

# DIAMETER IN ULTRA-SMALL SCALE-FREE RANDOM GRAPHS

FRANCESCO CARAVENNA<sup>1</sup>, ALESSANDRO GARAVAGLIA<sup>2</sup>, AND REMCO VAN DER HOFSTAD<sup>2</sup>

**ABSTRACT.** It is well known that many random graphs with infinite variance degrees are ultra-small. More precisely, for configuration models and preferential attachment models where the proportion of vertices of degree at least  $k$  is approximately  $k^{-(\tau-1)}$  with  $\tau \in (2, 3)$ , typical distances between pairs of vertices in a graph of size  $n$  are asymptotic to  $\frac{2 \log \log n}{|\log(\tau-2)|}$  and  $\frac{4 \log \log n}{|\log(\tau-2)|}$ , respectively. In this paper, we investigate the behavior of the diameter in such models. We show that the diameter is of order  $\log \log n$  precisely when the minimal forward degree  $d_{\text{fwd}}$  of vertices is at least 2. We identify the exact constant, which equals that of the typical distances plus  $2/\log d_{\text{fwd}}$ . Interestingly, the proof for both models follows identical steps, even though the models are quite different in nature.

## 1. INTRODUCTION AND RESULTS

In this paper, we study the diameter of two different random graph models: the *configuration model* and the *preferential attachment model*, when these two models have a power-law degree distribution with exponent  $\tau \in (2, 3)$ , so that the degrees have finite mean but infinite variance. In this first section, we give a brief introduction to the models, stating the main technical conditions required as well as the two main results proved in the paper.

Throughout the paper, we write “with high probability” to mean “with probability  $1 - o(1)$  as  $n \rightarrow \infty$ , or as  $t \rightarrow \infty$ ”, where  $n$  and  $t$  denote the number of vertices in the configuration model and in the preferential attachment model, respectively.

**1.1. Configuration model and main result.** The configuration model  $\text{CM}_n$  is a random graph with vertex set  $[n] := \{1, 2, \dots, n\}$  and with prescribed degrees. Let  $\mathbf{d} = (d_1, d_2, \dots, d_n)$  be a given *degree sequence*, i.e., a sequence of  $n$  positive integers with total degree

$$\ell_n = \sum_{i \in [n]} d_i, \quad (1.1)$$

assumed to be even. The configuration model (CM) on  $n$  vertices with degree sequence  $\mathbf{d}$  is constructed as follows: Start with  $n$  vertices and  $d_i$  half-edges adjacent to vertex  $i \in [n]$ . Randomly choose pairs of half-edges and match the chosen pairs together to form edges. Although self-loops may occur, these become rare as  $n \rightarrow \infty$  (see e.g. [Bol01, Theorem 2.16], [Jan09]). We denote the resulting multi-graph on  $[n]$  by  $\text{CM}_n$ , with corresponding edge set  $\mathcal{E}_n$ . We often omit the dependence on the degree sequence  $\mathbf{d}$ , and write  $\text{CM}_n$  for  $\text{CM}_n(\mathbf{d})$ .

**Regularity of vertex degrees.** Let us now describe our regularity assumptions. For each  $n \in \mathbb{N}$  we have a degree sequence  $\mathbf{d}^{(n)} = (d_1^{(n)}, \dots, d_n^{(n)})$ . To lighten notation, we omit the superscript  $(n)$  and write  $\mathbf{d}$  instead of  $\mathbf{d}^{(n)}$  or  $(\mathbf{d}^{(n)})_{n \in \mathbb{N}}$  and  $d_i$  instead of  $d_i^{(n)}$ . Let  $(p_k)_{k \in \mathbb{N}}$  be a probability mass function on  $\mathbb{N}$ . We introduce the *empirical degree distribution* of the graph as

$$p_k^{(n)} = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{\{d_i=k\}}. \quad (1.2)$$

---

<sup>1</sup>DIPARTIMENTO DI MATEMATICA E APPLICAZIONI, UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA, VIA COZZI 55, 20125 MILANO, ITALY

<sup>2</sup>DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, EINDHOVEN UNIVERSITY OF TECHNOLOGY, 5600 MB EINDHOVEN, THE NETHERLANDS

*E-mail addresses:* francesco.caravenna@unimib.it, a.garavaglia@tue.nl, rhofstad@win.tue.nl.

*Key words and phrases.* random graphs, diameter, scale free, ultra small, configuration model, preferential attachment model.

We can define now the *degree regularity conditions*:

**Condition 1.1** (Degree regularity conditions). *Let  $\text{CM}_n$  be a configuration model, then we say that  $\mathbf{d}$  satisfies the degrees regularity conditions (a), (b), with respect to  $(p_k)_{k \in \mathbb{N}}$  if:*

(a) *for every  $k \in \mathbb{N}$ , as  $n \rightarrow \infty$*

$$p_k^{(n)} \longrightarrow p_k. \quad (1.3)$$

(b)  *$\sum_k kp_k < \infty$ , and as  $n \rightarrow \infty$*

$$\sum_{k \in \mathbb{N}} kp_k^{(n)} \longrightarrow \sum_{k \in \mathbb{N}} kp_k. \quad (1.4)$$

As notation, we write that  $\mathbf{d}$  satisfies the d.r.c. (a), (b).

Let  $F_{\mathbf{d},n}$  be the distribution function of  $(p_k^{(n)})_{k \in \mathbb{N}}$ , that is, for  $k \in \mathbb{N}$ ,

$$F_{\mathbf{d},n}(k) = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{\{d_i \leq k\}}. \quad (1.5)$$

We suppose that  $\mathbf{d}$  satisfies the d.r.c. (a) and (b) with respect to some probability mass function  $(p_k)_{k \in \mathbb{N}}$ , corresponding to a distribution function  $F$ .

**Condition 1.2** (Polynomial distribution condition). *We say that  $\mathbf{d}$  satisfies the polynomial distribution condition with exponent  $\tau \in (2, 3)$  if for all  $\delta > 0$  there exist  $\alpha = \alpha(\delta) > \frac{1}{2}$ ,  $c_1(\delta) > 0$  and  $c_2(\delta) > 0$  such that, for every  $n \in \mathbb{N}$ , the lower bound*

$$1 - F_{\mathbf{d},n}(x) \geq c_1 x^{-(\tau-1+\delta)} \quad (1.6)$$

*holds for all  $x \leq n^\alpha$ , and the upper bound*

$$1 - F_{\mathbf{d},n}(x) \leq c_2 x^{-(\tau-1-\delta)} \quad (1.7)$$

*holds for all  $x \geq 1$ .*

There are two examples that explain Condition 1.2. Consider the case of i.i.d. degrees with  $\mathbb{P}(D_i > x) = cx^{-(\tau-1)}$ , then the degree sequence satisfies Condition 1.2 a.s. A second case is when the number of vertices of degree  $k$  is  $n_k = \lceil nF(k) \rceil - \lceil nF(k-1) \rceil$ , and  $1 - F(x) = cx^{-(\tau-1)}$ . Condition 1.2 allows for more flexible degree sequences than just these examples.

If we fix  $\beta < \min\{\alpha, \frac{1}{\tau-1+\delta}\}$ , the lower bound (1.6) ensures that the number of vertices of degree higher than  $x = n^\beta$  is at least  $n^{1-\beta(\tau-1+\delta)}$ , which diverges as a positive power of  $n$ . If we take  $\beta > \frac{1}{2}$ , these vertices with high probability form a complete graph. This will be essential for proving our main results. The precise value of  $\beta$  is irrelevant in the sequel of this paper.

For an asymptotic degree distribution with asymptotic probability mass function  $(p_k)_{k \in \mathbb{N}}$ , we say that

$$d_{\min} = \min \{k \in \mathbb{N}: p_k > 0\} \quad (1.8)$$

is the minimal degree of the probability given by  $(p_k)_{k \in \mathbb{N}}$ . With these technical requests, we can state the main result for the configuration model:

**Theorem 1.3** (Diameter of  $\text{CM}_n$  for  $\tau \in (2, 3)$ ). *Let  $\mathbf{d}$  be a sequence satisfying Condition 1.1 with asymptotic degree distribution  $(p_k)_k$  with  $d_{\min} \geq 3$ . Suppose that  $\mathbf{d}$  satisfies Condition 1.2 with  $\tau \in (2, 3)$  and  $d_i \geq d_{\min}$  for all  $i \in [n]$ . Then*

$$\frac{\text{diam}(\text{CM}_n)}{\log \log n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \frac{2}{\log(d_{\min} - 1)} + \frac{2}{|\log(\tau - 2)|}, \quad (1.9)$$

where  $\xrightarrow[n \rightarrow \infty]{\mathbb{P}}$  denotes convergence in probability as  $n \rightarrow \infty$ .

In fact, the result turns out to be false when  $p_1 + p_2 > 0$ , as shown by Fernholz and Ramachandran [FR07] (see also [HZ07b]), since then there are long strings of vertices with low degrees that are of logarithmic length.

**1.2. Preferential attachment model and main result.** The configuration model presented in the previous section is a *static model*, because the size  $n \in \mathbb{N}$  of the graph was fixed.

The preferential attachment model instead is a *dynamic model*, because, in this model, vertices are added sequentially with a number of edges connected to them. These edges are attached to a receiving vertex with a probability proportional to the degree of the receiving vertex at that time plus a constant, thus favoring vertices with high degrees.

The idea of the preferential attachment model is simple, and we start by defining it informally. We start with a single vertex with a self loop, which is the graph at time 1. At every time  $t \geq 2$ , we add a vertex to the graph. This new vertex has an edge incident to it, and we attach this edge to a random vertex already present in the graph, with probability proportional to the degree of the receiving vertex plus a constant  $\delta$ , which means that vertices with large degrees are favored. Clearly, at each time  $t$  we have a graph of size  $t$  with exactly  $t$  edges.

We can modify this model by changing the number of edges incident to each new vertex we add. If we start at time 1 with a single vertex with  $m \in \mathbb{N}$  self loops, and at every time  $t \geq 2$  we add a single vertex with  $m$  edges, then at time  $t$  we have a graph of size  $t$  but with  $mt$  edges, that we call  $\text{PA}_t(m, \delta)$ . When no confusion can arise, we omit the arguments  $(m, \delta)$  and abbreviate  $\text{PA}_t = \text{PA}_t(m, \delta)$ . We now give the explicit expression for the attachment probabilities.

**Definition 1.4** (Preferential attachment model). *Fix  $m \in \mathbb{N}$ ,  $\delta \in (-m, \infty)$ . Denote by  $\{t \xrightarrow{j} v\}$  the event that the  $j$ -th edge of vertex  $t \in \mathbb{N}$  is attached to vertex  $v \in [t]$  (for  $1 \leq j \leq m$ ). The preferential attachment model with parameters  $(m, \delta)$  is defined by the attachment probabilities*

$$\mathbb{P}\left(t \xrightarrow{j} v \mid \text{PA}_{t,j-1}\right) = \begin{cases} \frac{D_{t,j-1}(v) + 1 + j\delta/m}{c_{t,j}} & \text{for } v = t, \\ \frac{D_{t,j-1}(v) + \delta}{c_{t,j}} & \text{for } v < t, \end{cases} \quad (1.10)$$

where  $\text{PA}_{t,j-1}$  is the graph after the first  $j - 1$  edges of vertex  $t$  have been attached, and correspondingly  $D_{t,j-1}(v)$  is the degree of vertex  $v$ . The normalizing constant  $c_{t,j}$  in (1.10) is

$$c_{t,j} := [m(t-1) + (j-1)](2 + \delta/m) + 1 + \delta/m. \quad (1.11)$$

We refer to Section 4.1 for more details and explanations on the construction of the model (in particular, for the reason behind the factor  $j\delta/m$  in the first line of (1.10)).

Consider, as in (1.2), the empirical degree distribution of the graph, which we denote by  $P_k(t)$ , where in this case the degrees are random variables. It is known from the literature ([BRST01], [Hof17]) that, for every  $k \geq m$ , as  $t \rightarrow \infty$ ,

$$P_k(t) \xrightarrow[t \rightarrow \infty]{\mathbb{P}} p_k, \quad (1.12)$$

where  $p_k \sim ck^{-\tau}$ , and  $\tau = 3 + \delta/m$ . We focus on the case  $\delta \in (-m, 0)$ , so that  $\text{PA}_t$  has a power-law degree sequence with power-law exponent  $\tau \in (2, 3)$ .

For the preferential attachment model, our main result is the following:

**Theorem 1.5** (Diameter of the preferential attachment model). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Then*

$$\frac{\text{diam}(\text{PA}_t)}{\log \log t} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} \frac{2}{\log m} + \frac{4}{|\log(\tau - 2)|}, \quad (1.13)$$

where  $\tau = 3 + \delta/m \in (2, 3)$ .

In the proof of Theorem 1.5 we are also able to identify the typical distances in  $\text{PA}_t$ :

**Theorem 1.6** (Typical distance in the preferential attachment model). *Let  $V_1^t$  and  $V_2^t$  be two independent uniform random vertices in  $[t]$ . Denote the distance between  $V_1^t$  and  $V_2^t$  in  $\text{PA}_t$  by  $H_t$ . Then*

$$\frac{H_t}{\log \log t} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} \frac{4}{|\log(\tau - 2)|}. \quad (1.14)$$

Theorems 1.5–1.6 prove [HHZ07b, Conjecture 1.8].

**1.3. Structure of the paper and heuristics.** The proofs of our main results on the diameter in Theorems 1.3 and 1.5 have a surprisingly similar structure. We present a detailed outline in Section 2 below, where we split the proof into a lower bound (Section 2.1) and an upper bound (Section 2.2) on the diameter. Each of these bounds is then divided into 3 statements, that hold for each model. In Sections 3 and 4 we prove the lower bound for the configuration model and for the preferential attachment model, respectively, while in Sections 5 and 6 we prove the corresponding upper bounds. In [CGH16, Appendix], some proofs of technical results that are minor modifications of proofs in the literature are presented in detail.

Even though the configuration and preferential attachment models are quite different in nature, they are *locally* similar, because for both models the attachment probabilities are roughly proportional to the degrees. The core of our proof is a combination of *conditioning arguments* (which are particularly subtle for the preferential attachment model), that allow to combine local estimates in order to derive bounds on *global* quantities, such as the diameter.

Let us give a heuristic explanation of the proof (see Figure 1.1 for a graphical representation). For a quantitative outline, we refer to Section 2. We write  $\text{PA}_n$  instead of  $\text{PA}_t$  to simplify the exposition, and denote by  $d_{\text{fwd}}$  the minimal *forward degree*, that is  $d_{\text{fwd}} = d_{\min} - 1$  for the configuration model and  $d_{\text{fwd}} = m$  for the preferential attachment model.

- For the *lower bound* on the diameter, we prove that there are so-called *minimally-connected* vertices. These vertices are quite special, in that their neighborhoods up to distance  $k_n^- \approx \log \log n / \log d_{\text{fwd}}$  are *trees with the minimal possible degree*, given by  $d_{\text{fwd}} + 1$ . This explains the first term in the right hand sides of (1.9) and (1.13).

Pairs of minimally-connected vertices are good candidates for achieving the maximal possible distance, i.e., the diameter. In fact, the boundaries of their tree-like neighborhoods turn out to be at distance equal to the *typical distance*  $2\bar{k}_n$  between vertices in the graph, that is  $2\bar{k}_n \approx 2c_{\text{dist}} \log \log n / |\log(\tau - 2)|$ , where  $c_{\text{dist}} = 1$  for the configuration model and  $c_{\text{dist}} = 2$  for the preferential attachment model. This leads to the second term in the right hand sides of (1.9) and (1.13).

In the proof, we split the possible paths between the boundaries of two minimally connected vertices into bad paths, which are too short, and typical paths, which have the right number of edges in them, and then show that the contribution due to bad paths vanishes. The degrees along the path determine whether a path is bad or typical.

The strategy for the lower bound is depicted in the bottom part of Figure 1.1.

- For the *upper bound* on the diameter, we perform a lazy-exploration from every vertex in the graph and realize that the neighborhood up to a distance  $k_n^+$ , which is roughly the same as  $k_n^-$ , contains at least *as many vertices as the tree-like neighborhood of a minimally-connected vertex*. All possible other vertices in this neighborhood are ignored.

We then show that the vertices at the boundary of these lazy neighborhoods are with high probability *quickly* connected to the core, that is by a path of  $h_n = o(\log \log n)$  steps. By *core* we mean the set of all vertices with large degrees, which is known to be highly connected, with a diameter close to  $2\bar{k}_n$ , similar to the typical distances (see [HHZ07b] for the configuration model and [DHH10] for the preferential attachment model).

The proof strategy for the upper bound is depicted in the top part of Figure 1.1.

**1.4. Links to the literature and comments.** This paper studies the diameter in  $\text{CM}_n$  and  $\text{PA}_t$  when the degree power-law exponent  $\tau$  satisfies  $\tau \in (2, 3)$ , which means the degrees have finite mean but infinite variance. Both in (1.9) and (1.13), the explicit constant is the sum of two terms, one depending on  $\tau$ , and the other depending on the minimal forward degree (see (2.2)), which is  $d_{\min} - 1$  for  $\text{CM}_n$  and  $m$  for  $\text{PA}_t$ . We remark that the term depending on  $\tau$  is related to the typical distances, while the other is related to the periphery of the graph.

There are several other works that have already studied typical distances and diameters of such models. Van der Hofstad, Hooghiemstra and Znamenski [HZ07a] analyze typical distances

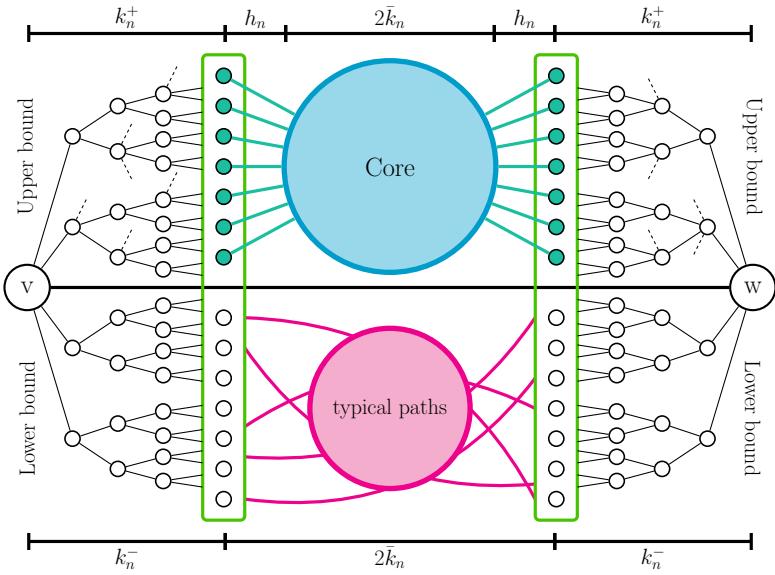


FIGURE 1.1. Structure of the proof in a picture

in  $\text{CM}_n$  for  $\tau \in (2, 3)$ , while Van der Hofstad, Hooghiemstra and Van Mieghem [HHVM05] study  $\tau > 3$ . They prove that for  $\tau \in (2, 3)$  typical distances are of order  $\log \log n$ , while for  $\tau > 3$  is of order  $\log n$ , and it presents the explicit constants of asymptotic growth. Van der Hofstad, Hooghiemstra and Znamensky [HHZ07b] shows for  $\tau > 2$  and when vertices of degree 1 or 2 are present, that with high probability the diameter of  $\text{CM}_n$  is bounded from below by a constant times  $\log n$ , while when  $\tau \in (2, 3)$  and the minimal degree is 3, the diameter is bounded from above by a constant times  $\log \log n$ . In [HK17], Van der Hofstad and Komjáthy investigate typical distances for configuration models and  $\tau \in (2, 3)$  in great generality, extending the results in [HHZ07b] beyond the setting of i.i.d. degrees. Interestingly, they also investigate the effect of truncating the degrees at  $n^{\beta_n}$  for values of  $\beta_n \rightarrow 0$ . It would be of interest to also extend our diameter results to this setting.

We significantly improve upon the result in [HHZ07b] for  $\tau \in (2, 3)$ . We do make use of similar ideas in our proof of the upper bound on the diameter. Indeed, we again define a core consisting of vertices with high degrees, and use the fact that the diameter of this core can be computed exactly (for a definition of the core, see (2.8)). The novelty of our current approach is that we quantify precisely how far the further vertex is from this core in the configuration model. It is a pair of such remote vertices that realizes the graph diameter.

Fernholz and Ramachandran [FR07] prove that the diameter of  $\text{CM}_n$  is equal to an explicit constant times  $\log n$  plus  $o(\log n)$  when  $\tau \in (2, 3)$  but there are vertices of degree 1 or 2 present in the graph, by studying the longest paths in the configuration model that are not part of the 2-core (which is the part of the graph for which all vertices have degree at least 2). Since our minimal degree is at least 3, the 2-core is whp the entire graph, and thus this logarithmic phase vanishes. Dereich, Möch and Mörters [DMM12] prove that typical distances in  $\text{PA}_t$  are asymptotically equal to an explicit constant times  $\log \log t$ , using path counting techniques. We use such path counting techniques as well, now for the lower bound on the diameters. Van der Hofstad [Hof18+] studies the diameter of  $\text{PA}_t$  when  $m = 1$ , and proves that the diameter still has logarithmic growth. Dommers, van der Hofstad and Hooghiemstra [DHH10] prove an upper bound on the diameter of  $\text{PA}_t$ , but the explicit constant is not sharp.

Again, we significantly improve upon that result. Our proof uses ideas from [DHH10], in the sense that we again rely on an appropriately chosen core for the preferential attachment model, but our upper bound now quantifies precisely how the further vertex is from this core, as for the configuration model, but now applied to the much harder preferential attachment model.

$\text{CM}_n$  and  $\text{PA}_t$  are two different models, in the sense that  $\text{CM}_n$  is a static model while  $\text{PA}_t$  is a dynamic model. It is interesting to notice that the main strategy to prove Theorems 1.3 and 1.5 is the same. In fact, all the statements formulated in Section 2 are general and hold for both models. Also the explicit constants appearing in (1.9) and (1.13) are highly similar, which reflects the same structure of the proofs. The differences consist in a factor 2 in the terms containing  $\tau$  and in the presence of  $d_{\min} - 1$  and  $m$  in the remaining term. The factor 2 can be understood by noting that in  $\text{CM}_n$  pairs of vertices with high degree are likely to be at distance 1, while in  $\text{PA}_t$  they are at distance 2. The difference in  $d_{\min} - 1$  and  $m$  is due to the fact that  $d_{\min} - 1$  and  $m$  play the same role in the two models, i.e., they are the minimal forward degree (or “number of children”) of a vertex that is part of a tree contained in the graph. We refer to Section 2 for more details.

While the structures of the proofs for both models are identical, the details of the various steps are significantly different. Pairings in the configuration model are uniform, making explicit computations easy, even when already many edges have been paired. In the preferential attachment model, on the other hand, the edge statuses are highly dependent, so that we have to rely on delicate conditioning arguments. These conditioning arguments are arguably the most significant innovation in this paper. This is formalized in the notion of factorizable events in Definition 4.4.

Typical distances and diameters have been studied for other random graphs models as well, showing  $\log \log$  behavior. Bloznelis [Blo09] investigates the typical distance in power-law intersection random graphs, where such distance, conditioning on being finite, is of order  $\log \log n$ , while results on diameter are missing. Chung and Lu [CL02, CL01] present results respectively for random graphs with given expected degrees and Erdős and Rényi random graphs  $G(n, p)$ , see also van den Esker, the last author and Hooghiemstra [EHH08] for the rank-1 setting. The setting of the configuration model with finite-variance degrees is studied in [FR07]. In [CL02], they prove that for the power-law regime with exponent  $\tau \in (2, 3)$ , the diameter is  $\Theta(\log n)$ , while typical distances are of order  $\log \log n$ . This can be understood from the existence of a positive proportion of vertices with degree 2, creating long, but thin, paths. In [CL01], the authors investigate the different behavior of the diameter according to the parameter  $p$ .

