\nu}{\pi}\left(\frac{\pi}{2}e^{R/2} - x'\right)e^{-\alpha R}\\ &\le \xi\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)e^{-(\alpha-\frac{1}{2})\hat{y}_{\text{right}}(p_0,p')} + \frac{\nu}{2}e^{R/2}e^{-\alpha h(y)}.\end{aligned}$$

Finally, by adding the results for $\mu(A_1)$ and $\mu(A_2)$ and using that $\xi e^{\frac{y}{2}} = \mu(\mathcal{B}_{\text{box}}(p_0))$ and $\xi e^{\frac{y'}{2}} = \mu(\mathcal{B}_{\text{box}}(p'))$, we get

$$\begin{aligned}\mathbb{E}[\mathcal{N}_{\text{box}}(p,p')] &\le \left(\mu(\mathcal{B}_{\text{box}}(p)) + \mu(\mathcal{B}_{\text{box}}(p'))\right)\left(e^{-(\alpha-\frac{1}{2})\hat{y}_{\text{left}}(p,p')} + e^{-(\alpha-\frac{1}{2})\hat{y}_{\text{right}}(p,p')}\right)\\ &\quad + \frac{3\nu}{2}e^{R/2}e^{-\alpha h(y)} + \nu e^{-(\alpha-\frac{1}{2})R}.\end{aligned}$$

To finish the argument we recall that $\hat{y}_{\text{left}}(p,p') \ge \hat{y}_{\text{right}}(p,p')$ for all $0 \le x' \le \frac{\pi}{2}e^{R/2}$ and that $h(y) = R - y + 2\log(\pi/\nu)$. Hence

$$\begin{aligned}&\mathbb{E}[\mathcal{N}_{\text{box}}(p,p')]\\ &\le \left(\mu(\mathcal{B}_{\text{box}}(p)) + \mu(\mathcal{B}_{\text{box}}(p'))\right)\left(2e^{-(\alpha-\frac{1}{2})\hat{y}_{\text{right}}(p,p')} + \frac{3\nu^{2\alpha+1}e^{-(\alpha-\frac{1}{2})R}e^{\alpha y}}{2\pi^{2\alpha}\left(e^{\frac{y}{2}} + e^{\frac{y'}{2}}\right)} + \frac{\nu e^{-(\alpha-\frac{1}{2})R}}{\xi(e^{\frac{y}{2}} + e^{\frac{y'}{2}})}\right)\end{aligned}$$

and the result follows by plugging in

$$\hat{y}_{\text{right}}(p,p') = 2\log\left(\frac{x'}{e^{\frac{y}{2}} + e^{\frac{y'}{2}}}\right),$$

and noting that x' is the same as $|x - x'|$, by our generalization step.

□

Degrees

We now return to the joint degree distribution of nodes in G_{box}. Recall the definition of $\mathcal{E}_\varepsilon(k_n)$

$$\mathcal{E}_\varepsilon(k_n) = \left\{(p,p') \in \mathcal{R}\times\mathcal{R} \,:\, y,y' \in \mathcal{K}_C(k_n) \text{ and } |x - x'|_n > k_n^{1+\varepsilon}\right\}.$$

The following result, which follows from Lemma 8.4, shows that on this set, the expected number of common neighbours is $o(k_n)$.

Lemma 8.5. *Fix $0 < \varepsilon < 1$ and let $\varepsilon' = \min\{\varepsilon(2\alpha - 1), \varepsilon\}$. Then for all $(p,p') \in \mathcal{E}_\varepsilon(k_n)$, as $n \to \infty$,*

$$\mu(\mathcal{B}_{box}(p) \cap \mathcal{B}_{box}(p')) = O\left(k_n^{1-\varepsilon'}\right).$$

Proof. Since for all $(p,p') \in \mathcal{E}_\varepsilon(k_n)$ we have $\mu\left(\mathcal{B}_{\text{box}}\left(p\right)\right), \mu\left(\mathcal{B}_{\text{box}}\left(p'\right)\right) = \Theta\left(k_n\right)$, Lemma 8.4 implies that

$$\mu\left(\mathcal{B}_{\text{box}}\left(p\right) \cap \mathcal{B}_{\text{box}}\left(p'\right)\right) \leq O\left(k_n\right) \phi_n(p,p'),$$

where

$$\phi_n(p,p') = 2\left(\frac{|x-x'|}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}\right)^{-(2\alpha-1)} + \frac{3\nu^{2\alpha+1}e^{-(\alpha-\frac{1}{2})R}e^{\alpha y}}{2\pi^{2\alpha}\left(e^{\frac{y}{2}}+e^{\frac{y'}{2}}\right)} + \frac{\nu e^{-(\alpha-\frac{1}{2})R}}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}.$$

We thus need to show that $\phi_n(p,p') = O\left(k_n^{-\varepsilon}\right)$. For $(p,p') \in \mathcal{E}_\varepsilon(k_n)$, it holds that $e^{y/2}, e^{y'/2} = \Theta\left(k_n\right)$ and $|x-x'| > k_n^{1+\varepsilon}$ and hence

$$2\left(\frac{|x-x'|}{e^{\frac{y}{2}}+e^{\frac{y'}{2}}}\right)^{-(2\alpha-1)} = O\left(k_n^{-\varepsilon(2\alpha-1)}\right).$$

For the second term in $\phi_n(p,p')$ we use that $e^{\alpha y^*} = \Theta\left(k_n^{2\alpha}\right)$ and $e^R = \Theta\left(n^2\right)$ to obtain

$$\frac{3\nu^{2\alpha+1}e^{-(\alpha-\frac{1}{2})R}e^{\alpha y}}{2\pi^{2\alpha}\left(e^{\frac{y}{2}}+e^{\frac{y'}{2}}\right)} = O\left(1\right) n^{-(2\alpha-1)} k_n^{2\alpha-1} = O\left(n^{-(\alpha-\frac{1}{2})}\right).$$

Finally, the third term satisfies $O\left(n^{-(2\alpha-1)}k_n^{-1}\right)$, and we conclude that

$$\phi_n(p,p') = O\left(k_n^{-\varepsilon(2\alpha-1)} + n^{-(\alpha-\frac{1}{2})} + n^{-(2\alpha-1)}k_n^{-1}\right) = O\left(k_n^{-\varepsilon'}\right),$$

where we used that $\varepsilon' = \min\{\varepsilon(2\alpha-1), \varepsilon\}$. □

As a corollary we get that on the set $\mathcal{E}_\varepsilon(k_n)$ the joint degree distribution in G_{box} is asymptotically equivalent to the product of the degree distributions. We shall however prove a slightly stronger result (Lemma 8.7) which also takes care of bounded shifts in the joint degree distribution $\rho_{\text{box}}(p,p',k_n-t,k_n-t')$, for some uniformly bounded $t,t' \in \mathbb{Z}$. For this we first need the following simple result for Poisson distributions.

Lemma 8.6. *Let $k_n \to \infty$ be a sequence of non-negative integers and $X = \text{Po}(\lambda_n)$ be a Poisson random variable with mean λ_n satisfying*

$$k_n - C\sqrt{k_n \log(k_n)} \leq \lambda_n \leq k_n + C\sqrt{k_n \log(k_n)}$$

for some $C > 0$. Then, for any $t_n, s_n = O(1)$, as $n \to \infty$,

$$\mathbb{P}\left(X = k_n - t_n\right) \sim \mathbb{P}\left(X = k_n - s_n\right).$$

Proof. Note that $k_n > t_n, s_n$ for large enough n. Hence, using Stirling's formula, as $n \to \infty$,

$$\begin{aligned}
\frac{\mathbb{P}\left(X = k_n - t_n\right)}{\mathbb{P}\left(X = k_n - s_n\right)} &= \frac{(k_n - t_n - (s_n - t_n))!}{(k_n-t_n)!}\lambda_n^{s_n-t_n} \\
&\sim \sqrt{\frac{k_n-s_n}{k_n-t_n}}\,\frac{(k_n-s_n)^{k_n-s_n}}{(k_n-t_n)^{k_n-t_n}}\,e^{t_n-s_n}\lambda_n^{s_n-t_n} \\
&= \sqrt{\ell_n}(\ell_n)^{k_n-t_n}e^{t_n-s_n}(k_n-s_n)^{t_n-s_n}\lambda_n^{s_n-t_n} \\
&= \sqrt{\ell_n}e^{(k_n-t_n)\log(\ell_n)+t_n-s_n}\left(\frac{k_n-s_n}{\lambda_n}\right)^{t_n-s_n}
\end{aligned}$$

where we wrote $\ell_n = (k_n-s_n)/(k_n-t_n)$. Note that $\ell_n \to 1$ and hence $\sqrt{\ell_n} \to 1$. Moreover, since $(k_n-s_n)/\lambda_n \to 1$ and $|s_n - t_n| = O\left(1\right)$ we have that $\left(\frac{k_n-s_n}{\lambda_n}\right)^{t_n-s_n} \sim 1$ Therefore it remains to show that

$$\lim_{n\to\infty} e^{(k_n-t_n)\log(\ell_n)+t_n-s_n} = 1.$$

For this we note that for any x, such that $|x| \le 1/2$, we have

$$x - x^2 \le \log(1+x) \le x.$$

Write $x_n = \ell_n - 1 = \frac{t_n - s_n}{k_n - t_n}$. Then by the assumptions of the lemma, $x_n \to 0$, and thus, for n large enough,

$$t_n - s_n - \frac{(t_n - s_n)^2}{k_n - t_n} \le (k_n - t)\log(\ell_n) \le t_n - s_n.$$

In particular

$$e^{-\frac{(t_n-s_n)^2}{k_n-t_n}} \le e^{(k_n-t_n)\log(\ell_n)+t_n-s_n} \le 1,$$

and the result follows since $\frac{(t_n-s_n)^2}{k_n-t_n} \to 0$. $\square$

We can now prove the main result of this section.

Lemma 8.7. *Let $0 < \varepsilon < 1$, $k_n \to \infty$ and let $t_n, t'_n, s_n, s'_n \in \mathbb{Z}$ be uniformly bounded. Then for any $(p, p') \in \mathcal{E}_\varepsilon(k_n)$, as $n \to \infty$,*

$$\rho_{\text{box}}(p, p', k_n - t_n, k_n - t'_n) = (1 + o(1))\rho_{\text{box}}(p, k_n - s_n)\rho_{\text{box}}(p', k_n - s'_n).$$

Proof. 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

$$\rho_{\text{box}}(p, p', k_n - t_n, k_n - t'_n) = \mathbb{P}\left(X_1(p,p') + Y(p,p') = k_n - t_n, X_2(p,p') + Y(p,p') = k_n - t'_n\right).$$

Since by Lemma 8.5 $\mu\left(\mathcal{B}_{\text{box}}(p) \cap \mathcal{B}_{\text{box}}(p')\right) = O\left(k_n^{1-\varepsilon'}\right)$, it follows from Lemma 8.3 that

$$\rho_{\text{box}}(p, p', k_n - t_n, k_n - t'_n) = (1 + o(1))\rho_{\text{box}}(p, k_n - t_n)\rho_{\text{box}}(p', k_n - t'_n).$$

The result then follows by applying Lemma 8.6 twice. $\square$

8.3 Concentration result 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 (72)) that for any $C > 0$

$$\widetilde{T}_{\text{box}}(k_n, C) = \sum_{p \in \mathcal{P}_n \cap \mathcal{K}_{C,n}(k_n)} \mathbb{1}_{\{\deg_{\text{box}}(p) = k\}} \widetilde{T}_{\text{box}}(p)$$

Then we have

$$\widetilde{T}_{\text{box}}(k_n, C)^2 = \sum_{p,p' \in \mathcal{P}_n \cap \mathcal{K}_C(k_n)} \mathbb{1}_{\{\deg_{\text{box}}(p),\, \deg_{\text{box}}(p') = k_n\}} \sum_{(p_1,p_2),(p'_1,p'_2) \in \mathcal{P}_n}^{\neq} \widetilde{T}_{\mathcal{P}}(p, p_1, p_2)\widetilde{T}_{\mathcal{P}}(p', p'_1, p'_2),$$

This expression can be written as the sum 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 $b \in \{0, 1, 2\}$,

$$I_{a,b} = \sum_{\substack{p,p' \in \mathcal{P}_n \cap \mathcal{K}_C(k) \\ |\{p\} \cap \{p'\}| = a}} \mathbb{1}_{\{\deg_{\text{box}}(p),\, \deg_{\text{box}}(p') = k_n\}} J_b(p, p'),$$

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 sum taken 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

$$\mathbb{E}\left[I_{0,0}\right] = (1+o(1))\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]^2 \tag{85}$$

and for all other combinations

$$\mathbb{E}\left[I_{a,b}\right] = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]^2\right). \tag{86}$$

Note $I_{1,2} = \widetilde{T}_{\text{box}}(k_n,C)$ and since (74) implies that $\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right] \to \infty$, it follows that (86) holds for $I_{1,2}$.

Recall that $\mathcal{R}(k_n,C) = [-I_n,I_n]\times\mathcal{K}_C(k_n)$ and (76)

$$\mathcal{E}_\varepsilon(k_n) = \left\{(p,p')\in\mathcal{R}\times\mathcal{R}\,:\, y,y'\in\mathcal{K}_C(k_n) \text{ and } |x-x'|_n > k_n^{1+\varepsilon}\right\}.$$

Let $\mathcal{E}_\varepsilon(k_n)^c$ be the same set but with $|x-x'|_n \le k_n^{1+\varepsilon}$ and denote by $I_{a,b}^*$ the the part of $I_{a,b}$ where $(p,p')\in\mathcal{E}_\varepsilon(k_n)$. Will split the analysis between $I_{a,b}^*$ and $I_{a,b}-I_{a,b}^*$. The idea for these two cases is that by Lemma 8.7 it follows that on the set $\mathcal{E}_\varepsilon(k_n)$ and for any uniformly bounded $t,t'\in\mathbb{Z}$, the joint degree distribution factorizes,

$$\rho_{\text{box}}(p,p',k_n+t,k_n+t') = (1+o(1))\rho_{\text{box}}(p,k_n)\rho_{\text{box}}(p,k_n).$$

In particular this allows us to prove that $\mathbb{E}\left[I_{0,0}^*\right] = (1+o(1))\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]^2$. On the other hand, the expected number of points in $\mathcal{E}_\varepsilon(k_n)^c$ is $O\left(k_n^{1+\varepsilon}k_n^{-2\alpha}\mathbb{E}\left[N_{\text{box}}(k_n)\right]\right) = o\left(\mathbb{E}\left[N_{\text{box}}(k_n)\right]^2\right)$, where the latter is the expected number of points in $\mathcal{R}(k_n,C)\times\mathcal{R}(k_n,C)$. Hence we expect the contributions coming from $\mathcal{E}_\varepsilon(k_n)^c$ to be negligible.

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 $\widetilde{D}_{\text{box}}(p,p',k,\ell)$ to denote the indicator that $|\mathcal{B}_{\text{box}}(p)\cap(\mathcal{P}\setminus\{p,p',p_1,p_2,p_1',p_2'\})| = k$ and $|\mathcal{B}_{\text{box}}(p')\cap(\mathcal{P}\setminus\{p,p',p_1,p_2,p_1',p_2'\})| = \ell$. Note that this also depend on $\{p_1,p_2,p_1',p_2'\}$ but we suppressed this to keep notation concise. Similarly we write $D_{\text{box}}(p,p',k,\ell)$ to denote the indicator that $|\mathcal{B}_{\text{box}}(p)\cap(\mathcal{P}\setminus\{p,p'\})| = k$ and $|\mathcal{B}_{\text{box}}(p')\cap(\mathcal{P}\setminus\{p,p'\})| = \ell$, which now only depends on p and p'. Then, by the Campbell-Mecke formula

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p)=k_n,\deg_{\text{box}}(p')=k_n\}}J_b(p,p')\right] \\
&= \mathbb{E}\left[\sum_{\substack{p_1,p_2,p_1',p_2'\in\mathcal{P}_n \\ |\{p_1,p_2\}\cap\{p_1',p_2'\}|=b}}^{\neq} \widetilde{D}_{\text{box}}(p,p',k_n-i,k_n-i')\,\widetilde{T}_{\text{box}}(p,p_1,p_2)\widetilde{T}_{\text{box}}(p',p_1',p_2')\right]
\end{aligned}$$

where the sum is over all distinct pairs (p_1,p_2) and (p_1',p_2'). We also know that

$$\mathbb{E}\left[T_{\mathcal{P}}(k_n)\right] = \Theta\left(nk_n^{-(2\alpha-1)}s_\alpha(k_n)\right).$$

We will now proceed to establish (85) and (86).

Computing $I_{0,0}$ We first show that

$$\mathbb{E}\left[I_{0,0} - I_{0,0}^*\right] = o\left(\mathbb{E}\left[T_{\text{box}}(k_n, C)\right]^2\right), \tag{87}$$

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 8.7.

For J_0 we have, using Lemma 8.7

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p)=k_n, \deg_{\text{box}}(p')=k_n\}} J_0(p,p')\right] \\
&= \mathbb{E}\left[\sum^{\neq}_{\substack{p_1,p_2,p_1',p_2' \in \mathcal{P}\setminus\{p,p'\} \\ |\{p_1,p_2\}\cap\{p_1',p_2'\}|=0}} \widetilde{D}_{\text{box}}(p,p',k_n-i,k_n-i')\, \widetilde{T}_{\text{box}}(p,p_1,p_2)\widetilde{T}_{\text{box}}(p',p_1',p_2')\right] \\
&= \mathbb{E}\left[D_{\text{box}}(p,p',k_n-j-2,k_n-j'-2) \sum^{\neq}_{p_1,p_2\in\mathcal{P}\setminus p} \widetilde{T}_{\text{box}}(p,p_1,p_2) \sum^{\neq}_{p_1',p_2'\in\mathcal{P}\setminus p'} \widetilde{T}_{\text{box}}(p',p_1',p_2')\right] \\
&= (1+o(1))\rho_{\text{box}}(p,p',k_n,k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle|\deg_{\text{box}}(p)=k_n\right]\mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\middle|\deg_{\text{box}}(p')=k_n\right],
\end{aligned}$$

Next we recall that for all $y' \in \mathcal{K}_C(k_n)$ (see (46)),

$$\mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\middle|\deg_{\text{box}}(p')=k_n\right] = \binom{k_n}{2}\mu\left(\mathcal{B}_{\text{box}}(p')\right)^{-2}\mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\right] = O(1)\, k_n^2 P(y'),$$

where $p' = (x', y')$ and we used that $\mathbb{E}\left[\widetilde{T}_{\text{box}}(p')\right] = (1+o(1))k_n^2 P(y')$, for all $y' \in \mathcal{K}_C(k_n)$. Therefore, using that $\rho_{\text{box}}(p,p',k_n,k_n) \le \rho_{\text{box}}(p,k_n)$,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p)=k_n, \deg_{\text{box}}(p')=k_n\}} J_0(p,p')\right] \\
&\le O\left(k_n^2\right)\rho_{\text{box}}(p,k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle|\deg_{\text{box}}(p)=k_n\right] P(y')
\end{aligned}$$

and thus

$$\begin{aligned}
&\mathbb{E}\left[I_{0,0} - I_{0,0}^*\right] \\
&= \int_{\mathcal{E}_\varepsilon(k_n)^c} \mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p), \deg_{\text{box}}(p')=k_n\}} J_0(p,p')\right] f(x,y)f(x',y')\, dx'\, dx\, dy'\, dy \\
&\le O\left(k_n^2\right) k_n^{1+\varepsilon}\left(\int_{a_n^-}^{a_n^+} P(y')e^{-\alpha y'}\, dy'\right)\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right] \\
&= O\left(k_n^{3+\varepsilon-2\alpha} s_\alpha(k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]\right) \\
&= o\left(n k_n^{-(2\alpha-1)} s_\alpha(k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]\right) = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right]^2\right),
\end{aligned}$$

which proves (87). Here we used that $k_n^{2+\varepsilon} = o(n)$ and

$$\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n,C)\right] = \Theta\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n)\right]\right) = \Theta\left(n k_n^{-(2\alpha-1)} s_\alpha(k_n)\right)$$

for the last line.