An interesting open problem is the study of fluctuations of the diameters in  $\text{CM}_n$  and  $\text{PA}_t$  around the asymptotic mean, i.e., the study of the difference between the diameter of the graph and the asymptotic behavior (for these two models, the difference between the diameter and the right multiple of  $\log \log n$ ). In [HZ07a], the authors prove that in graphs with i.i.d. power-law degrees with  $\tau \in (2, 3)$ , the difference  $\Delta_n$  between the typical distance and the asymptotic behavior  $2 \log \log n / |\log(\tau - 2)|$  does not converge in distribution, even though it is *tight* (i.e., for every  $\epsilon > 0$  there is  $M < \infty$  such that  $\mathbb{P}(|\Delta_n| \leq M) > 1 - \epsilon$  for all  $n \in \mathbb{N}$ ). These results have been significantly improved in [HK17].

In the literature results on fluctuations for the diameter of random graph models are rare. Bollobás in [Bol81], and, later, Riordan and Wormald in [RW10] give precise estimates on the diameter of the Erdős-Renyi random graph. It would be of interest to investigate whether the diameter has tight fluctuations around  $c \log \log n$  for the appropriate  $c$ .

## 2. GENERAL STRUCTURE OF THE PROOFS

We split the proof of Theorems 1.3 and 1.5 into a lower and an upper bound. Remarkably, the strategy is the same for both models despite the inherent difference in the models. In this section we explain the strategy in detail, formulating general statements that will be proved for each model separately in the next sections.

Throughout this section, we assume that the assumptions of Theorems 1.3 and 1.5 are satisfied and, to keep unified notation, we denote the size of the preferential attachment model by  $n \in \mathbb{N}$ , instead of the more usual  $t \in \mathbb{N}$ .

Throughout the paper, we treat real numbers as integers when we consider graph distances. By this, we mean that we round real numbers to the closest integer. To keep the notation light and make the paper easier to read, we omit the rounding operation.

**2.1. Lower bound.** We start with the structure of the proof of the lower bound in (1.9) and (1.13). The key notion is that of a *minimally- $k$ -connected vertex*, defined as a vertex whose  $k$ -neighborhood (i.e., the neighborhood up to distance  $k$ ) is essentially a *regular tree with the smallest possible degree*, equal to  $d_{\min}$  for the configuration model and to  $m+1$  for the preferential attachment model. Due to technical reasons, the precise definition of minimally- $k$ -connected vertex is slightly different for the two models (see Definitions 3.2 and 4.2).

Henceforth we fix  $\varepsilon > 0$  and define, for  $n \in \mathbb{N}$ ,

$$k_n^- = (1 - \varepsilon) \frac{\log \log n}{\log(d_{\text{fwd}})}, \quad (2.1)$$

where  $d_{\text{fwd}}$  denotes the *forward degree*, or “number of children”:

$$d_{\text{fwd}} = \begin{cases} d_{\min} - 1 & \text{for } \text{CM}_n; \\ m & \text{for } \text{PA}_n. \end{cases} \quad (2.2)$$

Our first goal is to prove that the number of minimally- $k_n^-$ -connected vertices is large enough, as formulated in the following statement:

**Statement 2.1** (Moments of  $M_{k_n^-}$ ). *Denote by  $M_{k_n^-}$  the number of minimally- $k_n^-$ -connected vertices in the graph (either  $\text{CM}_n$  or  $\text{PA}_n$ ). Then, as  $n \rightarrow \infty$ ,*

$$\mathbb{E}[M_{k_n^-}] \rightarrow \infty, \quad \text{Var}(M_{k_n^-}) = o\left(\mathbb{E}[M_{k_n^-}]^2\right), \quad (2.3)$$

where  $\text{Var}(X) := \mathbb{E}[X^2] - \mathbb{E}[X]^2$  denotes the variance of the random variable  $X$ .

The proof for the preferential attachment model makes use of conditioning arguments. Indeed, we describe as much information as necessary to be able to bound probabilities that vertices are minimally- $k$  connected. Particularly in the variance estimate, these arguments are quite delicate, and crucial for our purposes.

The bounds in (2.3) show that  $M_{k_n^-} \xrightarrow{\mathbb{P}} \infty$  as  $n \rightarrow \infty$ . This will imply that *there is a pair of minimally- $k_n^-$ -connected vertices with disjoint  $k_n^-$ -neighborhoods*,<sup>1</sup> hence the diameter of the graph is at least  $2k_n^-$ , which explains the first term in (1.9) and (1.13). Our next aim is to prove that these minimally connected trees are typically at distance  $2c_{\text{dist}} \log \log n / |\log(\tau - 2)|$ , where  $c_{\text{dist}} = 1$  for the configuration model and  $c_{\text{dist}} = 2$  for the preferential attachment model.

For this, let us now define

$$\bar{k}_n = (1 - \varepsilon) \frac{c_{\text{dist}} \log \log n}{|\log(\tau - 2)|}, \quad (2.4)$$

where

$$c_{\text{dist}} = \begin{cases} 1 & \text{for } \text{CM}_n; \\ 2 & \text{for } \text{PA}_n. \end{cases} \quad (2.5)$$

The difference in the definition of  $c_{\text{dist}}$  is due to fact that in  $\text{CM}_n$  vertices with high degree are likely at distance 1, while in  $\text{PA}_n$  are at distance 2. We explain the origin of this effect in more detail in the proofs.

It turns out that the distance between the  $k_n^-$ -neighborhoods of two minimally- $k_n^-$ -connected vertices is at least  $2\bar{k}_n$ . More precisely, we have the following statement:

**Statement 2.2** (Distance between neighborhoods). *Let  $W_1^n$  and  $W_2^n$  be two random vertices chosen independently and uniformly among the minimally- $k_n^-$ -connected ones. Denoting by  $\tilde{H}_n$  the distance between the  $k_n^-$ -neighborhoods of  $W_1^n$  and  $W_2^n$ , we have  $\tilde{H}_n \geq 2\bar{k}_n$  with high probability.*

<sup>1</sup>A justification for this fact is provided by the following Statement 2.2 (the randomly chosen vertices  $W_1^n$  and  $W_2^n$  have disjoint  $k_n^-$ -neighborhoods, because  $\tilde{H}_n > 0$  with high probability). For a more direct justification, see Remark 3.6 for the configuration model and Remark 4.7 for the preferential attachment model.

It follows immediately from Statement 2.2 that the distance between the vertices  $W_1^n$  and  $W_2^n$  is at least  $2k_n^- + 2\bar{k}_n$ , with high probability. This proves the lower bound in (1.9) and (1.13).

It is known from the literature that  $2\bar{k}_n$ , see (2.4), represents the *typical distance* between two vertices chosen independently and uniformly in the graph. In order to prove Statement 2.2, we collapse the  $k_n^-$ -neighborhoods of  $W_1^n$  and  $W_2^n$  into single vertices and show that their distance is roughly equal to the typical distance  $2\bar{k}_n$ . This is a delicate point, because the collapsed vertices have a relatively large degree and thus *could* be closer than the typical distance. The crucial point why they are not closer is that the degree of the boundary only grows polylogarithmically. The required justification is provided by the next statement:

**Statement 2.3** (Bound on distances). *Let us introduce the set*

$$V_n := \begin{cases} \{v \in [n]: d_v \leq \log n\} & \text{for CM}_n; \\ \{v \in [n]: v \geq \frac{n}{(\log n)^2}\} & \text{for PA}_n. \end{cases} \quad (2.6)$$

*Then, denoting the distance in the graph of size  $n$  by  $\text{dist}_n$ ,*

$$\max_{a,b \in V_n} \mathbb{P}(\text{dist}_n(a, b) \leq 2\bar{k}_n) = O\left(\frac{1}{(\log n)^2}\right). \quad (2.7)$$

The proof of Statement 2.3 is based on *path counting techniques*. These are different for the two models, but the idea is the same: We split the possible paths between the vertices  $a$  and  $b$  into two sets, called *good paths* and *bad paths*. Here *good* means that the degrees of vertices along the path increase, but *not too much*. We then separately and directly estimate the contribution of each set. The details are described in the proof.

**2.2. Upper bound.** We now describe the structure of the proof for the upper bound, which is based on two key concepts: the *core of the graph* and the  *$k$ -exploration graph* of a vertex.

We start by introducing some notation. First of all, fix a constant  $\sigma \in (1/(3 - \tau), \infty)$ . We define  $\text{Core}_n$  as the set of vertices in the graph of size  $n$  with degree larger than  $(\log n)^\sigma$ . More precisely, denoting by  $D_t(v) = D_{t,m}(v)$  the degree of vertex  $v$  in the preferential attachment model after time  $t$ , i.e. in the graph  $\text{PA}_t$  (see the discussion after (1.10)), we let

$$\text{Core}_n := \begin{cases} \{v \in [n]: d_v \geq (\log n)^\sigma\} & \text{for CM}_n; \\ \{v \in [n]: D_{n/2}(v) \geq (\log n)^\sigma\} & \text{for PA}_n. \end{cases} \quad (2.8)$$

The fact that we evaluate the degrees at time  $n/2$  for the PAM is quite crucial in the proof of Statement 2.4 below. In Section 6, we also give bounds on  $D_v(n)$  for  $v \in \text{Core}_n$ , as well as for  $v \notin \text{Core}_n$ , that show that the degrees cannot grow too much between time  $n/2$  and  $n$ . The first statement, that we formulate for completeness, upper bounds the diameter of  $\text{Core}_n$  and is already known from the literature for both models:

**Statement 2.4.** Define  $c_{\text{dist}}$  as in (2.5). Then, for every  $\varepsilon > 0$ , with high probability

$$\frac{\text{diam}(\text{Core}_n)}{\log \log n} \leq (1 + \varepsilon) \frac{2c_{\text{dist}}}{|\log(\tau - 2)|}. \quad (2.9)$$

Statement 2.4 for  $\text{CM}_n$  is [HHZ07b, Proposition 3.1], for  $\text{PA}_n$  it is [DHH10, Theorem 3.1].

Next we bound the distance between a vertex and  $\text{Core}_n$ . We define the  *$k$ -exploration graph* of a vertex  $v$  as a suitable subgraph of its  $k$ -neighborhood, built as follows: We consider the usual exploration process starting at  $v$ , but instead of exploring all the edges incident to a vertex, we only explore a *fixed* number of them, namely  $d_{\text{fwd}}$  defined in (2.2). (The choice of which edges to explore is a natural one, and it will be explained in more detail in the proofs.)

We stress that it is possible to explore vertices that have already been explored, leading to what we call a *collision*. If there are no collisions, then the  $k$ -exploration graph is a tree. In presence of collisions, the  $k$ -exploration graph is not a tree, and it is clear that every collision reduces the number of vertices in the  $k$ -exploration graph.

Henceforth we fix  $\varepsilon > 0$  and, in analogy with (2.1), we define, for  $n \in \mathbb{N}$ ,

$$k_n^+ = (1 + \varepsilon) \frac{\log \log n}{\log(d_{\text{fwd}})}. \quad (2.10)$$

Our second statement for the upper bound shows that the  $k_n^+$ -exploration graph of *every* vertex in the graph either intersects  $\text{Core}_n$ , or it has a bounded number of collisions:

**Statement 2.5** (Bound on collisions). *There is a constant  $c < \infty$  such that, with high probability, the  $k_n^+$ -exploration graph of every vertex in the graph has at most  $c$  collisions before hitting  $\text{Core}_n$ . As a consequence, for some constant  $s > 0$ , the  $k_n^+$ -exploration graph of every vertex in the graph either intersects  $\text{Core}_n$ , or its boundary has cardinality at least*

$$s(d_{\text{fwd}})^{k_n^+} = (\log n)^{1+\varepsilon+o(1)}. \quad (2.11)$$

With a bounded number of collisions, the  $k_n^+$ -exploration graph is not far from being a tree, which explains the lower bound (2.11) on the cardinality of its boundary. Having enough vertices on its boundary, the  $k_n^+$ -exploration is likely to be connected to  $\text{Core}_n$  *fast*, which for our purpose means in  $o(\log \log n)$  steps. This is the content of our last statement:

**Statement 2.6.** *There are constants  $B, C < \infty$  such that, with high probability, the  $k_n^+$ -exploration graph of every vertex in the graph is at distance at most  $h_n = \lceil B \log \log \log n + C \rceil$  from  $\text{Core}_n$ .*

The proof for this is novel. For example, for the configuration model, we grow the  $k_n^+ + h_n$  neighborhood of a vertex, and then show that there are so many half-edges at its boundary that it is very likely to connect immediately to the core. The proof for the preferential attachment model is slightly different, but the conclusion is the same. This shows that the vertex is indeed at most at distance  $k_n^+ + h_n$  away from the core.

In conclusion, with high probability, the diameter of the graph is at most

$$(k_n^+ + h_n) + \text{diam}(\text{Core}_n) + (k_n^+ + h_n),$$

which gives us the expressions in (1.9) and (1.13) and completes the proof of the upper bound.

### 3. LOWER BOUND FOR CONFIGURATION MODEL

In this section we prove Statements 2.1, 2.2 and 2.3 for the configuration model. By the discussion in Section 2.1, this completes the proof of the lower bound in Theorem 1.3.

In our proof, it will be convenient to choose a particular order to pair the half-edges. This is made precise in the following remark:

**Remark 3.1** (Exchangeability in half-edge pairing). Given a sequence  $\mathbf{d} = (d_1, \dots, d_n)$  such that  $\ell_n = d_1 + \dots + d_n$  is even, the configuration model  $\text{CM}_n$  can be built iteratively as follows:

- ▷ start with  $d_i$  half-edges attached to each vertex  $i \in [n] = \{1, 2, \dots, n\}$ ;
- ▷ choose an *arbitrary* half-edge and pair it to a uniformly chosen half-edge;
- ▷ choose an *arbitrary* half-edge, among the  $\ell_n - 2$  that are still unpaired, and pair it to a uniformly chosen half-edge; and so on.

The *order* in which the arbitrary half-edges are chosen does not matter in the above, by exchangeability (see also [Hof17, Chapter 7]).

**3.1. Proof of Statement 2.1.** With a slight abuse of notation (see (1.8)), in this section we set

$$d_{\min} = \min\{d_1, \dots, d_n\}.$$

Given a vertex  $v \in [n]$  and  $k \in \mathbb{N}$ , we denote the set of vertices at distance at most  $k$  from  $v$  (in the graph  $\text{CM}_n$ ) by  $U_{\leq k}(v)$  and we call it the *k-neighborhood of v*.

**Definition 3.2** (Minimally- $k$ -connected vertex). For  $k \in \mathbb{N}_0$ , a vertex  $v \in [n]$  is called *minimally- $k$ -connected* when all the vertices in  $U_{\leq k}(v)$  have minimal degree, i.e.,

$$d_i = d_{\min} \quad \text{for all } i \in U_{\leq k}(v),$$

and furthermore there are no self-loops, multiple edges or cycles in  $U_{\leq k}(v)$ . Equivalently,  $v$  is *minimally- $k$ -connected* when the graph  $U_{\leq k}(v)$  is a regular tree with degree  $d_{\min}$ .

We denote the (random) set of minimally- $k$ -connected vertices by  $\mathcal{M}_k \subseteq [n]$ , and its cardinality by  $M_k = |\mathcal{M}_k|$ , i.e.  $M_k$  denotes the number of minimally- $k$ -connected vertices.

**Remark 3.3** (The volume of the  $k$ -neighborhood of  $k$ -minimally connected vertices). For a minimally- $k$ -connected vertex  $v$ , since  $U_{\leq k}(v)$  is a tree with degree  $d_{\min}$ , the number of edges inside  $U_{\leq k}(v)$  equals (assuming  $d_{\min} \geq 2$ )

$$i_k = \sum_{l=1}^k d_{\min} (d_{\min} - 1)^{l-1} = \begin{cases} d_{\min} k & \text{if } d_{\min} = 2; \\ d_{\min} \frac{(d_{\min} - 1)^k - 1}{d_{\min} - 2} & \text{if } d_{\min} \geq 3. \end{cases} \quad (3.1)$$

Moreover, the number of vertices inside  $U_{\leq k}(v)$  equals  $i_k + 1$ . By (3.1), it is clear why  $d_{\min} > 2$ , or  $d_{\min} \geq 3$ , is crucial. Indeed, this implies that the volume of neighborhoods of minimally- $k$ -connected vertices grows exponentially in  $k$ .

**Remark 3.4** (Collapsing minimally- $k$  connected trees). By Remarks 3.1 and 3.3, conditionally on the event  $\{v \in \mathcal{M}_k\}$  that a given vertex  $v$  is minimally- $k$ -connected, the random graph obtained from  $\text{CM}_n$  by collapsing  $U_{\leq k}(v)$  to a single vertex, called  $a$ , is still a configuration model with  $n - i_k$  vertices and with  $\ell_n$  replaced by  $\ell_n - 2i_k$ , where the new vertex  $a$  has degree  $d_{\min}(d_{\min} - 1)^k$ .

Analogously, conditionally on the event  $\{v \in \mathcal{M}_k, w \in \mathcal{M}_m, U_{\leq k}(v) \cap U_{\leq m}(w) = \emptyset\}$  that two given vertices  $v$  and  $w$  are minimally- $k$  and minimally- $m$ -connected with disjoint neighborhoods, collapsing  $U_{\leq k}(v)$  and  $U_{\leq m}(w)$  to single vertices  $a$  and  $b$  yields again a configuration model with  $n - i_k - i_m$  vertices, where  $\ell_n$  is replaced by  $\ell_n - 2i_k - 2i_m$  and where the new vertices  $a$  and  $b$  have degrees equal to  $d_{\min}(d_{\min} - 1)^k$  and  $d_{\min}(d_{\min} - 1)^m$ , respectively.

We denote the number of vertices of degree  $k$  in the graph by  $n_k$ , i.e.,

$$n_k = \sum_{i \in [n]} \mathbf{1}_{\{d_i=k\}}. \quad (3.2)$$

We now study the first two moments of  $M_k$ , where we recall that the total degree  $\ell_n$  is defined by (1.1):

**Proposition 3.5** (Moments of  $M_k$ ). Let  $\text{CM}_n$  be a configuration model such that  $d_{\min} \geq 2$ . Then, for all  $k \in \mathbb{N}$ ,

$$\mathbb{E}[M_k] = n_{d_{\min}} \prod_{i=1}^{i_k} \frac{d_{\min}(n_{d_{\min}} - i)}{\ell_n - 2i + 1}, \quad (3.3)$$

where  $i_k$  is defined in (3.1). When, furthermore,  $\ell_n > 4i_k$ ,

$$\mathbb{E}[M_k^2] \leq \mathbb{E}[M_k]^2 + \mathbb{E}[M_k] \left( (i_k + 1) + i_{2k} d_{\min} \frac{n_{d_{\min}}}{\ell_n - 4i_k} \right). \quad (3.4)$$

Before proving Proposition 3.5, let us complete the proof of Statement 2.1 subject to it. We are working under the assumptions of Theorem 1.3, hence  $d_{\min} \geq 3$  and the degree sequence  $\mathbf{d}$  satisfies the degree regularity condition Condition 1.1, as well as the polynomial distribution condition Condition 1.2 with exponent  $\tau \in (2, 3)$ . Recalling (1.1)-(1.2), we can write  $n_{d_{\min}} = n p_{d_{\min}}^{(n)}$  and  $\ell_n = n \sum_{k \in \mathbb{N}} kp_k^{(n)}$ , so that, as  $n \rightarrow \infty$ ,

$$n_{d_{\min}} = n p_{d_{\min}} (1 + o(1)), \quad \ell_n = n \mu (1 + o(1)), \quad \text{with} \quad p_{d_{\min}} > 0, \quad \mu := \sum_{k \in \mathbb{N}} kp_k < \infty. \quad (3.5)$$

Recalling the definition (2.1) of  $k_n^-$  and (3.1), for  $k = k_n^-$ ,

$$i_{k_n^-} = d_{\min} \frac{(d_{\min} - 1)^{k_n^-} - 1}{d_{\min} - 2} = \frac{d_{\min}}{d_{\min} - 2} (\log n)^{1-\varepsilon} (1+o(1)), \quad \text{hence} \quad i_{2k_n^-} = O((\log n)^{2(1-\varepsilon)}). \quad (3.6)$$

Bounding  $\mathbb{E}[M_k] \leq n$ , it follows by (3.4) that

$$\text{Var}[M_{k_n^-}] \leq \mathbb{E}[M_{k_n^-}] \left( O(i_{k_n^-}) + O(i_{2k_n^-}) \right) \leq n O((\log n)^{2(1-\varepsilon)}) = n^{1+o(1)}. \quad (3.7)$$

On the other hand, applying (3.3), for some  $c \in (0, 1)$  one has

$$\mathbb{E}[M_{k_n^-}] \geq n p_{d_{\min}} \left( \frac{d_{\min} p_{d_{\min}}}{\mu} + o(1) \right)^{i_{k_n^-}} \geq n p_{d_{\min}} c^{(\log n)^{1-\varepsilon}} = n^{1-o(1)}. \quad (3.8)$$

Relations (3.7) and (3.8) show that (2.3) holds, completing the proof of Statement 2.1.  $\square$

**Remark 3.6** (Disjoint neighborhoods). Let us show that, with high probability, there are vertices  $v, w \in \mathcal{M}_{k_n^-}$  with  $U_{\leq k_n^-}(v) \cap U_{\leq k_n^-}(w) = \emptyset$ . We proceed by contradiction: fix  $v \in \mathcal{M}_{k_n^-}$  and assume that, for every vertex  $w \in \mathcal{M}_{k_n^-}$ , one has  $U_{\leq k_n^-}(v) \cap U_{\leq k_n^-}(w) \neq \emptyset$ . Then, for any  $w \in \mathcal{M}_{k_n^-}$  there must exist a *self-avoiding path from v to w of length  $\leq 2k_n^-$  which only visits vertices with degree  $d_{\min}$*  (recall that  $U_{\leq k_n^-}(v)$  and  $U_{\leq k_n^-}(w)$  are regular trees). However, for fixed  $v$ , the number of such paths is  $O((d_{\min} - 1)^{2k_n^-}) = O((\log n)^{2(1-\varepsilon)})$ , see (2.1), while by Statement 2.1 the number of vertices  $w \in \mathcal{M}_{k_n^-}$  is much larger, since  $M_{k_n^-} \sim \mathbb{E}[M_{k_n^-}] = n^{1-o(1)}$ , see (3.8).

*Proof of Proposition 3.5.* To prove (3.3) we write

$$M_k = \sum_{v \in [n]: d_v = d_{\min}} \mathbb{1}_{\{v \in \mathcal{M}_k\}}, \quad (3.9)$$

and since every vertex in the sum has the same probability of being minimally- $k$ -connected,

$$\mathbb{E}[M_k] = n_{d_{\min}} \mathbb{P}(v \in \mathcal{M}_k). \quad (3.10)$$

A vertex  $v$  with  $d_v = d_{\min}$  is in  $\mathcal{M}_k$  when all the half-edges in  $U_{\leq k}(v)$  are paired to half-edges incident to distinct vertices having minimal degree, without generating cycles. By Remark 3.1, we can start pairing a half-edge incident to  $v$  to a half-edge incident to another vertex of degree  $d_{\min}$ . Since there are  $n_{d_{\min}} - 1$  such vertices, this event has probability

$$\frac{d_{\min}(n_{d_{\min}} - 1)}{\ell_n - 1}$$

We iterate this procedure, and suppose that we have already successfully paired  $(i - 1)$  couples of half-edges; then the next half-edge can be paired to a distinct vertex of degree  $d_{\min}$  with probability

$$\frac{d_{\min}(n_{d_{\min}} - i)}{\ell_n - 2(i - 1) - 1} = \frac{d_{\min}(n_{d_{\min}} - i)}{\ell_n - 2i + 1}. \quad (3.11)$$

Indeed, every time that we use a half-edge of a vertex of degree  $d_{\min}$ , we cannot use its remaining half-edges, and every step we make reduces the total number of possible half-edges by two. By (3.1), exactly  $i_k$  couples of half-edges need to be paired for  $v$  to be minimally- $k$ -connected, so that

$$\mathbb{E}[M_k] = n_{d_{\min}} \mathbb{P}(v \in \mathcal{M}_k) = n_{d_{\min}} \prod_{i=1}^{i_k} \frac{d_{\min}(n_{d_{\min}} - i)}{\ell_n - 2i + 1}. \quad (3.12)$$

which proves (3.3). If  $n_{d_{\min}} \leq i_k$  the right hand side vanishes, in agreement with the fact that there cannot be any minimally- $k$ -connected vertex in this case (recall (3.1)).

To prove (3.4), we notice that

$$\mathbb{E}[M_k^2] = \sum_{v, w \in [n]: d_v = d_w = d_{\min}} \mathbb{P}(v, w \in \mathcal{M}_k). \quad (3.13)$$

We distinguish different cases: the  $k$ -neighborhoods of  $v$  and  $w$  might be disjoint or they may overlap, in which case  $w$  can be included in  $U_{\leq k}(v)$  or not. Introducing the events

$$A_{v,w} = \{U_{\leq k}(v) \cap U_{\leq k}(w) \neq \emptyset\}, \quad B_{v,w} = \{w \in U_{\leq k}(v)\}, \quad (3.14)$$

we can write the right hand side of (3.13) as

$$\sum_{\substack{v,w \in [n] \\ d_v=d_w=d_{\min}}} [\mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}^c) + \mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}, B_{v,w}) + \mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}, B_{v,w}^c)]. \quad (3.15)$$

Let us look at the first term in (3.15). By Remarks 3.3 and 3.4, conditionally on  $\{v \in \mathcal{M}_k\}$ , the probability of  $\{w \in \mathcal{M}_k, A_{v,w}^c\}$  equals the probability that  $w$  is minimally- $k$ -connected in a new configuration model, with  $\ell_n$  replaced by  $\ell_n - 2i_k$  and with the number of vertices with minimal degree reduced from  $n_{d_{\min}}$  to  $n_{d_{\min}} - (i_k + 1)$ . Then, by the previous analysis (see (3.12)),

$$\mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}^c) = \prod_{i=1}^{i_k} \frac{d_{\min}(n_{d_{\min}} - i - i_k - 1)}{\ell_n - 2i - 2i_k + 1} \mathbb{P}(v \in \mathcal{M}_k). \quad (3.16)$$

By direct computation, the ratio in the right hand side of (3.16) is always maximized for  $i_k = 0$  (provided  $\ell_n \geq 2n_{d_{\min}} - 3$ , which is satisfied since  $\ell_n \geq d_{\min} n_{d_{\min}} \geq 3n_{d_{\min}}$  by assumption). Therefore, setting  $i_k = 0$  in the ratio and recalling (3.12), we get the upper bound

$$\mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}^c) \leq \left[ \prod_{i=1}^{i_k} \frac{d_{\min}(n_{d_{\min}} - i)}{\ell_n - 2i + 1} \right] \mathbb{P}(v \in \mathcal{M}_k) = \mathbb{P}(v \in \mathcal{M}_k)^2. \quad (3.17)$$

Since there are at most  $n_{d_{\min}}^2$  pairs of vertices of degree  $d_{\min}$ , it follows from (3.17) that

$$\sum_{\substack{v,w \in [n] \\ d_v=d_w=d_{\min}}} \mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}^c) \leq n_{d_{\min}}^2 \mathbb{P}(v \in \mathcal{M}_k)^2 = \mathbb{E}[M_k]^2, \quad (3.18)$$

which explains the first term in (3.4).

For the second term in (3.15),  $v$  and  $w$  are minimally- $k$ -connected with overlapping neighborhoods, and  $w \in U_{\leq k}(v)$ . Since  $\{v, w \in \mathcal{M}_k\} \cap A_{v,w} \cap B_{v,w} \subseteq \{v \in \mathcal{M}_k\} \cap B_{v,w}$ , we can bound

$$\sum_{\substack{v,w \in [n] \\ d_v=d_w=d_{\min}}} \mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}, B_{v,w}) \leq \mathbb{E} \left[ \sum_{v \in [n]: d_v=d_{\min}} \mathbb{1}_{\{v \in \mathcal{M}_k\}} \sum_{w \in [n]: d_w=d_{\min}} \mathbb{1}_{B_{v,w}} \right], \quad (3.19)$$

and note that  $\sum_{w \in [n]} \mathbb{1}_{B_{v,w}} = |U_{\leq k}(v)| = i_k + 1$ , by Remark 3.3. Therefore

$$\sum_{\substack{v,w \in [n] \\ d_v=d_w=d_{\min}}} \mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}, B_{v,w}) \leq \mathbb{E}[M_k] (i_k + 1), \quad (3.20)$$

which explains the second term in (3.4).

For the third term in (3.15),  $v$  and  $w$  are minimally- $k$ -connected vertices with overlapping neighborhoods, but  $w \notin U_{\leq k}(v)$ . This means that  $\text{dist}(v, w) = l + 1$  for some  $l \in \{k, \dots, 2k - 1\}$ , so that  $U_{\leq k}(v) \cap U_{\leq l-k}(w) = \emptyset$  and, moreover, a half-edge on the boundary of  $U_{\leq(l-k)}(w)$  is paired to a half-edge on the boundary of  $U_{\leq k}(v)$ , an event that we call  $F_{v,w;l,k}$ . Therefore

$$\{w \in \mathcal{M}_k\} \cap A_{v,w} \cap B_{v,w}^c \subseteq \bigcup_{l=k}^{2k-1} \{w \in \mathcal{M}_{l-k}\} \cap \{U_{\leq k}(v) \cap U_{\leq l-k}(w) = \emptyset\} \cap F_{v,w;l,k}. \quad (3.21)$$

and we stress that in the right hand side  $w$  is only minimally- $(l - k)$ -connected (in case  $l = k$  this just means that  $d_w = d_{\min}$ ). Then

$$\mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}, B_{v,w}^c) \leq \sum_{l=k}^{2k-1} \mathbb{E} \left[ \mathbb{1}_{\{v \in \mathcal{M}_k, w \in \mathcal{M}_{l-k}, U_{\leq k}(v) \cap U_{\leq l-k}(w) = \emptyset\}} \mathbb{1}_{F_{v,w;l,k}} \right]. \quad (3.22)$$

By Remark 3.4, conditionally on  $\{v \in \mathcal{M}_k, w \in \mathcal{M}_{l-k}, U_{\leq k}(v) \cap U_{\leq l-k}(w) = \emptyset\}$ , we can collapse  $U_{\leq k}(v)$  and  $U_{\leq l-k}(w)$  to single vertices  $a$  and  $b$  with degrees respectively  $d_{\min}(d_{\min} - 1)^k$  and  $d_{\min}(d_{\min} - 1)^{l-k}$ , getting a new configuration model with  $\ell_n$  replaced by  $\ell_n - 2i_k - 2i_{l-k}$ . Bounding the probability that a half-edge of  $a$  is paired to a half-edge of  $b$ , we get

$$\begin{aligned} & \mathbb{P}(F_{v,w;l,k} \mid v \in \mathcal{M}_k, w \in \mathcal{M}_{l-k}, U_{\leq k}(v) \cap U_{\leq l-k}(w) = \emptyset) \\ & \leq \frac{d_{\min}(d_{\min} - 1)^k d_{\min}(d_{\min} - 1)^{l-k}}{\ell_n - 2i_k - 2i_{l-k} - 1} \leq \frac{d_{\min}^2(d_{\min} - 1)^l}{\ell_n - 4i_k}, \end{aligned} \quad (3.23)$$

because  $l \leq 2k - 1$  and, consequently,  $i_{l-k} \leq i_{k-1} \leq i_k - 1$ . Plugging (3.23) into (3.22), and then forgetting the event  $\{w \in \mathcal{M}_{l-k}, U_{\leq k}(v) \cap U_{\leq l-k}(w) = \emptyset\}$ , leads to

$$\begin{aligned} \sum_{\substack{v,w \in [n] \\ d_v = d_w = d_{\min}}} \mathbb{P}(v, w \in \mathcal{M}_k, A_{v,w}, B_{v,w}^c) & \leq \left( \sum_{l=k}^{2k-1} \frac{d_{\min}^2(d_{\min} - 1)^l}{\ell_n - 4i_k} \right) \sum_{\substack{v,w \in [n] \\ d_v = d_w = d_{\min}}} \mathbb{P}(v \in \mathcal{M}_k) \\ & \leq \frac{d_{\min}(d_{\min} - 1)}{\ell_n - 4i_k} i_{2k-1} n_{d_{\min}} \mathbb{E}[M_k], \end{aligned} \quad (3.24)$$

where we have used the definition (3.1) of  $i_{2k-1}$ . Since  $(d_{\min} - 1)i_{2k-1} \leq i_{2k}$ , again by (3.1), we have obtained the third term in (3.4).  $\square$

**3.2. Proof of Statement 2.2.** We recall that  $W_1^n$  and  $W_2^n$  are two independent random vertices chosen uniformly in  $\mathcal{M}_{k_n^-}$  (the set of minimally- $k_n^-$ -connected vertices), assuming that  $\mathcal{M}_{k_n^-} \neq \emptyset$  (which, as we have shown, occurs with high probability). Our goal is to show that

$$\lim_{n \rightarrow \infty} \mathbb{P}(E_n) = 0, \quad (3.25)$$

where we set

$$E_n := \left\{ \text{dist}\left(U_{\leq k_n^-}(W_1^n), U_{\leq k_n^-}(W_2^n)\right) \leq 2\bar{k}_n \right\} = \left\{ \text{dist}(W_1^n, W_2^n) \leq 2k_n^- + 2\bar{k}_n \right\}. \quad (3.26)$$

We know from Statement 2.1 that as  $n \rightarrow \infty$

$$\mathbb{P}\left(M_{k_n^-} \leq \frac{1}{2}\mathbb{E}[M_{k_n^-}]\right) \leq \mathbb{P}\left(|M_{k_n^-} - \mathbb{E}[M_{k_n^-}]| > \frac{1}{2}\mathbb{E}[M_{k_n^-}]\right) \leq \frac{\text{Var}[M_{k_n^-}]}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} = o(1). \quad (3.27)$$

Therefore,

$$\begin{aligned} \mathbb{P}(E_n) &= \mathbb{P}\left(E_n \cap \{M_{k_n^-} > \frac{1}{2}\mathbb{E}[M_{k_n^-}]\}\right) + o(1) \\ &= \mathbb{E}\left[\sum_{v_1, v_2 \in [n]} \mathbb{1}_{\{W_1^n = v_1, W_2^n = v_2\}} \mathbb{1}_{\{\text{dist}(v_1, v_2) \leq 2k_n^- + 2\bar{k}_n\}} \mathbb{1}_{\{M_{k_n^-} > \frac{1}{2}\mathbb{E}[M_{k_n^-}]\}}\right] + o(1) \\ &\leq \mathbb{E}\left[\sum_{v_1, v_2 \in [n]} \frac{\mathbb{1}_{\{v_1 \in \mathcal{M}_{k_n^-}, v_2 \in \mathcal{M}_{k_n^-}\}}}{M_{k_n^-}^2} \mathbb{1}_{\{\text{dist}(v_1, v_2) \leq 2k_n^- + 2\bar{k}_n\}} \mathbb{1}_{\{M_{k_n^-} > \frac{1}{2}\mathbb{E}[M_{k_n^-}]\}}\right] + o(1) \\ &\leq \sum_{v_1, v_2 \in [n]} \frac{\mathbb{P}\left(v_1, v_2 \in \mathcal{M}_{k_n^-}, \text{dist}(v_1, v_2) \leq 2k_n^- + 2\bar{k}_n\right)}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} + o(1). \end{aligned} \quad (3.28)$$

In analogy with (3.14), we introduce the event

$$A_{v_1, v_2} := \left\{ U_{\leq k_n^-}(v_1) \cap U_{\leq k_n^-}(v_2) \neq \emptyset \right\},$$

and show that it gives a negligible contribution. Recalling the proof of Proposition 3.5, in particular (3.20) and (3.24), the sum restricted to  $A_{v_1, v_2}$  leads precisely to the second term in the

right hand side of (3.4):

$$\begin{aligned} \sum_{v_1, v_2 \in [n]} \frac{\mathbb{P}(v_1, v_2 \in \mathcal{M}_{k_n^-}, A_{v_1, v_2})}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} &\leq \frac{\mathbb{E}[M_{k_n^-}] \left( (i_{k_n^-} + 1) + i_{2k_n^-} \frac{d_{\min} n d_{\min}}{\ell_n - 4i_{k_n^-}} \right)}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} \\ &= \frac{O(i_{k_n^-}) + O(i_{2k_n^-})}{\mathbb{E}[M_{k_n^-}]} = \frac{O((\log n)^2)}{n^{1-o(1)}} = o(1), \end{aligned} \quad (3.29)$$

where we have used (3.6) and (3.8) (see also (3.5)).

We can thus focus on the event  $A_{v_1, v_2}^c = \{U_{\leq k_n^-}(v_1) \cap U_{\leq k_n^-}(v_2) = \emptyset\}$ . By Remark 3.4,

$$\mathbb{P}(\text{dist}(v_1, v_2) \leq 2k_n^- + 2\bar{k}_n \mid v_1, v_2 \in \mathcal{M}_{k_n^-}, A_{v_1, v_2}^c) = \hat{\mathbb{P}}(\text{dist}(a, b) \leq 2\bar{k}_n), \quad (3.30)$$

where  $\hat{\mathbb{P}}$  is the law of the new configuration model which results from collapsing the neighborhoods  $U_{\leq k_n^-}(v_1)$  and  $U_{\leq k_n^-}(v_2)$  to single vertices  $a$  and  $b$ , with degrees  $d_{\min}(d_{\min}-1)^{k_n^-} = O(\log n)$  (recall (2.1)-(2.2)). The degree sequence  $\hat{\mathbf{d}}$  of this new configuration model is a slight modification of the original degree sequence  $\mathbf{d}$ : two new vertices of degree  $O(\log n)$  have been added, while  $2(i_{k_n^-} + 1) = O(\log n)$  vertices with degree  $d_{\min}$  have been removed (recall (3.6)). Consequently  $\hat{\mathbf{d}}$  still satisfies the assumptions of Theorem 1.3, hence Statement 2.3 (to be proved in Section 3.3) holds for  $\hat{\mathbb{P}}$  and we obtain

$$\hat{\mathbb{P}}(\text{dist}(a, b) \leq 2\bar{k}_n) = o(1). \quad (3.31)$$

We are ready to conclude the proof of Statement 2.2. By (3.28)-(3.29)-(3.30),

$$\begin{aligned} \mathbb{P}(E_n) &= \sum_{v_1, v_2 \in [n]} \frac{\mathbb{P}(v_1, v_2 \in \mathcal{M}_{k_n^-}, \text{dist}(v_1, v_2) \leq 2k_n^- + 2\bar{k}_n, A_{v_1, v_2}^c)}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} + o(1) \\ &\leq \hat{\mathbb{P}}(\text{dist}(a, b) \leq 2\bar{k}_n) \sum_{v_1, v_2 \in [n]} \frac{\mathbb{P}(v_1, v_2 \in \mathcal{M}_{k_n^-})}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} + o(1) \\ &= \hat{\mathbb{P}}(\text{dist}(a, b) \leq 2\bar{k}_n) \frac{\mathbb{E}[(M_{k_n^-})^2]}{\frac{1}{4}\mathbb{E}[M_{k_n^-}]^2} + o(1). \end{aligned}$$

Observe that  $\mathbb{E}[(M_{k_n^-})^2] = \mathbb{E}[M_{k_n^-}]^2 + \text{Var}(M_{k_n^-}) = O(\mathbb{E}[M_{k_n^-}]^2)$ , by the second relation in (2.3). Applying (3.31), it follows that  $\mathbb{P}(E_n) = o(1)$ , completing the proof of Statement 2.2.  $\square$

**3.3. Proof of Statement 2.3.** In this section, we give a self-contained proof of Statement 2.3 for  $\text{CM}_n$ , as used in the proof of Statement 2.2.

Given two vertices  $a, b \in [n]$ , let  $\mathcal{P}_k(a, b)$  be the set of all self-avoiding paths of length  $k$  from  $a$  to  $b$ , i.e. of all sequences  $(\pi_0, \pi_1, \dots, \pi_k) \in [n]^{k+1}$  with  $\pi_0 = a$ ,  $\pi_k = b$  and such that  $(\pi_{i-1}, \pi_i)$  is an edge in the graph, for all  $i = 1, \dots, k$ . Analogously, let  $\mathcal{P}_k(a) = \cup_{b \in [n]} \mathcal{P}_k(a, b)$  denote the set of all paths of length  $k$  starting at  $a$ .

Let us fix an arbitrary increasing sequence  $(g_l)_{l \in \mathbb{N}_0}$  (that will be specified later). Define, for  $a, b \in \mathbb{R}$ ,  $a \wedge b := \min\{a, b\}$ . We say that a path  $\pi \in \mathcal{P}_k(a, b)$  is *good* when  $d_{\pi_l} \leq g_l \wedge g_{k-l}$  for every  $l = 0, \dots, k$ , and *bad* otherwise. In other words, a path is good when the degrees along the path do not increase too much from  $\pi_0$  to  $\pi_{k/2}$ , and similarly they do not increase too much in the backward direction, from  $\pi_k$  to  $\pi_{k/2}$ .

For  $k \in \mathbb{N}_0$ , we introduce the event

$$\mathcal{E}_k(a, b) = \{\exists \pi \in \mathcal{P}_k(a, b) : \pi \text{ is a good path}\}. \quad (3.32)$$

To deal with bad paths, we define

$$\mathcal{F}_k(a) = \{\exists \pi \in \mathcal{P}_k(a) : d_{\pi_k} > g_k \text{ but } d_{\pi_i} \leq g_i \forall i \leq k-1\}. \quad (3.33)$$

If  $\text{dist}_{\text{CM}_n}(a, b) \leq 2\bar{k}$ , then there must be a path in  $\mathcal{P}_k(a, b)$  for some  $k \leq \bar{k}$ , and this path might be good or bad. This leads to the simple bound

$$\mathbb{P}(\text{dist}_{\text{CM}_n}(a, b) \leq 2\bar{k}) \leq \sum_{k=0}^{2\bar{k}} \mathbb{P}(\mathcal{E}_k(a, b)) + \sum_{k=0}^{\bar{k}} [\mathbb{P}(\mathcal{F}_k(a)) + \mathbb{P}(\mathcal{F}_k(b))]. \quad (3.34)$$

We give explicit estimates for the two sums in the right hand side. We introduce the *size-biased distribution function*  $F_n^*$  associated to the degree sequence  $\mathbf{d} = (d_1, \dots, d_n)$  by

$$F_n^*(t) = \frac{1}{\ell_n} \sum_{v \in [n]} d_v \mathbf{1}_{\{d_v \leq t\}}. \quad (3.35)$$

If we choose uniformly one of the  $\ell_n$  half-edges in the graph, and call  $D_n^*$  the degree of the vertex incident to this half-edge, then  $F_n^*(t) = \mathbb{P}(D_n^* \leq t)$ . We also define the truncated mean

$$\nu_n(t) = \mathbb{E}[(D_n^* - 1)\mathbf{1}_{\{D_n^* \leq t\}}] = \frac{1}{\ell_n} \sum_{v \in [n]} d_v(d_v - 1)\mathbf{1}_{\{d_v \leq t\}}. \quad (3.36)$$

Now we are ready to bound (3.34).

**Proposition 3.7** (Path counting for configuration model). *Fix  $\mathbf{d} = (d_1, \dots, d_n)$  (such that  $\ell_n = d_1 + \dots + d_n$  is even) and an increasing sequence  $(g_l)_{l \in \mathbb{N}_0}$ . For all distinct vertices  $a, b \in [n]$  with  $d_a \leq g_0$ ,  $d_b \leq g_0$ , and for all  $\bar{k} \in \mathbb{N}$ ,*

$$\begin{aligned} \mathbb{P}(\text{dist}_{\text{CM}_n}(a, b) \leq 2\bar{k}) &\leq \frac{d_a d_b}{\ell_n} \sum_{k=1}^{2\bar{k}} \left(1 - \frac{2k}{\ell_n}\right)^{-k} \prod_{l=1}^{k-1} \nu_n(g_l \wedge g_{h-l}) \\ &\quad + (d_a + d_b) \sum_{k=1}^{\bar{k}} \left(1 - \frac{2k}{\ell_n}\right)^{-k} (1 - F_n^*(g_k)) \prod_{l=1}^{k-1} \nu_n(g_l). \end{aligned} \quad (3.37)$$

*Proof.* Fix an arbitrary sequence of vertices  $\pi = (\pi_i)_{0 \leq i \leq k} \in [n]^{k+1}$ . The probability that vertex  $\pi_0$  is connected to  $\pi_1$  is at most

$$\frac{d_{\pi_0} d_{\pi_1}}{\ell_n - 1},$$

because there are  $d_{\pi_0} d_{\pi_1}$  ordered couples of half-edges, each of which can be paired with probability  $1/(\ell_n - 1)$  (recall Remark 3.1), and we use the union bound. By similar arguments, conditionally on a specific half-edge incident to  $\pi_0$  being paired to a specific half-edge incident to  $\pi_1$ , the probability that another half-edge incident to  $\pi_1$  is paired to a half-edge incident to  $\pi_2$  is by the union bound bounded from above by

$$\frac{(d_{\pi_1} - 1)d_{\pi_2}}{\ell_n - 3}.$$

Iterating the argument, the probability that  $\pi$  is a path in  $\text{CM}_n$  is at most

$$\frac{d_{\pi_0} d_{\pi_1}}{\ell_n - 1} \frac{(d_{\pi_1} - 1)d_{\pi_2}}{\ell_n - 3} \frac{(d_{\pi_2} - 1)d_{\pi_3}}{\ell_n - 5} \dots \frac{(d_{\pi_{k-1}} - 1)d_{\pi_k}}{\ell_n - (2k - 1)}. \quad (3.38)$$

Let us now fix  $a, b \in [n]$  with  $a \neq b$ . Recalling (3.32)-(3.36), choosing  $\pi_0 = a$ ,  $\pi_k = b$  and summing (3.38) over all vertices  $\pi_1, \dots, \pi_{k-1}$  satisfying  $d_{\pi_i} \leq g_i \wedge g_{k-i}$  yields

$$\mathbb{P}(\mathcal{E}_k(a, b)) \leq d_a d_b \frac{(\ell_n - 2k - 1)!!}{(\ell_n - 1)!!} \left( \prod_{i=1}^{k-1} \ell_n \nu_n(g_i \wedge g_{k-i}) \right). \quad (3.39)$$

Bounding  $(\ell_n - 2k - 1)!!/(\ell_n - 1)!! \leq (\ell_n - 2k)^{-k}$  yields the first term in the right hand side of (3.37). The bound for  $\mathbb{P}(\mathcal{F}_k(a))$  is similar. Recalling (3.33)-(3.35), choosing  $\pi_0 = a$  and summing (3.38) over vertices  $\pi_1, \dots, \pi_{k-1}, \pi_k$  such that  $d_{\pi_i} \leq g_i$  for  $i \leq k - 1$  while  $d_{\pi_k} > g_k$  gives

$$\mathbb{P}(\mathcal{F}_k(a)) \leq d_a \frac{(\ell_n - 2k - 1)!!}{(\ell_n - 1)!!} \left( \prod_{i=1}^{k-1} \ell_n \nu_n(g_i) \right) \{ \ell_n (1 - F_n^*(g_k)) \}, \quad (3.40)$$

and the same holds for  $\mathbb{P}(\mathcal{F}_k(b))$ . Plugging (3.39) and (3.40) into (3.34) proves (3.37).  $\square$

In order to exploit (3.37), we need estimates on  $F_n^*$  and  $\nu_n$ , provided by the next lemma:

**Lemma 3.8** (Tail and truncated mean bounds for  $D_n^*$ ). *Assume that Condition 1.2 holds. Fix  $\eta > 0$ , then there exist two constants  $C_1 = C_1(\eta)$  and  $C_2 = C_2(\eta)$  such that, for every  $x \geq 0$ ,*

$$1 - F_n^*(x) \leq C_1 x^{-(\tau-2-\eta)}, \quad \nu_n(x) \leq C_2 x^{(3-\tau+\eta)}. \quad (3.41)$$

*Proof.* For every  $x \geq 0$  and  $t \geq 0$  we can see that

$$1 - F_n^*(x) = \frac{1}{\ell_n} \sum_{v \in [n]} d_v \mathbb{1}_{\{d_v > x\}} = \frac{n}{\ell_n} \left[ \frac{1}{n} \sum_{v \in [n]} d_v \mathbb{1}_{\{d_v > x\}} \right] = \frac{n}{\ell_n} \mathbb{E} [D_n \mathbb{1}_{\{D_n > x\}}], \quad (3.42)$$

where we recall that  $D_n$  is the degree of a uniformly chosen vertex. This means that

$$\begin{aligned} \frac{n}{\ell_n} \mathbb{E}[D_n \mathbb{1}_{\{D_n > x\}}] &= \frac{n}{\ell_n} \sum_{j=0}^{\infty} \mathbb{P}(D_n \mathbb{1}_{\{D_n > x\}} > j) = \frac{n}{\ell_n} \sum_{j=0}^{\infty} \mathbb{P}(D_n > j, D_n > x) \\ &= \frac{n}{\ell_n} \sum_{j=0}^{\infty} \mathbb{P}(D_n > j \vee x) = \frac{n}{\ell_n} \sum_{j=0}^{\infty} (1 - F_{d,n}(j \vee x)) \\ &= \frac{n}{\ell_n} \left[ x(1 - F_{d,n}(x)) + \sum_{j=x}^{\infty} (1 - F_{d,n}(j)) \right] \\ &\leq \frac{n}{\ell_n} C \left[ x^{-(\tau-2-\eta)} + \sum_{j=x}^{\infty} j^{-(\tau-1-\eta)} \right] \leq C_1 x^{-(\tau-2-\eta)}, \end{aligned} \quad (3.43)$$

where we have used Condition 1.2 in the second last step (recall that  $2 < \tau < 3$ ).