We will now show that

$$\mathbb{E}\left[I_{0,0}^*\right] = (1+o(1))\mathbb{E}\left[\mathbb{E}\left[T_{\text{box}}(k_n,C)\right]^2\right].$$

Recall the result from Lemma 8.7, that for $(p,p') \in \mathcal{E}_\varepsilon(k_n)$ and any two uniformly bounded $t, t' \in \mathbb{Z}$,

$$\rho_{\text{box}}(p,p',k_n+t,k_n+t') = (1+o(1))\rho_{\text{box}}(p,k_n)\rho_{\text{box}}(p,k_n).$$

Therefore, by defining $h(y) = \mathbb{E}\left[\widetilde{T}_{\text{box}}(y)\middle| \deg_{\text{box}}(y) = k_n\right]$

$$\mathbb{E}\left[I^*_{0,0}\right] = (1 + o\,(1)) \int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p, k_n)\rho_{\text{box}}(p', k_n)h(y)h(y')f(x, y)f(x', y')\,dx'\,dx\,dy'\,dy.$$

The difference with $\mathbb{E}\left[\mathbb{E}\left[T_{\text{box}}(k_n, C)\right]^2\right]$ is in that the above integral is over $\mathcal{E}_\varepsilon(k_n)$ instead of $\mathcal{R}(k_n, C) \times \mathcal{R}(k_n, C)$. Since the difference between the two sets is $\mathcal{E}_\varepsilon(k_n)^c$ and $nk_n^{1+\varepsilon} = o\left(n^2\right)$ it follows that

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[T_{\text{box}}(k_n, C)\right]^2\right] - \int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p, k_n)\rho_{\text{box}}(p', k_n)h(y)h(y')f(x, y)f(x', y')\,\mathrm{d}x'\,\mathrm{d}x\,\mathrm{d}y'\,\mathrm{d}y \\
&= \int_{\mathcal{E}_\varepsilon(k_n)^c} \rho_{\text{box}}(p, k_n)\rho_{\text{box}}(p', k_n)h(y)h(y')f(x, y)f(x', y')\,\mathrm{d}x'\,\mathrm{d}x\,\mathrm{d}y'\,\mathrm{d}y \\
&= O\left(k_n^{1+\varepsilon}n\right)\left(\int_{\mathcal{K}_C(k_n)} h(y)\rho_{\text{box}}(y, k_n)\alpha e^{-\alpha y}\,\mathrm{d}y\right)^2 = o\left(\mathbb{E}\left[\mathbb{E}\left[T_{\text{box}}(k_n, C)\right]^2\right]\right).
\end{aligned}$$

Thus we conclude that $\mathbb{E}\left[I^*_{0,0}\right] = (1 + o\,(1))\mathbb{E}\left[\mathbb{E}\left[T_{\text{box}}(k_n, C)\right]^2\right]$, which finishes the proof of (85).

Computing $\mathbb{E}\left[I_{0,1}\right]$ We first write

$$\mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p)=k_n, \deg_{\text{box}}(p')=k_n\}} J_1\right] \leq O\,(1)\,k_n\rho_{\text{box}}(p, p', k_n, k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle| \deg_{\text{box}}(p) = k_n\right] \tag{88}$$

Then, using that $\rho_{\text{box}}(p, p', k_n, k_n) \leq \rho_{\text{box}}(p, k_n)$,

$$\begin{aligned}
&\mathbb{E}\left[I_{0,1} - I^*_{0,1}\right] \\
&\int_{\mathcal{E}_\varepsilon(k_n)^c} \mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p), \deg_{\text{box}}(p')=k_n\}} J_1(p, p')\right] f(x, y)f(x', y')\,dx'\,dx\,dy'\,dy \\
&= k_n \int_{\mathcal{K}_C(k_n)^2} \mathbb{1}_{\{|x-x'| \leq k_n^{1+\varepsilon}\}} \rho_{\text{box}}(p, k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle| \deg_{\text{box}}(p) = k_n\right] f(x, y)f(x', y')\,dx'\,dx\,dy'\,dy \\
&\leq O\left(k_n^{2+\varepsilon}\right)\left(\int_{a_n^-}^{a_n^+} e^{-\alpha y'}\,dy'\right)\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] \\
&= O\left(k_n^{2+\varepsilon-2\alpha}\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right).
\end{aligned}$$

Recall that $\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] = \Theta\left(nk_n^{-(2\alpha-1)}s(k_n)\right)$. Therefore to show that $\mathbb{E}\left[I_{0,1} - I^*_{0,1}\right] = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]^2\right)$ it suffices to show that $k_n^{2+\varepsilon-2\alpha} = o\left(nk_n^{-(2\alpha-1)}s(k_n)\right)$. When $\frac{1}{2} < \alpha \leq \frac{3}{4}$ we have

$$\frac{4\alpha - 1 + \varepsilon}{2\alpha + 1} < 1,$$

for ε small enough. Hence

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

When $\alpha \geq \frac{3}{4}$,

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

for ε small enough.

For $(p, p') \in \mathcal{E}_\varepsilon(k_n)$ we assume without loss of generality that $p'_1 = p_1 = (x_1, y_1)$, i.e.

$$J_{0,1} = \sum_{(p_1,p_2)\in\mathcal{P}\setminus\{p\}}^{\neq} \widetilde{T}_{\text{box}}(p, p_1, p_2) \sum_{p'_2\in\mathcal{P}\setminus\{p',p_1\}} \widetilde{T}_{\text{box}}(p', p_1, p'_2).$$

Now let $Z_{0,1}$ denote the part of $J_{0,1}$ where $y_1 \leq 4\log(k_n)$ and $y_2, y_2' \leq \varepsilon \log(k_n)$.

We first analyze $\mathbb{E}\left[Z_{0,1} | \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right]$. When $y_1 \leq 4\log(k_n)$ and both $y_2, y_2' \leq \varepsilon \log(k_n)$ we have that

$$|x_2 - x_2'| \leq |x_1 - x_2| + |x_1 - x_2'| \leq e^{\frac{y_1}{2}} \left(e^{\frac{y_2}{2}} + e^{\frac{y_2'}{2}}\right) \leq 2k_n^{2+\varepsilon}$$

whenever $\widetilde{T}_{\text{box}}(p, p_1, p_2)\widetilde{T}_{\text{box}}(p', p_1, p_2') > 0$ while both $|x - x_2|, |x' - x_2'| = O\left(k_n^{1+\varepsilon}\right)$. Hence it follows that $\widetilde{T}_{\text{box}}(p, p_1, p_2)\widetilde{T}_{\text{box}}(p', p_1, p_2') > 0$ implies that

$$|x - x'| \leq |x - x_2| + |x_2 - x_2'| + |x_2' - x'| = O\left(k_n^{2+\varepsilon}\right).$$

Next, by integrating only over x_2' and y_2' we get

$$\begin{aligned}
\mathbb{E}\left[Z_{0,1} | \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right] &= O\left(e^{\frac{y'}{2}} \mathbb{1}_{\left\{|x-x'| \leq O(1)k_n^{2+\varepsilon}\right\}} \mathbb{E}\left[\widetilde{T}_{\text{box}}(p) \middle| \deg_{\text{box}}(p) = k_n\right]\right) \\
&= O\left(k_n \mathbb{E}\left[\widetilde{T}_{\text{box}}(p) \middle| \deg_{\text{box}}(p) = k_n\right]\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&\int_{\mathcal{E}_\varepsilon(k_n)} \rho_{\text{box}}(p, p', k_n, k_n) \mathbb{E}\left[Z_{0,1} | \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right] f(x, y) f(x', y') \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}x' \, \mathrm{d}y' \\
&= O\left(k_n^{3+\varepsilon}\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] \int_{\mathcal{K}_C(k_n)} \rho_{\text{box}}(y', k_n) e^{-\alpha y'} \, \mathrm{d}y' \\
&= O\left(k_n^{2+\varepsilon} k_n^{-2\alpha} \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right) = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]^2\right),
\end{aligned}$$

where the last line follows from the analysis done for $\mathbb{E}\left[I_{0,0} - I_{0,0}^*\right]$.

It now remains to consider $J_{0,1} - Z_{0,1} := Z_{0,1}^*$. We will show that

$$\mathbb{E}\left[Z_{0,1}^* \middle| \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right] = o\left(k_n^4 s(k_n)^2\right). \tag{89}$$

Using that the joint degree distribution factorizes on $\mathcal{E}_\varepsilon(k_n)$ this then implies that

$$\begin{aligned}
\mathbb{E}\left[I_{0,1}^*\right] &= o\left(k_n^4 s(k_n)^2\right) \left(\int_{\mathcal{R}(k_n, C)} \rho_{\text{box}}(y, k_n) f(x, y) \, \mathrm{d}x \, \mathrm{d}y\right)^2 \\
&= o\left(\left(n s(k_n) k_n^{-2\alpha+1}\right)^2\right) = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]^2\right),
\end{aligned}$$

which finished the proof of (86) for $a = 0, b = 1$.

We first consider the part with $y_1 > 4\log(k_n)$. Since the integration over x_1, x_2 and x_2' of $\mathbb{E}\left[Z_{0,1}^* \middle| \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right]$ is bounded by $O\left(e^y e^{\frac{y'}{2}}\right)$ we get that the contribution to $\mathbb{E}\left[Z_{0,1}^* \middle| \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right]$ with $y > 4\log(k_n)$ and $(p, p') \in \mathcal{E}_\varepsilon(k_n)$ is

$$\begin{aligned}
O\left(e^y e^{\frac{y'}{2}} \int_{4\log(k_n)}^R e^{-(\alpha - \frac{1}{2})y_1} \, dy_1\right) &= O\left(k_n^3 \int_{4\log(k_n)}^R e^{-(\alpha - \frac{1}{2})y_1} \, dy_1\right) \\
&= O\left(k_n^{3-(4\alpha-2)}\right) = o\left(k_n^4 s_\alpha(k_n)^2\right).
\end{aligned}$$

Here the last step follows since for $\frac{1}{2} < \alpha < \frac{3}{4}$

$$k_n^{3-(4\alpha-2)-4} s(k_n)^{-2} = k_n^{3-(4\alpha-2)-4+2(4\alpha-2)} = k_n^{-5+4\alpha} = o(1),$$

while for $\alpha = \frac{3}{4}$
$$k_n^{3-(4\alpha-2)-4} s(k_n)^{-2} = O\left(\log(k_n)^{-2}\right) k_n^{3-(4\alpha-2)-2} = O\left(\log(k_n)^{-2}\right) = o\left(1\right),$$
and for $\alpha > \frac{3}{4}$
$$k_n^{3-(4\alpha-2)-4} s(k_n)^{-2} = k_n^{3-(4\alpha-2)-2} = o\left(1\right).$$

Next we consider the case where $y_1 \le 4\log(k_n)$ and at least one of y_2, y_2' is larger than $\varepsilon \log(k_n)$. Due to symmetry it is enough to consider the case with $y_2 > \varepsilon \log(k_n)$. Here the contribution to $\mathbb{E}\left[Z_{0,1}^* \middle| \deg_{\text{box}}(p), \deg_{\text{box}}(p') = k_n\right]$ is

$$\begin{aligned}\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] O\left(e^{\frac{y'}{2}} \int_{\varepsilon\log(k_n)}^{R} e^{-(\alpha-\frac{1}{2})y_2}\, dy_2\right) &= O\left(k_n^{1-\varepsilon(\alpha-\frac{1}{2})}\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\right] \\ &= O\left(k_n^{3-\varepsilon(\alpha-\frac{1}{2})} s(k_n)\right) = o\left(k_n^4 s(k_n)^2\right).\end{aligned}$$

The last line follows since $k_n^{-1} = o\left(s(k_n)\right)$ for $\frac{1}{2} < \alpha < \frac{3}{4}$ and $k_n^{-1} = O\left(s(k_n)\right)$ for $\alpha \ge \frac{3}{4}$.

Computing $\mathbb{E}\left[I_{0,2}\right]$ In this case we have

$$\mathbb{E}\left[\mathbb{1}_{\{\deg_{\text{box}}(p)=k_n, \deg_{\text{box}}(p')=k_n\}} J_2\right] = (1+o(1))\rho_{\text{box}}(p,p',k_n,k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}(p)\middle| \deg_{\text{box}}(p) = k_n\right].$$

We then use that $\rho_{\text{box}}(p,p',k_n,k_n) \le \rho_{\text{box}}(p,k_n)$ to obtain

$$\begin{aligned}\mathbb{E}\left[I_{0,2} - I_{0,2}^*\right] &= O\left(k_n^{1+\varepsilon}\right)\left(\int_{\mathcal{K}_C(k_n)} e^{-\alpha y'}\, \mathrm{d}y'\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] \\ &= O\left(k_n^{\varepsilon-(2\alpha-1)}\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]\right)\end{aligned}$$

were the last line follows since $\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] = \Theta\left(n k_n^{-(2\alpha-1)} s(k_n)\right)$ and $k_n^{\varepsilon} n^{-1} = o\left(s(k_n)\right)$.

For the other term we use the fact that the degree distribution factorizes;

$$\begin{aligned}\mathbb{E}\left[I_{0,2}^*\right] &= O\left(1\right)\left(\int_{\mathcal{R}(k_n,C)} \rho_{\text{box}}(y',k_n) f(x',y')\, \mathrm{d}x'\, \mathrm{d}y'\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] \\ &= O\left(n k_n^{-(2\alpha+1)}\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right] = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]^2\right),\end{aligned}$$

where we also used that $k_n^{-2} = o\left(s(k_n)\right)$.

Computing $\mathbb{E}\left[I_{1,1}\right]$ Using (89) we get

$$\begin{aligned}\mathbb{E}\left[I_{1,1}\right] &= O\left(k_n\right) \int_{\mathcal{R}(k_n,C)} \rho_{\text{box}}(y,k_n)\mathbb{E}\left[\widetilde{T}_{\text{box}}\middle| \deg_{\text{box}}(p) = k_n\right] f(x,y)\, \mathrm{d}x\, \mathrm{d}y \\ &= O\left(k_n\right) \mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right].\end{aligned}$$

Now observe that for $\frac{1}{2} < \alpha < \frac{3}{4}$

$$k_n n^{-1} k_n^{(2\alpha-1)} s(k_n)^{-1} = k_n^{6\alpha-2} n^{-1} = O\left(n^{\frac{4\alpha-3}{2\alpha+1}}\right) = o\left(1\right),$$

while for $\alpha \ge \frac{3}{4}$

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

We conclude that $k_n = o\left(n k^{-(2\alpha-1)} s(k_n)\right)$ and hence $\mathbb{E}\left[I_{1,1}\right] = o\left(\mathbb{E}\left[\widetilde{T}_{\text{box}}(k_n, C)\right]^2\right)$.

□

9 Equivalence for local clustering in G_{Po} and G_{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. As in the previous section we write $|\cdot|_n$ for the norm $|\cdot|_{\pi e^{R/2}}$.

Recall the map Ψ from (7)

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

and that $\mathcal{B}(p)$ denotes the image under Ψ of the ball of hyperbolic radius R around the point $\Psi^{-1}(p)$. Under the coupling between the hyperbolic random graph and the finite box model, described in Section 2.4, two points $p = (x, y)$ and $p' = (x', y')$ are connected if and only if

$$|x - x'|_n \leq \Phi(y, y') = \frac{1}{2} e^{R/2} \arccos\left(\frac{\cosh(R-y)\cosh(R-y') - \cosh R}{\sinh(R-y)\sinh(R-y')}\right),$$

see (8). We will often use the result from Lemma 2.2 to approximate the function Φ, for $y+y' < R$, by

$$e^{\frac{1}{2}(y+y')} - K e^{\frac{3}{2}(y+y')-R} \leq \Phi(R-y, R-y') \leq e^{\frac{1}{2}(y+y')} + K e^{\frac{3}{2}(y+y')-R},$$

where K is a constant determined by the lemma.

9.1 Some results on the hyperbolic geometric graph

We start with some basic results for the hyperbolic random geometric graph. Recall that $\mathcal{B}_\infty(p) = \{p' \in \mathbb{R} \times \mathbb{R}_+ : |x - x'| \leq e^{(y+y')/2}\}$ and observe that (10) from Lemma 2.2 implies the following.

Corollary 9.1. *For sufficiently large n and $p \in \mathcal{R}$,*

$$\mathcal{B}_\infty(p) \cap \mathcal{R}([K, R]) \subseteq \mathcal{B}(p) \cap \mathcal{R}([K, R]),$$

where K is the constant from Lemma 2.2.