For  $\nu_n$ , we can instead write

$$\begin{aligned} \nu_n(x) &= \frac{1}{\ell_n} \sum_{v \in [n]} d_v (d_v - 1) \mathbb{1}_{\{d_v \leq x\}} = \frac{n}{\ell_n} \left[ \frac{1}{n} \sum_{v \in [n]} d_v (d_v - 1) \mathbb{1}_{\{d_v \leq x\}} \right] \\ &= \frac{n}{\ell_n} \mathbb{E} [D_n(D_n - 1) \mathbb{1}_{\{D_n \leq x\}}] \leq \frac{n}{\ell_n} \mathbb{E} [D_n^2 \mathbb{1}_{\{D_n \leq x\}}], \end{aligned} \quad (3.44)$$

where  $D_n$  is again the degree of a uniformly chosen vertex. The claim now follows from

$$\begin{aligned} \frac{n}{\ell_n} \mathbb{E} [D_n^2 \mathbb{1}_{\{D_n \leq x\}}] &= \frac{n}{\ell_n} \sum_{j=0}^{\infty} (2j+1) \mathbb{P}(D_n \mathbb{1}_{\{D_n \leq x\}} > j) \\ &= \frac{n}{\ell_n} \sum_{j=0}^{\infty} (2j+1) \mathbb{P}(D_n > j, D_n \leq x) \leq \frac{n}{\ell_n} \sum_{j=0}^{x-1} (2j+1) \mathbb{P}(D_n > j) \\ &= \frac{n}{\ell_n} \sum_{j=0}^{x-1} (2j+1) [1 - F_{d,n}(j)] \leq \frac{n}{\ell_n} \sum_{j=0}^{x-1} C j^{-(\tau-2-\eta)} \leq \frac{n}{\ell_n} C_2 x^{3-\tau+\eta}. \end{aligned} \quad (3.45)$$

We are finally ready to complete the proof of Statement 2.3:

*Proof of Statement 2.3.* As in (2.4), we take

$$\bar{k}_n = (1 - \varepsilon) \frac{\log \log n}{|\log(\tau - 2)|}, \quad (3.46)$$

and our goal is to show that, as  $n \rightarrow \infty$ ,

$$\max_{a,b \in [n]: d_a, d_b \leq \log n} \mathbb{P}(\text{dist}_{\text{CM}_n}(a, b) \leq 2\bar{k}_n) \longrightarrow 0. \quad (3.47)$$

We stress that  $\tau \in (2, 3)$  and  $\varepsilon > 0$  are fixed. Then we choose  $\eta > 0$  so small that

$$2\eta < \tau - 2 \quad \text{and} \quad \frac{|\log(\tau - 2 - 2\eta)|}{|\log|\log(\tau - 2)||} \leq \frac{1 - \varepsilon/2}{1 - \varepsilon}. \quad (3.48)$$

We use the inequality (3.37) given by Proposition 3.7, with the following choice of  $(g_k)_{k \in \mathbb{N}_0}$ :

$$g_k := (g_0)^{p^k}, \quad \text{where} \quad \begin{cases} g_0 := (\log n)^{\log \log n}; \\ p := \frac{1}{\tau - 2 - 2\eta} > 1. \end{cases} \quad (3.49)$$

Let us focus on the first term in the right hand side of (3.37), that is

$$\frac{d_a d_b}{\ell_n} \sum_{k=1}^{2\bar{k}} \left(1 - \frac{2k}{\ell_n}\right)^{-k} \prod_{l=1}^{k-1} \nu_n(g_l \wedge g_{h-l}). \quad (3.50)$$

Since  $\ell_n = \mu n(1 + o(1))$  by (3.5), for  $k \leq 2\bar{k}_n$  we have

$$\left(1 - \frac{2k}{\ell_n}\right)^{-k} \leq \left(1 - \frac{4\bar{k}_n}{\ell_n}\right)^{-2\bar{k}_n} = 1 + O\left(\frac{\bar{k}_n^2}{\ell_n}\right) = 1 + O\left(\frac{(\log \log n)^2}{n}\right) = 1 + o(1). \quad (3.51)$$

Then observe that, by Lemma 3.8 and (3.49), for  $k \leq 2\bar{k}_n$

$$\begin{aligned} \prod_{l=1}^{k-1} \nu_n(g_l \wedge g_{k-l}) &= \prod_{l=1}^{k/2} \nu_n(g_l)^2 \leq C_2^{k/2} \prod_{l=1}^{k/2} (g_l)^{2(3-\tau+\eta)} = C_2^{k/2} (g_0)^{2(3-\tau+\eta) \sum_{l=1}^{k/2} p^l} \\ &\leq C_2^{\bar{k}_n} (g_0)^{2(3-\tau+\eta) C p^{\bar{k}_n}}, \end{aligned} \quad (3.52)$$

with  $C = \frac{p}{p-1}$ . Note that  $C_2^{\bar{k}_n} = O((\log n)^c)$  for some  $c \in (0, \infty)$ , see (3.46), while by (3.48)

$$p^{\bar{k}_n} = \exp\left(|\log(\tau - 2 - 2\eta)|(1 - \varepsilon) \frac{\log \log n}{|\log(\tau - 2)|}\right) = (\log n)^{(1 - \varepsilon) \frac{|\log(\tau - 2 - 2\eta)|}{|\log(\tau - 2)|}} \leq (\log n)^{(1 - \varepsilon/2)}, \quad (3.53)$$

hence the right hand side of (3.52) is  $n^{o(1)}$  (since  $g_0 = (\log n)^{\log \log n}$ ). Then, for  $d_a, d_b \leq \log n$ ,

$$(3.50) \leq \frac{(\log n)^2}{\ell_n} (2\bar{k}_n) (1 + o(1)) n^{o(1)} = O\left(\frac{(\log n)^2}{n} (\log \log n) n^{o(1)}\right) = o(1).$$

It remains to look at the second sum in (3.37):

$$(d_a + d_b) \sum_{k=1}^{\bar{k}_n} \left(1 - \frac{2k}{\ell_n}\right)^{-k} (1 - F_n^*(g_k)) \prod_{l=1}^{k-1} \nu_n(g_l). \quad (3.54)$$

By Lemma 3.8, we can bound  $1 - F_n^*(g_k) \leq C_1(g_k)^{-(\tau-2-\eta)}$ . By (3.51) and  $C_1^{\bar{k}_n} = O((\log n)^c)$  for some  $c \in (0, \infty)$ , see (3.46), bounding the product in (3.54) like we did in (3.52) yields

$$O((\log n)^c) (d_a + d_b) \sum_{k=1}^{\bar{k}_n} (g_k)^{-(\tau-2-\eta)} (g_0)^{(3-\tau+\eta) C p^{k-1}}, \quad (3.55)$$

where  $p = 1/(\tau - 2 - 2\eta)$  and  $C = \frac{p}{p-1}$ . By (3.49)

$$(g_k)^{-(\tau-2-\eta)} (g_0)^{-\frac{p}{p-1}(3-\tau+\eta)p^{k-1}} = (g_{k-1})^{-p(\tau-2-\eta)} (g_{k-1})^{\frac{p}{p-1}(3-\tau+\eta)}, \quad (3.56)$$

where

$$p(\tau - 2 - \eta) = \frac{\tau - 2 - \eta}{\tau - 2 - 2\eta} > 1, \quad \text{and} \quad \frac{p}{p-1}(3 - \tau + \eta) = \frac{3 - \tau + \eta}{3 - \tau + 2\eta} < 1. \quad (3.57)$$

This means that, setting  $D := p(\tau - 2 - \eta) - \frac{p}{p-1}(3 - \tau + \eta) > 0$ , by (3.49),

$$(3.55) = O((\log n)^c) (d_a + d_b) \sum_{k=1}^{\bar{k}_n} (g_0)^{-D p^{k-1}} \leq O((\log n)^c) \frac{d_a + d_b}{(g_0)^D}. \quad (3.58)$$

Since  $g_0 = (\log n)^{\log \log n}$  while  $d_a, d_b \leq \log n$ , the right hand side of (3.58) is  $o(1)$ .  $\square$

#### 4. LOWER BOUND FOR PREFERENTIAL ATTACHMENT MODEL

In this section we prove Statements 2.1, 2.2 and 2.3 for the preferential attachment model. By the discussion in Section 2.1, this completes the proof of the lower bound in Theorem 1.5.

We recall that, given  $m \in \mathbb{N}$  and  $\delta \in (-m, \infty)$ , the preferential attachment model  $\text{PA}_t$  is a random graph with vertex set  $[t] = \{1, 2, \dots, t\}$ , where each vertex  $w$  has  $m$  outgoing edges, which are attached to vertices  $v \in [w]$  with probabilities given in (1.10). In the next subsection we give a more detailed construction using random variables. This equivalent reformulation will be used in a few places, when we need to describe carefully some complicated events. However, for most of the exposition we will stick to the intuitive description given in Section 1.2.

**4.1. Alternative construction of the preferential attachment model.** We introduce random variables  $\xi_{w,j}$  to represent the vertex to which the  $j$ -th edge of vertex  $w$  is attached, i.e.

$$\xi_{w,j} = v \iff w \xrightarrow{j} v. \quad (4.1)$$

The graph  $\text{PA}_t$  is a *deterministic* function of these random variables: two vertices  $v, w \in [t]$  with  $v \leq w$  are connected in  $\text{PA}_t$  if and only if  $\xi_{w,j} = v$  for some  $j \in [m]$ . In particular, the degree of a vertex  $v$  after the  $k$ -th edge of vertex  $t$  has been attached, denoted by  $D_{t,k}(v)$ , is

$$D_{t,k}(v) := \sum_{(s,i) \leq (t,k)} (\mathbb{1}_{\{\xi_{s,i}=v\}} + \mathbb{1}_{\{s=v\}}), \quad (4.2)$$

where we use the natural order relation

$$(s, i) \leq (t, j) \iff s < t \quad \text{or} \quad s = t, \quad i \leq j.$$

Defining the preferential attachment model amounts to giving a joint law for the sequence  $\xi = (\xi_{w,j})_{(w,j) \in \mathbb{N} \times [m]}$ . In agreement with (1.10), we set  $\xi_{1,j} = 1$  for all  $j \in [m]$ , and for  $t \geq 2$

$$\mathbb{P}(\xi_{t,j} = v | \xi_{\leq(t,j-1)}) = \begin{cases} \frac{D_{t,j-1}(v) + 1 + j\delta/m}{c_{t,j}} & \text{if } v = t; \\ \frac{D_{t,j-1}(v) + \delta}{c_{t,j}} & \text{if } v < t, \end{cases} \quad (4.3)$$

where  $\xi_{\leq(t,j-1)}$  is a shorthand for the vector  $(\xi_{s,i})_{(s,i) \leq (t,j-1)}$  (and we agree that  $(t, 0) := (t-1, m)$ ). The normalizing constant  $c_{t,j}$  in (4.3) is indeed given by (1.11), because by (4.2),

$$\sum_{v \in [t]} D_{t,j-1}(v) = \sum_{(s,i) \leq (t,j-1)} (1+1) = 2((t-1)m + (j-1)).$$

The factor  $j\delta/m$  in the first line of (4.3) is commonly used in the literature (instead of the possibly more natural  $\delta$ ). The reason is that, with such a definition, the graph  $\text{PA}_t(m, \delta)$  can be obtained from the special case  $m = 1$ , where every vertex has only one outgoing edge: one first generates the random graph  $\text{PA}_{mt}(1, \delta/m)$ , whose vertex set is  $[mt]$ , and then collapses the block of vertices  $[m(i-1) + 1, mi]$  into a single vertex  $i \in [t]$  (see also [Hof17, Chapter 8]).

**Remark 4.1.** It is clear from the construction that  $\text{PA}_t$  is a *labeled directed graph*, because any edge connecting sites  $v, w$ , say with  $v \leq w$ , carries a label  $j \in [m]$  and a direction, from the newer vertex  $w$  to the older one  $v$  (see (4.1)). Even though our final result, the asymptotic behavior of the diameter, only depends on the underlying undirected graph, it will be convenient to exploit the labeled directed structure of the graph in the proofs.

**4.2. Proof of Statement 2.1.** We denote by  $U_{\leq k}(v)$  the  $k$ -neighborhood in  $\text{PA}_t$  of a vertex  $v \in [t]$ , i.e. the set of vertices at distance at most  $k$  from  $v$ , viewed as a labeled directed subgraph (see Remark 4.1). We denote by  $D_t(v) = D_{t,m}(v)$  the degree of vertex  $v$  after time  $t$ , i.e. in the graph  $\text{PA}_t$  (recall (4.2)).

We define the notion of *minimally- $k$ -connected vertex* in analogy with the configuration model (see Definition 3.2), up to minor technical restrictions made for later convenience.

**Definition 4.2** (Minimally- $k$ -connected vertex). *For  $k \in \mathbb{N}_0$ , a vertex  $v \in [t] \setminus [t/2]$  is called minimally- $k$ -connected when  $D_t(v) = m$ , all the other vertices  $i \in U_{\leq k}(v)$  are in  $[t/2] \setminus [t/4]$  and have degree  $D_t(i) = m + 1$ , and there are no self-loops, multiple edges or cycles in  $U_{\leq k}(v)$ . The graph  $U_{\leq k}(v)$  is thus a tree with degree  $m + 1$ , except for the root  $v$  which has degree  $m$ .*

We denote the (random) set of minimally- $k$ -connected vertices by  $\mathcal{M}_k \subseteq [t] \setminus [t/2]$ , and its cardinality by  $M_k = |\mathcal{M}_k|$ .

For the construction of a minimally- $k$ -connected neighborhood in the preferential attachment model we remind that the vertices are added to the graph at different times, so that the vertex degrees change while the graph grows. The relevant degree for Definition 4.2 is the one at the final time  $t$ . To build a minimally- $k$ -connected neighborhood, we need

$$i_k = 1 + \sum_{i=1}^k m^i = \frac{m^{k+1} - 1}{m - 1} \quad (4.4)$$

many vertices. The center  $v$  of the neighborhood is the youngest vertex in  $U_{\leq k}(v)$ , and it has degree  $m$ , while all the other vertices have degree  $m + 1$ .

Our first goal is to evaluate the probability  $\mathbb{P}(v \in \mathcal{M}_k)$  that a given vertex  $v \in [t] \setminus [t/2]$  is minimally- $k$ -connected. The analogous question for the configuration model could be answered quite easily in Proposition 3.5, because the configuration model can be built exploring its vertices in an arbitrary order, in particular starting from  $v$ , see Remark 3.1. This is no longer true for the preferential attachment model, whose vertices have an order, the chronological one, along which the conditional probabilities take the explicit form (1.10) or (4.3). This is why the proofs for the preferential attachment model are harder than for the configuration model.

As it will be clear in a moment, to get explicit formulas it is convenient to evaluate the probability  $\mathbb{P}(v \in \mathcal{M}_k, U_{\leq k}(v) = H)$ , where  $H$  is a fixed *labeled directed* subgraph, i.e. it comes with the specification of which edges are attached to which vertices. To avoid trivialities, we restrict to those  $H$  for which the probability does not vanish, i.e. which satisfy the constraints in Definition 4.2, and we call them *admissible*.

Let us denote by  $H^o := H \setminus \partial H$  the set of vertices in  $H$  that are not on the boundary (i.e. they are at distance at most  $k - 1$  from  $v$ ). With this notation, we have the following result:

**Lemma 4.3.** *Let  $\{\text{PA}_t\}_{t \in \mathbb{N}}$  be a preferential attachment model. For any vertex  $v \in [t] \setminus [t/2]$  and any directed labeled graph  $H$  which is admissible,*

$$\mathbb{P}(v \in \mathcal{M}_k, U_{\leq k}(v) = H) = L_1(H) L_2(H), \quad (4.5)$$

where

$$L_1(H) := \prod_{u \in H^o} \prod_{j=1}^m \frac{m + \delta}{c_{u,j}}, \quad (4.6)$$

$$L_2(H) := \prod_{u \notin H^o} \prod_{j=1}^m \left[ 1 - \frac{D_{u-1}(H) + |H \cap [u-1]| \delta}{c_{u,j}} \right], \quad (4.7)$$

and  $D_{u-1}(H) = \sum_{w \in H} D_{u-1,m}(w)$  is the total degree of  $H$  before vertex  $u$  is added to the graph, and the normalization constant  $c_{u,j}$  is defined in (1.11).

*Proof.* We recall that  $\{a \xrightarrow{i} b\}$  denotes the event that the  $i$ -th edge of  $a$  is attached to  $b$  (see (4.1)). Since  $H$  is an admissible labeled directed subgraph, for all  $u \in H^o$  and  $j \in [m]$ , the  $j$ -th

edge of  $u$  is connected to a vertex in  $H$ , that we denote by  $\theta_j^H(u)$ . We can then write

$$\{v \in \mathcal{M}_k, U_{\leq k}(v) = H\} = \left( \bigcap_{u \in H^o} \bigcap_{j=1}^m \{u \xrightarrow{j} \theta_j^H(u)\} \right) \cap \left( \bigcap_{u \notin H^o} \bigcap_{j=1}^m \{u \not\xrightarrow{j} H\} \right), \quad (4.8)$$

where of course  $\{u \not\xrightarrow{j} H\} := \bigcup_{w \notin H} \{u \xrightarrow{j} w\}$ . The first term in (4.8) is exactly the event that the edges present in  $H$  are connected in  $\text{PA}_t$  as they should be. The second term is the event that the vertices  $u \notin H^o$  are not attached to  $H$ , so that  $U_{\leq k}(v) = H$ . Notice that in (4.8) every vertex and every edge of the graph appears. For a vertex  $u \in H^o$ , by (1.10)

$$\mathbb{P}\left(u \xrightarrow{j} \theta_j^H(u) \mid \text{PA}_{u,j-1}\right) = \frac{m + \delta}{c_{u,j}}, \quad (4.9)$$

because the vertex  $\theta_j^H(u)$  has degree precisely  $m$  (when  $u$  is not already present in the graph). For  $u \notin H^o$ , we have to evaluate the probability that its edges do no attach to  $H$ , which is

$$\mathbb{P}\left(u \not\xrightarrow{j} H \mid \text{PA}_{u-1,j-1}\right) = 1 - \frac{D_{u-1}(H) + |H \cap [u-1]| \delta}{c_{u,j}}. \quad (4.10)$$

Using conditional expectation iteratively, we obtain (4.9) or (4.10) for every edge in the graph, depending on whether the edge is part of  $H$  or not. This proves (4.6) and (4.7).  $\square$

The event  $\{v \in \mathcal{M}_k, U_{\leq k}(v) = H\}$  is an example of a class of events, called *factorizable*, that will be used throughout this section and Section 6. For this reason we define it precisely.

It is convenient to use the random variable  $\xi_{w,j}$ , introduced in Section 4.1, to denote the vertex to which the  $j$ -th edge of vertex  $w$  is attached (see (4.1)). Any event  $A$  for  $\text{PA}_t$  can be characterized iteratively, specifying a set  $A_{s,i} \subseteq [s]$  of values for  $\xi_{s,i}$ , for all  $(s,i) \leq (t,m)$ :

$$A = \bigcap_{(s,i) \leq (t,m)} \{\xi_{s,i} \in A_{s,i}\}.$$

Of course, the set  $A_{s,i}$  is allowed to depend on the ‘‘past’’, i.e.  $A_{s,i} = A_{s,i}(\xi_{\leq(s,i-1)})$ , or equivalently  $A_{s,i} = A_{s,i}(\text{PA}_{s,i-1})$ . Let us set  $A_{\leq(s,i)} := \bigcap_{(u,j) \leq (s,i)} A_{u,j}$ .

**Definition 4.4** (Factorizable events). *An event  $A$  for  $\text{PA}_t$  is called factorizable when the conditional probabilities of the events  $\{\xi_{s,i} \in A_{s,i}\}$ , given the past, are deterministic. More precisely, for any  $(s,i)$  there is a (non-random)  $p_{s,i} \in [0,1]$  such that*

$$\mathbb{P}(\xi_{s,i} \in A_{s,i} \mid \xi_{\leq(s,i-1)}) = p_{s,i} \quad (4.11)$$

on the event  $\xi_{\leq(s,i-1)} \in A_{\leq(s,i-1)}$ . As a consequence, the chain rule for probabilities yields

$$\mathbb{P}(A) = \prod_{(s,i) \leq (t,m)} p_{s,i}.$$

**Remark 4.5.** Relations (4.9) and (4.10) show that  $A = \{v \in \mathcal{M}_k, U_{\leq k}(v) = H\}$  is a factorizable event. In fact,  $A_{s,i}$  is either the single vertex  $\theta_i^H(s)$  (if  $s \in H^o$ ) or the set  $[s-1] \setminus H$  (if  $s \notin H^o$ ). In both cases, the set  $A_{s,i} \subseteq [s-1]$  has a fixed total degree and a fixed cardinality, hence the conditional probabilities (4.11) are specified in a deterministic way (recall (4.3)).

Note that the event  $\{v \in \mathcal{M}_k\}$  is not factorizable. This is the reason for specifying the realization of the  $k$ -neighborhood  $U_{\leq k}(v) = H$ .

Henceforth we fix  $\varepsilon > 0$ . We recall that  $k_n^-$  was defined in (2.1). Using the more customary  $t$  instead of  $n$ , we have

$$k_t^- = (1 - \varepsilon) \frac{\log \log t}{\log m}. \quad (4.12)$$

We recall that  $M_{k_t^-} = |\mathcal{M}_{k_t^-}|$  denotes the number of minimally- $k_t^-$ -connected vertices in  $\text{PA}_t$  (see Definition 4.2). We can now prove half of Statement 2.1 for the preferential attachment model, more precisely the first relation in equation (2.3).

**Proposition 4.6** (First moment of  $M_{k_t^-}$ ). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Then, for  $k_t^-$  as in (4.12), as  $t \rightarrow \infty$ ,*

$$\mathbb{E}[M_{k_t^-}] \rightarrow \infty. \quad (4.13)$$

*Proof.* Similarly to the proof of (3.3), we write

$$\mathbb{E}[M_k] = \sum_{v \in [t] \setminus [t/2]} \mathbb{P}(v \in \mathcal{M}_k) = \sum_{v \in [t] \setminus [t/2]} \sum_{H \subseteq [t] \setminus [t/4]} \mathbb{P}(v \in \mathcal{M}_k, U_{\leq k}(v) = H), \quad (4.14)$$

where the sum is implicitly restricted to admissible  $H$  (i.e., to  $H$  that are possible realizations of  $U_{\leq k}(v)$ ).

Since we will use (4.5), we need a lower bound on (4.6) and (4.7). Recalling (1.11), it is easy to show, since the number of vertices in  $H^o$  equals  $i_k - m^k = i_{k-1}$ , and  $u \leq v$  for  $u \in H^o$ ,

$$L_1(H) \geq \left[ \frac{m + \delta}{v(2m + \delta) + 1 + \delta/m} \right]^{mi_{k-1}}. \quad (4.15)$$

Note that for  $u \leq t/4$  all the factors in the product in (4.7) equal 1, because  $H \subseteq [t] \setminus [t/4]$ . Restricting to  $u > t/4$  and bounding  $D_{u-1}(H) + |H \cap [u-1]|/\delta \leq (m+1+\delta)i_k$ , we get

$$L_2(H) \geq \left[ 1 - \frac{(m+1+\delta)i_k}{\frac{t}{4}(2m+\delta)} \right]^{3mt/4}. \quad (4.16)$$

Let us write  $H = \{v\} \cup H'$  where  $H'$  is a subset of  $[t/2] \setminus [t/4]$  with  $|H'| = i_k - 1$ . Clearly, for any such subset there is at least one way to order the vertices to generate an admissible  $H$ . The number of possible subsets in  $[t/2] \setminus [t/4]$  is at least  $\binom{t/4}{i_k-1}$ . Then, we obtain

$$\mathbb{E}[M_k] \geq \sum_{v \in [t] \setminus [t/2]} \binom{t/4}{i_k-1} \left[ \frac{m + \delta}{v(2m + \delta) + 1 + \delta/m} \right]^{mi_{k-1}} \left[ 1 - \frac{(m+1+\delta)i_k}{\frac{t}{4}(2m+\delta)} \right]^{3mt/4}. \quad (4.17)$$

Recalling that

$$\binom{t/4}{i_k-1} = \frac{t^{i_k}}{4^{i_k}(i_k-1)!} (1 + o(1)), \quad (4.18)$$

since  $mi_{k-1} \leq i_k$ , we obtain

$$\mathbb{E}[M_k] \geq \frac{t}{2} \frac{t^{i_k}}{4^{i_k}(i_k-1)!} \left[ \frac{m + \delta}{t(2m + \delta) + 1 + \delta/m} \right]^{i_k} \left[ 1 - \frac{(m+1+\delta)i_k}{\frac{t}{4}(2m+\delta)} \right]^{3mt/4}. \quad (4.19)$$

Choosing  $k = k_t^-$  as in (4.12) and bounding  $1 - x \geq e^{-2x}$  for  $x$  small, as well as  $m+1 \leq 2m$ , we obtain

$$\mathbb{E}[M_{k_t^-}] \geq \frac{t}{2} \frac{t^{i_{k_t^-}}}{4^{i_{k_t^-}} i_{k_t^-}!} \left( \frac{m}{Ct} \right)^{i_{k_t^-}} \exp(-3cmi_{k_t^-}) \geq \frac{1}{(C')^{i_{k_t^-}}} \frac{t}{2i_{k_t^-}!} \exp(-3cmi_{k_t^-}), \quad (4.20)$$

where  $C$  is a constant and  $C' = 4C/m$ . Recalling that  $i_k$  is given by (4.4), and  $k_t^-$  by (4.12), hence  $i_{k_t^-} = \frac{m}{m-1} m^{k_t^-} (1 + o(1)) \leq 2(\log t)^{1-\varepsilon}$ , hence

$$i_{k_t^-}! \leq \lfloor 2(\log t)^{1-\varepsilon} \rfloor! \leq [2(\log t)^{1-\varepsilon}]^{2(\log t)^{1-\varepsilon}} = t^{o(1)}, \quad (4.21)$$

and also  $(C'e^{3cm})^{i_{k_t^-}} = t^{o(1)}$ . This implies that  $\mathbb{E}[M_k] \rightarrow \infty$ , as required.  $\square$

**Remark 4.7** (Disjoint neighborhoods for minimally  $k$ -connected pairs). We observe that, on the event  $\{v, w \in \mathcal{M}_k\}$  with  $v \neq w$ , necessarily

$$U_{\leq k}(v) \cap U_{\leq k}(w) = \emptyset,$$

because if a vertex  $x$  is in  $U_{\leq k}(v) \cap U_{\leq k}(w)$  and  $x \neq v, w$ , this means that  $D_x(t) = m+2$ , because in addition to its original  $m$  outgoing edges, vertex  $x$  has one incident edge from a younger vertex

in  $U_{\leq k}(v)$  and one incident edge from a younger vertex in  $U_{\leq k}(u)$ , which gives a contradiction. Similar arguments apply when  $x = v$  or  $x = w$ .

We use the previous remark to prove the second relation in Statement 2.1 for the preferential attachment model.

**Proposition 4.8** (Second moment of  $M_{k_t^-}$ ). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Then, for  $k \in \mathbb{N}$ ,*

$$\mathbb{E}[M_k^2] \leq \exp(32mi_k^2/t) \mathbb{E}[M_k]^2 + \mathbb{E}[M_k]. \quad (4.22)$$

Consequently, for  $k = k_t^-$  as in (4.12), as  $t \rightarrow \infty$ ,

$$\mathbb{E}[M_{k_t^-}^2] \leq (1 + o(1)) \mathbb{E}[M_{k_t^-}]^2. \quad (4.23)$$

*Proof.* We write

$$\mathbb{E}[M_k^2] = \sum_{v,w \in [t] \setminus [t/2]} \mathbb{P}(v, w \in \mathcal{M}_k) = \sum_{v \neq w} \mathbb{P}(v, w \in \mathcal{M}_k) + \mathbb{E}[M_k]. \quad (4.24)$$

By Remark 4.7, for  $v \neq w$  we can write

$$\mathbb{P}(v, w \in \mathcal{M}_k) = \sum_{H_v \cap H_w = \emptyset} \mathbb{P}(v, w \in \mathcal{M}_k, U_{\leq k}(v) = H_v, U_{\leq k}(w) = H_w). \quad (4.25)$$

The crucial observation is that the event  $\{v, w \in \mathcal{M}_k, U_{\leq k}(v) = H_v, U_{\leq k}(w) = H_w\}$  is factorizable (recall Definition 4.4 and Remark 4.5). More precisely, in analogy with (4.6) and (4.7):

$$\mathbb{P}(v, w \in \mathcal{M}_k, U_{\leq k}(v) = H_v, U_{\leq k}(w) = H_w) = L_1(H_v, H_w)L_2(H_v, H_w), \quad (4.26)$$

where now

$$L_1(H_v, H_w) = \prod_{x \in H_v^o \cup H_w^o} \prod_{j=1}^m \frac{m+\delta}{c_{x,j}}, \quad (4.27)$$

$$L_2(H_v, H_w) = \prod_{x \notin H_v^o \cup H_w^o} \prod_{j=1}^m \left[ 1 - \frac{D_{x-1}(H_v \cup H_w) + |(H_v \cup H_w) \cap [x-1]| \delta}{c_{x,j}} \right]. \quad (4.28)$$

To prove (4.26), notice that in (4.27) and (4.28) every edge and every vertex of the graph appear. Further, (4.27) is the probability of the event  $\{U_{\leq k}(v) = H_v, U_{\leq k}(w) = H_w\}$ , while (4.28) is the probability that all vertices not in the two neighborhoods do not attach to the two trees.

A look at (4.6) shows that  $L_1(H_v, H_w) = L_1(H_v)L_1(H_w)$ . We now show that analogous factorization holds approximately also for  $L_2$ . Since, for every  $a, b \in [0, 1]$ , with  $a + b < 1$ , it is true that  $1 - (a + b) \leq (1 - a)(1 - b)$ , we can bound

$$\begin{aligned} & \left[ 1 - \frac{D_{x-1}(H_v \cup H_w) + |(H_v \cup H_w) \cap [x-1]| \delta}{c_{x,j}} \right] \\ & \leq \left[ 1 - \frac{D_{x-1}(H_v) + |H_v \cap [x-1]| \delta}{c_{x,j}} \right] \left[ 1 - \frac{D_{x-1}(H_w) + |H_w \cap [x-1]| \delta}{c_{x,j}} \right]. \end{aligned} \quad (4.29)$$

When we plug (4.29) into (4.28), we obtain  $L_2(H_v)L_2(H_w)$  (recall (4.7)) times the following terms:

$$\left( \prod_{x \in H_w^o} \left[ 1 - \frac{D_{x-1}(H_v) + |H_v \cap [x-1]| \delta}{c_{x,j}} \right] \right)^{-1} \left( \prod_{x \in H_v^o} \left[ 1 - \frac{D_{x-1}(H_w) + |H_w \cap [x-1]| \delta}{c_{x,j}} \right] \right)^{-1}. \quad (4.30)$$

We can bound  $D_{x-1}(H_v) + |H_v \cap [x-1]| \delta \leq D_{x-1}(H_v) \leq (m+1)i_k$  (recall that  $\delta < 0$ ) and analogously for  $H_w$ . The square brackets in (4.30) equal 1 for  $x \leq t/4$  (since  $H_v, H_w \subseteq [t] \setminus [t/4]$

by construction), and for  $x > t/4$  we have  $c_{x,j} \geq \frac{t}{4}(2m + \delta) \geq \frac{m}{4}t$  by (1.11) and  $\delta > -m$ . We can thus write

$$\begin{aligned} L_2(H_v, H_w) &\leq L_2(H_v) L_2(H_w) \prod_{x \in H_v^o \cup H_w^o} \prod_{j=1}^m \left[ 1 - \frac{(m+1)i_k}{\frac{m}{4}t} \right]^{-1} \\ &\leq L_2(H_v) L_2(H_w) \exp \left( 2(2i_k)m \frac{(m+1)i_k}{\frac{m}{4}t} \right), \end{aligned} \quad (4.31)$$

where we have used the bound  $1 - z \geq e^{-2z}$  for small  $z > 0$ . Since  $m+1 \leq 2m$ , we obtain

$$\begin{aligned} &\sum_{v \neq w} \left[ \sum_{H_v \cap H_w = \emptyset} \mathbb{P}(v, w \in \mathcal{M}_k, U_{\leq k}(v) = H_v, U_{\leq k}(w) = H_w) \right] \\ &\leq \exp(32mi_k^2/t) \sum_{v \in [t] \setminus [t/2]} \sum_{H_v} L_1(H_v) L_2(H_v) \sum_{w \in [t] \setminus [t/2]} \sum_{H_w} L_1(H_w) L_2(H_w) \\ &= \exp(32mi_k^2/t) \mathbb{E}[M_k]^2. \end{aligned} \quad (4.32)$$

Substituting (4.32) in (4.24) completes the proof of (4.22).

Finally, for  $k = k_t^-$  as in (4.12) we have  $i_{k_t^-} \leq 2(\log t)^{1-\varepsilon}$  (recall that  $i_k$  is given by (4.4)). We have already shown in Proposition 4.6 that  $\mathbb{E}[M_{k_t^-}] \rightarrow \infty$ , hence (4.23) follows.  $\square$

Together, Propositions 4.6 and 4.8 prove Statement 2.1. This means, as for the configuration model, since  $\text{Var}(M_{k_t^-}^2) = o(\mathbb{E}[M_{k_t^-}]^2)$ , that  $M_{k_t^-}/\mathbb{E}[M_{k_t^-}] \xrightarrow[t \rightarrow \infty]{\mathbb{P}} 1$ , so in particular  $M_{k_t^-} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} \infty$ .  $\square$

### 4.3. Proof of Statement 2.3.

Fix  $\varepsilon > 0$  and define, as in (2.4),

$$\bar{k}_t = (1 - \varepsilon) \frac{2 \log \log t}{|\log(\tau - 2)|}. \quad (4.33)$$

Statement 2.3 follows from the following result on distances between not too early vertices:

**Proposition 4.9** (Lower bound on distances). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Then, there exists a constant  $p > 0$  such that*

$$\max_{x,y \geq \frac{t}{(\log t)^2}} \mathbb{P}(\text{dist}_{\text{PA}_t}(x, y) \leq 2\bar{k}_t) \leq \frac{p}{(\log t)^2}. \quad (4.34)$$

Inequality (4.34) is an adaptation of a result proved in [DMM12, Section 4.1]. Consequently we just give a sketch of the proof (the complete proof can be found in [CGH16, Appendix A]).

Let us denote by  $u \leftrightarrow v$  the event that vertices  $u, v$  are neighbors in  $\text{PA}_t$ , that is

$$\{u \leftrightarrow v\} = \bigcup_{j=1}^m \left( \{u \xrightarrow{j} v\} \cup \{v \xrightarrow{j} u\} \right).$$

(As a matter of fact,  $\{v \xrightarrow{j} u\}$  is only possible if  $v > u$ , while  $\{u \xrightarrow{j} v\}$  is only possible if  $v < u$ .) Given a sequence  $\pi = (\pi_0, \pi_1, \dots, \pi_k) \in [t]^{k+1}$  of distinct vertices, we denote by  $\{\pi \subseteq \text{PA}_t\}$  the event that  $\pi$  is a path in  $\text{PA}_t$ , that is

$$\{\pi \subseteq \text{PA}_t\} = \{\pi_0 \leftrightarrow \pi_1 \leftrightarrow \pi_2 \cdots \leftrightarrow \pi_k\} = \bigcap_{i=1}^k \{\pi_{i-1} \leftrightarrow \pi_i\}.$$

The proof of Proposition 4.9 requires the following bound on the probability of connection between two vertices from [DHH10, Lemma 2.2]: for  $\gamma = m/(2m + \delta) \in (\frac{1}{2}, 1)$ , there exists  $c \in (0, \infty)$  such that, for all vertices  $u, v \in [t]$ .

$$\mathbb{P}(u \leftrightarrow v) \leq c(u \vee v)^{\gamma-1} (u \wedge v)^{-\gamma}. \quad (4.35)$$

From [DHH10, Corollary 2.3] we know, for any sequence  $\pi = (\pi_0, \pi_1, \dots, \pi_k) \in [t]^{k+1}$  of distinct vertices,

$$\mathbb{P}(\pi \subseteq \text{PA}_t) \leq p(\pi_0, \pi_1, \dots, \pi_k) := \prod_{i=0}^{k-1} \frac{Cm}{(\pi_i \wedge \pi_{i+1})^\gamma (\pi_i \vee \pi_{i+1})^{1-\gamma}}, \quad (4.36)$$

where  $C$  is an absolute constant. The history of (4.36) is that it was first proved by Bollobás and Riordan [BR04] for  $\delta = 0$  (so that  $\gamma = 1 - \gamma = 1/2$ ), and the argument was extended to all  $\delta$  in [DHH10, Corollary 2.3].

**Remark 4.10.** *Proposition 4.9 holds for every random graphs that satisfies (4.36).*

We proceed in a similar way as in Section 3.3. Given two vertices  $x, y \in [t]$ , we consider paths  $\pi = (\pi_0, \pi_1, \dots, \pi_k)$  between  $x = \pi_0$  and  $y = \pi_k$ . We fix a decreasing sequence of numbers  $(g_l)_{l \in \mathbb{N}_0}$  that serve as truncation values for the *age* of vertices along the path (rather than the degrees as for the configuration model). We say that a path  $\pi$  is *good* when  $\pi_l \geq g_l \wedge g_{k-l}$  for every  $l = 0, \dots, k$ , and *bad* otherwise. In other words, a path is good when the age of vertices does not decrease too much from  $\pi_0$  to  $\pi_{k/2}$  and, backwards, from  $\pi_k$  to  $\pi_{k/2}$ . Intuitively, this also means that their degrees do not grow too fast. This means that

$$\mathbb{P}(\text{dist}_{\text{PA}_t}(x, y) \leq 2\bar{k}_t) \leq \sum_{k=1}^{2\bar{k}_t} \mathbb{P}(\mathcal{E}_k(x, y)) + \sum_{k=1}^{\bar{k}_t} [\mathbb{P}(\mathcal{F}_k(x)) + \mathbb{P}(\mathcal{F}_k(y))], \quad (4.37)$$

where  $\mathcal{E}_k(x, y)$  is the event of there being a good path of length  $k$ , as in (3.32), while  $\mathcal{F}_k(x)$  is the event of there being a path  $\pi$  with  $\pi_i \geq g_i$  for  $i \leq k-1$  but  $\pi_k < g_k$ , in analogy with (3.33).

Recalling the definition of  $p(\pi_0, \pi_1, \dots, \pi_k)$  in (4.36), we define for  $l \in \mathbb{N}$ ,

$$f_{l,t}(x, w) = \mathbb{1}_{\{x \geq g_0\}} \sum_{\pi_1=g_1}^t \sum_{\pi_2=g_2}^t \cdots \sum_{\pi_{l-1}=g_{l-1}}^t p(x, \pi_1, \dots, \pi_{l-1}, w), \quad (4.38)$$

setting  $f_{0,t}(x, w) = \mathbb{1}_{\{x \geq g_0\}}$  and  $f_{1,t}(x, w) = \mathbb{1}_{\{x \geq g_0\}} p(x, w)$ . From (4.37) we then obtain

$$\begin{aligned} \mathbb{P}(\text{dist}_{\text{PA}_t}(x, y) \leq 2\bar{k}_t) &\leq \sum_{k=1}^{2\bar{k}_t} \sum_{l=g_{\lfloor k/2 \rfloor}}^t f_{\lfloor k/2 \rfloor, t}(x, l) f_{\lceil k/2 \rceil, t}(y, l) \\ &\quad + \sum_{k=1}^{\bar{k}_t} \sum_{l=1}^{g_k-1} f_{k,t}(x, l) + \sum_{k=1}^{\bar{k}_t} \sum_{l=1}^{g_k-1} f_{k,t}(y, l). \end{aligned} \quad (4.39)$$

This is the starting point of the proof of Proposition 4.9.

We will show in [CGH16, Appendix A] that the following recursive bound holds

$$f_{k,t}(x, l) \leq \alpha_k l^{-\gamma} + \mathbb{1}_{\{l > g_{k-1}\}} \beta_k l^{\gamma-1}, \quad (4.40)$$

for suitable sequences  $(\alpha_k)_{k \in \mathbb{N}}$ ,  $(\beta_k)_{k \in \mathbb{N}}$  and  $(g_k)_{k \in \mathbb{N}}$  (see [CGH16, Definition A.2]). We will prove recursive bounds on these sequences that guarantee that the sums in (4.39) satisfy the required bounds. We omit further details at this point, and refer the interested reader to [CGH16, Appendix A].

**4.4. Proof of Statement 2.2.** Consider now two independent random vertices  $W_1^t$  and  $W_2^t$  that are uniformly distributed in the set of minimally- $k_t^-$ -connected vertices  $\mathcal{M}_{k_t^-}$ . We set

$$E_t := \left\{ \text{dist}(U_{\leq k_t^-}(W_1^t), U_{\leq k_t^-}(W_2^t)) \leq 2\bar{k}_t \right\} = \left\{ \text{dist}(W_1^t, W_2^t) \leq 2k_t^- + 2\bar{k}_t \right\} \quad (4.41)$$

and, in analogy with Section 3.2, our goal is to show that

$$\lim_{t \rightarrow \infty} \mathbb{P}(E_t) = 0. \quad (4.42)$$

We know from Statement 2.1 that, as  $t \rightarrow \infty$ ,

$$\mathbb{P}\left(M_{k_t^-} \leq \frac{1}{2}\mathbb{E}[M_{k_t^-}]\right) \leq \mathbb{P}\left(|M_{k_t^-} - \mathbb{E}[M_{k_t^-}]| > \frac{1}{2}\mathbb{E}[M_{k_t^-}]\right) \leq \frac{\text{Var}(M_{k_t^-})}{\frac{1}{4}\mathbb{E}[M_{k_t^-}]^2} = o(1). \quad (4.43)$$

We also define the event

$$B_t := \left\{ \max_{v \in [t]} D_t(v) \leq \sqrt{t} \right\} \quad (4.44)$$

and note that it is known (see [Hof17, Theorem 8.13]) that  $\lim_{t \rightarrow \infty} \mathbb{P}(B_t) = 1$ . Therefore,

$$\begin{aligned} \mathbb{P}(E_t) &= \mathbb{P}\left(E_t \cap \{M_{k_t^-} > \frac{1}{2}\mathbb{E}[M_{k_t^-}]\} \cap B_t\right) + o(1) \\ &= \mathbb{E}\left[\sum_{v_1, v_2 \in [t]} \mathbb{1}_{\{W_1^t = v_1, W_2^t = v_2\}} \mathbb{1}_{\{\text{dist}(v_1, v_2) \leq 2k_t^- + 2\bar{k}_t\}} \mathbb{1}_{\{M_{k_t^-} > \frac{1}{2}\mathbb{E}[M_{k_t^-}]\}} \mathbb{1}_{B_t}\right] + o(1) \\ &\leq \mathbb{E}\left[\sum_{v_1, v_2 \in [t] \setminus [t/2]} \frac{\mathbb{1}_{\{v_1 \in \mathcal{M}_{k_t^-}, v_2 \in \mathcal{M}_{k_t^-}\}}}{M_{k_t^-}^2} \mathbb{1}_{\{\text{dist}(v_1, v_2) \leq 2k_t^- + 2\bar{k}_t\}} \mathbb{1}_{\{M_{k_t^-} > \frac{1}{2}\mathbb{E}[M_{k_t^-}]\}} \mathbb{1}_{B_t}\right] + o(1) \\ &\leq \sum_{v_1, v_2 \in [t] \setminus [t/2]} \frac{\mathbb{P}(v_1, v_2 \in \mathcal{M}_{k_t^-}, \text{dist}(v_1, v_2) \leq 2k_t^- + 2\bar{k}_t, B_t)}{\frac{1}{4}\mathbb{E}[M_{k_t^-}]^2} + o(1). \end{aligned} \quad (4.45)$$