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 $x' \geq 0$ meets the right boundary of $\mathcal{R}$, that is, the line $x' = \frac{\pi}{2} e^{R/2}$ at $y' = R - y + 2\ln\frac{\pi}{2}$. Hence, any point $p' \in \mathcal{R}([R - y + 2\ln\frac{\pi}{2}, R])$ is included in $\mathcal{B}_\infty(p)$. In other words,

$$\mathcal{B}_\infty(p) \cap \mathcal{R}([R - y + 2\ln\frac{\pi}{2}, R]) = \mathcal{R}([R - y + 2\ln\frac{\pi}{2}, R]).$$

This together with the fact that for any $u' = (r', \theta')$,

$$r' < y = R - r \Rightarrow d_{\mathbb{H}}(\Psi^{-1}(p), u') \leq R$$

implies that

$$(\mathcal{B}(p) \triangle \mathcal{B}_\infty(p)) \cap \mathcal{R}([R - y + 2\ln\frac{\pi}{2}, R]) = \emptyset, \tag{90}$$

where $A \triangle B$ denotes the symmetric difference of the sets A and B. 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 $0 \leq y_n < R$ be such that $R - y_n \to \infty$ and write $p_n = (x_n, y_n)$. Then we have, as $n \to \infty$,*

$$\mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)) = \Theta(1) \cdot \begin{cases} e^{(1/2-\alpha)R + \alpha y_n}, & \text{if } \alpha < 3/2 \\ (R - y_n) e^{3y/2 - R}, & \text{if } \alpha = 3/2 \\ e^{3y_n/2 - R}, & \text{if } \alpha > 3/2 \end{cases}.$$

Proof. Let $r_n := R - y$. 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

$$|x_n - x| = \Theta(1) \cdot e^{\frac{3}{2}(y_n+y)-R}.$$

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

$$|x_n - x|_n = \frac{\pi}{2}e^{R/2} - e^{\frac{1}{2}(y_n+y)}.$$

Finally, (90) implies that no point in $\mathcal{R}([r_n + 2\ln\frac{\pi}{2}, R])$ 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 $R - y \le r_n$. The result depends on the value of α, yielding the following three cases

$$\begin{aligned}
\mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n) \cap \mathcal{R}([0, r_n])) &= \Theta(1) \cdot e^{3y_n/2-R} \int_0^{r_n} e^{(3/2-\alpha)y}\, dy \\
&= \Theta(1) \cdot \begin{cases} e^{(1/2-\alpha)R+\alpha y_n}, & \text{if } \alpha < 3/2 \\ (R - y_n)e^{3y_n/2-R}, & \text{if } \alpha = 3/2 \\ e^{3y_n/2-R}, & \text{if } \alpha > 3/2 \end{cases}.
\end{aligned}$$

Next we compute the number of remaining points in $\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)$,

$$\begin{aligned}
\mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n) \cap \mathcal{R}([r_n, R])) &= \frac{\nu\alpha}{\pi} \int_{r_n}^{r_n+2\ln\frac{\pi}{2}} \left(\frac{\pi}{2}e^{R/2} - e^{\frac{1}{2}(y_n+y)}\right) e^{-\alpha y}\, dy \\
&= O(1) \cdot e^{R/2} \int_{r_n}^{r_n+2\ln\frac{\pi}{2}} e^{-\alpha y}\, dy = O(1) \cdot e^{R/2}e^{-\alpha r_n} \\
&= O(1) \cdot e^{(1/2-\alpha)R+\alpha y_n}.
\end{aligned}$$

Now note that for any $\alpha > 3/2$, we have

$$((1/2 - \alpha)R + \alpha y_n) - (3y_n/2 - R) = (3/2 - \alpha)(R - y_n) \to -\infty,$$

by our assumption on y_n. For $\alpha = 3/2$, these two quantities are equal. From these observations, we deduce that

$$\mu(\mathcal{B}(p_n) \triangle \mathcal{B}_\infty(p_n)) = \Theta(1) \cdot \begin{cases} e^{(1/2-\alpha)R+\alpha y_n}, & \text{if } \alpha < 3/2 \\ r_n e^{3y_n/2-R}, & \text{if } \alpha = 3/2 \\ e^{3y_n/2-R}, & \text{if } \alpha > 3/2 \end{cases}.$$

□

9.2 Equivalence clustering G_{Po} and G_{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 $(k_n)_{n\ge1}$ be a sequence such that $k_n = O\left(n^{1/(2\alpha+1)}\right)$. Then*

$$\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] = \Theta(1)\, n k_n^{-(2\alpha+1)}, \tag{91}$$

and

$$\mathbb{E}\left[N_{\mathrm{box}}(k_n)\right] = \Theta(1)\, n k_n^{-(2\alpha+1)}. \tag{92}$$

Moreover,

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]}{\mathbb{E}\left[N_{\mathrm{box}}(k_n)\right]} = 1. \tag{93}$$

Proof. By the Campbell-Mecke formula

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

Then by Lemma 6.5

$$\begin{aligned}\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] &= (1 + o(1)) \int_{\mathcal{R}} \rho(y, k_n) f(x, y)\, \mathrm{d}x\, \mathrm{d}y \\ &= (1 + o(1)) n \int_0^R \rho(y, k_n) f(x, y)\, \mathrm{d}x\, \mathrm{d}y = \Theta(1)\, n k_n^{-(2\alpha+1)}.\end{aligned}$$

Similarly,

$$\mathbb{E}\left[N_{\mathrm{box}}(k_n)\right] = (1 + o(1)) \int_{\mathcal{R}} \rho(y, k_n) f(x, y)\, \mathrm{d}x\, \mathrm{d}y$$

From which the results follow. ◻

Recall that Proposition 5.3 states

$$\lim_{n\to\infty} s(k_n)^{-1}\, \mathbb{E}\left[|c^*(k_n; G_{\mathrm{Po}}) - c^*(k_n; G_{\mathrm{box}})|\right] = 0.$$

Next recall the definition of $\mathcal{K}_C(k_n)$

$$\mathcal{K}_C(k_n) = \left\{ y \in \mathbb{R}_+ : \frac{k_n - C\sqrt{k_n \log(k_n)}}{\xi} \vee 0 \leq e^{\frac{y}{2}} \leq \frac{k_n + C\sqrt{k_n \log(k_n)}}{\xi} \wedge e^{R/2} \right\},$$

and (73)

$$\widetilde{c}_{\mathrm{box}}(k_n) = \frac{\widetilde{T}_{\mathrm{box}}(k_n, C)}{\binom{k_n}{2} \mathbb{E}\left[N_{\mathrm{box}}(k_n)\right]},$$

where $\widetilde{T}_{\mathrm{box}}(k_n, C)$ counts for all nodes $p = (x, y)$ with $y \in \mathcal{K}_C(k_n)$ the pairs (p_1, p_2) that form a triangle with p, with the exception that it considers $p_2 \in \mathcal{B}_\infty(p_1) \cap \mathcal{R}$ instead of $\mathcal{B}_{\mathrm{box}}(p_1)$. Then using Corollary 8.2 we get

$$\mathbb{E}\left[|c^*(k_n; G_{\mathrm{Po}}) - c^*(k_n; G_{\mathrm{box}})|\right] \leq \mathbb{E}\left[|c^*(k_n; G_{\mathrm{Po}}) - \widetilde{c}_{\mathrm{box}}(k_n)|\right] + o\left(s(k_n)\right),$$

and hence it is enough to prove that

$$\lim_{n\to\infty} s(k_n)^{-1} \mathbb{E}\left[|c^*(k_n; G_{\mathrm{Po}}) - \widetilde{c}_{\mathrm{box}}(k_n)|\right] = 0.$$

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 non-negative 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 the second statement of Lemma 6.5

$$\begin{aligned}\int_{\mathcal{K}_C} e^{ty} \hat{\rho}_n(y, k_n - r) e^{-\alpha y}\, \mathrm{d}y &= \Theta\left(k_n^{2t}\right) \int_{\mathcal{K}_C} \hat{\rho}_n(y, k_n - r) e^{-\alpha y}\, \mathrm{d}y \\ &= O\left(k_n^{2t}\right) (k_n - r)^{-(2\alpha+1)} = O(1)\, k_n^{-2\alpha-1+2t}.\end{aligned}$$

◻

Proof of Proposition 5.3. To keep notation concise we abbreviate $\mathbb{E}[N_{\mathrm{Po}}(k_n)]$ and $\mathbb{E}[N_{\mathrm{box}}(k_n)]$ by $\overline{n}_{\mathrm{Po}}(k_n)$ and $\overline{n}_{\mathrm{box}}(k_n)$, respectively. We will also suppress the subscript n in most expressions regarding the graphs G_{Po} and G_{box}. Finally we will write

$$T_{\mathrm{Po}}(p) = \sum_{(p_1,p_2)\in\mathcal{P}\setminus\{p\}}^{\neq} T_{\mathrm{Po}}(p, p_1, p_2),$$

with

$$T_{\mathrm{Po}}(p, p_1, p_2) = \mathbb{1}_{\{p_1\in\mathcal{B}(p)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(p)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(p_1)\}}$$

to denote the triangle count function for p in G_{Po}. Then we have

$$\begin{aligned}
\mathbb{E}\left[|c^*(k_n; G_{\mathrm{Po}}) - \widetilde{c}_{\mathrm{box}}(k_n)|\right] &= \binom{k_n}{2}^{-1} \mathbb{E}\left[\left|\sum_{p\in\mathcal{P}} \frac{\mathbb{1}_{\{\deg_{\mathrm{Po}}(p)=k_n\}}}{\overline{n}_{\mathrm{Po}}(k_n)} T_{\mathrm{Po}}(p) - \frac{\mathbb{1}_{\{\deg_{\mathrm{box}}(p)=k_n\}}}{\overline{n}_{\mathrm{box}}(k_n)} \widetilde{T}_{\mathrm{box}}(p)\right|\right] \\
&\leq \binom{k_n}{2}^{-1} \overline{n}_{\mathrm{Po}}(k_n)^{-1} \mathbb{E}\left[\left|\sum_{p\in\mathcal{P}} \mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} T_{\mathrm{Po}}(p) - \mathbb{1}_{\{\deg_{\mathrm{box}}(p)=k_n\}} \widetilde{T}_{\mathrm{box}}(p)\right|\right] \\
&\quad + \binom{k_n}{2}^{-1} \left|\frac{1}{\overline{n}_{\mathrm{Po}}(k_n)} - \frac{1}{\overline{n}_{\mathrm{box}}(k_n)}\right| \mathbb{E}\left[\sum_{p\in\mathcal{P}} \mathbb{1}_{\{\deg_{\mathrm{box}}(p)=k_n\}} \widetilde{T}_{\mathrm{box}}(p)\right]
\end{aligned}$$

The last term can be rewritten as

$$\left|1 - \frac{\overline{n}_{\mathrm{Po}}(k_n)}{\overline{n}_{\mathrm{box}}(k_n)}\right| \mathbb{E}\left[\widetilde{c}_{\mathrm{box}}(k_n)\right] = \left|1 - \frac{\overline{n}_{\mathrm{Po}}(k_n)}{\overline{n}_{\mathrm{box}}(k_n)}\right| \gamma(k_n)(1 + o(1)),$$

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

$$\left|1 - \frac{\overline{n}_{\mathrm{Po}}(k_n)}{\overline{n}_{\mathrm{box}}(k_n)}\right| \mathbb{E}\left[\widetilde{c}_{\mathrm{box}}(k_n)\right] = o\left(s(k_n)\right),$$

and therefore we are left to analyze the other term. By the Campbell-Mecke formula we have that

$$\begin{aligned}
&\mathbb{E}\left[\left|\sum_{p\in\mathcal{P}} \mathbb{1}_{\{\deg_{\mathrm{Po}}(p)=k_n\}} T_{\mathrm{Po}}(p) - \mathbb{1}_{\{\deg_{\mathrm{box}}(p)=k_n\}} \widetilde{T}_{\mathrm{box}}(p)\right|\right] \\
&= \int_{\mathcal{R}} \mathbb{E}\left[\left|\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} T_{\mathrm{Po}}(y) - \mathbb{1}_{\{\deg_{\mathrm{box}}(y)=k_n\}} \widetilde{T}_{\mathrm{box}}(y)\right|\right] f(x, y)\, \mathrm{d}y\, \mathrm{d}x.
\end{aligned}$$

Since

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}}}{\overline{n}_{\mathrm{Po}}(k_n)} T_{\mathrm{Po}}(y)\right] &\leq \binom{k_n}{2} \rho_{\mathrm{Po}}(y, k_n) \overline{n}_{\mathrm{Po}}(k_n)^{-1} \\
&= \binom{k_n}{2} \rho_{\mathrm{Po}}(y, k_n) \Theta\left(\overline{n}_{\mathrm{box}}(k_n)^{-1}\right) \\
&= \Theta\left(n^{-1} k_n^{2\alpha+3}\right) \rho_{\mathrm{Po}}(y, k_n)
\end{aligned}$$

and similar for the other term, it follows that

$$\begin{aligned}
&\mathbb{E}\left[\left|\frac{\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}}}{\overline{n}_{\mathrm{Po}}(k_n)} T_{\mathrm{Po}}(y) - \frac{\mathbb{1}_{\{\deg_{\mathrm{box}}(y)=k_n\}}}{\overline{n}_{\mathrm{Po}}(k_n)} \widetilde{T}_{\mathrm{box}}(y)\right|\right] \\
&\qquad \leq \Theta\left(n^{-1} k_n^{2\alpha+3}\right) \left(\rho_{\mathrm{Po}}(y, k_n) + \rho_{\mathrm{box}}(y, k_n)\right).
\end{aligned}$$

Therefore, by a concentration of heights argument (c.f. first statement of Lemma 6.5), it is enough to consider the integral

$$n \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left|\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} T_{\mathrm{Po}}(y) - \mathbb{1}_{\{\deg_{\mathrm{box}}(y)=k_n\}} \widetilde{T}_{\mathrm{box}}(y)\right|\right] e^{-\alpha y}\, \mathrm{d}y, \tag{94}$$

where we also used that $f(x,y)$ is simply a constant multiple of the function $e^{-\alpha y}$. Since $\binom{k_n}{2}\overline{n}_{\mathrm{Po}}(k_n) = \Theta\left(n k_n^{-(2\alpha-1)}\right)$ we have to show that

$$\lim_{n\to\infty} k_n^{2\alpha-1} s(k_n)^{-1} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left|\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} T_{\mathrm{Po}}(y) - \mathbb{1}_{\{\deg_{\infty}(y)=k_n\}} \widetilde{T}_{\mathrm{box}}(y)\right|\right] e^{-\alpha y}\, \mathrm{d}y = 0.$$

For $\alpha > 3/4$, $s_{3/4}(k_n) = \log(k_n)^{-1} s_\alpha(k_n) = o\left(s_\alpha(k_n)\right)$ and thus it suffices to prove the following two cases:

1. if $1/2 < \alpha \leq 3/4$, then

$$\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left|\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} T_{\mathrm{Po}}(y) - \mathbb{1}_{\{\deg_{\mathrm{box}}(y)=k_n\}} \widetilde{T}_{\mathrm{box}}(y)\right|\right] e^{-\alpha y}\, \mathrm{d}y = 0,$$

2. if $3/4 < \alpha$, then

$$\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left|\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} T_{\mathrm{Po}}(y) - \mathbb{1}_{\{\deg_{\mathrm{box}}(y)=k_n\}} \widetilde{T}_{\mathrm{box}}(y)\right|\right] e^{-\alpha y}\, \mathrm{d}y = 0.$$

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 an expression such as $\mathcal{B}(y)$. In addition we write $D_{\mathrm{Po}}(y,k_n;\mathcal{P})$ for the indicator which is equal to 1 if and only if $\mathcal{B}(y)$ contains k_n points from $\mathcal{P}\setminus\{(0,y)\}$. We define $D_{\mathrm{box}}(y,k_n;\mathcal{P})$ analogously for the ball $\mathcal{B}_{\mathrm{box}}(y)$. It is important to note that for any $p' \in \mathcal{R}$ it holds that $p' \in \mathcal{B}_{\mathrm{box}}(y) \iff p' \in \mathcal{B}_\infty(y)$.

We need to split the integrand over several terms and then analyze each of these separately. Applying the Campbel-Mecke formula yields

$$\begin{aligned}
&\mathbb{E}\left[\left|\mathbb{1}_{\{\deg_{\mathrm{Po}}(y)=k_n\}} P_{\mathrm{Po}}(y) - \mathbb{1}_{\{\deg_{\infty}(y)=k_n\}} \widetilde{T}_{\mathrm{box}}(y)\right|\right] \leq \\
&\qquad \mathbb{E}\Bigg[\sum_{(p_1,p_2)\in\mathcal{P}\setminus\{(0,y)\}}^{\neq} \big|D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) T_{\mathrm{Po}}(y,p_1,p_2) \\
&\qquad\qquad\qquad - D_{\mathrm{box}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \widetilde{T}_{\mathrm{box}}(y,p_1,p_2)\big|\Bigg],
\end{aligned}$$

where the sum ranges over all distinct pairs of points in $\mathcal{P}\setminus\{(0,y)\}$. In what follows, we will set $\mathcal{B}_{\mathrm{Po}\triangle\infty}(p') = \mathcal{B}(p') \triangle (\mathcal{B}_\infty(p') \cap \mathcal{R})$ and $\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(p') = \mathcal{B}(p') \cap \mathcal{B}_{\mathrm{box}}(p')$ and observe that $\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y) = \mathcal{B}(y) \cap \mathcal{B}_\infty(y)$. We will now bound the sum that is inside the expectation. We will split the sum into different parts, depending on combinations of $p_1, p_2 \in \mathcal{P}\setminus\{(0,y)\}$ for which only one of the two terms of the difference is non-zero. Clearly, for this we need that either $p_1 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)$ or $p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(p_1)$. We will consider the following four cases:

1. $p_1 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)$
   (a) $y_1, y_2 < (1-\varepsilon)R \wedge (R-y)$
   (b) $y_1 \geq (1-\varepsilon)R \wedge (R-y)$
2. $p_1 \in \mathcal{B}(y) \setminus \mathcal{B}_\infty(y)$ with $y_1 < K$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)$.

3. $p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)$ with $y_1 \geq K$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)$,

where K in the last two cases is the constant from Lemma 2.2.

Observe that when $y_1 < (1-\varepsilon)R\wedge(R-y)$ and $y_2 \geq (1-\varepsilon)R\wedge(R-y)$ it follows from Corollary 9.1 that $p_2 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(p_1)$ and thus we do not have to consider this case when $p_1 \in \mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)$. Similarly, when $y_1 \geq K$ and $p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)$ Corollary 9.1 implies that $p_1 \in \mathcal{B}(y) \setminus \mathcal{B}_\infty(y)$ which explains the setting of case 2.

We can now bound the sum by the following expression:

$$
\begin{aligned}
&\sum^{\neq}_{(p_1,p_2)\in\mathcal{P}\setminus\{(0,y)\}} \Big| D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})T_{\mathrm{Po}}(y,p_1,p_2) \\
&\qquad\qquad\qquad\qquad - D_{\mathrm{box}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\widetilde{T}_{\mathrm{box}}(y,p_1,p_2)\Big| \\
&\leq \sum^{\neq}_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y_1,y_2<(1-\varepsilon)R\wedge(R-y)}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \qquad (95)\\
&\quad + \sum^{\neq}_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y_1,y_2<(1-\varepsilon)R\wedge(R-y)}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\}}\, D_{\mathrm{box}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \qquad (96)\\
&\quad + \sum^{\neq}_{\substack{p_1,p_2\ \in\mathcal{P}\setminus\{(0,y)\}\\ y_1\geq(1-\varepsilon)R\wedge(R-y)}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \qquad (97)\\
&\quad + \sum^{\neq}_{\substack{p_1,p_2\ \in\mathcal{P}\setminus\{(0,y)\}\\ y_1\geq(1-\varepsilon)R\wedge(R-y)}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathrm{box}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \qquad (98)\\
&\quad + \sum^{\neq}_{\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_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \qquad (99)\\
&\quad + \sum^{\neq}_{\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_{\mathrm{box}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}) \qquad (100)\\
&\quad + \sum^{\neq}_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y(p_1)<K}} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}. \qquad (101)
\end{aligned}
$$

In the following paragraphs we will give upper bounds on the expected values of each one of these partial sums.

The sums (95) and (96) We will analyze (95). The analysis of the other sum (96) is similar. Note first that for any two points p_1, p_2 the following holds: $p_1 \in \mathcal{B}(y)$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1) \cap \mathcal{B}(y)$, then $p_2 \in \mathcal{B}(y)$ and $p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_2) \cap \mathcal{B}(y)$. 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

$$
\begin{aligned}
\mathcal{D}_1 &= \{(p_1,p_2)\in\mathcal{D}\cap\mathcal{P} \ : \ y \leq y_1 \leq R/2,\ y_\omega \leq y_2 \leq y_1\},\\
\mathcal{D}_2 &= \{(p_1,p_2)\in\mathcal{D}\cap\mathcal{P} \ : \ y_1 \leq R/2,\ y_2 \leq y_\omega\} \text{ and}\\
\mathcal{D}_3 &= \{(p_1,p_2)\in\mathcal{D}\cap\mathcal{P} \ : R/2 < y_1 \leq R-y,\ y_2 \leq y_1\}.
\end{aligned}
$$

Note that $\mathcal{D} \subseteq \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$. Hence, we can write

$$\begin{aligned} &\mathbb{E}\left[\sum_{\substack{p_1,p_2\in\mathcal{P}\setminus\{(0,y)\}\\ y_1,y_2\le(1-\varepsilon)R\wedge(R-y)}} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\ &\le \sum_{i=1}^{3}\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_i} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\cdot D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right]. \end{aligned} \tag{102}$$

We bound each one of the above three summands as follows:

$$\begin{aligned} &\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_1} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\ &\le \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_1} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] := \mathcal{I}_n^{(1)}(y), \end{aligned} \tag{103}$$

$$\begin{aligned} &\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_2} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\ &\le \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_2} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] := \mathcal{I}_n^{(2)}(y) \end{aligned} \tag{104}$$

and

$$\begin{aligned} &\mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_3} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] \\ &\le \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_3} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}\, D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right] := \mathcal{I}_n^{(3)}(y). \end{aligned} \tag{105}$$

We will bound each term using the Campbell-Mecke formula and show for $i = 1, 2, 3$ that for $1/2 < \alpha < 3/4$

$$\lim_{n\to\infty} k_n^{6\alpha-3}\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(i)}(y)e^{-\alpha}\,\mathrm{d}y = 0, \tag{106}$$

and for $\alpha \ge 3/4$

$$\lim_{n\to\infty} k_n^{2\alpha}\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(i)}(y)e^{-\alpha}\,\mathrm{d}y = 0. \tag{107}$$

For the first term (103), we note that

$$\mathbb{E}\left[D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\}\right] = \rho_{\mathrm{Po}}(y,k_n-2).$$

and hence $\mathcal{I}_n^{(1)}(y)$ becomes

$$\rho_{\mathrm{Po}}(y,k_n-2)\int_{-I_n}^{I_n}\int_{y}^{R/2}\int_{-I_n}^{I_n}\int_{y_\omega}^{y_1} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}(y)\}}\,\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}e^{-\alpha(y_1+y_2)}\,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1. \tag{108}$$

Next, Lemma 2.2 implies that for $y' \le R - y$, 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:

$$\mathbb{E}\left[\sum_{p_1,p_2\in\mathcal{D}_1} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\cap\mathrm{box}}((0,y))\}}\cdot\mathbb{1}_{\{p_2\in\mathcal{B}(y)\}}\cdot D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right]$$

$$= \rho_{\mathrm{Po}}(y, k_n - 2)e^{y} \int_{y}^{R/2} e^{y_1/2} \int_{y_\omega}^{y_1} e^{y_2/2} e^{-\alpha y_2} \cdot e^{-\alpha y_1} dy_2 dy_1.$$

Now, the double integral becomes

$$\begin{aligned}
&\int_{y}^{R/2} e^{y_1/2} \int_{y_\omega}^{y_1} e^{y_2/2} e^{-\alpha y_2} \cdot e^{-\alpha y_1} dy_2 dy_1 = \\
&O(1) \cdot \int_{y}^{R/2} e^{y_1/2 - \alpha y_1} \cdot e^{(1/2-\alpha)y_\omega} dy_1 \\
&= O(1) \cdot e^{(1/2-\alpha)y_\omega} \cdot \int_{y}^{R/2} e^{y_1/2-\alpha y_1} dy_1 \\
&= O(1) \cdot e^{(1/2-\alpha)y_\omega + (1/2-\alpha)y} \\
&\ll e^{(1-2\alpha)y},
\end{aligned} \tag{109}$$

since $y_\omega = y + \omega$ and $\omega \to \infty$. We then deduce that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{p_1,p_2 \in \mathcal{D}_1} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\cap \mathrm{box}}((0,y))\}} \cdot \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\}} \cdot D_{\mathrm{Po}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\})\right] \\
&\qquad \ll \rho_{\mathrm{Po}}(y, k_n - 2)e^{(1-2\alpha)y}.
\end{aligned} \tag{110}$$

We now integrate this with respect to y and determine its contribution to (94);

$$\begin{aligned}
&\int_{\mathcal{K}_C(k_n)} \rho_{\mathrm{Po}}(y, k_n - 2)e^{(1-2\alpha)y} e^{-\alpha y} \, \mathrm{d}y \, \mathrm{d}x \\
&= O\left(k_n^{-6\alpha+1}\right)
\end{aligned}$$

where we used Lemma 9.4 with $t = 1 - 2\alpha$.

Since $1 - 6\alpha + \min\{6\alpha - 3, 2\alpha\} < 0$ for all $\alpha > 1/2$ we deduce that for $1/2 < \alpha < 3/4$

$$\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y} \, \mathrm{d}y = 0,$$

while for $\alpha \geq 3/4$

$$\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y} \, \mathrm{d}y = 0.$$

We will now bound the term in (104). Using similar observations as for the previous term we get that $\mathcal{I}_n^{(2)}(y)$ equals

$$\rho_{\mathrm{Po}}(y, k_n - 2) \int_{-I_n}^{I_n} \int_{0}^{R/2} \int_{-I_n}^{I_n} \int_{0}^{y_\omega} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}((0,y))\}} e^{-\alpha(y_1+y_2)} \, \mathrm{d}y_2 \, \mathrm{d}x_2 \, \mathrm{d}y_1 \, \mathrm{d}x_1.$$

Now, Lemma 2.2 implies that for $y_2 \leq R - y_1$, we have that if $(x_2, y_2) \in \mathcal{B}_{\mathrm{Po}\triangle\infty}((x_1, y_1))$, then x_2 lies in an interval of length $Ke^{3y/2+3y'/2-R}$, where $K > 0$ is again the constant in Lemma 2.2. Using these observations we obtain:

$$\mathcal{I}_n^{(2)}(y) = \rho_{\mathrm{Po}}(y, k_n - 2)e^{y/2} \int_{0}^{R/2} e^{y_1/2+3y_1/2} \int_{0}^{y_\omega} e^{3y_2/2-R} e^{-\alpha y_2} \cdot e^{-\alpha y_1} \, \mathrm{d}y_2 \, \mathrm{d}y_1. \tag{111}$$

The integrals satisfy

$$e^{-R}\left(\int_{0}^{R/2} e^{(2-\alpha)y_1} \, \mathrm{d}y_1\right)\left(\int_{0}^{y_\omega} e^{(3/2-\alpha)y_2} \, \mathrm{d}y_2\right)$$

$$
= O(1)\, e^{-R} \left(\begin{cases} e^{(1-\alpha/2)R} & \text{if } \frac{1}{2} < \alpha < 2 \\ R & \text{if } \alpha \geq 2 \end{cases} \right) \left(\begin{cases} e^{(3/2-\alpha)y_\omega} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ y & \text{if } \alpha \geq \frac{3}{2} \end{cases} \right)
$$

$$
= O(1) \begin{cases} e^{-\frac{\alpha}{2}R} e^{(3/2-\alpha)y} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ (y+\omega(n)) e^{-\frac{\alpha}{2}R} & \text{if } \frac{3}{2} \leq \alpha < 2 \\ (y+\omega(n)) R e^{-R} & \text{if } \alpha \geq 2 \end{cases} .
$$

Since $y_\omega := y + \omega(n) \leq R = O(\log(n))$ we conclude that on $\mathcal{K}_C(k_n)$

$$
\mathcal{I}_n^{(2)}(y) = O(1)\, \rho_{\mathrm{Po}}(y, k_n - 2) \begin{cases} n^{-\alpha} k_n^{3-2\alpha} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ n^{-\alpha} \log(n) & \text{if } \frac{3}{2} \leq \alpha < 2 \\ n^{-2} \log(n)^2 & \text{if } \alpha \geq 2 \end{cases} ,
$$

and hence

$$
\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) e^{-\alpha y} \, dy = O(1)\, k_n^{-(2\alpha+1)} \begin{cases} n^{-\alpha} k_n^{3-2\alpha} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ n^{-\alpha} \log(n) & \text{if } \frac{3}{2} \leq \alpha < 2 \\ n^{-2} \log(n)^2 & \text{if } \alpha \geq 2 \end{cases} ,
$$

$$
= O(1) \begin{cases} n^{-\alpha} k_n^{2-4\alpha} & \text{if } \frac{1}{2} < \alpha < \frac{3}{2} \\ n^{-\alpha} \log(n) k_n^{-(2\alpha+1)} & \text{if } \frac{3}{2} \leq \alpha < 2 \\ n^{-2} \log(n)^2 k_n^{-(2\alpha+1)} & \text{if } \alpha \geq 2 \end{cases} .
$$

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

$$
k_n^{6\alpha-3} n^{-\alpha} k_n^{2-4\alpha} = n^{-\alpha} k_n^{2\alpha-1} = O\left(n^{-\alpha+\frac{2\alpha-1}{2\alpha+1}}\right) = O\left(k_n^{-\frac{4\alpha^2-\alpha+1}{2\alpha+1}}\right) = o(1),
$$

from which we deduce that

$$
\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) e^{-\alpha y} \, dy = 0.
$$

For $\alpha \geq 3/4$ we have that both $n^{-\alpha}\log(n)k_n^{-1}$ and $n^{-2}\log(n)^2 k_n^{-1}$ converge to zero as $n \to \infty$ and hence in this case

$$
\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) e^{-\alpha y} \, dy = 0.
$$

We will now consider the term in (105). Recall that $\mathcal{D}_3$ consists of all pairs $(p_1, p_2) \in \mathcal{D}$ such that $R/2 < y_1 \leq (1-\varepsilon)R \wedge (R-y)$ and $y_1 \leq y_\omega$ with the property that $p_1 \in \mathcal{B}(y)$ and $p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1) \cap \mathcal{B}(y)$. 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 $p_1 = (x_1, y_1) \in \mathbb{R} \times \mathbb{R}_+$,

$$
\check{\mathcal{B}}_{\mathrm{Po}}(p_1) := \{(x', y') \ : \ y' < R - y_1, \ |x_1 - x'| < (1+K) e^{\frac{1}{2}(y_1 + y')}\}. \tag{112}
$$

It is an implication of Lemma 2.2 that

$$
(\mathcal{B}(p_1) \cup \mathcal{B}_\infty(p_1)) \cap \mathcal{R}([0, R-y_1]) \subseteq \check{\mathcal{B}}_{\mathrm{Po}}(p_1)
$$

Therefore, any point $p_2 = (x_2, y_2) \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1) \cap \mathcal{B}(y)$ with $y_2 \leq R - y_1$ must belong to $\check{\mathcal{B}}_{\mathrm{Po}}(p_1) \cap \check{\mathcal{B}}_{\mathrm{Po}}(y)$.

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 $\check{\mathcal{B}}_{\mathrm{Po}}((0, y))$ is given by the equation $x' = (1-K)e^{\frac{1}{2}(y+y')}$ whereas the right boundary of $\check{\mathcal{B}}_{\mathrm{Po}}(p_1)$ 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 $y_1 > R/2$ and since $y \in \mathcal{K}_C(k_n)$, it follows that for sufficiently large n, $y \le (1+\varepsilon)R/(2\alpha+1)$. 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^2 denote the multiplicative term $1+K+o(1)$, which appears in the above. The above yields

$$\hat{y} = \left(2\log(|x_1|e^{-y_1/2}) - \log c_K\right) \vee 0 := \hat{y}(x_1, y_1). \tag{113}$$

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 $x_1 \le R - y$, we also have that $|x_1| \le e^{(y+y_1)/2}(1+o(1))$. This upper bound on $|x_1|$ together with (113), 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

$$p' \in \check{\mathcal{B}}_{\mathrm{Po}}(y) \cap \check{\mathcal{B}}_{\mathrm{Po}}((x_1,y_1)) \Rightarrow y' \ge \hat{y}(x_1,y_1),$$

which implies

$$\mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} \le \mathbb{1}_{\left\{y_2 \ge \hat{y}(x_1,y_1),\, p_2 \in \check{\mathcal{B}}_{\mathrm{Po}}((0,y))\right\}}. \tag{114}$$

If we integrate this over x_2, y_2 we get

$$\begin{aligned}
\int_{-I_n}^{I_n}\int_0^{y_1} \mathbb{1}_{\{p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha y_2} dy_2 dx_2 &\le \int_{-I_n}^{I_n}\int_0^{y_1} \mathbb{1}_{\left\{y_2 \ge \hat{y}(x_1,y_1),\, p_2 \in \check{\mathcal{B}}_{\mathrm{Po}}(y)\right\}} e^{-\alpha y_2} dy_2 dx_2 \\
&\le (1+K)\cdot e^{y/2} \int_{\hat{y}(x_1,y_1)}^{y_1} e^{y_2/2 - \alpha y_2} dy_2 \\
&= O(1)\cdot e^{y/2 + (1/2-\alpha)\hat{y}(x_1,y_1)}.
\end{aligned}$$

Note also that

$$\mathbb{E}\left[D_{\mathrm{Po}}(y, k_n - 2; \mathcal{P}\setminus\{p_1,p_2\})\right] = \rho_{\mathrm{Po}}(y, k_n - 2),$$

uniformly over all $(p_1, p_2) \in \mathcal{D}_3$. Hence the Campbell-Mecke formula yields that $\mathcal{I}_n^{(3)}(y)$ equals:

$$\begin{aligned}
&O(1)\rho_{\mathrm{Po}}(y, k_n - 2)\, e^{y/2} \int_{-I_n}^{I_n}\int_{R/2}^{(R-y)\wedge(1-\varepsilon)R} \mathbb{1}_{\{p_1 \in \mathcal{B}(y)\}} e^{(1/2-\alpha)\hat{y}(x_1,y_1) - \alpha y_1} dy_1 dx_1 \\
&= O(1)\rho_{\mathrm{Po}}(y, k_n - 2)\, e^{y/2} \int_{-I_n}^{I_n}\int_{R/2}^{(R-y)\wedge(1-\varepsilon)R} \mathbb{1}_{\left\{p_1 \in \check{\mathcal{B}}_{\mathrm{Po}}(y)\right\}} e^{(1/2-\alpha)\hat{y}(x_1,y_1) - \alpha y_1} dy_1 dx_1.
\end{aligned}$$

Due to the symmetry of $\check{\mathcal{B}}_{\mathrm{Po}}(y)$, the integration over x_1 is:

$$O(1)\cdot e^{y/2}\cdot \int_0^{(1+K)e^{y/2+y_1/2}} e^{\hat{y}(x_1,y_1)(1/2-\alpha)} dx_1$$

We will split this integral into two parts according to the value of $\hat{y}(x_1, y_1)$:

$$\int_0^{(1+K)e^{y/2+y_1/2}} e^{\hat{y}(x_1,y_1)(1/2-\alpha)} dx_1 = \int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} e^{\hat{y}(x_1,y_1)(1/2-\alpha)} dx_1 + \int_0^{c_K e^{y_1/2}} dx_1.$$

The first integral becomes:

$$\int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} e^{\hat{y}(x_1,y_1)(1/2-\alpha)} dx_1 = \int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} e^{\hat{y}(x_1,y_1)/2(1-2\alpha)} dx_1$$
$$= O(1) \cdot \int_{c_K e^{y_1/2}}^{(1+K)e^{y/2+y_1/2}} x_1^{1-2\alpha} e^{-\frac{y_1}{2}(1-2\alpha)} dx_1$$
$$= O(1) \cdot e^{-y_1/2+\alpha y_1} \cdot e^{\frac{(y+y_1)}{2}2(1-\alpha)}$$
$$= O(1) \cdot e^{y_1/2+y(1-\alpha)}.$$

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

$$e^{y/2} \cdot \int_0^{(1+K)e^{y/2+y_1/2}} e^{\hat{y}(x_1,y_1)(1/2-\alpha)} dx_1 = O(1) \cdot e^{y_1/2+y(3/2-\alpha)}.$$

Now, we integrate this with respect to y_1 and get

$$e^{y(3/2-\alpha)} \int_{R/2}^{R-y} e^{(1/2-\alpha)y_1} dy_1 = O(1) \cdot e^{y(3/2-\alpha)} e^{(1/2-\alpha)R/2} = O(1) \cdot n^{1/2-\alpha} \cdot e^{y(3/2-\alpha)},$$

from which we deduce

$$\mathcal{I}_n^{(3)}(y) = O(1) \cdot n^{1/2-\alpha} e^{y(3/2-\alpha)} \rho_{\mathrm{Po}}(y, k_n - 2). \tag{115}$$

We now apply Lemma 9.4 with $t = \frac{3}{2} - \alpha$ and get

$$\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y} \,\mathrm{d}y = O\left(1\right) n^{-(\alpha - \frac{1}{2})} \int_{\mathcal{K}_C(k_n)} e^{(3/2-\alpha)y} \rho_{\mathrm{Po}}(y, k_n - 2) e^{-\alpha y} \,\mathrm{d}y$$
$$= O\left(n^{-(\alpha-\frac{1}{2})} k_n^{2-4\alpha}\right).$$

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 $k_n^{6\alpha-3} k_n^{2-4\alpha} n^{-(\alpha-1/2)} = o\left(1\right)$ and hence for $1/2 < \alpha < 3/4$.

$$\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y} \,\mathrm{d}x \,\mathrm{d}y = 0,$$

For $\alpha \geq 3/4$ we observe that $2\alpha^2 + 2\alpha - 5/2 > 0$. Hence,

$$k_n^{2\alpha} n^{-(\alpha-\frac{1}{2})} k_n^{2-4\alpha} = O\left(n^{-(\alpha-1/2)} n^{\frac{2-2\alpha}{2\alpha+1}}\right) = O\left(n^{-\frac{2\alpha^2+2\alpha-5/2}{2\alpha+1}}\right) = o\left(1\right).$$

and we get for $\alpha \geq 3/4$

$$\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y} \,\mathrm{d}x \,\mathrm{d}y = 0.$$

The sums (97) and (98) Again, we will only consider (97) 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 $y_1 \geq (R - y) \wedge (1 - \varepsilon)R$, and $p_1 \in \mathcal{B}(y)$, $p_2 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1) \cap \mathcal{B}(y)$. We split this into three sub-domains: i) $y_2 \geq R-y$; ii) $R-y_1 \leq y_2 \leq R-y$ and iii) $y_2 < R-y_1$. Similar to the analysis above we define

$$
\begin{aligned}
\mathcal{D}_1 &:= \{(p_1,p_2) \ : \ p_1,p_2 \ \in \mathcal{P} \setminus \{(0,y)\}, \ y_1 \geq (1-\varepsilon)R \wedge (R-y), \ R-y \leq y_2 \leq R\} \\
\mathcal{D}_2 &:= \{(p_1,p_2) \ : \ p_1,p_2 \ \in \mathcal{P} \setminus \{(0,y)\}, \ y_1 \geq (1-\varepsilon)R \wedge (R-y), \ R-y_1 \leq y_2 \leq R-y\} \\
\mathcal{D}_3 &:= \{(p_1,p_2) \ : \ p_1,p_2 \ \in \mathcal{P} \setminus \{(0,y)\}, \ y_1 \geq (1-\varepsilon)R \wedge (R-y), \ y_2 \leq R-y_1\}
\end{aligned}
$$

and write, for $i=1,2,3$,

$$
\mathcal{I}_n^{(i)}(y) := \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_i} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \cdot \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} \cdot D_{\mathrm{Po}}(y,k_n-2;\mathcal{P}\setminus\{p_1,p_2\})\right].
$$

In the first case, note that for $y \in \mathcal{K}_C(k_n)$ we have, for small enough ε and sufficiently large n, $2y \leq 2(1+\varepsilon)\frac{R}{2\alpha+1} = o(R)$. Thus $y_1+y_2 \geq 2(R-y) = \Omega(R)$ and thus $p_2 \in \mathcal{B}(p_1)$ for large enough n. Furthermore, $y_2 > R-y_1+2\ln(\pi/2)$, 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:

$$
\begin{aligned}
\mathcal{I}_n^{(1)}(y) = O(1) \int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \times \\
\int_{-I_n}^{I_n} \int_{R-y}^{R} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} \rho_{\mathrm{Po}}(y,k_n-2) \cdot e^{-\alpha(y_2+y_1)} \,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1.
\end{aligned}
$$