The contribution of the terms with  $v_1 = v_2$  is negligible, since it gives

$$\frac{\sum_{v_1 \in [t] \setminus [t/2]} \mathbb{P}(v_1 \in \mathcal{M}_{k_t^-})}{\frac{1}{4}\mathbb{E}[M_{k_t^-}]^2} = \frac{4}{\mathbb{E}[M_{k_t^-}]} = o(1),$$

because  $\mathbb{E}[M_{k_t^-}] \rightarrow \infty$  by Proposition 4.6. Henceforth we restrict the sum in (4.45) to  $v_1 \neq v_2$ . Summing over the realizations  $H_1$  and  $H_2$  of the random neighborhoods  $U_{\leq k_t^-}(v_1)$  and  $U_{\leq k_t^-}(v_2)$ , and over paths  $\pi$  from an arbitrary vertex  $x \in \partial H_1$  to an arbitrary vertex  $y \in \partial H_2$ , we obtain

$$\begin{aligned} \mathbb{P}(E_t) &\leq \frac{4}{\mathbb{E}[M_{k_t^-}]^2} \sum_{v_1, v_2 \in [t] \setminus [t/2]} \sum_{H_1, H_2 \subseteq [t] \setminus [t/4]} \sum_{x \in \partial H_1, y \in \partial H_2} \sum_{\substack{\pi: x \rightarrow y \\ |\pi| \leq 2\bar{k}_t}} \\ &\quad \mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2, \pi \subseteq \text{PA}_t, B_t\right) + o(1). \end{aligned} \quad (4.46)$$

The next proposition, proved below, decouples the probability appearing in the last expression:

**Proposition 4.11.** *There is a constant  $q \in (1, \infty)$  such that, for all  $v_1, v_2, H_1, H_2$  and  $\pi$ ,*

$$\begin{aligned} &\mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2, \pi \subseteq \text{PA}_t, B_t\right) \\ &\leq q \mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2\right) \mathbb{P}(\pi \subseteq \text{PA}_t). \end{aligned} \quad (4.47)$$

The proof of Proposition 4.11 reveals that we can take  $q = 2$  for  $t$  sufficiently large. Using (4.47) in (4.46), we obtain

$$\begin{aligned} \mathbb{P}(E_t) &\leq \frac{4q}{\mathbb{E}[M_{k_t^-}]^2} \sum_{v_1, v_2 \in [t] \setminus [t/2]} \sum_{H_1, H_2 \subseteq [t] \setminus [t/4]} \mathbb{P}(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2) \\ &\quad \times \left\{ \sum_{x \in \partial H_1, y \in \partial H_2} \sum_{\substack{\pi: x \rightarrow y \\ |\pi| \leq 2\bar{k}_t}} \mathbb{P}(\pi \subseteq \text{PA}_t) \right\}. \end{aligned} \quad (4.48)$$

If we bound  $\mathbb{P}(\pi \subseteq \text{PA}_t) \leq p(\pi)$  in (4.48), as in (4.36), the sum over  $\pi$  can be rewritten as the right hand side of (4.39) (recall (4.37)-(4.38)). We can thus apply Proposition 4.9 —because the proof of Proposition 4.9 really gives a bound on (4.39)— concluding that the sum over  $\pi$  is at

most  $p/(\log t)^2$ , where the constant  $p$  is defined in Proposition 4.9. Since  $|\partial H_1| = |\partial H_2| = m^{k_t^-} = (\log t)^{1-\varepsilon}$  (recall (4.12)), we finally obtain

$$\mathbb{P}(E_t) \leq \frac{4q}{\mathbb{E}[M_{k_t^-}]^2} \frac{p(\log t)^{2(1-\varepsilon)}}{(\log t)^2} \mathbb{E}[M_{k_t^-}^2] = (1 + o(1)) \frac{4pq}{(\log t)^{2\varepsilon}}, \quad (4.49)$$

where the last step uses Proposition 4.8. This completes the proof that  $\mathbb{P}(E_t) = o(1)$ .  $\square$

*Proof of Proposition 4.11.* We recall that  $H_1 \subseteq [t] \setminus [t/4]$  is a labeled directed subgraph containing  $v_1$ , such that it is an admissible realization of the neighborhood  $U_{\leq k_t^-}(v_1)$  of the minimally- $k_t^-$ -connected vertex  $v_1$  (recall Definition 4.2); in particular,  $H_1 \setminus \{v_1\} \subseteq [t/2] \setminus [t/4]$ . We also recall that, for all  $u \in H_1^o := H_1 \setminus \partial H_1$  and  $j \in [m]$ , the  $j$ -th edge of  $u$  is connected to a well specified vertex in  $H_1$ , denoted by  $\theta_j^{H_1}(u)$ . Analogous considerations apply to  $H_2$ .

We have to bound the probability

$$\mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2, \pi \subseteq \text{PA}_t, B_t\right), \quad (4.50)$$

where  $\pi = (\pi_0, \pi_1, \dots, \pi_k) \in [t]^{k+1}$  is a given sequence of vertices with  $\pi_0 \in \partial H_1$  and  $\pi_k \in \partial H_2$ . The event in (4.50) is not factorizable, because the degrees of the vertices in the path  $\pi$  are not specified, hence it is not easy to evaluate its probability. To get a factorizable event, we need to give more information. For a vertex  $v \in [t]$ , define its *incoming neighborhood*  $\mathcal{N}(v)$  by

$$\mathcal{N}(v) := \{(u, j) \in [t] \times [m] : u \xrightarrow{j} v\}. \quad (4.51)$$

The key observation is that *the knowledge of  $\mathcal{N}(v)$  determines the degree  $D_s(v)$  at any time  $s \leq t$*  (for instance, at time  $t$  we simply have  $D_t(v) = |\mathcal{N}(v)| + m$ ).

We are going to fix the incoming neighborhoods  $\mathcal{N}(\pi_1) = K_1, \dots, \mathcal{N}(\pi_{k-1}) = K_{k-1}$  of all vertices in the path  $\pi$ , except the extreme ones  $\pi_0$  and  $\pi_k$  (note that  $\mathcal{N}(\pi_0)$  and  $\mathcal{N}(\pi_k)$  reduce to single points in  $H_1^o$  and  $H_2^o$ , respectively, because  $\pi_0 \in \partial H_1$  and  $\pi_k \in \partial H_2$ ). We emphasize that such incoming neighborhoods allow us to determine whether  $\pi = (\pi_0, \dots, \pi_k)$  is a path in  $\text{PA}_t$ . Recalling the definition of the event  $B_t$  in (4.44), we restrict to

$$|K_i| \leq \sqrt{t}, \quad \text{for } i \in [k-1], \quad (4.52)$$

and simply drop  $B_t$  from (4.50). We will then prove the following relation: for all  $v_1, v_2, H_1, H_2, \pi = (\pi_0, \dots, \pi_k)$ , and for all  $K_1, \dots, K_{k-1}$  satisfying (4.52), we have

$$\begin{aligned} & \mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2, \{\mathcal{N}(\pi_1) = K_1, \dots, \mathcal{N}(\pi_{k-1}) = K_{k-1}\}\right) \\ & \leq q \mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2\right) \mathbb{P}(\mathcal{N}(\pi_1) = K_1, \dots, \mathcal{N}(\pi_{k-1}) = K_{k-1}). \end{aligned} \quad (4.53)$$

Our goal (4.47) follows by summing this relation over all  $K_1, \dots, K_{k-1}$  for which  $\pi \subseteq \text{PA}_t$ .

The first line of (4.53) is the probability of a factorizable event. In fact, setting for short

$$R := (H_1^o \times [m]) \cup (H_2^o \times [m]) \cup K_1 \cup \dots \cup K_{k-1},$$

the event in the first line of (4.53) is the intersection of the following four events (see (4.8)):

$$\begin{aligned} & \bigcap_{u \in H_1^o} \bigcap_{j=1}^m \{u \xrightarrow{j} \theta_j^{H_1}(u)\}, \quad \bigcap_{u \in H_2^o} \bigcap_{j=1}^m \{u \xrightarrow{j} \theta_j^{H_2}(u)\}, \quad \bigcap_{i=1}^{k-1} \bigcap_{(u,j) \in K_i} \{u \xrightarrow{j} \pi_i\}, \\ & \bigcap_{(u,j) \in [t] \times [m] \setminus R} \{u \not\xrightarrow{j} (H_1 \cup H_2 \cup \pi^o)\}, \end{aligned}$$

where we set  $\pi^o := \pi \setminus \{\pi_0, \pi_k\} = (\pi_1, \dots, \pi_{k-1})$ . Generalizing (4.9)-(4.10), we can rewrite the first line of (4.53) as follows, recalling (1.10):

$$\begin{aligned} & \mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2, \{\mathcal{N}(\pi_1) = K_1, \dots, \mathcal{N}(\pi_{k-1}) = K_{k-1}\}\right) \\ &= \left\{ \prod_{u \in H_1^o} \prod_{j=1}^m \frac{m+\delta}{c_{u,j}} \right\} \left\{ \prod_{u \in H_2^o} \prod_{j=1}^m \frac{m+\delta}{c_{u,j}} \right\} \left\{ \prod_{i=1}^{k-1} \prod_{(u,j) \in K_i} \frac{D_{u,j-1}(\pi_i) + \delta}{c_{u,j}} \right\} \\ & \quad \left\{ \prod_{(u,j) \in [t] \times [m] \setminus R} \left( 1 - \frac{D_{u,j-1}(H_1 \cup H_2 \cup \pi^o) + |(H_1 \cup H_2 \cup \pi^o) \cap [u-1]| \delta}{c_{u,j}} \right) \right\}. \end{aligned} \quad (4.54)$$

We stress that  $D_{u,j-1}(\pi_i)$  is *non-random*, because it is determined by  $K_i$ . Analogous considerations apply to  $D_{u,j-1}(H_1 \cup H_2 \cup \pi^o)$ . We have thus obtained a factorizable event.

Next we evaluate the second line of (4.53). Looking back at (4.26)-(4.28), we have

$$\begin{aligned} & \mathbb{P}\left(U_{\leq k_t^-}(v_1) = H_1, U_{\leq k_t^-}(v_2) = H_2\right) = \left\{ \prod_{u \in H_1^o} \prod_{j=1}^m \frac{m+\delta}{c_{u,j}} \right\} \left\{ \prod_{u \in H_2^o} \prod_{j=1}^m \frac{m+\delta}{c_{u,j}} \right\} \\ & \quad \left\{ \prod_{(u,j) \in [t] \times [m] \setminus (H_1^o \cup H_2^o) \times [m]} \left( 1 - \frac{D_{u,j-1}(H_1 \cup H_2) + |(H_1 \cup H_2) \cap [u-1]| \delta}{c_{u,j}} \right) \right\}. \end{aligned} \quad (4.55)$$

On the other hand,

$$\begin{aligned} & \mathbb{P}(\mathcal{N}(\pi_1) = K_1, \dots, \mathcal{N}(\pi_{k-1}) = K_{k-1}) = \left\{ \prod_{i=1}^{k-1} \prod_{(u,j) \in K_i} \frac{D_{u,j-1}(\pi_i) + \delta}{c_{u,j}} \right\} \\ & \quad \left\{ \prod_{(u,j) \in [t] \times [m] \setminus K_1 \cup \dots \cup K_{k-1}} \left( 1 - \frac{D_{u,j-1}(\pi^o) + |\pi^o \cap [u-1]| \delta}{c_{u,j}} \right) \right\}. \end{aligned} \quad (4.56)$$

Using the bound  $(1 - (a+b)) \leq (1-a)(1-b)$  in the second line of (4.54), and comparing with (4.55)-(4.56), we only need to take into account the missing terms in the product in the last lines. This shows that relation (4.53) holds if one sets  $q = C_1 C_2$  therein, where

$$\begin{aligned} C_1 &:= \left\{ \prod_{(u,j) \in K_1 \cup \dots \cup K_{k-1}} \left( 1 - \frac{D_{u,j-1}(H_1 \cup H_2) + |(H_1 \cup H_2) \cap [u-1]| \delta}{c_{u,j}} \right) \right\}^{-1}, \\ C_2 &:= \left\{ \prod_{(u,j) \in (H_1^o \cup H_2^o) \times [m]} \left( 1 - \frac{D_{u,j-1}(\pi^o) + |\pi^o \cap [u-1]| \delta}{c_{u,j}} \right) \right\}^{-1}. \end{aligned}$$

To complete the proof, it is enough to give uniform upper bounds on  $C_1$  and  $C_2$ , that does not depend on  $H_1$ ,  $H_2$ ,  $\pi$ . We start with  $C_1$ . In the product we may assume  $u > t/4$ , because the terms with  $u \leq t/4$  are identically one, since  $H_1, H_2 \subseteq [t] \setminus [t/4]$ . Moreover, for  $u > t/4$  we have  $c_{u,j} \geq t(2m+\delta)/4 \geq mt/4$  by (1.11) and  $\delta > -m$ . Since  $D_{u,j-1}(H_1 \cup H_2) \leq 2(m+1)i_k$ , using  $1-x \geq e^{-2x}$  for  $x$  small and recalling that  $\delta < 0$ , it follows that

$$C_1^{-1} \geq \prod_{(u,j) \in K_1 \cup \dots \cup K_{k-1}} \left( 1 - \frac{2(m+1)i_k}{\frac{mt}{4}} \right) \geq e^{-\frac{8(m+1)}{tm}|K_{[k-1]}|i_k}, \quad (4.57)$$

where  $K_{[k-1]} = K_1 \cup \dots \cup K_{k-1}$ . Since  $i_k$  is given by (4.4), for  $k = k_t^-$  as in (4.12) we have  $i_k = \frac{m}{m-1} m^{k_t^-} (1 + o(1)) \leq 2(\log t)^{1-\varepsilon}$ . Recalling also (4.52) and bounding  $m+1 \leq 2m$ , we obtain

$$C_1 \leq e^{\frac{8(m+1)}{tm}|K_{[k-1]}|i_k} \leq e^{16k i_k / \sqrt{t}} = e^{O(\log t / \sqrt{t})} = 1 + o(1).$$

For  $C_2$ , since  $D_{u,j-1}(\pi^o) \leq D_t(\pi^o) = |K_{[k-1]}| \leq k\sqrt{t}$ , again by (4.52), we get

$$C_2^{-1} \geq \prod_{(u,j) \in (H_1^o \cup H_2^o) \times [m]} \left(1 - \frac{k\sqrt{t}}{\frac{m}{4}t}\right) \geq e^{-\frac{8}{m} \frac{k}{\sqrt{t}} |H_1^o \cup H_2^o| m} \geq e^{-16kik/\sqrt{t}} = 1 - o(1). \quad (4.58)$$

It follows that  $C_1 C_2$  is bounded from above by some constant  $q$ . This completes the proof.  $\square$

**4.5. Proof of Theorem 1.6.** Dereich, Möch and Mörters [DMM12] have already proved the upper bound. For the lower bound we use Proposition 4.9. In fact, for  $\bar{k}_t$  as in (4.33),

$$\mathbb{P}(H_t \leq 2\bar{k}_t) = \sum_{v_1, v_2 \in [t]} \mathbb{P}(V_1 = v_1, V_2 = v_2, \text{dist}(v_1, v_2) \leq 2\bar{k}_t). \quad (4.59)$$

If  $v_1$  and  $v_2$  are both larger or equal than  $g_0 = \lceil \frac{t}{(\log t)^2} \rceil$ , then we can apply Proposition 4.9. The probability that  $V_1 < g_0$  or  $V_2 < g_0$  is

$$\mathbb{P}(\{V_1 < g_0\} \cup \{V_2 < g_0\}) \leq 2g_0/t = o(1), \quad (4.60)$$

hence we get

$$\frac{1}{t^2} \sum_{v_1, v_2 \in [t] \setminus [g_0]} \mathbb{P}(\text{dist}(v_1, v_2) \leq 2\bar{k}_t) + o(1) \leq \frac{(t-g_0)^2}{t^2} \frac{p}{(\log t)^2} + o(1) = o(1),$$

and this completes the proof of Theorem 1.6.  $\square$

## 5. UPPER BOUND FOR CONFIGURATION MODEL

In this section we prove Statements 2.5 and 2.6 for the configuration model. By the discussion in Section 2.2, this completes the proof of the upper bound in Theorem 1.3, because the proof of Statement 2.4 is already known in the literature, as explained below Statement 2.4.

Throughout this section, the assumptions of Theorem 1.3 apply. In particular, we work on a configuration model  $\text{CM}_n$ , with  $\tau \in (2, 3)$  and  $d_{\min} \geq 3$ .

**5.1. Proof of Statement 2.5.** We first recall what  $\text{Core}_n$  is, and define the  $k$ -exploration graph.

Recall from (2.8) that, for  $\text{CM}_n$ ,  $\text{Core}_n$  is defined as

$$\text{Core}_n = \{i \in [n] \text{ such that } d_i > (\log n)^\sigma\},$$

where  $\sigma > 1/(3-\tau)$ . Since the degrees  $d_i$  are fixed in the configuration model,  $\text{Core}_n$  is a deterministic subset.

For any  $v \in [n]$ , we recall that  $U_{\leq k}(v) \subseteq [n]$  denotes the subgraph of  $\text{CM}_n$  consisting of the vertices at distance at most  $k$  from  $v$ . We next consider the  $k$ -exploration graph  $\widehat{U}_{\leq k}(v)$  as a modification of  $U_{\leq k}(v)$ , where we only explore  $d_{\min}$  half-edges of the starting vertex  $v$ , and only  $d_{\min} - 1$  for the following vertices:

**Definition 5.1** ( $k$ -exploration graph in  $\text{CM}_n$ ). *The  $k$ -exploration graph of a vertex  $v$  is the subgraph  $\widehat{U}_{\leq k}(v)$  built iteratively as follows:*

- ▷ Starting from  $\widehat{U}_{\leq 0}(v) = \{v\}$ , we consider the first  $d_{\min}$  half-edges of  $v$  and we pair them, one by one, to a uniformly chosen unpaired half-edge (see Remark 3.1), to obtain  $\widehat{U}_{\leq 1}(v)$ .
- ▷ Assume that we have built  $\widehat{U}_{\leq \ell}(v)$ , for  $\ell \geq 1$ , and set  $\widehat{U}_{=\ell}(v) := \widehat{U}_{\leq \ell}(v) \setminus \widehat{U}_{\leq (\ell-1)}(v)$ . For each vertex in  $\widehat{U}_{=\ell}(v)$ , we consider the first  $d_{\min} - 1$  unpaired half-edges and we pair them, one by one, to a uniformly chosen unpaired half-edge, to obtain  $\widehat{U}_{\leq (\ell+1)}(v)$ . (Note that, by construction, each vertex in  $\widehat{U}_{=\ell}(v)$  has at least one already paired half-edge.)

**Definition 5.2** (Collision). *In the process of building the  $k$ -exploration graph  $\widehat{U}_{\leq k}(v)$ , we say that there is a collision when a half-edge is paired to a vertex already included in the  $k$ -exploration graph.*

We now prove Statement 2.5. Let us fix  $\varepsilon > 0$  and set

$$k_n^+ = (1 + \varepsilon) \frac{\log \log n}{\log(d_{\min} - 1)}. \quad (5.1)$$

**Proposition 5.3** (At most one collision). *Under the assumption of Theorem 1.3, the following holds with high probability: the  $k_n^+$ -exploration graph of every vertex either intersects  $\text{Core}_n$ , or it has at most one collision.*

*Proof.* Let us fix a vertex  $v \in [n]$ . We are going to estimate the probability

$$q_n(v) := \mathbb{P}\left(\text{there are at least 2 collisions in } \hat{U}_{\leq k_n^+}(v) \text{ and } \hat{U}_{\leq k_n^+}(v) \cap \text{Core}_n = \emptyset\right).$$

If we show that  $\sup_{v \in [n]} q_n(v) = o(1/n)$ , then it follows that  $\sum_{v \in [n]} q_n(v) = o(1)$ , completing the proof.

Starting from the vertex  $v$ , we pair successively one half-edge after the other, as described in Definition 5.1 (recall also Remark 3.1). In order to build  $\hat{U}_{\leq k_n^+}(v)$ , we need to make a number of pairings, denoted by  $\mathcal{N}$ , which is *random*, because collisions may occur. In fact, when there are no collisions,  $\mathcal{N}$  is deterministic and takes its maximal value given by  $i_{k_n^+}$  in (3.1), therefore

$$\mathcal{N} \leq i_{k_n^+} \leq \frac{d_{\min}}{d_{\min} - 2} (d_{\min} - 1)^{k_n^+} \leq 3 (\log n)^{1+\varepsilon}. \quad (5.2)$$

Introducing the event  $C_i := \text{"there is a collision when pairing the } i\text{-th half-edge"}$ , we can write

$$\begin{aligned} q_n(v) &\leq \mathbb{E}\left[\sum_{1 \leq i < j \leq \mathcal{N}} \mathbb{1}_{\{C_i, C_j, \hat{U}_{\leq k_n^+}(v) \cap \text{Core}_n = \emptyset\}}\right] \\ &= \sum_{1 \leq i < j \leq 3(\log n)^{1+\varepsilon}} \mathbb{P}(C_i, C_j, j \leq \mathcal{N}, \hat{U}_{\leq k_n^+}(v) \cap \text{Core}_n = \emptyset). \end{aligned} \quad (5.3)$$

Let  $E_\ell$  be the event that the first  $\ell$  half-edges are paired to vertices with degree  $\leq (\log n)^\sigma$  (i.e., the graph obtained after pairing the first  $\ell$  half-edges is disjoint from  $\text{Core}_n$ ). Then

$$\begin{aligned} \mathbb{P}(C_i, C_j, j \leq \mathcal{N}, \hat{U}_{\leq k_n^+}(v) \cap \text{Core}_n = \emptyset) &\leq \mathbb{P}(C_i, C_j, E_{j-1}) \\ &= \mathbb{P}(E_{i-1}) \mathbb{P}(C_i | E_{i-1}) \mathbb{P}(C_j | C_i, E_{j-1}). \end{aligned} \quad (5.4)$$

On the event  $E_{i-1}$ , before pairing the  $i$ -th half-edge, the graph is composed by at most  $i-1$  vertices, each with degree at most  $(\log n)^\sigma$ , hence, for  $i \leq 3(\log n)^{1+\varepsilon}$ ,

$$\mathbb{P}(C_i | E_{i-1}) \leq \frac{(i-1)(\log n)^\sigma}{\ell_n - 2i + 1} \leq \frac{3(\log n)^{1+\varepsilon}(\log n)^\sigma}{\ell_n - 6(\log n)^{1+\varepsilon}} \leq c \frac{(\log n)^{\sigma+1+\varepsilon}}{n},$$

for some  $c \in (0, \infty)$ , thanks to  $\ell_n = n\mu(1 + o(1))$  (recall (3.5)). The same arguments show that

$$\mathbb{P}(C_j | C_i, E_{j-1}) \leq c \frac{(\log n)^{\sigma+1+\varepsilon}}{n}.$$

Looking back at (5.3)-(5.4), we obtain

$$\sup_{v \in [n]} q_n(v) \leq \sum_{1 \leq i < j \leq 3(\log n)^{1+\varepsilon}} c^2 \frac{(\log n)^{2(\sigma+1+\varepsilon)}}{n^2} \leq 9c^2 \frac{(\log n)^{2\sigma+4(1+\varepsilon)}}{n^2} = o\left(\frac{1}{n}\right),$$

which completes the proof.  $\square$

**Corollary 5.4** (Large boundaries). *Under the assumptions of Theorem 1.3 and on the event  $\hat{U}_{\leq k_n^+}(v) \cap \text{Core}_n = \emptyset$ , with high probability, the boundary  $\hat{U}_{\leq k_n^+}(v)$  of the  $k_n^+$ -exploration graph of any vertex  $v \in [n]$  contains at least  $(d_{\min} - 2)(d_{\min} - 1)^{k_n^+-1} \geq \frac{1}{2}(\log n)^{1+\varepsilon}$  vertices, each one with at least two unpaired half-edges.*

*Proof.* By Proposition 5.3, with high probability, every  $k_n^+$ -exploration graph has at most one collision before hitting  $\text{Core}_n$ . The worst case is when the collision happens immediately, i.e. a half-edge incident to  $v$  is paired to another half-edge incident to  $v$ : in this case, removing both half-edges, the  $k_n^+$ -exploration graph becomes a tree with  $(d_{\min} - 2)(d_{\min} - 1)^{k_n^+-1}$  vertices on its boundary, each of which has at least  $(d_{\min} - 1) \geq 2$  yet unpaired half-edges. Since  $(d_{\min} - 2)/(d_{\min} - 1) \geq \frac{1}{2}$  for  $d_{\min} \geq 3$ , and moreover  $(d_{\min} - 1)^{k_n^+} = (\log n)^{1+\varepsilon}$  by (5.1), we obtain the claimed bound.