We proceed to bound the integral:

$$
\begin{aligned}
&\int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \int_{-I_n}^{I_n} \int_{R-y}^{R} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)} \,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1 \\
&\leq \int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \int_{-I_n}^{I_n} \int_{R-y}^{R} e^{-\alpha(y_1+y_2)} \,\mathrm{d}y_2\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}x_1 \\
&= \left(\int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} e^{-\alpha y_1}\,\mathrm{d}y_1\,\mathrm{d}x_1\right)\left(\int_{-I_n}^{I_n} \int_{R-y}^{R} e^{-\alpha y_2}\,\mathrm{d}y_2\,\mathrm{d}x_2\right).
\end{aligned}
$$

We evaluate

$$
\int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} e^{-\alpha y_1} dy_1 dx_1 = O(1)\cdot n \cdot e^{-\alpha R+((\varepsilon R)\vee y))\alpha} = O(1)\cdot n\cdot e^{-\alpha R+\alpha y+\alpha\varepsilon R}
$$

and

$$
\int_{-I_n}^{I_n} \int_{R-y}^{R} e^{-\alpha y_2} dy_2 dx_2 = O(1)\cdot n\cdot e^{-\alpha R+\alpha y}.
$$

Also, $n \cdot e^{-\alpha R} = O(1)\cdot e^{(1/2-\alpha)R}$, whereby we deduce that

$$
\begin{aligned}
&\int_{\mathcal{D}_1} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)} dy_2 dx_2 dy_1 dx_1 \\
&\qquad = O(1)\cdot e^{(1-2\alpha)R+2\alpha y+\alpha\varepsilon R} = O(1)\cdot n^{2(1-2\alpha)+2\alpha\varepsilon}\cdot e^{2\alpha y}.
\end{aligned}
$$

With these computations we obtain

$$
\begin{aligned}
\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y &= O(1) n^{2(1-2\alpha)+2\alpha\varepsilon} \int_{\mathcal{K}_C(k_n)} e^{2\alpha y} \rho_{\mathrm{Po}}(y,k_n-2) e^{-\alpha y}\,\mathrm{d}y\,\mathrm{d}x \\
&= O(1) n^{2(1-2\alpha)+2\alpha\varepsilon}\, k_n^{2\alpha-1}.
\end{aligned}
$$

Thus, for $1/2 < \alpha < 3/4$, we have

$$k_n^{6\alpha-3}\, n^{2(1-2\alpha)+2\alpha\varepsilon}\, k_n^{2\alpha-1} = n^{2\alpha\varepsilon} \left(\frac{k_n^2}{n}\right)^{2(2\alpha-1)} = o(1),$$

provided that $\varepsilon = \varepsilon(\alpha) > 0$ is small enough, and hence for such ε

$$\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y}\, \mathrm{d}x\, \mathrm{d}y = 0.$$

When $\alpha \geq 3/4$ we have $2(1-2\alpha) < 1/2(4\alpha-1)$ and we get

$$k_n^{2\alpha}\, n^{2(1-2\alpha)+2\alpha\varepsilon} \cdot k_n^{2\alpha-1} \leq k_n^{4\alpha-1}\, n^{2(1-2\alpha)} n^{2\alpha\varepsilon} = o(1),$$

provided that ε is small enough, depending on α, so that

$$\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y) e^{-\alpha y}\, \mathrm{d}x\, \mathrm{d}y = 0.$$

We now consider the second sub-domain $\mathcal{D}_2$. The Campbell-Mecke formula yields that:

$$\begin{aligned}
\mathcal{I}_n^{(2)}(y) &= \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_2} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} D_{\mathrm{Po}}(y, k_n-2; \mathcal{P}\setminus\{p_1\})\right] \\
&= O(1)\rho_{\mathrm{Po}}(y, k_n-2) \cdot \int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \times \\
&\quad \int_{-I_n}^{I_n} \int_{R-y_1}^{R-y} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}x_2\, \mathrm{d}y_1\, \mathrm{d}x_1.
\end{aligned}$$

We bound the integral as follows:

$$\begin{aligned}
&\int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \int_{-I_n}^{I_n} \int_{R-y_1}^{R-y} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}x_2\, \mathrm{d}y_1\, \mathrm{d}x_1 \\
&\leq \int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \int_{-I_n}^{I_n} \int_{R-y_1}^{R-y} \mathbb{1}_{\{p_2\in\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)}\, \mathrm{d}y_2\, \mathrm{d}x_2\, \mathrm{d}y_1\, \mathrm{d}x_1.
\end{aligned}$$

Now, by Lemma 2.2,

$$\begin{aligned}
&\int_{-I_n}^{I_n} \int_{R-y_1}^{R-y} \mathbb{1}_{\{p_2\in\mathcal{B}(y)\}} \cdot e^{-\alpha y_2} dy_2 dx_2 = O(1) \cdot e^{y/2} \int_{R-y_1}^{R-y} e^{(1/2-\alpha)y_2} dy_2 \\
&= O(1) \cdot e^{y/2+(1/2-\alpha)(R-y_1)}.
\end{aligned}$$

We then integrate with respect to y_1:

$$\begin{aligned}
O(1) \cdot e^{y/2} \cdot &\int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} e^{(1/2-\alpha)(R-y_1)} e^{-\alpha y_1} dy_1 dx_1 \\
&\leq O(1) \cdot e^{y/2+(1/2-\alpha)R} \cdot \int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} e^{(\alpha-1/2)y_1} e^{-\alpha y_1} dy_1 dx_1 \\
&= O(1) \cdot e^{y/2+(1-\alpha)R-((1-\varepsilon)R\wedge(R-y))/2} \\
&= O(1) \cdot e^{y/2+(1/2-\alpha)R+((\varepsilon R)\vee y)/2} \\
&= O(1) \cdot e^{y+(1/2-\alpha)R+\varepsilon R} = O(1) \cdot n^{1-2\alpha+\varepsilon} \cdot e^{y}.
\end{aligned}$$

Therefore we get

$$
\begin{aligned}
&\int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y \\
&= O\left(n^{1-2\alpha+\varepsilon}\right) \int_{\mathcal{K}_C(k_n)} \rho_{\mathrm{Po}}(y, k_n-2) e^{y} e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y \\
&= O\left(1\right) n^{1-2\alpha+\varepsilon} k_n^{-2\alpha+1},
\end{aligned}
$$

where we used Lemma 9.4 with $t=1$.

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

$$
\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.
$$

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

$$
\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y) e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.
$$

For the third sub-domain $\mathcal{D}_3$ we shall use (114) which states that if $p_2 = (x_2, y_2) \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1) \cap \mathcal{B}(y)$ and $y_2 \le R - y_1$, then $y_2 \ge \hat{y}(x_1, y_1)$, where $\hat{y}(x_1, y_1) = \left(2\log(|x_1| e^{-y_1/2}) - \log c_K\right) \vee 0$ (cf. (118)). Moreover, $p_2 \in \check{\mathcal{B}}_{\mathrm{Po}}(p_1)$.

Again, we will use the Campbell-Mecke formula:

$$
\begin{aligned}
\mathcal{I}_n^{(3)}(y) &= \mathbb{E}\left[\sum_{(p_1,p_2)\in\mathcal{D}_3} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \cdot \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} \cdot D_{\mathrm{Po}}(y, k_n-2; \mathcal{P}\setminus\{p_1,p_2\})\right] \\
&= O(1)\rho_{\mathrm{Po}}(y, k_n-2) \int_{-I_n}^{I_n} \int_{(1-\varepsilon)R\wedge(R-y)}^{R} \mathbb{1}_{\{p_1\in\mathcal{B}(y)\}} \times \\
&\qquad \int_{-I_n}^{I_n} \int_{0}^{R-y_1} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha(y_1+y_2)} dy_2 dx_2 dy_1 dx_1
\end{aligned}
$$

The inner integral with respect to $p_2 := (x_2, y_2)$ is

$$
\begin{aligned}
&\int_{-I_n}^{I_n} \int_{0}^{R-y_1} \mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}} e^{-\alpha y_2} dy_2 dx_2 \\
&\qquad \le \int_{-I_n}^{I_n} \int_{0}^{R-y_1} \mathbb{1}_{\left\{y_2\ge\hat{y}(x_1,y_1),\, p_2\in\check{\mathcal{B}}_{\mathrm{Po}}((0,y))\right\}} e^{-\alpha y_2} dy_2 dx_2 \\
&\qquad = O(1) e^{y/2} \int_{\hat{y}(x_1,y_1)}^{R-y_1} e^{y_2/2-\alpha y_2} dy_2 \\
&\qquad = O(1) e^{y/2+(1/2-\alpha)\hat{y}(x_1,y_1)}.
\end{aligned}
$$

Thus, we get

$$\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R\wedge(R-y)}^{R}\mathbb{1}_{\{p_1\in\mathcal{B}(y)\}}\int_{-I_n}^{I_n}\int_{0}^{R-y_1}\mathbb{1}_{\{p_2\in\mathcal{B}_{\mathrm{Po}\triangle\infty}(p_1)\cap\mathcal{B}(y)\}}\times$$
$$e^{-\alpha(y_1+y_2)}dy_2dx_2dy_1dx_1$$
$$\leq O(1)\int_{-I_n}^{I_n}\int_{(1-\varepsilon)R\wedge(R-y)}^{R}e^{y/2+(1/2-\alpha)\hat{y}(x_1,y_1)}e^{-\alpha y_1}dy_1dx_1.$$

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 $c(x_1,y_1)$:

$$\int_0^{I_n}e^{\hat{y}(x_1,y_1)(1/2-\alpha)}dx_1=\int_{c_Ke^{y_1/2}}^{I_n}e^{c(x_1,y_1)(1/2-\alpha)}dx_1+\int_0^{c_Ke^{y_1/2}}dx_1.$$

The first integral becomes:

$$\int_{c_Ke^{y_1/2}}^{I_n}e^{\hat{y}(x_1,y_1)(1/2-\alpha)}dx_1=O(1)\cdot\int_{c_Ke^{y_1/2}}^{I_n}x_1^{1-2\alpha}e^{-\frac{y_1}{2}(1-2\alpha)}dx_1$$
$$=\begin{cases}O(R)\cdot e^{-y_1/2+\alpha y_1}\cdot e^{\frac{R}{2}2(1-\alpha)} & \text{if } \alpha\leq 1\\ O(1)\cdot e^{-y_1/2+\alpha y_1+2(1-\alpha)y_1/2} & \text{if } \alpha>1\end{cases}$$
$$=\begin{cases}O(R)\cdot e^{(\alpha-1/2)y_1}\cdot n^{2(1-\alpha)} & \text{if } \alpha\leq 1\\ O(1)\cdot e^{y_1/2} & \text{if } \alpha>1\end{cases}.$$

The second integral trivially gives:

$$\int_0^{c_Ke^{y_1/2}}dx_1=O(1)\cdot e^{y_1/2}.$$

Putting these two together we conclude that

$$e^{y/2}\cdot\int_0^{I_n}e^{\hat{y}(x_1,y_1)(1/2-\alpha)}dx_1=O(1)\cdot e^{y_1/2+y(3/2-\alpha)}.$$

Now, we integrate these with respect to y_1:

$$n^{2(1-\alpha)}\cdot\int_{(1-\varepsilon)R\wedge(R-y)}^{R}e^{(\alpha-1/2)y_1-\alpha y_1}dy_1=O(1)\cdot n^{2(1-\alpha)}\cdot e^{-R/2+\varepsilon R/2+y/2}$$
$$=O(1)\cdot n^{1-2\alpha+\varepsilon}\cdot e^{y/2}.$$

Therefore, we conclude that

$$\mathcal{I}_n^{(3)}(y)=O\left(R\right)n^{1-2\alpha+\varepsilon(2\alpha-1)}\,e^{y/2}\rho_{\mathrm{Po}}(y,k_n-2)$$

and hence, using again Lemma 9.4,

$$\int_{\mathcal{K}_C(k_n)}\mathcal{I}_n^{(3)}(y)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y=O\left(R\right)n^{1-2\alpha+\varepsilon(2\alpha-1)}\int_{\mathcal{K}_C(k_n)}e^{y/2}\rho_{\mathrm{Po}}(y,k_n-2)e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y$$
$$=O\left(R\right)n^{1-2\alpha+\varepsilon(2\alpha-1)}k_n^{-2\alpha+1}.$$

It follows that for $\varepsilon=\varepsilon(\alpha)$ small enough

$$k_n^{6\alpha-3}Rn^{1-2\alpha+\varepsilon(2\alpha-1)}k_n^{-2\alpha+1}=Rn^{\varepsilon(2\alpha-1)}\left(\frac{k_n^2}{n}\right)^{2\alpha-1}=o\left(1\right)$$

and hence for $\alpha > 1/2$,
$$\lim_{n\to\infty} k_n^{6\alpha-3} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.$$
Since $2\alpha - 1 \geq 1/2$ when $\alpha \geq 3/4$ it immediately follows that
$$\lim_{n\to\infty} k_n^{2\alpha} \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y) e^{-\alpha y}\,\mathrm{d}x\,\mathrm{d}y = 0.$$

The sums (99) **and** (100) Again, the analysis for both terms are similar and we shall analyze (99). Let us set $p = (0, y)$. Recall that $\mathcal{B}_{\mathrm{Po}\triangle\infty}(y) \cap \mathcal{R}([R - y + 2\log\left(\frac{\pi}{2}\right), R]) = \emptyset$. Thus, the summand in (99) is equal to 0, when $y_1 > R - y + 2\log(\pi/2)$.

Recall the definition of the extended ball $\breve{\mathcal{B}}_{\mathrm{Po}}(p)$ around p (112) that contains both $\mathcal{B}(p)$ and $\mathcal{B}_\infty(p)$
$$\breve{\mathcal{B}}_{\mathrm{Po}}(y) := \{p' : y' < R - y,\ |x'| < (1+K)e^{\frac{1}{2}(y+y')}\},$$
and that we have $\mathbb{E}\left[D_{\mathrm{Po}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\})\right] = \rho_{\mathrm{Po}}(y, k_n - 2)$.

Further, observe that,
$$\mathcal{B}(y) \cap \mathcal{R}([0, R - y)) \subseteq \breve{\mathcal{B}}_{\mathrm{Po}}(y)$$
and
$$\mathcal{B}(y) \cap \mathcal{R}([R - y, R]) = \mathcal{R}([R - y, R]).$$
We thus conclude that
$$\mathcal{B}(y) \subseteq \breve{\mathcal{B}}_{\mathrm{Po}}(y) \cup \mathcal{R}([R - y, R]). \tag{116}$$
Hence, if we set
$$h_y(p_1) := \mathbb{1}_{\{p_1 \in \mathcal{B}(p) \setminus \mathcal{B}_\infty(y)\}} \cdot \left(\mu\left(\breve{\mathcal{B}}_{\mathrm{Po}}(p_1) \cap \breve{\mathcal{B}}_{\mathrm{Po}}(y)\right) + \mu\left(\mathcal{R}([R - y, R])\right)\right),$$
then
$$\begin{aligned}
&\mathbb{1}_{\{p_1 \in \mathcal{B}(p) \setminus \mathcal{B}_\infty(y)\}} \cdot \mathbb{E}\left[\left(\sum_{p_2 \in \mathcal{P} \setminus \{p, p_1\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y) \cap \mathcal{B}_\infty(p_1)\}}\right) \cdot D_{\mathrm{Po}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\})\right] \\
&= O(1) \cdot \mathbb{1}_{\{p_1 \in \mathcal{B}(y) \setminus \mathcal{B}_\infty(y)\}} \cdot \mu(\mathcal{B}(y) \cap \mathcal{B}(p_1)) \rho_{\mathrm{Po}}(y, k_n - 2) \\
&\leq O(1) \cdot h_y(p_1) \rho_{\mathrm{Po}}(y, k_n - 2).
\end{aligned}$$

To calculate the expectation of the above function we need to approximate the intersection of the two balls $\breve{\mathcal{B}}_{\mathrm{Po}}(y)$ and $\breve{\mathcal{B}}_{\mathrm{Po}}(p_1)$, where $p_1 = (x_1, y_1)$. Let us assume without loss of generality that $x_1 > 0$. The right boundary of $\breve{\mathcal{B}}_{\mathrm{Po}}(y)$ is given by the equation $x = x(y') = (1+K)e^{\frac{1}{2}(y+y')}$ whereas the left boundary of $\breve{\mathcal{B}}_{\mathrm{Po}}(p_1)$ is given by the curve $x = x(y') = x_1 - (1+K)e^{\frac{1}{2}(y_1+y')}$.

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 $p_1 = (x_1, y_1) \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p)$, we also have $x_1 > e^{\frac{y+y_1}{2}}$. Therefore,
$$e^{\hat{y}/2} > \frac{1}{1+K}\,\frac{e^{\frac{y+y_1}{2}}}{e^{y/2} + e^{y_1/2}} \geq \frac{1}{2(1+K)}\,\frac{e^{\frac{y_1+y}{2}}}{e^{(y\vee y_1)/2}} > \frac{1}{2(1+K)}\,e^{(y\wedge y_1)/2}. \tag{117}$$
The above yields
$$\hat{y} > (y \wedge y_1) - 2\log(2(1+K)) := \hat{y}(y_1, y). \tag{118}$$

which, in turn, implies the following

$$p \in \check{\mathcal{B}}_{\mathrm{Po}}((0,y)) \cap \check{\mathcal{B}}_{\mathrm{Po}}(p_1) \Rightarrow y(p) \geq \hat{y}(y_1, y). \tag{119}$$

We thus conclude that

$$\mathcal{B}\left(p_1\right) \cap \mathcal{B}\left(p\right) \subseteq \left(\check{\mathcal{B}}_{\mathrm{Po}}(p) \cap \mathcal{R}([\hat{y}(y_1,y), R])\right) \cup \mathcal{R}([R-y, R]),$$

which in turn implies that

$$\mu\left(\check{\mathcal{B}}_{\mathrm{Po}}(p_1) \cap \mathcal{B}\left(p\right)\right) \leq \mu\left(\check{\mathcal{B}}_{\mathrm{Po}}(p) \cap \mathcal{R}([\hat{y}(y_1,y), R]\right) + \mu(\mathcal{R}([R-y, R])).$$

Therefore,

$$\begin{aligned} h_y(p_1, \mathcal{P}) \leq \mathbb{1}_{\{p_1 \in \mathcal{B}(p) \setminus \mathcal{B}_\infty(p)\}} \mu\left(\check{\mathcal{B}}_{\mathrm{Po}}(p) \cap \mathcal{R}([\hat{y}(y_1,y), R])\right) \\ + \mathbb{1}_{\{p_1 \in \mathcal{B}(p) \setminus \mathcal{B}_\infty(p)\}} \mu\left(\mathcal{R}([R-y, R])\right). \end{aligned}$$

Now, the Campbell-Mecke formula gives

$$\begin{aligned} &\mathbb{E}\left[\sum_{\substack{p_1,p_2 \in \mathcal{P}\setminus\{(0,y)\} \\ y(p_1) \geq K}} \mathbb{1}_{\{p_1 \in \mathcal{B}(y) \setminus \mathcal{B}_\infty(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y) \cap \mathcal{B}_\infty(y)\}}\, D_{\mathrm{Po}}(y, k_n - 2; \mathcal{P} \setminus \{p_1, p_2\})\right] \\ &\leq \mathbb{E}\left[\left(\sum_{p_1 \in \mathcal{P}} h_y(p_1, \mathcal{P} \setminus \{p_1\})\right)\right] \\ &= \frac{\nu\alpha}{\pi} \int_{\mathcal{R}} \mathbb{E}\left[h_y(p_1, \mathcal{P} \setminus \{p_1\})\right] e^{-\alpha y_1}\, \mathrm{d}x_1\, \mathrm{d}y_1 \\ &\leq \frac{\nu\alpha}{\pi} \int_{\mathcal{R}} \mathbb{1}_{\{p_1 \in \mathcal{B}(p) \setminus \mathcal{B}_\infty(p)\}} \mu\left(\check{\mathcal{B}}_{\mathrm{Po}}(p) \cap \mathcal{R}([\hat{y}(y_1,y), R])\right) e^{-\alpha y_1}\, \mathrm{d}x_1\, \mathrm{d}y_1 \qquad (120) \\ &\quad + \frac{\nu\alpha}{\pi} \int_{\mathcal{R}} \mathbb{1}_{\{p_1 \in \mathcal{B}(p) \setminus \mathcal{B}_\infty(p)\}} \mu\left(\mathcal{R}([R-y, R])\right) e^{-\alpha y_1}\, \mathrm{d}x_1\, \mathrm{d}y_1. \qquad (121) \end{aligned}$$

Recall that $(\mathcal{B}_{\mathrm{Po}\triangle\infty}\left((0,y)\right)) \cap \mathcal{R}([R-y+2\log\left(\frac{\pi}{2}\right), R]) = \emptyset$. We will first calculate the measures μ appearing in (120) and (121). The first one is:

$$\begin{aligned} \mu\left(\check{\mathcal{B}}_{\mathrm{Po}}(y) \cap \mathcal{R}([c(y_1,y), R])\right) &\leq (1+K)\frac{\nu\alpha}{\pi} \cdot e^{y/2} \int_{\hat{y}(y_1,y)}^{R} e^{-(\alpha-\frac{1}{2})y'}\, dy' \\ &= O\left(e^{\frac{y}{2}-(\alpha-\frac{1}{2})(y \wedge y_1)}\right). \end{aligned}$$

The second term is:

$$\mu\left(\mathcal{R}([R-y, R])\right) = \frac{\nu\alpha}{\pi} \int_{R-y}^{R} \pi e^{\frac{R}{2}} e^{-\alpha y'}\, dy' = O\left(e^{\frac{R}{2}} e^{-\alpha(R-y)}\right) = O\left(e^{\alpha y - (\alpha-\frac{1}{2})R}\right).$$

Using these, we get

$$\begin{aligned} &\int_{\mathcal{R}([0,R-y_n+2\ln\frac{\pi}{2}])} \mathbb{E}\left[h_y(p_1, \mathcal{P} \setminus \{p_1\})\right] e^{-\alpha y_1}\, dx_1\, dy_1 \\ &= O\left(1\right) \int_{\mathcal{R}([0,R-y+2\ln\frac{\pi}{2}])} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(p)\}} e^{\frac{y}{2}-(\alpha-\frac{1}{2})(y\wedge y_1)-\alpha y_1}\, dx_1\, dy_1 \qquad (122) \\ &\quad + O\left(1\right) \int_{\mathcal{R}([0,R-y+2\ln\frac{\pi}{2}])} \mathbb{1}_{\{p_1 \in \mathcal{B}((0,y))\}} e^{\alpha y - (\alpha-\frac{1}{2})R - \alpha y_1}\, dx_1\, dy_1. \qquad (123) \end{aligned}$$

Now, Lemma 2.2 implies that for any $y_1 \in [0, R-y+2\ln\frac{\pi}{2}]$, we have

$$\int_{-I_n}^{I_n} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}}\, dx_1 \le 2Ke^{\frac{3}{2}(y_1+y)-R}.$$

Therefore, (122) is

$$\begin{aligned}
&O(1)\cdot e^{2y-R}\int_0^{R-y+2\ln\frac{\pi}{2}} e^{\frac{3y_1}{2}-(\alpha-\frac{1}{2})(y_1\wedge y)-\alpha y_1}\, dy_1\\
&\quad = O(1)\cdot e^{2y-R}\left(\int_0^{y} e^{\frac{3y_1}{2}-(2\alpha-\frac{1}{2})y_1}\, dy_1 + e^{-(\alpha-\frac{1}{2})y}\int_y^{R-y+2\ln\frac{\pi}{2}} e^{(\frac{3}{2}-\alpha)y_1}\, dy_1\right)\\
&\quad = O(1)\left(\begin{cases} e^{(4-2\alpha)y-R}, & \text{if } \alpha<1\\ R\cdot e^{2y-R}, & \text{if } \alpha\ge 1\end{cases} + \begin{cases} e^{-(\alpha-\frac{1}{2})R+y}, & \text{if } \alpha<3/2\\ R\cdot e^{2(2-\alpha)y-R}, & \text{if } \alpha\ge 3/2\end{cases}\right).
\end{aligned}$$

Similarly, for (123) we have

$$\begin{aligned}
&\int_{\mathcal{R}([0,R-y+2\ln\frac{\pi}{2}])} \mathbb{1}_{\{p_1\in\mathcal{B}_{\mathrm{Po}\triangle\infty}((0,y))\}} e^{\alpha y-(\alpha-\frac{1}{2})R-\alpha y_1}\, dx_1\, dy_1\\
&\quad = e^{\frac{3y}{2}-R+\alpha y-(\alpha-\frac{1}{2})R}\cdot\int_0^{R-y+2\ln\frac{\pi}{2}} e^{\frac{3y_1}{2}-\alpha y_1}\, dy_1\\
&\quad = O(1)\cdot\begin{cases} e^{\frac{3y}{2}-R+\alpha y-(\alpha-\frac{1}{2})R+(\frac{3}{2}-\alpha)(R-y)}, & \text{if } \alpha<3/2\\ R\cdot e^{(\frac{3}{2}+\alpha)y-(\alpha+\frac{1}{2})R}, & \text{if } \alpha\ge 3/2\end{cases}\\
&\quad = O(1)\cdot\begin{cases} e^{-(2\alpha-1)R+2\alpha y}, & \text{if } \alpha<3/2\\ R\cdot e^{(\frac{3}{2}+\alpha)y-(\alpha+\frac{1}{2})R}, & \text{if } \alpha\ge 3/2\end{cases}.
\end{aligned}$$

We thus conclude, using $2(2-\alpha)y\le y$ for $\alpha>3/2$, that

$$\mathbb{E}\left[\left(\sum_{p_1\in\mathcal{P}\setminus\{p\}} h_y(p_1)\right)\right] \le O(1)\cdot\left(\mathcal{I}_n^{(1)}(y)+\mathcal{I}_n^{(2)}(y)+\mathcal{I}_n^{(3)}(y)\right), \tag{124}$$

where

$$\begin{aligned}
\mathcal{I}_n^{(1)}(y) &= \begin{cases} e^{(4-2\alpha)y-R}, & \text{if } \alpha<1\\ R\cdot e^{2y-R}, & \text{if } \alpha\ge 1\end{cases},\\
\mathcal{I}_n^{(2)}(y) &= \begin{cases} e^{-(\alpha-\frac{1}{2})R+y}, & \text{if } \alpha<3/2\\ R\cdot e^{y-R}, & \text{if } \alpha\ge 3/2\end{cases}\\
\mathcal{I}_n^{(3)}(y) &= \begin{cases} e^{-(2\alpha-1)R+2\alpha y}, & \text{if } \alpha<3/2\\ R\cdot e^{(\frac{3}{2}+\alpha)y-(\alpha+\frac{1}{2})R}, & \text{if } \alpha\ge 3/2\end{cases}.
\end{aligned}$$

We proceed to calculate:

$$\int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left(\sum_{p_1\in\mathcal{P}} h_y(p_1,\mathcal{P}\setminus\{p_1\})\right)\right]\cdot\rho_{\mathrm{Po}}(y,k_n-2)e^{-\alpha y}\,\mathrm{d}y.$$

For this we define

$$M_i = \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(i)}(y)\rho_{\mathrm{Po}}(y,k_n-1)e^{-\alpha y}\,\mathrm{d}y$$

so that

$$\int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\left(\sum_{p_1\in\mathcal{P}\setminus\{(0,y)\}} h_y(p_1)\right)\right] \rho_{\mathrm{Po}}(y,k_n-1)e^{-\alpha y}\,\mathrm{d}y = O\left(M_1+M_2+M_3\right).$$

Computing each of the integral separately we obtain, using Lemma 9.4 and the fact that $n=\nu e^{R/2}$,

$$M_1 := \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(1)}(y)\rho_{\mathrm{Po}}(y,k_n-1)e^{-\alpha y}\,\mathrm{d}y = O(1)\cdot\begin{cases}\frac{k_n^{7-6\alpha}}{n^2}, & \text{if } \alpha<1\\ R\frac{k_n^{3-2\alpha}}{n^2}, & \text{if } \alpha\geq 1\end{cases}.$$

$$M_2 := \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(2)}(y)\rho_{\mathrm{Po}}(y,k_n-1)e^{-\alpha y}dy = O(1)\cdot\begin{cases}\frac{k_n^{1-2\alpha}}{n^{2\alpha-1}}, & \text{if } \alpha<3/2\\ R\frac{k_n^{1-2\alpha}}{n^2}, & \text{if } \alpha\geq 3/2\end{cases}$$

and finally

$$M_3 := \int_{\mathcal{K}_C(k_n)} \mathcal{I}_n^{(3)}(y)\rho_{\mathrm{Po}}(y,k_n-1)e^{-\alpha y}dy = O(1)\cdot\begin{cases}\frac{k_n^{2\alpha-1}}{n^{4\alpha-2}}, & \text{if } \alpha<3/2\\ R\cdot\frac{k_n^{2}}{n^{2\alpha+1}}, & \text{if } \alpha\geq 3/2\end{cases}.$$

Now, we will consider the two cases according to the value of α. First we note that $R=O\left(\log(n)\right)$ and since $k_n=O(n^{\frac{1}{2\alpha+1}})$ and $\alpha>1/2$ we have that $Rk_n^2n^{-1}=o\left(1\right)$. Assume first that $1/2<\alpha<3/4$. In this case, we want to show that

$$\lim_{n\to\infty} k_n^{6\alpha-3}(M_1+M_2+M_3)=0. \tag{125}$$

Using the above expression for M_i, we have

$$k_n^{6\alpha-3}(M_1+M_2+M_3) = O(1)\cdot k_n^{6\alpha-3}\left(\frac{k_n^{7-6\alpha}}{n^2}+\frac{k_n^{1-2\alpha}}{n^{2\alpha-1}}+\frac{k_n^{2\alpha-1}}{n^{4\alpha-3}}.\right)$$

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

$$k_n^{6\alpha-3}\frac{k_n^{7-6\alpha}}{n^2} = \left(\frac{k_n^2}{n}\right)^2 = o\left(1\right).$$

The second term yields:

$$k_n^{6\alpha-3}\frac{k_n^{-2\alpha+1}}{n^{2\alpha-1}} = \left(\frac{k_n^2}{n}\right)^{2\alpha-1} = o\left(1\right).$$

Finally, the third one yields:

$$k_n^{6\alpha-3}\cdot\frac{k_n^{2\alpha-1}}{n^{4\alpha-2}} = \left(\frac{k_n^2}{n}\right)^{4\alpha-2} = o\left(1\right).$$

For $\alpha\geq 3/4$, we would like to show that

$$\lim_{n\to\infty} k_n^{2\alpha}\cdot(M_1+M_2+M_3)=0. \tag{126}$$

Firstly, we note that each M_i is as above if $3/4<\alpha<1$. Therefore, since for this range $2\alpha<6\alpha-3$ the result follows from the above analysis. Next we consider the case $1\leq\alpha<3/2$. Here, only the value of M_1 changes and we compute that

$$k_n^{6\alpha-3}M_1 = O\left(1\right)\log(n)n^{-2}k_n^{4\alpha} \leq O\left(\log(n)\right)\left(\frac{k_n^2}{n}\right)^2 = o\left(1\right),$$

so that (126) holds for $3/4 < \alpha < 1$.

Proceeding with the case $\alpha \geq 3/2$, it is only M_2 and M_3 that change values. In particular, for any $\alpha \geq 3/2$ we have

$$\frac{k_n}{n} M_2 = O(1) R \frac{k_n}{n^2} = o(1).$$

Also,

$$k_n^{2\alpha} M_3 = O(1) R \frac{k_n^{2\alpha+2}}{n^{2\alpha+1}} = Ro\left(\frac{n^{\alpha+1}}{n^{2\alpha+1}}\right) = o(1),$$

since $k_n = o(n^{1/2})$ and hence (126) holds. This finished the proof for (99).

The sum of (101) Using the Campbell-Mecke formula, we write

$$\begin{aligned}
&\mathbb{E}\left[\sum_{p_1,p_2 \in \mathcal{P}\setminus\{(0,y)\},\ y_1<K}^{\neq} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right] \\
&\quad \leq \int_0^K \int_{-I_n}^{I_n} \int_0^R \int_{-I_n}^{I_n} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}} e^{-\alpha y_2} e^{-\alpha y_1} \, \mathrm{d}x_2 \, \mathrm{d}y_2 \, \mathrm{d}x_1 \, \mathrm{d}y_1 \\
&\quad \leq \mu(\mathcal{B}(y)) \cdot \int_{-I_n}^{I_n} \int_0^K \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}} e^{-\alpha y_1} dx_1 dy_1.
\end{aligned}$$

Recall that by Lemma 6.9, $\mu(\mathcal{B}(y)) = O(1)e^{y/2}$. We bound the integral using Lemma 2.2. In particular, (9) implies that if $p_1 = (x_1, y_1) \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)$, then because $y_1 < K$

$$|x_1 - e^{(y+y_1)/2}| \leq e^{(y+y_1)/2} \cdot K e^{y+y_1-R} = O(1) e^{(y+y_1)/2} \cdot e^{y-R}.$$

Therefore,

$$\int_{-I_n}^{I_n} \int_0^K \mathbb{1}_{\{(x_1,y_1) \in \mathcal{B}_{\mathrm{Po}\triangle\infty}((0,y))\}} e^{-\alpha y_1} dx_1 dy_1 = O(1) \cdot e^{y-R} \cdot \int_0^K e^{(y+y_1)/2} e^{-\alpha y_1} dy_1 = O(1) \cdot e^{3y/2-R},$$

and hence

$$\mathbb{E}\left[\sum_{\substack{p_1,p_2 \in \mathcal{P}\setminus\{(0,y)\} \\ y_1<K}}^{\neq} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right] = O(1) \cdot e^{2y-R}.$$

Now, we integrate this over y to obtain that

$$\begin{aligned}
&\int_{\mathcal{K}_C(k_n)} \mathbb{E}\left[\sum_{\substack{p_1,p_2 \in \mathcal{P}\setminus\{(0,y)\} \\ y_1<K}}^{\neq} \mathbb{1}_{\{p_1 \in \mathcal{B}_{\mathrm{Po}\triangle\infty}(y)\}} \mathbb{1}_{\{p_2 \in \mathcal{B}(y)\cap\mathcal{B}_\infty(y)\}}\right] e^{-\alpha y} \, \mathrm{d}y \\
&\quad = O(1) e^{-R} \int_{\mathcal{K}_C(k_n)} e^{2y-\alpha y} \, \mathrm{d}y = O(1) n^{-2} \begin{cases} k_n^{4-2\alpha}, & \text{if } \alpha < 2 \\ \log k_n, & \text{if } \alpha = 2 \\ 1, & \text{if } \alpha > 2 \end{cases}.
\end{aligned}$$

To finish the argument assume first that $1/2 < \alpha \leq 3/4$. In this case,

$$k_n^{6\alpha-3} n^{-2} k_n^{4-2\alpha} = n^{-2} \cdot k_n^{4\alpha+1} = o(1).$$

For $3/4 \leq \alpha < 2$ we use that $2\alpha < 6\alpha - 3$, so that $k_n^{2\alpha} n^{-2} k_n^{4-2\alpha} = o(1)$. Finally, when $\alpha \geq 2$, we have that

$$k_n^{2\alpha}(\log(k_n) \wedge 1) n^{-2} \leq k_n^{2\alpha+1} n^{-2} = O\left(n^{-1}\right) = o(1).$$

which completes the proof for (101) and thus the proof of Proposition 5.3. $\square$

9.3 Coupling G_n to G_{Po}

Now that we have established the equivalence of the clustering function between the Poissonized KPKVB graph G_{Po} and the finite box graph G_{box} the final step is to relate the clustering function in G_{Po} to the KPKVB graph G_n. As mentioned in Section 5.2, this is done by moving from $c(k_n; G_n)$ to the adjusted clustering function $c^*(k_n; G_n)$ (Lemma 5.1) and then to $c^*(k_n; G_{\mathrm{Po}})$ (Proposition 5.2). We start with a technical lemma on the difference between the number of vertices with degree k_n in both models. Then we give the proof of Proposition 5.2 and after that we prove Lemma 5.1.

Let $N_n(k_n)$ and $N_{\mathrm{Po}}(k_n)$ denote, respectively, the number of vertices with degree k_n in G_n and G_{Po}. We will first show that $|N_n(k_n) - N_{\mathrm{Po}}(k_n)|$ is small. Recall that we consider the standard coupling between the binomial and Poisson process to couple G_n and G_{Po}. That is, we take a sequence of i.i.d. random elements $u_1, u_2, \dots$ uniformly on the hyperbolic disk of radius R, i.e. according to the distribution (1). Then the KPKVB graph consists of the first n points and the poissonized version of the first $N \stackrel{d}{=} \mathrm{Po}(n)$ points (N is a Poisson random variable with expectation n, independent of $u_1, u_2, \dots$). Under this coupling $N_n(k) = \sum_{j=1}^n \mathbb{1}_{\{\deg_n(u_j)=k\}}$ denotes the (random) number of degree k vertices in the KPKVB graph G_n model with n vertices and $N_{\mathrm{Po}}(k) = \sum_{j=1}^N \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_j)=k\}}$ denotes the (random) number of degree k vertices in the poissonized KPKVB graph G_{Po}.

Lemma 9.5. *Let $(k_n)_{n\geq 1}$ be sequence of natural numbers with $k_n = o(n^{\frac{1}{2\alpha+1}})$. Then, on the coupling space described above,*

$$\mathbb{E}\left[|N_n(k_n) - N_{\mathrm{Po}}(k_n)|\right] = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right) = o\left(nk^{-(2\alpha+1)}\right),$$

and in particular,

$$\mathbb{E}\left[N_n(k_n)\right] = \Theta\left(nk_n^{-(2\alpha+1)}\right).$$

Proof. The second claim follows immediately from the first and the fact that $\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] = \Theta\left(nk_n^{-(2\alpha+1)}\right)$, see Lemma 9.3.