If the collision happens at a later stage, i.e. for a half-edge incident to a vertex different from the starting vertex  $v$ , then we just remove the branch from  $v$  to that vertex, getting a tree with  $(d_{\min} - 1)(d_{\min} - 1)^{k_n^+-1}$  vertices on its boundary. The conclusion follows.  $\square$

Together, Proposition 5.3 and Corollary 5.4 prove Statement 2.5.  $\square$

**5.2. Proof of Statement 2.6.** Consider the  $k_n^+$ -exploration graph  $\widehat{U} = \widehat{U}_{\leq k_n^+}(v)$  of a fixed vertex  $v \in [n]$ , as in Definition 5.1, and let  $x_1, \dots, x_N$  be the (random) vertices on its boundary. We stress that, by Corollary 5.4, with high probability  $N \geq \frac{1}{2}(\log n)^{1+\varepsilon}$ . Set

$$h_n = \lceil B \log \log \log n + C \rceil, \quad (5.5)$$

where  $B, C$  are fixed constants, to be determined later on.

Henceforth we fix a realization  $H$  of  $\widehat{U} = \widehat{U}_{\leq k_n^+}(v)$  and we work *conditionally* on the event  $\{\widehat{U} = H\}$ . By Remark 3.1, we can complete the construction of the configuration model  $\text{CM}_n$  by pairing uniformly all the yet unpaired half-edges. We do this as follows: for each vertex  $x_1, \dots, x_N$  on the boundary of  $\widehat{U}$ , we explore its neighborhood, looking for *fresh* vertices with higher and higher degree, up to distance  $h_n$  (we call a vertex *fresh* if it is connected to the graph for the first time, hence it only has one paired half-edge). We now describe this procedure in detail:

**Definition 5.5** (Exploration procedure). *Let  $x_1, \dots, x_N$  denote the vertices on the boundary of a  $k_n^+$ -exploration graph  $\widehat{U} = \widehat{U}_{\leq k_n^+}(v)$ . We start the exploration procedure from  $x_1$ .*

- ▷ *Step 1. We set  $v_0^{(1)} := x_1$  and we pair all its unpaired half-edges. Among the fresh vertices to which  $v_0^{(1)}$  has been connected, we call  $v_1$  the one with maximal degree.*
- ▷ *When there are no fresh vertices at some step, the procedure for  $x_1$  stops.*
- ▷ *Step 2. Assuming we have built  $v_1^{(1)}$ , we pair all its unpaired half-edges: among the fresh connected vertices, we denote by  $v_2^{(1)}$  the vertex with maximal degree.*
- ▷ *We continue in this way for (at most)  $h_n$  steps, defining  $v_j^{(1)}$  for  $0 \leq j \leq h_n$  (recall (5.5)).*

*After finishing the procedure for  $x_1$ , we perform the same procedure for  $x_2, x_3, \dots, x_N$ , defining the vertices  $v_0^{(i)}, v_1^{(i)}, \dots, v_{h_n}^{(i)}$  starting from  $v_0^{(i)} = x_i$ .*

**Definition 5.6** (Success). *Let  $x_1, \dots, x_N$  be the vertices on the boundary of a  $k_n^+$ -exploration graph  $\widehat{U} = \widehat{U}_{\leq k_n^+}(v)$ . We define the event  $S_{x_i} := \text{"}x_i \text{ is a success"}$  by*

$$S_{x_i} := \{\{v_0^{(i)}, v_1^{(i)}, \dots, v_{h_n}^{(i)}\} \cap \text{Core}_n \neq \emptyset\} = \{d_{v_j^{(i)}} > (\log n)^\sigma \text{ for some } 0 \leq j \leq h_n\}.$$

Here is the key result, proved below:

**Proposition 5.7** (Hitting the core quickly). *There exists a constant  $\eta > 0$  such that, for every  $n \in \mathbb{N}$  and for every realization  $H$  of  $\widehat{U}$ ,*

$$\mathbb{P}(S_{x_1} \mid \widehat{U} = H) \geq \eta, \quad (5.6)$$

and, for each  $i = 2, \dots, N$ ,

$$\mathbb{P}(S_{x_i} \mid \widehat{U} = H, S_{x_1}^c, \dots, S_{x_{i-1}}^c) \geq \eta. \quad (5.7)$$

This directly leads to the proof of Statement 2.6, as the following corollary shows:

**Corollary 5.8** (Distance between periphery and  $\text{Core}_n$ ). *Under the hypotheses of Theorem 1.3, with high probability, the distance of every vertex in the graph from  $\text{Core}_n$  is at most*

$$(1 + \varepsilon) \frac{\log \log n}{\log(d_{\min} - 1)} + o(\log \log n). \quad (5.8)$$

*Proof.* By Corollary 5.4, with high probability, every vertex  $v \in [n]$  either is at distance at most  $k_n^+$  from  $\text{Core}_n$ , or has a  $k_n^+$ -exploration graph  $\hat{U} = \hat{U}_{\leq k_n^+}(v)$  with at least  $N \geq \frac{1}{2}(\log n)^{1+\varepsilon}$  vertices on its boundary. It suffices to consider the latter case. Conditionally on  $\hat{U} = H$ , the probability that none of these vertices is a success can be bounded by Proposition 5.7:

$$\begin{aligned} \mathbb{P}(S_{x_1}^c \cap \dots \cap S_{x_N}^c \mid \hat{U} = H) &= \mathbb{P}(S_{x_1}^c \mid \hat{U} = H) \prod_{j=2}^N \mathbb{P}(S_{x_j}^c \mid \hat{U} = H, S_{x_1}^c, \dots, S_{x_{j-1}}^c) \\ &\leq (1 - \eta)^N \leq (1 - \eta)^{\frac{1}{2}(\log n)^{1+\varepsilon}} = o(1/n). \end{aligned} \quad (5.9)$$

This is uniform over  $H$ , hence the probability that no vertex is a success, without conditioning, is still  $o(1/n)$ . It follows that, with high probability, every  $v \in [n]$  has at least one successful vertex on the boundary of its  $k_n^+$ -exploration graph. This means that the distance of every vertex  $v \in [n]$  from  $\text{Core}_n$  is at most  $k_n^+ + h_n = k_n^+ + o(\log \log n)$ , by (5.5). Recalling (5.1), we have completed the proof of Corollary 5.8 and thus of Statement 2.6.  $\square$

To prove Proposition 5.7, we need the following technical (but simple) result:

**Lemma 5.9** (High-degree fresh vertices). *Consider the process of building a configuration model  $\text{CM}_n$  as described in Remark 3.1. Let  $\mathcal{G}_l$  be the random graph obtained after  $l$  pairings of half-edges and let  $V_l$  be the random vertex incident to the half-edge to which the  $l$ -th half-edge is paired. For all  $l, n \in \mathbb{N}$  and  $z \in [0, \infty)$  such that*

$$l \leq \frac{n}{4}(1 - F_{d,n}(z)), \quad (5.10)$$

*the following holds:*

$$\mathbb{P}(d_{V_{l+1}} > z, V_{l+1} \notin \mathcal{G}_l \mid \mathcal{G}_l) \geq z[1 - F_{d,n}(z)] \frac{n}{2\ell_n}. \quad (5.11)$$

*In particular, when Conditions 1.1 and 1.2 hold, for every  $\zeta > 0$  there are  $c > 0$ ,  $n_0 < \infty$  such that*

$$\forall n \geq n_0, \quad 0 \leq z \leq n^{1/3}, \quad l \leq n^{1/3} : \quad \mathbb{P}(d_{V_{l+1}} > z, V_{l+1} \notin \mathcal{G}_l \mid \mathcal{G}_l) \geq \frac{c}{z^{\tau-2+\zeta}}. \quad (5.12)$$

*Proof.* By definition of  $\text{CM}_n$ , the  $(l+1)$ -st half-edge is paired to a uniformly chosen half-edge among the  $\ell_n - 2l - 1$  that are not yet paired. Consequently

$$\mathbb{P}(d_{V_{l+1}} > z, V_{l+1} \notin \mathcal{G}_l \mid \mathcal{G}_l) = \frac{1}{\ell_n - 2l - 1} \sum_{v \notin \mathcal{G}_l} d_v \mathbb{1}_{\{d_v > z\}}. \quad (5.13)$$

Since  $|\mathcal{G}_l| \leq 2l \leq \frac{n}{2}(1 - F_{d,n}(z))$  by (5.10), we obtain

$$\frac{1}{\ell_n - 2l - 1} \sum_{v \notin \mathcal{G}_l} d_v \mathbb{1}_{\{d_v > z\}} \geq \frac{z}{\ell_n} (n(1 - F_{d,n}(z)) - |\mathcal{G}_l|) \geq z(1 - F_{d,n}(z)) \frac{n}{2\ell_n}, \quad (5.14)$$

which proves (5.11).

Assuming Conditions 1.1 and 1.2, we have  $\ell_n = \mu n(1 + o(1))$ , with  $\mu \in (0, \infty)$ , see (3.5), and there are  $c_1 > 0$  and  $\alpha > 1/2$  such that  $1 - F_{d,n}(z) \geq c_1 z^{-(\tau-1)}$  for  $0 \leq z \leq n^\alpha$ . Consequently, for  $0 \leq z \leq n^{1/3}$ , the right hand side of (5.10) is at least  $\frac{n}{4} \frac{c_1}{n^{(\tau-1)/3}}$ . Note that  $(\tau-1)/3 < 2/3$  (because  $\tau < 3$ ), hence we can choose  $n_0$  so that  $\frac{n}{4} \frac{c_1}{n^{(\tau-1)/3}} \geq n^{1/3}$  for all  $n \geq n_0$ . This directly leads to (5.12).  $\square$

With Lemma 5.9 in hand, we are able to prove Proposition 5.7:

*Proof of Proposition 5.7.* We fix  $v \in [n]$  and a realization  $H$  of  $\widehat{U} = \widehat{U}_{\leq k_n^+}(v)$ . We abbreviate

$$\mathbb{P}^*(\cdot) := \mathbb{P}(\cdot | \widehat{U} = H). \quad (5.15)$$

The vertices on the boundary of  $\widehat{U}$  are denoted by  $x_1, \dots, x_N$ . We start proving (5.6), hence we focus on  $x_1$  and we define  $v_0^{(1)}, v_1^{(1)}, \dots, v_{h_n}^{(1)}$  as in Definition 5.5, with  $v_0^{(1)} = x_1$ .

We first fix some parameters. Since  $2 < \tau < 3$ , we can choose  $\zeta, \gamma > 0$  small enough so that

$$\xi := 1 - e^\gamma(\tau - 2 + \zeta) > 0. \quad (5.16)$$

Next we define a sequence  $(g_\ell)_{\ell \in \mathbb{N}_0}$  that grows *doubly exponentially* fast:

$$g_\ell := 2^{e^{\gamma\ell}} = \exp((\log 2) \exp(\gamma \ell)). \quad (5.17)$$

Then we fix  $B = 1/\gamma$  and  $C = \log(\sigma/\log 2)$  in (5.5), where  $\sigma$  is the same constant as in  $\text{Core}_n$ , see (2.8). With these choices, we have

$$g_{h_n} = e^{\sigma e^{\lceil \log \log \log n \rceil}} > e^{\sigma \log \log n} = (\log n)^\sigma, \quad \text{while} \quad g_{h_n-1} < (\log n)^\sigma. \quad (5.18)$$

Roughly speaking, the idea is to show that, with positive probability, one has  $d_{v_j^{(1)}} > g_j$ . As a consequence,  $d_{v_{h_n}^{(1)}} > g_{h_n} \geq (\log n)^\sigma$ , that is  $v_{h_n}^{(1)}$  belongs to  $\text{Core}_n$  and  $x_1$  is a success. The situation is actually more involved, since we can only show that  $d_{v_j^{(1)}} > g_j$  before reaching  $\text{Core}_n$ .

Let us make the above intuition precise. Recalling (5.15), let us set

$$H_{-1} := \emptyset, \quad H_0 := H, \quad H_k := H \cup \{v_1^{(1)}, \dots, v_k^{(1)}\} \quad \text{for } 1 \leq k \leq h_n.$$

Then we introduce the events

$$T_\ell := \bigcup_{k=0}^{\ell} \{d_{v_k^{(1)}} > (\log n)^\sigma\}, \quad W_\ell := \bigcap_{k=0}^{\ell} \{d_{v_k^{(1)}} > g_k, v_k^{(1)} \notin H_{k-1}\}. \quad (5.19)$$

In words, the event  $T_\ell$  means that one of the vertices  $v_0^{(1)}, \dots, v_\ell^{(1)}$  has already reached  $\text{Core}_n$ , while the event  $W_\ell$  means that the degrees of vertices  $v_0^{(1)}, \dots, v_\ell^{(1)}$  grow at least like  $g_0, \dots, g_\ell$  and, furthermore, each  $v_k$  is a fresh vertex (this is actually already implied by Definition 5.5, otherwise  $v_k$  would not even be defined). We finally set

$$E_0 := W_0, \quad E_j := T_{j-1} \cup W_j \quad \text{for } 1 \leq j \leq h_n.$$

Note that  $T_{h_n}$  coincides with  $S_{x_1}$  = “ $x_1$  is a success”. Also note that  $W_{h_n} \subseteq \{d_{v_{h_n}^{(1)}} > (\log n)^\sigma\}$ , because  $d_{v_{h_n}^{(1)}} > g_{h_n} > (\log n)^\sigma$  by (5.18), hence

$$E_{h_n} = T_{h_n-1} \cup W_{h_n} \subseteq T_{h_n-1} \cup \{d_{v_{h_n}^{(1)}} > (\log n)^\sigma\} = T_{h_n} = S_{x_1}.$$

Consequently, if we prove that  $\mathbb{P}^*(E_{h_n}) \geq \eta$ , then our goal  $\mathbb{P}^*(S_{x_1}) \geq \eta$  follows (recall (5.6)).

The reason for working with the events  $E_j$  is that their probabilities can be controlled by an induction argument. Recalling (5.15), we can write

$$\begin{aligned} \mathbb{P}^*(E_{j+1}) &= \mathbb{P}^*(T_j) + \mathbb{P}^*(T_j^c \cap W_{j+1}) \\ &= \mathbb{P}^*(T_j) + \mathbb{P}(d_{v_{j+1}^{(1)}} > g_{j+1}, v_{j+1}^{(1)} \notin H_j | \{\widehat{U} = H\} \cap T_j^c \cap W_j) \mathbb{P}^*(T_j^c \cap W_j). \end{aligned} \quad (5.20)$$

The key point is the following estimate on the conditional probability, proved below:

$$\mathbb{P}(d_{v_{j+1}^{(1)}} > g_{j+1}, v_{j+1}^{(1)} \notin H_j | \{\widehat{U} = H\} \cap T_j^c \cap W_j) \geq 1 - \varepsilon_j, \quad \text{where } \varepsilon_j := e^{-c(g_j)\xi/2}, \quad (5.21)$$

with  $\xi > 0$  is defined in (5.16) and  $c > 0$  is the constant appearing in relation (5.12). This yields

$$\begin{aligned} \mathbb{P}^*(E_{j+1}) &\geq \mathbb{P}^*(T_j) + (1 - \varepsilon_j) \mathbb{P}^*(T_j^c \cap W_j) \geq (1 - \varepsilon_j)(\mathbb{P}^*(T_j) + \mathbb{P}^*(T_j^c \cap W_j)) \\ &= (1 - \varepsilon_j) \mathbb{P}^*(T_j \cup W_j) \geq (1 - \varepsilon_j) \mathbb{P}^*(T_{j-1} \cup W_j) \\ &= (1 - \varepsilon_j) \mathbb{P}^*(E_j), \end{aligned}$$

which leads us to

$$\mathbb{P}^*(E_{h_n}) \geq \mathbb{P}^*(E_0) \prod_{j=0}^{h_n-1} (1 - \varepsilon_j) \geq \mathbb{P}^*(E_0) \prod_{j=0}^{\infty} (1 - \varepsilon_j) =: \eta.$$

Since  $\sum_{j \geq 0} \varepsilon_j < \infty$  and  $\varepsilon_j < 1$  for every  $j \geq 0$ , by (5.21) and (5.17), the infinite product is strictly positive. Also note that  $\mathbb{P}^*(E_0) = \mathbb{P}^*(d_{v_0^{(1)}} \geq 2) = 1$ , because  $g_0 = 2$  and  $d_{v_0^{(1)}} \geq d_{\min} \geq 3$ . Then  $\eta > 0$ , as required.

It remains to prove (5.21). To lighten notation, we rewrite the left hand side of (5.21) as

$$q_{j+1} := \mathbb{P}(d_{v_{j+1}^{(1)}} > g_{j+1}, v_{j+1}^{(1)} \notin H_j \mid D_j), \quad \text{where } D_j := \{\widehat{U} = H\} \cap T_j^c \cap W_j. \quad (5.22)$$

Note that, on the event  $D_j \subseteq W_j$ , vertex  $v_j^{(1)}$  is fresh (i.e., it is connected to the graph for the first time), hence it has  $m = d_{v_j^{(1)}} - 1$  unpaired half-edges. These are paired uniformly, connecting  $v_j^{(1)}$  to (not necessarily distinct) vertices  $w^{(1)}, \dots, w^{(m)}$ . Let us introduce for  $1 \leq \ell \leq m$  the event

$$C_\ell := \bigcap_{k=1}^{\ell} \{d_{w^{(k)}} > g_{j+1}, w^{(k)} \notin H_j\}^c. \quad (5.23)$$

By Definition 5.5,  $v_{j+1}^{(1)}$  is the *fresh* vertex with maximal degree among them, hence

$$\{d_{v_{j+1}^{(1)}} > g_{j+1}, v_{j+1}^{(1)} \notin H_j\}^c = C_m.$$

Since  $m = d_{v_j^{(1)}} - 1 > g_j - 1$  on  $W_j \subseteq D_j$ , the left hand side of (5.21) can be estimated by

$$\begin{aligned} q_{j+1} &= 1 - \mathbb{P}(C_m \mid D_j) \geq 1 - \prod_{k=1}^{g_j-1} \mathbb{P}(C_k \mid D_j \cap C_{k-1}) \\ &= 1 - \prod_{k=1}^{g_j-1} \left(1 - \mathbb{P}(d_{w^{(k)}} > g_{j+1}, w^{(k)} \notin H_j \mid D_j \cap C_{k-1})\right). \end{aligned} \quad (5.24)$$

We claim that we can apply relation (5.12) from Lemma 5.9 to each of the probabilities in the last line of (5.24). To justify this claim, we need to look at the conditioning event  $D_j \cap C_{k-1}$ , recalling (5.23), (5.22) and (5.19). In order to produce it, we have to do the following:

- ▷ First we build the  $k_n^+$ -exploration graph  $\widehat{U}_{\leq k_n^+}(v) = H$ , which requires to pair at most  $O((d_{\min} - 1)^{k_n^+}) = O((\log n)^{1+\varepsilon})$  half-edges (recall Definition 5.1);
- ▷ Next, starting from the boundary vertex  $x_1$ , we generate the fresh vertices  $v_0^{(1)}, \dots, v_j^{(1)}$  *all outside*  $\text{Core}_n$ , because we are on the event  $T_j^c$ , and this requires to pair a number of half-edges which is at most  $(\log n)^\sigma j \leq (\log n)^\sigma h_n = O((\log n)^{\sigma+1})$ ;
- ▷ Finally, in order to generate  $w^{(1)}, \dots, w^{(k-1)}$ , we pair exactly  $k - 1$  half-edges, and note that  $k - 1 \leq g_j - 1 \leq g_{h_n} - 1 = O((\log n)^\sigma)$  (always because  $v_j \notin \text{Core}_n$ ).

It follows that the conditioning event  $D_j \cap C_{k-1}$  is in the  $\sigma$ -algebra generated by  $\mathcal{G}_l$  for  $l \leq O((\log n)^{1+\sigma+\varepsilon})$  (we use the notation of Lemma 5.9). In particular,  $l \leq n^{1/3}$ . Also note that  $z = g_{j+1} \leq g_{h_n} = O((\log n)^\sigma)$ , see (5.18), hence also  $z \leq n^{1/3}$ . Applying (5.12), we get

$$\begin{aligned} q_{j+1} &\geq 1 - \left(1 - \frac{c}{(g_{j+1})^{\tau-2+\zeta}}\right)^{g_j-1} \geq 1 - \exp\left(-c \frac{g_j - 1}{(g_{j+1})^{\tau-2+\zeta}}\right) \\ &\geq 1 - \exp\left(-\frac{c}{2} \frac{g_j}{(g_{j+1})^{\tau-2+\zeta}}\right) \end{aligned} \quad (5.25)$$

because  $1 - x \leq e^{-x}$  and  $n - 1 \geq n/2$  for all  $n \geq 2$  (note that  $g_j \geq g_0 = 2$ ). Since  $g_{j+1} = (g_j)^{e^\gamma}$ , by (5.17), we finally arrive at

$$q_{j+1} \geq 1 - \exp\left(-\frac{c}{2}(g_j)^{1-e^\gamma(\tau-2+\zeta)}\right) = 1 - e^{-c(g_j)^{\xi/2}}, \quad (5.26)$$

which is precisely (5.21). This completes the proof of (5.6).

In order to prove (5.7), we proceed in the same way: for any fixed  $2 \leq i \leq N$ , we start from the modification of (5.15) given by  $\mathbb{P}^*(\cdot) := \mathbb{P}(\cdot | \widehat{U} = H, S_{x_1}^c, \dots, S_{x_{i-1}}^c)$  and we follow the same proof, working with the vertices  $v_1^{(i)}, \dots, v_{h_n}^{(i)}$  instead of  $v_1^{(1)}, \dots, v_{h_n}^{(1)}$  (recall Definition 5.5). We leave the details to the reader.  $\square$

## 6. UPPER BOUND FOR PREFERENTIAL ATTACHMENT MODEL

In this section we prove Statements 2.5 and 2.6 for the preferential attachment model. By the discussion in Section 2.2, this completes the proof of the upper bound in Theorem 1.5, because the proof of Statement 2.4 is already known in the literature, as explained below Statement 2.4.

**6.1. Proof of Statement 2.5.** Recall the definition of  $\text{Core}_t$  in (2.8). It is crucial that in  $\text{Core}_t$ , we let  $D_{t/2}(v)$  be large. We again continue to define what a  $k$ -exploration graph and its collisions are, but this time for the preferential attachment model:

**Definition 6.1** ( $k$ -exploration graph). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model. For  $v \in [t]$ , we call the  $k$ -exploration graph of  $v$  to be the subgraph of  $\text{PA}_t$ , where we consider the  $m$  edges originally incident to  $v$ , and the  $m$  edges originally incident to any other vertex that is connected to  $v$  in this procedure, up to distance  $k$  from  $v$ .*

**Definition 6.2** (Collision). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model with  $m \geq 2$ , and let  $v$  be a vertex. We say that we have a collision in the  $k$ -exploration graph of  $v$  when one of the  $m$  edges of a vertex in the  $k$ -exploration graph of  $v$  is connected to a vertex that is already in the  $k$ -exploration graph of  $v$ .*

Now we want to show that every  $k$ -exploration graph has at most a finite number of collisions before hitting the  $\text{Core}_t$ , as we did for the configuration model. The first step is to use [DHH10, Lemma 3.9]:

**Lemma 6.3** (Early vertices have large degree). *Fix  $m \geq 1$ . There exists  $a > 0$  such that*

$$\mathbb{P}\left(\min_{i \leq t^a} D_t(i) \geq (\log t)^\sigma\right) \rightarrow 1 \quad (6.1)$$

for some  $\sigma > 1/(3 - \tau)$ . As consequence,  $[t^a] \subseteq \text{Core}_t$  with high probability.