To prove the first statement let $u = (r, \theta) \in \mathbb{H}$. Slightly abusing notation we write $B_{\mathbb{H}}(u)$ for those u' such that $d_{\mathbb{H}}(u, u') \leq R$ and $\mu_{\mathbb{H}}(y)$ for the measure of $B_{\mathbb{H}}(u)$ with respect to the (α, R)-quasi uniform measure $g(r, \theta)$ defined in (6). Note that if u' is sampled according to g, then $\mu_{\mathbb{H}}(u)/n$ denotes the probability that $u' \in B_{\mathbb{H}}(u)$. In particular, if $G_n = G(n; \alpha, \nu)$ is a KPKVB graph and $u \in \mathcal{V}_n$, then the degree distribution of u, conditioned on its coordinates, is Binomial on $n-1$ trials with success probability $\mu_{\mathbb{H}}(u)/n$. Let us define for any $n \geq 1$, $\deg_n(u) = \mathrm{Bin}(n-1, \mu_{\mathbb{H}}(u)/n)$ then if U is sampled according to the (α, R)-quasi uniform distribution (1), $\mathbb{P}\left(\deg_n(U) = k_n\right)$ denotes the degree distribution in G_n and $\mathbb{E}\left[N_n(k_n)\right] = n\mathbb{P}\left(\deg_n(U) = k_n\right)$.

Recall that we couple the KPKVB graph G_n with the Poisson version G_{Po} by considering an infinite supply of i.i.d. points $u_1, u_2, \dots$ sampled on $\mathbb{H}$ according to the (α, R)-quasi uniform distribution (1), take $N \stackrel{d}{=} \mathrm{Po}(n)$, independently of $u_1, u_2, \dots$, and set $\mathcal{V}_n = \{u_1, \dots, u_n\}$ and $\mathcal{V}_{\mathrm{Po}} = \{u_1, \dots, u_N\}$. Observe that $\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right] = \mathbb{E}\left[N\mathbb{P}\left(\deg_{N-1}(U) = k_n\right)\right]$, where U is sampled from the (α, R)-quasi uniform distribution (1). It is important to note that under this coupling the only difference between G_n and G_{Po} is in the number of vertices.

Denote by $\mathcal{V}_n(k_n)$ and $\mathcal{V}_{\mathrm{Po}}(k_n)$ the set of points u_i with degree k_n in G_n and G_{Po}, respectively. Then the result follows if we can show that

$$\mathbb{E}\left[\sum_{i=1}^{n\wedge N} \mathbb{1}_{\{u_i \in \mathcal{V}_n(k_n) \Delta \mathcal{V}_{\mathrm{Po}}(k_n)\}} \middle| N\right] = o\left(n\right)\mathbb{P}\left(\deg_{N-1}(U) = k_n \middle| N\right)$$

We first use a Chernoff based large deviation result for a Poisson random variable (132), which implies that for any $C > 0$ we have $N \in [n - C\sqrt{n\log n}, n + C\sqrt{n\log n}]$ with probability $1 - n^{-C^2/2}$.

Since we can select $C > 0$ arbitrarily large it follows that we only need to prove the result conditional on the event $|N-n| \leq C\sqrt{n\log(n)}$. Note that if $N = n$ then

$$\sum_{i=1}^{n\wedge N} \mathbb{1}_{\{u_i \in \mathcal{V}_n(k_n)\Delta\mathcal{V}_{\mathrm{Po}}(k_n)\}} = 0.$$

Hence if we denote by A_n^- the event that $n - C\sqrt{n\log(n)} \leq N < n$ and by A_n^+ the event that $n < N \leq n + C\sqrt{n\log(n)}$, we need to show that

$$\mathbb{E}\left[\sum_{i=1}^{n\wedge N} \mathbb{1}_{\{u_i \in \mathcal{V}_n(k_n)\Delta\mathcal{V}_{\mathrm{Po}}(k_n)\}} \middle| A_n^+\right] = o(n)\,\mathbb{P}\left(\deg_N(U) = k_n \middle| A_n^+\right)$$

and likewise with A_n^+ replaced by A_n^-.

We first consider the case where $n < N \leq n + C\sqrt{n\log(n)}$, i.e. the case where we have more vertices in G_{Po} than in G_n. The proof for the other case is similar and we omit it. Let $W = \{u_{n+1}, \ldots, u_N\}$ be the set of vertices in G_{Po} that are not in G_n. To ease notation we let $\mathbb{E}_n$ and $\mathbb{P}_n$ denote the conditional expectation and probability on the event A_n^+. First we note that

$$\begin{aligned}
\mathbb{E}_n\left[\sum_{i=n+1}^{N} \mathbb{1}_{\{u_i \in \mathcal{V}_n(k_n)\Delta\mathcal{V}_{\mathrm{Po}}(k_n)\}}\right] &= \mathbb{E}_n\left[\sum_{i=n+1}^{N} \mathbb{1}_{\{\deg_{Po}(u_i) = k_n\}}\right] \\
&= \frac{(N-n)}{n}\mathbb{1}_{\{A_n^+\}} n\mathbb{P}_n\left(\deg_N(U) = k_n\right) \\
&\leq C\sqrt{\frac{\log(n)}{n}} n\mathbb{P}_n\left(\deg_N(U) = k_n\right) \\
&= o(n)\, n\mathbb{P}_n\left(\deg_N(U) = k_n\right).
\end{aligned}$$

Hence we are left to consider the sum over all vertices in $\mathcal{V}_n$.

Let $u_i \in \mathcal{V}_n(k_n)$. Then since $N > n$ we have that $u_i \in \mathcal{V}_{\mathrm{Po}}(k_n)$ and hence $u_i \in \mathcal{V}_n(k_n)\Delta\mathcal{V}_{\mathrm{Po}}(k_n)$ if and only if $|B_{\mathbb{H}}(u_i) \cap W| \geq 1$ and either $u_i \in \mathcal{V}_n(k_n)$ or $u_i \in \mathcal{V}_{\mathrm{Po}}(k_n)$. From this it follows that

$$\sum_{i=1}^{n} \mathbb{1}_{\{u_i \in \mathcal{V}_n(k_n)\Delta\mathcal{V}_{\mathrm{Po}}(k_n)\}} = \sum_{i=1}^{n} \left(\mathbb{1}_{\{\deg_n(u_i) = k_n\}} + \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_i) = k_n\}}\right) \mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W| \geq 1\}}$$

We will bound this sum further. For any $u \in \mathbb{H}$ let $Q_n(u)$ denote the event that $|\mu_{\mathbb{H}}(u) - k_n| \leq C\sqrt{k_n \log(k_n)}$ and denote by $Q_n(u)^c$ its complement. Then,

$$\begin{aligned}
&\sum_{i=1}^{n} \left(\mathbb{1}_{\{\deg_n(u_i) = k_n\}} + \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_i) = k_n\}}\right) \mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W| \geq 1\}} \\
&\leq \sum_{i=1}^{n} \left(\mathbb{1}_{\{\deg_n(u_i) = k_n\}} + \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_i) = k_n\}}\right) \mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W| \geq 1\}} \mathbb{1}_{\{Q_n(u_i)\}} \\
&\quad + \sum_{i=1}^{n} \left(\mathbb{1}_{\{\deg_n(u_i) = k_n\}} + \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_i) = k_n\}}\right) \mathbb{1}_{\{Q_n(u_i)^c\}}.
\end{aligned}$$

The second sum splits into two parts. The expectation of the first part yields

$$\begin{aligned}
\mathbb{E}_n\left[\sum_{i=1}^{n} \mathbb{1}_{\{\deg_n(u_i) = k_n\}} \mathbb{1}_{\{Q_n(u_i)^c\}}\right] &= n\,\mathbb{P}\left(\mathrm{Bin}\left(n-1, \frac{\mu_{\mathbb{H}}(U)}{n}\right) = k_n, Q_n(U)^c\right) \\
&= O\left(nk_n^{-\frac{C^2}{3}}\right) = o\left(nk_n^{-(2\alpha+1)}\right),
\end{aligned}$$

where we used Lemma D.3, for any $C^2/3 > 2\alpha + 1$. Similarly, Lemma D.3 implies that the expectation of the second part is

$$\begin{aligned}
\mathbb{E}_n\left[\sum_{i=1}^{n} \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_i)=k_n\}}\mathbb{1}_{\{Q_n(u_i)^c\}}\right] &= n\,\mathbb{P}\left(\mathrm{Bin}\left(N-1, \frac{\mu_{\mathbb{H}}(U)}{n}\right) = k_n, Q_n(U)^c\right)\\
&= O\left(nk_n^{-\frac{C^2}{3}}\right) = o\left(nk_n^{-(2\alpha+1)}\right).
\end{aligned}$$

Hence it remains to show that

$$\mathbb{E}_n\left[\sum_{i=1}^{n}\left(\mathbb{1}_{\{\deg_n(u_i)=k_n\}} + \mathbb{1}_{\{\deg_{\mathrm{Po}}(u_i)=k_n\}}\right)\mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W|\geq 1\}}\mathbb{1}_{\{Q_n(u_i)\}}\right] = o\left(nk_n^{-(2\alpha+1)}\right).$$

We will show that

$$\mathbb{E}_n\left[\sum_{i=1}^{n}\mathbb{1}_{\{\deg_n(u_i)=k_n\}}\mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W|\geq 1\}}\mathbb{1}_{\{Q_n(u_i)\}}\right] = o\left(nk_n^{-(2\alpha+1)}\right).$$

The other term follows using almost identical arguments and is omitted. To prove the above result we note that since $N > n$ we have $\mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W|\geq 1\}} \leq \sum_{j=n+1}^{N}\mathbb{1}_{\{u_j\in B_{\mathbb{H}}(u_i)\}}$. Thus

$$\begin{aligned}
&\mathbb{E}_n\left[\sum_{i=1}^{n}\mathbb{1}_{\{\deg_n(u_i)=k_n\}}\mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W|\geq 1\}}\mathbb{1}_{\{Q_n(u_i)\}}\right]\\
&\leq \sum_{i=1}^{n}\sum_{j=n+1}^{N}\mathbb{E}_n\left[\mathbb{1}_{\{\deg_n(u_i)=k_n\}}\mathbb{1}_{\{Q_n(u_i)\}}\mathbb{1}_{\{u_j\in B_{\mathbb{H}}(u_i)\}}\right]\\
&= n(N-n)\mathbb{P}_n\left(\mathrm{Bin}\left(n-2, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n, Q_n(U_1), U_2 \in B_{\mathbb{H}}(U_1)\right)\\
&= n(N-n)\mathbb{E}_n\left[\mathbb{1}_{\{Q_n(U_1)\}}\mathbb{P}\left(\mathrm{Bin}\left(n-2, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n, U_2 \in B_{\mathbb{H}}(U_1)\middle| U_1\right)\right]\\
&= n(N-n)\mathbb{E}_n\left[\mathbb{1}_{\{Q_n(U_1)\}}\mathbb{P}\left(\mathrm{Bin}\left(n-2, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n\middle| U_1\right)\frac{\mu_{\mathbb{H}}(U_1)}{n}\right],
\end{aligned}$$

where the last step follows since, conditional on U_1, the events $\mathrm{Bin}\left(n-2, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n$ and $U_2 \in B_{\mathbb{H}}(U_1)$ are independent and $\mathbb{P}\left(U_2 \in B_{\mathbb{H}}(U_1)|\, U_1\right) = \mu_{\mathbb{H}}(U_1)/n$. In addition, on the event $Q_n(U_1)$ we have that $\mu_{\mathbb{H}}(U_1) \leq k_n + C\sqrt{k_n\log(k_n)}$ and thus the second statement of Lemma D.3 implies that

$$\begin{aligned}
&\mathbb{E}_n\left[\mathbb{1}_{\{Q_n(U_1)\}}\mathbb{P}\left(\mathrm{Bin}\left(n-2, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n\middle| U_1\right)\right]\\
&= (1+o(1))\mathbb{E}_n\left[\mathbb{1}_{\{Q_n(U_1)\}}\mathbb{P}\left(\mathrm{Bin}\left(N-1, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n\middle| U_1\right)\right]\\
&\leq O(1)\,\mathbb{P}_n\left(\mathrm{Bin}\left(N-1, \frac{\mu_{\mathbb{H}}(U_1)}{n}\right) = k_n\right)\\
&= O(1)\,\mathbb{P}_n\left(\deg_N(U_1) = k_n\right).
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
&\mathbb{E}_n\left[\sum_{i=1}^{n}\mathbb{1}_{\{\deg_n(u_i)=k_n\}}\mathbb{1}_{\{|B_{\mathbb{H}}(u_i)\cap W|\geq 1\}}\mathbb{1}_{\{Q_n(u_i)\}}\right]\\
&\leq (N-n)\left(k_n + C\sqrt{k_n\log(k_n)}\right)\mathbb{P}_n\left(\deg_N(U_1) = k_n\right)
\end{aligned}$$

$$= (1 + o(1))\sqrt{n\log(n)}k_n\mathbb{P}_n\left(\deg_N(U_1) = k_n\right)$$
$$= o\left(n\mathbb{P}_n\left(\deg_N(U_1) = k_n\right)\right) = o\left(\mathbb{E}_n\left[N_{\mathrm{Po}}(k_n)\right]\right),$$

where we again used that $\sqrt{n\log(n)}k_n = o(n)$. $\square$

With this result we can now prove Proposition 5.2, which states

$$\lim_{n\to\infty} s(k_n)\mathbb{E}\left[|c^*(k_n; G_n) - c^*(k_n; G_{\mathrm{Po}})|\right] = 0.$$

Proof of Proposition 5.2. First we note that Proposition 5.3, 5.4 and 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

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

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 again denote by $\mathcal{V}_n(k_n)$ the set of degree k_n vertices in G_n and by $\mathcal{V}_{\mathrm{Po}}(k_n)$ the set of degree k_n vertices in G_{Po}. If a vertex is contained in both sets, it must have the same degree in both the Poisson and KPKVB 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 $\mathcal{V}_n(k_n)\Delta\mathcal{V}_{\mathrm{Po}}(k_n)$. By Lemma 9.5 the expected number vertices in this set is $\mathbb{E}\left[|N_n(k_n) - N_{\mathrm{Po}}(k_n)|\right] = o\left(\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]\right)$. Therefore we have that

$$\mathbb{E}\left[|c^*(k_n; G_n) - c^*(k_n; G_{\mathrm{Po}})|\right] \le \frac{\mathbb{E}\left[|N_n(k_n) - N_{\mathrm{Po}}(k_n)|\right]}{(1 + o(1))\mathbb{E}\left[N_{\mathrm{Po}}(k_n)\right]}\mathbb{E}\left[c^*(k_n; G_{\mathrm{Po}})\right] = o(1)\,\mathbb{E}\left[c^*(k_n; G_{\mathrm{Po}})\right],$$

which finishes the proof. $\square$

Finally we prove 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

$$A_n = \left\{|N_n(k_n) - \mathbb{E}\left[N_n(k_n)\right]| \le \mathbb{E}\left[N_n(k_n)\right]^{\frac{1+\delta}{2}}\right\}.$$

Since $N_n(k_n) = \sum_{i=1}^n \mathbb{1}_{\{\deg_n(u_i)=k_n\}}$, with u_i sampled according to (1), it follows from Lemma D.1, with $c = \mathbb{E}\left[N_n(k_n)\right]^{-\frac{1-\delta}{2}}$, that

$$\mathbb{P}\left(A_n\right) \ge 1 - O\left(e^{-\frac{\mathbb{E}[N_n(k_n)]^{\delta}}{2}}\right) = 1 - O\left(e^{-\frac{n^{\delta}k_n^{-\delta(2\alpha+1)}}{2}}\right), \tag{127}$$

where for the last part we used that $\mathbb{E}\left[N_n(k_n)\right] = \Theta\left(nk_n^{-(2\alpha+1)}\right)$ (cf. Lemma 9.5).

On the event A_n

$$\left|\frac{\mathbb{E}\left[N_n(k_n)\right]}{N_n(k_n)} - 1\right| \le \frac{\mathbb{E}\left[N_n(k_n)\right]^{\frac{1+\delta}{2}}}{\mathbb{E}\left[N_n(k_n)\right] + \mathbb{E}\left[N_n(k_n)\right]^{\frac{1+\delta}{2}}} \le \mathbb{E}\left[N_n(k_n)\right]^{-\frac{1-\delta}{2}}.$$

Therefore we have

$$\begin{aligned}
\mathbb{E}\left[|c^*(k_n;G_n) - c(k_n;G_n)|\right] &\le \mathbb{E}\left[|c^*(k_n;G_n) - c(k_n;G_n)|\,\mathbb{1}_{\{A_n\}}\right] + O\left(1 - \mathbb{P}\left(A_n\right)\right) \\
&= \mathbb{E}\left[c^*(k_n;G_n)\left|\frac{\mathbb{E}\left[N_n(k_n)\right]}{N_n(k_n)} - 1\right|\mathbb{1}_{\{A_n\}}\right] + O\left(e^{-\frac{n^{\delta}k_n^{-\delta(2\alpha+1)}}{2}}\right) \\
&\le \mathbb{E}\left[c^*(k_n;G_n)\right]\mathbb{E}\left[N_n(k_n)\right]^{-\frac{1-\delta}{2}} + O\left(e^{-\frac{n^{\delta}k_n^{-\delta(2\alpha+1)}}{2}}\right).
\end{aligned}$$

The second term is clearly $o\left(s(k_n)\right)$. The first term is clearly $o\left(\mathbb{E}\left[c^*(k_n;G_n)\right]\right)$ and since Propositions 5.2-5.5 imply that

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

the first term is also $o\left(s(k_n)\right)$. □

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 in this model and his helpful insights during discussions of the topic.

References

[1] Nist digital library of mathematical functions. Release 1.0.22 of 2019-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[2] Mohammed Amin Abdullah, Nikolaos Fountoulakis, and Michel Bode. Typical distances in a geometric model for complex networks. *Internet Mathematics*, aug 2017.

[3] George E Andrews, Richard Askey, and Ranjan Roy. *Special functions*, volume 71. Cambridge university press, 2000.

[4] Thomas Bläsius, Tobias Friedrich, and Anton Krohmer. Hyperbolic random graphs: Separators and treewidth. In *24th Annual European Symposium on Algorithms (ESA 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[5] Thomas Bläsius, Tobias Friedrich, and Anton Krohmer. Cliques in hyperbolic random graphs. *Algorithmica*, 80(8):2324–2344, 2018.

[6] Thomas Bläsius, Tobias Friedrich, Anton Krohmer, and Sören Laue. Efficient embedding of scale-free graphs in the hyperbolic plane. *IEEE/ACM Transactions on Networking*, 26(2):920–933, 2018.

[7] 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.

[8] Michel Bode, Nikolaos Fountoulakis, and Tobias Müller. The probability of connectivity in a hyperbolic model of complex networks. *Random Structures & Algorithms*, 49(1):65–94, 2016.

[9] Marián Boguná, Fragkiskos Papadopoulos, and Dmitri Krioukov. Sustaining the internet with hyperbolic mapping. *Nature communications*, 1:62, 2010.

[10] Elisabetta Candellero and Nikolaos Fountoulakis. Clustering and the hyperbolic geometry of complex networks. *Internet Mathematics*, 12(1-2):2–53, 2016.

[11] Hanshuang Chen, Chuansheng Shen, Gang He, Haifeng Zhang, and Zhonghuai Hou. Critical noise of majority-vote model on complex networks. *Physical Review E*, 91(2):022816, 2015.

[12] 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.

[13] Brian Davies. *Integral transforms and their applications*, volume 41. Springer Science & Business Media, 2012.

[14] Arthur Erdélyi, Wilhelm Magnus, Fritz Oberhettinger, Francesco G Tricomi, et al. Higher transcendental functions, vol. 1, 1953.

[15] Nikolaos Fountoulakis. On the evolution of random graphs on spaces of negative curvature. *arXiv*, 2012.

[16] 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.

[17] 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.

[18] Tobias Friedrich and Anton Krohmer. On the diameter of hyperbolic random graphs. *SIAM Journal on Discrete Mathematics*, 32(2):1314–1334, 2018.

[19] Izrail Solomonovich Gradshteyn and Iosif Moiseevich Ryzhik. *Table of Integrals, Series, and Products*. Elsevier, 2015.

[20] 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.

[21] Marcos Kiwi and Dieter Mitsche. A bound for the diameter of random hyperbolic graphs. In *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics*, pages 26–39. Society for Industrial and Applied Mathematics, 2015.

[22] Marcos Kiwi and Dieter Mitsche. On the second largest component of random hyperbolic graphs. *arXiv*, 2017.

[23] Marcos Kiwi and Dieter Mitsche. Spectral gap of random hyperbolic graphs and related parameters. *The Annals of Applied Probability*, 28(2):941–989, 2018.

[24] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010.

[25] Günter Last and Mathew Penrose. *Lectures on the Poisson process*, volume 7. Cambridge University Press, 2017.

[26] Yudell L Luke. *Mathematical functions and their approximations*. Academic Press, 2014.

[27] Cornelis Simon Meijer. On the G-function. I-VIII. In *Nederl. Akad. Wetensch., Proc.*, volume 49, 1946.

[28] M Carmen Miguel, Jack T Parley, and Romualdo Pastor-Satorras. Effects of heterogeneous social interactions on flocking dynamics. *Physical review letters*, 120(6):068303, 2018.

[29] Tobias Müller and Merlijn Staps. The diameter of kpkvb random graphs. *arXiv*, 2017.

[30] Takashi Owada and D Yogeshwaran. Sub-tree counts on hyperbolic random geometric graphs. *arXiv*, 2018.

[31] Mathew Penrose. *Random geometric graphs*. Number 5. Oxford university press, 2003.

[32] Clara Stegehuis, Remco van der Hofstad, and Johan SH van Leeuwaarden. Scale-free network clustering in hyperbolic and other random graphs. *Journal of Physics A: Mathematical and Theoretical*, 2019.

[33] John Stillwell. *Geometry of surfaces*. Springer Science & Business Media, 2012.

[34] Nico M Temme. *Special functions: An introduction to the classical functions of mathematical physics*. John Wiley & Sons, 2011.

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 [27] is defined as

$$G_{p,q}^{m,\ell}\left(z \middle| \begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right) = \frac{1}{2\pi\iota}\int_L \frac{\prod_{j=1}^m \Gamma(b_j - t)\prod_{j=1}^{\ell}\Gamma(1 - a_j + t)}{\prod_{j=m+1}^{q}\Gamma(1 - b_j + t)\prod_{j=\ell+1}^{p}\Gamma(a_j - t)}\, z^t \,\mathrm{d}t, \tag{128}$$

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 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 $G_{p,q}^{m,\ell}\left(z \middle| \begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right)$ see [19, 26].

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 $\xi, s > 0$,*

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

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 $e^{-x} = \frac{1}{2\pi\iota}\int_{c-\iota\infty}^{c+\iota\infty}\Gamma(p)x^{-p}dp$ for $c > 0$ (see [13, p.196]). Applying the change of variable $p(r) = q - r$ yields $e^{-x} = \frac{1}{2\pi\iota}\int_{c+q-\iota\infty}^{c+q+\iota\infty}\Gamma(q-r)x^{r-q}dr$, then multiplying both sides with $-x^{q-1}$ gives $-x^{q-1}e^{-x} = -\frac{1}{2\pi\iota}\int_{c+q-\iota\infty}^{c+q+\iota\infty}\Gamma(q-r)x^{r-1}dr$. Now, integrating both sides gives $\int_x^\infty t^{q-1}e^{-t}dt = \frac{1}{2\pi\iota}\int_{c+q-\iota\infty}^{c+q+\iota\infty}\frac{\Gamma(q-r)}{-r}x^r dr$. 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. $\Gamma^+(q, x) = G_{1,2}^{2,0}\left(x \middle| \begin{matrix}1\\ q, 0\end{matrix}\right)$. The claim follows by plugging in $q = -a-1$ and $x = \frac{\xi}{s}$. □

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 [34] (or see [1, Section 8.17 (ii)]). 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 $c = a+b$ [3] (or see [1, Section 15.4(ii)]). 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
$$\lim_{z\to 0} F(a, b, c, 1-z) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \tag{129}$$
if $c = a+b$ then
$$\lim_{z\to 0} \frac{F(a, b, c, 1-z)}{\log(z)} = -\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, \tag{130}$$
and finally, when $c-a-b < 0$
$$\lim_{z\to 0} \frac{F(a, b, c, 1-z)}{z^{c-a-b}} = \frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}. \tag{131}$$

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 (130), together with the fact that $(1-z)^{3/2} \sim 1$ as $z \to 0$, 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 (131) we get
$$\lim_{z\to 0} z^{4\alpha-3} B^-(1-z, 2\alpha, 3-4\alpha) = \frac{1}{2\alpha}\frac{\Gamma(2\alpha+1)\Gamma(4\alpha-3)}{\Gamma(2\alpha)\Gamma(4\alpha-2)} = \frac{\Gamma(4\alpha-3)}{\Gamma(4\alpha-2)} = \frac{1}{4\alpha-3}.$$

$\square$

C Some results on functions

Lemma C.1. *For any $0 < \lambda < 1$ there exists a $K > 0$, such that for all $0 < x \leq (1-\lambda)2$*

$$\frac{1}{2}\arccos(1-x)\left(1 - \frac{(1+\sqrt{2})x}{1+x}\right) \leq \frac{x}{\sqrt{1-(1-x)^2}} \leq \frac{1}{2}\arccos(1-x)\left(1 + \frac{(K+1)x}{1-x}\right).$$

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 $0 < x < 2$

$$0 < \sqrt{2x}\left(1 - \frac{x}{\sqrt{8}}\right) \leq \arccos(1-x) \leq \sqrt{2x}\left(1 + \frac{x}{\sqrt{8}}\right)$$

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

Here we summarize several known results for random variables and provide one technical lemma for Binomial random variables.

We start with the following concentration result which follows from [17, 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*

$$\mathbb{P}\left(|X_n - \mathbb{E}\left[X_n\right]| > c\mathbb{E}\left[X_n\right]\right) \leq 2e^{-\frac{c^2\mathbb{E}[X_n]^2}{2}}.$$

Next we recall two Chernoff bound for Poisson and Binomial random variables. The following Chernoff bound can be found in [31, Lemma 1.2].

Lemma D.2. *Let* $\mathrm{Po}(\lambda)$ *denote a Poisson random variable with mean* λ *and let* $H(x) = x\log(x) - x + 1$*. Then*

$$\begin{aligned} \mathbb{P}\left(\mathrm{Po}(\lambda) \geq k\right) &\leq e^{-\lambda H(k/\lambda)} \quad \textit{for all } k \geq \lambda \\ \mathbb{P}\left(\mathrm{Po}(\lambda) \leq k\right) &\leq e^{-\lambda H(k/\lambda)} \quad \textit{for all } k \leq \lambda. \end{aligned}$$

It follows from the above lemma that

$$\mathbb{P}\left(|\mathrm{Po}(\lambda) - \lambda| \geq x\right) \leq 2e^{-\frac{x^2}{2(\lambda+x)}}. \tag{132}$$

In particular, if $\lambda_n \to \infty$, then, for any $C > 0$,

$$\mathbb{P}\left(|\mathrm{Po}(\lambda_n) - \lambda_n| \geq C\sqrt{\lambda_n \log(\lambda_n)}\right) \leq 2e^{-\frac{C^2\lambda_n \log(\lambda_n}{2\left(\lambda_n + C\sqrt{\lambda_n \log(\lambda_n)}\right)}} = O\left(\lambda_n^{-\frac{C^2}{2}}\right). \tag{133}$$

Let $\mathrm{Bin}(n,p)$ denote a Binomial random variable with n trials and success probability p, and $0 < \delta < 1$. Then we have the following well-known Chernoff bound.

$$\mathbb{P}\left(|\mathrm{Bin}(n,p) - np| > \delta np\right) \leq e^{-\frac{\delta^2 np}{3}}. \tag{134}$$

The following technical lemma establishes two results that are important in Section 9.

Lemma D.3. *Let* $\mathrm{Bin}(m,p)$ *denote a Binomial random variable with* m *trials and success probability* p*, let* $k_n \to \infty$ *be a sequence of integers such that* $k_n = o(n)$ *and fix some* $C > 0$*. Then the following holds for any* $s, t > 0$*:*

1. *for any* $0 < p_n < 1$ *such that* $|np_n - k_n| > C\sqrt{k_n \log(k_n)}$,

$$\mathbb{P}\left(\mathrm{Bin}(n-t, p_n) = k_n\right) = O\left(k_n^{-\frac{C^2}{3}}\right).$$

2. *for any* $0 < p_n < 1$ *such that* $|np_n - k_n| \leq C\sqrt{k_n \log(k_n)}$ *and any sequence of integers* m_n *such that* $|m_n - n| \leq C\sqrt{n \log(n)}$

$$\mathbb{P}\left(\mathrm{Bin}(n-t, p_n) = k_n\right) = (1 + o(1))\mathbb{P}\left(\mathrm{Bin}(m_n - s, p_n) = k_n\right).$$

Proof. First we observe that

$$\frac{\partial}{\partial x}\mathbb{P}\left(\mathrm{Bin}(m,x) = k\right) = \binom{m}{k}\left(kx^{k-1}(1-x)^{m-k} - (m-k)x^k(1-x)^{m-k-1}\right).$$

Hence, the function $x \mapsto \mathbb{P}\left(\mathrm{Bin}(m,x) = k\right)$ attains it maximum at $x = k/m$ and is strictly increasing on $(0, k/m]$ and strictly decreasing on $[k/m, 1)$.

We proceed with proving the first statement. Consider the case where $np_n < k_n - C\sqrt{k_n \log(k_n}$. Define

$$q_n = \frac{k_n - C\sqrt{k_n \log(k_n)}}{n-t}$$

and set $Y_n = \mathrm{Bin}(n-t, q_n)$. Then, since $q_n < k_n/(n-t)$ we get

$$\mathbb{P}\left(\mathrm{Bin}(n-t, p_n) = k_n\right) \leq \mathbb{P}\left(Y_n = k_n\right) \leq \mathbb{P}\left(Y_n > k_n - 1\right) = \mathbb{P}\left(Y_n > (1+\delta_n)(n-t)q_n\right),$$

with

$$\delta_n = \frac{k_n - 1 - nq_n}{(n-t)q_n} = \frac{C\sqrt{k_n \log(k_n)} - 1}{k_n - C\sqrt{k_n \log(k_n)}}.$$

By a Chernoff bound we get

$$\mathbb{P}\left(\mathrm{Bin}(n-t,p_n)=k_n\right) \le \mathbb{P}\left(Y_n > (1+\delta_n)nq_n\right) \le e^{-\frac{\delta_n^2 nq_n}{2}} \le e^{-\frac{\delta_n^2 nq_n}{3}}.$$

The result now follows by observing that $\delta_n^2 nq_n = \Omega\left(C^2\log(k_n)\right)$.

For the case $np_n > k_n - C\sqrt{k_n\log(k_n}$ we set $q_n = (k_n + C\sqrt{k_n\log(k_n)})/n > k_n/n$ and

$$\delta_n = \frac{nq_n-(k_n+1)}{nq_n} = \frac{C\sqrt{k_n\log(k_n)}-1}{k_n+C\sqrt{k_n\log(k_n)}}.$$

Then, $p_n > q_n > k_n/n$ and $\delta_n^2 nq_n = \Omega\left(C^2\log(k_n)\right)$. Thus by another Chernoff bound

$$\begin{aligned}\mathbb{P}\left(\mathrm{Bin}(n-t,p_n)=k_n\right) &\le \mathbb{P}\left(Y_n = k_n\right) \le \mathbb{P}\left(Y_n < k_n+1\right)\\ &= \mathbb{P}\left(Y_n < (1-\delta_n)nq_n\right) \le e^{-\frac{\delta_n^2 nq_n}{3}} = O\left(k_n^{-\frac{C^2}{3}}\right).\end{aligned}$$

Now let us prove the second statement. For this we first note that by Stirling's formula

$$\binom{n-t}{k_n} \sim (2\pi k_n)^{-1/2}\left(\frac{k_n}{n-t}\right)^{-k_n}\left(1-\frac{k_n}{n-t}\right)^{k_n-(n-t)},$$

and similarly for $\binom{m_n-s}{k_n}$. Therefore

$$\binom{n-t}{k_n}\binom{m_n-s}{k_n}^{-1} = (1+o(1))\left(\frac{n-t}{m_n-s}\right)^{k_n}\left(1-\frac{k_n}{n-t}\right)^{k_n-(n-t)}\left(1-\frac{k_n}{m_n-s}\right)^{m_n-s-k_n}$$

and thus

$$\frac{\mathbb{P}\left(\mathrm{Bin}(n-t,p_n)=k_n\right)}{\mathbb{P}\left(\mathrm{Bin}(m_n-s,p_n)=k_n\right)} = (1+o(1))\left(\frac{n-t-k_n}{n-t-p_n}\right)^{k_n-(n-t)}\left(\frac{m_n-s-k_n}{m_n-s-p_n}\right)^{k_n-(m_n-s)}$$

Let us rewrite the first multiplicative term as

$$\left(\frac{n-t-k_n}{n-t-p_n}\right)^{k_n-(n-t)} = \left(1-\frac{k_n-p_n}{n-t-p_n}\right)^{k_n-(n-t)} := (1-x_n)^{k_n-(n-t)}.$$

Then, using that for all $-1/2 \le x \le 1/2$, $-x-x^2 \le \log(1-x) \le -x$, we get

$$e^{-(k_n-(n-t))\left(x_n+x_n^2\right)} \le \left(\frac{n-t-k_n}{n-t-p_n}\right)^{k_n-(n-t)} \le e^{-(k_n-(n-t))x_n}.$$

Similarly,

$$e^{-(m_n-s-k_n)\left(y_n+y_n^2\right)} \le \left(\frac{m_n-s-k_n}{m_n-s-p_n}\right)^{m_n-s-k_n} \le e^{-(m_n-s-k_n)y_n},$$

where

$$y_n := \frac{k_n-p_n}{m_n-s-p_n}.$$

We first show that

$$\lim_{n\to\infty} e^{-(k_n-(n-t))x_n}e^{-(m_n-s-k_n)y_n} = 1. \tag{135}$$

With some algebra we get

$$\begin{aligned}&-(k_n-(n-t))x_n-(m_n-s-k_n)y_n\\ &\quad= \left(1-\frac{k_n-p_n}{n-t-p_n}\right)(k_n-p_n) - \left(1-\frac{k_n-p_n}{m_n-s-p_n}\right)(k_n-p_n)\end{aligned}$$

$$= (k_n - p_n)^2 \left(\frac{1}{m_n - s - p_n} - \frac{1}{n - t - p_n} \right)$$

$$= (k_n - p_n)^2 \frac{n - m_n - t + s}{(m_n - s - p_n)(n - t - p_n)}.$$

Now by our assumptions

$$(k_n - p_n)^2 = \Theta\left(k_n \log(k_n)\right)$$

while

$$m_n - s - p_n = \Theta\left(n - t - p_n\right) = \Theta\left(n\right)$$

and

$$-C\sqrt{n\log(n)} \le n - m_n \le C\sqrt{n\log(n)}.$$

Therefore we conclude that

$$(k_n - p_n)^2 \frac{n - m_n - t + s}{(m_n - s - p_n)(n - t - p_n)} = \Theta\left(\frac{k_n \log(k_n)\sqrt{n\log(n)}}{n^2}\right)$$

from which (135) follows.

In a similar way it follows that

$$\lim_{n\to\infty} e^{-(k_n-(n-t))x_n^2} e^{-(m_n-s-k_n)y_n^2} = 1,$$

which implies that

$$\lim_{n\to\infty} \frac{\mathbb{P}\left(\mathrm{Bin}(n-t, p_n) = k_n\right)}{\mathbb{P}\left(\mathrm{Bin}(m_n - s, p_n) = k_n\right)} = 1,$$

as thus finishes the proof. □

E Code for the simulations

The simulations of the clustering coefficient and function in the KPKVB model were done using Wolfram Mathematica 11.1. The simulation dots for the clustering coefficient in Figure 2 were generated by the following code (where in the second line, the entire script was also run for the values `nu=1` and `nu=0.5`):

```
1  n=10000;
2  nu=2;
3  R=2*Log[n/nu];
4  plotpoints=20;
5  reps=100;
6  Plotingdataalpha = ConstantArray[0,{plotpoints,2}];
7  SeedRandom[1];
8  For[z=1,z<=plotpoints,z++,a=0.4+z (4.6/plotpoints); sum=0;
9      For[r=1,r<=reps,r++,V = ConstantArray[0,{n,2}];
10         For[i=1,i<=n,i++,
11             V[[i,1]]=RandomReal[{-Pi ,Pi }];
12             V[[i,2]]=ArcCosh[RandomReal[{0,1}](Cosh[a*R]-1)+1]/a];
13         A= ConstantArray[0,{n,n}];
14         For[i=1,i<=n,i++,
15             For[j=1,j<=n,j++,
16                 If[Cosh[V[[i,2]]]Cosh[V[[j,2]]] - Sinh[V[[i,2]]]Sinh[V[[
                       j,2]]]Cos[Abs[V[[i,1]]-V[[j,1]]]] <= Cosh[R] && i
                       != j,A[[i,j]]=1,A[[i,j]]=0]]];
17         g = AdjacencyGraph[A];
```

```
        sum=sum+MeanClusteringCoefficient[g]];
    Plotingdataalpha[[z,1]]=a;
    Plotingdataalpha[[z,2]]=1.0*sum/reps;]
Print[Plotingdataalpha]
```

The simulation dots for the clustering function in Figure 3 were generated by the following code (where in the third line, the entire script was also run for the values `nu=1` and `nu=0.5`):

```
1  n=10000;
2  a=0.8;
3  nu=2;
4  R=2*Log[n/nu];
5  plotpoints=24;
6  reps=100;
7  Plotingdatak = ConstantArray[0,{reps,plotpoints,2}];
8  SeedRandom[1];
9  For[r=1,r<=reps,r++,V = ConstantArray[0,{n,2}];
10     For[i=1,i<=n,i++,
11         V[[i,1]]=RandomReal[{-Pi ,Pi }];
12         V[[i,2]]=ArcCosh[RandomReal[{0,1}](Cosh[a*R]-1)+1]/a];
13     A= ConstantArray[0,{n,n}];
14     For[i=1,i<=n,i++,
15         For[j=1,j<=n,j++,
16             If[Cosh[V[[i,2]]] Cosh[V[[j,2]]] - Sinh[V[[i,2]]] Sinh[V[[j
                   ,2]]] Cos[Abs[V[[i,1]]-V[[j,1]]]]  <= Cosh[R]  && i != j,A
                   [[i,j]]=1,A[[i,j]]=0]]];
17     g = AdjacencyGraph[A];
18     For[k=1,k<=plotpoints,k++,
19         sum=0;
20         result=0;
21         nrdegk=0;
22         For[v =1,v<=n,v++;
23             If[VertexDegree[g,v]==k+1,
24                 result=result+LocalClusteringCoefficient[g,v];nrdegk
                       ++]];
25         Plotingdatak[[r,k,1]]=k+1;
26         If[nrdegk>0,Plotingdatak[[r,k,2]]=1.0*result/nrdegk]];]
27 Print[Mean[Plotingdatak]];
```