In agreement with (2.10) (see also (4.12)), we set

$$k_t^+ = (1 + \varepsilon) \frac{\log \log t}{\log m}. \quad (6.2)$$

We want to prove that the exploration graph  $\widehat{U}_{\leq k_t^+}(v)$  has at most a finite number of collisions before hitting  $\text{Core}_t$ , similarly to the case of  $\text{CM}_n$ , now for  $\text{PA}_t$ . As it is possible to see from (2.8),  $\text{Core}_t \subseteq [t/2]$ , i.e., is a subset defined in  $\text{PA}_t$  when the graph has size  $t/2$ . As a consequence, we do not know the degree of vertices in  $[t/2]$  when the graph has size  $t$ . However, in [DHH10, Appendix A.4] the authors prove that at time  $t$  all the vertices  $t/2 + 1, \dots, t$  have degree smaller than  $(\log t)^\sigma$ .

We continue by giving a bound on the degree of vertices that are not in  $\text{Core}_t$ . For vertices  $i \in [t/2] \setminus \text{Core}_t$  we know that  $D_{t/2}(i) < (\log t)^\sigma$ , see (2.8), but in principle their degree  $D_t(i)$  at time  $t$  could be quite high. We need to prove that this happens with very small probability. Precisely, we prove that, for some  $B > 0$ ,

$$\mathbb{P}\left(\max_{i \in [t/2] \setminus \text{Core}_t} D_t(i) \geq (1 + B)(\log t)^\sigma\right) = o(1). \quad (6.3)$$

This inequality implies that when a degree is at most  $(\log t)^\sigma$  at time  $t/2$ , then it is unlikely to grow by  $B(\log t)^\sigma$  between time  $t/2$  and  $t$ . This provides a bound on the cardinality of incoming neighborhoods that we can use in the definition of the exploration processes that we will rely on, in order to avoid  $\text{Core}_t$ . We prove (6.3) in the following lemma that is an adaptation of the proof of [DHH10, Lemma A.4]. Its proof is deferred to [CGH16, Appendix B]:

**Lemma 6.4** (Old vertex not in  $\text{Core}_t$ ). *There exists  $B \in (0, \infty)$  such that, for every  $i \in [t/2]$ ,*

$$\mathbb{P}(D_t(i) \geq (1 + B)(\log t)^\sigma \mid D_{t/2}(i) < (\log t)^\sigma) = o(1/t). \quad (6.4)$$

We can now get to the core of the proof of Statement 2.5, that is we show that there are few collisions before reaching  $\text{Core}_t$ :

**Lemma 6.5** (Few collisions before hitting the core). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Fix  $a \in (0, 1)$  and  $l \in \mathbb{N}$  such that  $l > 1/a$ . With  $k_t^+$  as in (6.2), the probability that there exists a vertex  $v \in [t]$  such that its  $k_t^+$ -exploration graph has at least  $l$  collisions before hitting  $\text{Core}_t \cup [t^a]$  is  $o(1)$ .*

Next we give a lower bound on the number of vertices on the boundary of a  $k_n^+$ -exploration graph. First of all, for any fixed  $a \in (0, 1)$ , we notice that the probability of existence of a vertex in  $[t] \setminus [t^a]$ , that has only self loops is  $o(1)$ . Indeed, the probability that a vertex  $s$  has only self-loops is  $O(\frac{1}{s^m})$ . Thus, the probability that there exists a vertex in  $[t] \setminus [t^a]$  that has only self-loops is bounded above by

$$\sum_{s > t^a} O\left(\frac{1}{s^m}\right) = O(t^{-a(m-1)}) = o(1), \quad (6.5)$$

since we assume that  $m \geq 2$ . We can thus assume that no vertex in  $[t] \setminus [t^a]$  has only self-loops. This leads us formulate the following Lemma, whose proof is also deferred to [CGH16, Appendix B].

**Lemma 6.6** (Lower bound on boundary vertices). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . For  $a \in (0, 1)$ , consider a vertex  $v \in [t] \setminus (\text{Core}_t \cup [t^a])$  and its  $k$ -exploration graph. If there are at most  $l$  collisions in the  $k$ -exploration graph, and no vertex in  $[t] \setminus [t^a]$  has only self loops, then there exists a constant  $s = s(m, l) > 0$  such that the number of vertices in the boundary of the  $k$ -exploration graph is at least  $s(m, l)m^k$ .*

Together, Lemmas 6.3, 6.5 and 6.6 complete the proof of Statement 2.5.

The rest of this section is devoted to the proof of Lemma 6.5. We first need to introduce some notation, in order to be able to express the probability of collisions. We do this in the next subsection.

#### 6.1.1. Ulam-Harris notation for trees. Define

$$W_\ell := [m]^\ell, \quad W_{\leq k} := \bigcup_{\ell=0}^k W_\ell,$$

where  $W_0 := \emptyset$ . We use  $W_{\leq k}$  as a universal set to label any regular tree of depth  $k$ , where each vertex has  $m$  children. This is sometimes called the *Ulam-Harris notation* for trees.

Given  $y \in W_\ell$  and  $z \in W_m$ , we denote by  $(y, z) \in W_{\ell+m}$  the concatenation of  $y$  and  $z$ . Given  $x, y \in W_{\leq k}$ , we write  $y \succeq x$  if  $y$  is a descendant of  $x$ , that is  $y = (x, z)$  for some  $z \in W_{\leq k}$ .

Given a finite number of points  $z_1, \dots, z_m \in W_{\leq k}$ , abbreviate  $\vec{z}_m = (z_1, \dots, z_m)$ , and define  $W_{\leq k}^{(\vec{z}_m)}$  to be the tree obtained from  $W_{\leq k}$  by cutting the branches starting from any of the  $z_i$ 's (including the  $z_i$ 's themselves):

$$W_{\leq k}^{(\vec{z}_k)} := \{x \in W_{\leq k} : x \not\succeq z_1, \dots, x \not\succeq z_m\}. \quad (6.6)$$

**Remark 6.7** (Total order). The set  $W_{\leq k}$  comes with a natural total order relation, called *shortlex order*, in which shorter words precede longer ones, and words with equal length are ordered lexicographically. More precisely, given  $x \in W_\ell$  and  $y \in W_m$ , we say that  $x$  precedes  $y$  if either  $\ell < m$ , or if  $\ell = m$  and  $x_i \leq y_i$  for all  $1 \leq i \leq \ell$ . We stress that this is a *total* order relation, unlike the descendant relation  $\succeq$  which is only a partial order. (Of course, if  $y \succeq x$ , then  $x$  precedes  $y$ , but not vice versa).

6.1.2. *Collisions.* We recall that, given  $z \in [t]$  and  $j \in [m]$ , the  $j$ -th half-edge starting from vertex  $z$  in  $\text{PA}_t$  is attached to a random vertex, denoted by  $\xi_{z,j}$ . We can use the set  $W_{\leq k}$  to label the exploration graph  $\widehat{U}_{\leq k}(v)$ , as follows:

$$\widehat{U}_{\leq k}(v) = \{V_z\}_{z \in W_{\leq k}}, \quad (6.7)$$

where  $V_\emptyset = v$  and, iteratively,  $V_z = \xi_{V_x,j}$  for  $z = (x, j)$  with  $x \in W_{\leq k-1}$  and  $j \in [m]$ .

The first vertex generating a *collision* is  $V_{Z_1}$ , where the random index  $Z_1 \in W_{\leq k}$  is given by

$$Z_1 := \min \{z \in W_{\leq k} : V_z = V_y \text{ for some } y \text{ which precedes } z\},$$

where “min” refers to the total order relation on  $W_{\leq k}$  as defined in Remark 6.7.

Now comes a tedious observation. Since  $V_{Z_1} = V_y$  for some  $y$  which precedes  $Z_1$ , by definition of  $Z_1$ , then all descendants of  $Z_1$  will coincide with the corresponding descendants of  $y$ , that is  $V_{(Z_1,r)} = V_{(y,r)}$  for all  $r$ . In order not to over count collisions, in defining the second collision index  $Z_2$ , we avoid exploring the descendants of index  $Z_1$ , that is we only look at indices in  $W_{\leq k}^{(Z_1)}$ , see (6.6). The second vertex representing a (true) collision is then  $V_{Z_2}$ , where we define

$$Z_2 := \min \{z \in W_{\leq k}^{(Z_1)} : z \text{ follows } Z_1, \text{ i.e., } V_z = V_y \text{ for some } y \text{ which precedes } z\},$$

Iteratively, we define

$$Z_{i+1} := \min \{z \in W_{\leq k}^{(Z_i)} : z \text{ follows } Z_i, \text{ i.e., } V_z = V_y \text{ for some } y \text{ which precedes } z\},$$

so that  $V_{Z_i}$  is the  $i$ -th vertex that represents a collision. The procedure stops when there are no more collisions. Denoting by  $\mathcal{C}$  the (random) number of collisions, we have a family

$$Z_1, Z_2, \dots, Z_{\mathcal{C}}$$

of random elements of  $W_{\leq k}$ , such that  $(V_{Z_i})_{1 \leq i \leq \mathcal{C}}$  are the vertices generating the collisions.

6.1.3. *Proof of Lemma 6.5.* Recalling (6.7) and (6.6), given arbitrarily  $z_1, \dots, z_l \in W_{\leq k}$ , we define

$$\widehat{U}_{\leq k}^{(z_l)}(v) = \{V_z\}_{z \in W_{\leq k}^{(z_l)}}, \quad (6.8)$$

that is, we consider a subset of the full exploration graph  $\widehat{U}_{\leq k}(v)$ , consisting of vertices  $V_z$  whose indexes  $z \in W_{\leq k}$  are not descendants of  $z_1, \dots, z_l$ . The basic observation is that

$$\widehat{U}_{\leq k}(v) = \widehat{U}_{\leq k}^{(z_l)}(v) \quad \text{on the event } \{\mathcal{C} = l, Z_1 = z_1, \dots, Z_l = z_l\}. \quad (6.9)$$

In words, this means that to recover the full exploration graph  $\widehat{U}_{\leq k}(v)$ , it is irrelevant to look at vertices  $V_z$  for  $z$  that is a descendant of a collision index  $z_1, \dots, z_l$ .

We will bound the probability that there are  $l$  collisions before reaching  $\text{Core}_t \cup [t^a]$ , occurring at specified indices  $z_1, \dots, z_l \in W_{\leq k}$ , for  $k = k_t^+$  as in (6.2), as follows:

$$\mathbb{P}(\mathcal{C} = l, Z_1 = z_1, \dots, Z_l = z_l, \widehat{U}_{\leq k}(v) \cap (\text{Core}_t \cup [t^a]) = \emptyset) \leq \alpha(t)^l, \quad (6.10)$$

where, for the constant  $B$  given by Lemma 6.4, we define

$$\alpha(t) = \frac{4(1+B)}{m} \frac{(\log t)^{\sigma+1+\varepsilon}}{t^a}. \quad (6.11)$$

Summing (6.10) over  $z_1, \dots, z_l \in W_{\leq k}$  we get

$$\mathbb{P}(\mathcal{C} = l, \widehat{U}_{\leq k}(v) \cap (\text{Core}_t \cup [t^a]) = \emptyset) \leq \alpha(t)^l |W_{\leq k}|^l.$$

Since, for  $k = k_t^+$  as in (6.2), we can bound

$$|W_{\leq k}| = \frac{m^{k+1} - 1}{m - 1} \leq 2m^k \leq 2(\log t)^{1+\varepsilon}, \quad (6.12)$$

the probability of having at least  $l$  collisions, before reaching  $\text{Core}_t \cup [t^a]$ , is  $O(\alpha(t)^l (\log t)^{2l}) = o(1/t)$ , because  $l > 1/a$  by assumption. This completes the proof of Lemma 6.5. It only remains to show that (6.10) holds true.

**6.1.4. Proof of (6.10): case  $l = 1$ .** We start proving (6.10) for one collision. By (6.9), we can replace  $\widehat{U}_{\leq k}(v)$  by  $\widehat{U}_{\leq k}^{(z_1)}(v)$  in the left hand side of (6.10), i.e., we have to prove that

$$\mathbb{P}(\mathcal{C} = 1, Z_1 = z_1, \widehat{U}_{\leq k}^{(z_1)}(v) \cap (\text{Core}_t \cup [t^a]) = \emptyset) \leq \alpha(t). \quad (6.13)$$

Since  $v, k$  and  $z_1$  are fixed, let us abbreviate, and recalling (6.8),

$$\mathcal{W} := W_{\leq k}^{(z_1)}(v), \quad \widehat{U} := \widehat{U}_{\leq k}^{(z_1)}(v) = \{V_z\}_{z \in \mathcal{W}}. \quad (6.14)$$

Note that  $V_{z_1}$  is the only collision precisely when  $\widehat{U}$  is a tree and  $V_{z_1} \in \widehat{U}$ . Then (6.13) becomes

$$\mathbb{P}(\widehat{U} \text{ is a tree}, V_{z_1} \in \widehat{U}, \widehat{U} \cap (\text{Core}_t \cup [t^a]) = \emptyset) \leq \alpha(t). \quad (6.15)$$

We will actually prove a stronger statement: for any fixed *deterministic* labeled directed tree  $H \subseteq [t]$  and for any  $y \in H$ ,

$$\mathbb{P}(\widehat{U} = H, V_{z_1} = y, H \cap (\text{Core}_t \cup [t^a]) = \emptyset) \leq \frac{\alpha(t)}{2(\log t)^{1+\varepsilon}} \mathbb{P}(\widehat{U} = H, V_{z_1} \notin H). \quad (6.16)$$

This yields (6.15) by summing over  $y \in H$  —note that  $|H| \leq |W_{\leq k}| \leq 2(\log t)^{1+\varepsilon}$  by (6.12)— and then summing over all possible realizations of  $H$ .

It remains to prove (6.16). We again use the notion of a *factorizable event*, as in the proof of the lower bound. Since the events in (6.16) are not factorizable, we will specify the incoming neighborhood  $\mathcal{N}(y)$  (recall (4.51)) of all  $y \in H$ . More precisely, by labeling the vertices of  $H$ , see (6.14), as

$$H = \{v_s\}_{s \in \mathcal{W}} \quad \text{and} \quad y = v_{\bar{s}}, \quad \text{for some } \bar{s} \in \mathcal{W}, \quad (6.17)$$

we can consider the events  $\{\mathcal{N}(v_s) = N_{v_s}\}$  where  $N_{v_s}$  are (deterministic) disjoint subsets of  $[t] \times [m]$ . We say that the subsets  $(N_{v_s})_{s \in \mathcal{W}}$  are *compatible* with the tree  $H$  when  $(v_{s'}, j) \in N_{v_s}$  whenever  $s = (s', j)$  with  $s, s' \in \mathcal{W}, j \in [m]$ . Then we can write

$$\{\widehat{U} = H\} = \bigcup_{\text{compatible } (N_{v_s})_{s \in \mathcal{W}}} \{\mathcal{N}(v_s) = N_{v_s} \text{ for every } s \in \mathcal{W}\}. \quad (6.18)$$

Since the degree of vertex  $v_s$  equals  $D_t(v_s) = m + |N_{v_s}|$ , we can ensure that  $H \cap (\text{Core}_t \cup [t^a]) = \emptyset$  by restricting the union in (6.18) to those  $N_{v_s}$  satisfying the constraints

$$v_s > t^a \quad \text{and} \quad |N_{v_s}| \leq (1 + B)(\log t)^\sigma - m, \quad \forall s \in \mathcal{W}. \quad (6.19)$$

Finally, if we write

$$z_1 = (x, j) \quad \text{for some } x \in \mathcal{W}, j \in [m], \quad (6.20)$$

then, since  $V_{z_1} = \xi_{V_x, j}$ , the event  $\{V_{z_1} = v_{\bar{s}}\}$  amounts to require that<sup>2</sup>

$$(v_x, j) \in N_{v_{\bar{s}}}. \quad (6.21)$$

Let us summarize where we now stand: When we fix a family of  $(N_{v_s})_{s \in \mathcal{W}}$  that is compatible and satisfies the constraints (6.19) and (6.21), in order to prove (6.16) it is enough to show that

$$\begin{aligned} & \mathbb{P}(\mathcal{N}(v_s) = N_{v_s} \text{ for every } s \in \mathcal{W}) \\ & \leq \frac{\alpha(t)}{2(\log t)^{1+\varepsilon}} \mathbb{P}(\mathcal{N}(v_s) = N_{v_s} \text{ for every } s \in \mathcal{W} \setminus \{\bar{s}\}, \mathcal{N}(v_{\bar{s}}) = N_{v_{\bar{s}}} \setminus \{(v_x, j)\}). \end{aligned} \quad (6.22)$$

<sup>2</sup>Incidentally, we observe that the constraint (6.21) is not included in the requirement that  $(N_{v_s})_{s \in \mathcal{W}}$  are compatible, because  $z_1 = (x, j) \notin \mathcal{W}$  by definition (6.14) of  $\mathcal{W}$ .

Let us set

$$N := \bigcup_{s \in \mathcal{W}} N_{v_s} \subseteq [t] \times [m]. \quad (6.23)$$

The probability on the left-hand side of (6.22) can be factorized, using conditional expectations and the tower property, as a product of two kinds of terms:

- ▷ For every edge  $(u, r) \in N$  —say  $(u, r) \in N_{v_s}$ , with  $s \in \mathcal{W}$ — we have the term

$$\frac{D_{u,r-1}(v_s) + \delta}{c_{u,r}} \quad (6.24)$$

corresponding to the fact that the edge needs to be connected to  $v_s$ ;

- ▷ On the other hand, for every edge  $(u, r) \notin N$ , we have the term

$$1 - \frac{D_{u,r-1}(H) + |H \cap [u-1]| \delta}{c_{u,r}}, \quad (6.25)$$

corresponding to the fact that the edge may not connect to any vertex in  $H$ .

(We emphasize that all the degrees  $D_{\cdot,\cdot}(\cdot)$  appearing in (6.24) and (6.25) are *deterministic*, since they are fully determined by the realizations of the incoming neighborhoods  $(N_{v_s})_{s \in \mathcal{W}}$ .)

We can obtain the right-hand side in (6.22) by replacing some terms in the product.

- ▷ Among the edges  $(u, r) \in N$ , whose contribution is (6.24), we have the one that creates the collision, namely  $(v_x, j)$ . If we want this edge to be connected *outside*  $H$ , as in the right-hand side in (6.22), we need to divide the left hand side of (6.22) by

$$\left( \frac{D_{v_x,j-1}(v_{\bar{s}}) + \delta}{c_{v_x,j}} \right) \left( 1 - \frac{D_{v_x,j-1}(H) + |H \cap [v_x-1]| \delta}{c_{v_x,j}} \right)^{-1}. \quad (6.26)$$

We also have to replace some other terms corresponding to edges  $(u, r) \in N_{v_{\bar{s}}}$ , because the degree of vertex  $v_{\bar{s}}$  is decreased by one after connecting  $(v_x, j)$  outside  $H$ . More precisely, for every edge  $(u, r) \in N_{v_{\bar{s}}}$  that is younger than  $(v_x, j)$ , that is  $(u, r) > (v_x, j)$ , we can reduce the degree of  $v_{\bar{s}}$  by one by dividing the left-hand side of (6.22) by

$$\prod_{(u,r) \in N_{v_{\bar{s}}}, (u,r) > (v_x,j)} \frac{D_{u,r-1}(v_{\bar{s}}) + \delta}{D_{u,r-1}(v_{\bar{s}}) - 1 + \delta} = \frac{D_t(v_{\bar{s}}) + \delta}{D_{v_x,j-1}(v_{\bar{s}}) + \delta}. \quad (6.27)$$

Finally, the contribution of the edges  $(u, r) \in N_{v_s}$  for  $s \neq \bar{s}$  is unchanged.

- ▷ For every edge  $(u, r) \notin N$ , the probability that such edge is not attached to  $H$ , after we reconnect the edge  $(v_x, j)$ , becomes larger, since the degree of  $H$  is reduced by one.

It follows that the inequality (6.22) holds with  $\alpha(t)/(2(\log t)^{1+\varepsilon})$  replaced by  $\beta$ , defined by

$$\begin{aligned} \beta &= \left( \frac{D_{v_x,j-1}(v_{\bar{s}}) + \delta}{c_{v_x,j}} \right) \left( 1 - \frac{D_{v_x,j-1}(H) + |H \cap [v_x-1]| \delta}{c_{v_x,j}} \right)^{-1} \frac{D_t(v_{\bar{s}}) + \delta}{D_{v_x,j-1}(v_{\bar{s}}) + \delta} \\ &= \left( \frac{D_t(v_{\bar{s}}) + \delta}{c_{v_x,j}} \right) \left( 1 - \frac{D_{v_x,j-1}(H) + |H \cap [v_x-1]| \delta}{c_{v_x,j}} \right)^{-1} \\ &\leq \left( \frac{D_t(v_{\bar{s}})}{c_{v_x,j}} \right) \left( 1 - \frac{D_{v_x,j-1}(H)}{c_{v_x,j}} \right)^{-1} =: \beta', \end{aligned} \quad (6.28)$$

because  $\delta \leq 0$ . We only need to show that  $\beta' \leq \alpha(t)/(2(\log t)^{1+\varepsilon})$ .

Since  $c_{v_x,j} \geq m(v-1)$ , the first relation in (6.19) yields

$$c_{v_x,j} \geq t^a.$$

Hence, since  $D_t(v_{\bar{s}}) \leq (1+B)(\log t)^\sigma$  by the second relation in (6.19), we can bound

$$\left( \frac{D_t(v_{\bar{s}})}{c_{v_x,j}} \right) \leq \frac{(1+B)(\log t)^\sigma}{mt^a}.$$

Likewise, since  $D_t(H) \leq |H|(1+B)(\log t)^\sigma$ , for  $k = k_t^+$  we get, by (6.12),

$$\left(1 - \frac{D_{v_x,j-1}(H)}{c_{v_x,j}}\right)^{-1} \leq \left(1 - \frac{2(\log t)^{1+\varepsilon}(1+B)(\log t)^\sigma}{t^a}\right)^{-1} \leq 2,$$

where the last inequality holds for  $t$  large enough. Recalling (6.11),

$$\beta' \leq 2 \frac{(1+B)(\log t)^\sigma}{mt^a} = \frac{\alpha(t)}{2(\log t)^{1+\varepsilon}}.$$

This completes the proof of (6.22), and hence of (6.10), in the case where  $l = 1$ .  $\square$

**6.1.5. Proof of (6.10): general case  $l \geq 2$ .** The proof for the general case is very similar to that for  $l = 1$ , so we only highlight the (minor) changes.

In analogy with (6.13), we can replace  $\widehat{U}_{\leq k}(v)$  by  $\widehat{U}_{\leq k}^{(\bar{z}_l)}(v)$  in the left-hand side of (6.10), thanks to (6.9). Then, as in (6.14), we write

$$\mathcal{W} := W_{\leq k}^{(\bar{z}_l)}(v), \quad \widehat{U} := \widehat{U}_{\leq k}^{(\bar{z}_l)}(v) = \{V_z\}_{z \in \mathcal{W}}. \quad (6.29)$$

The extension of (6.16) becomes that for any fixed *deterministic* labeled directed tree  $H \subseteq [t]$  and for all  $y_1, \dots, y_l \in H$ ,

$$\begin{aligned} & \mathbb{P}(\widehat{U} = H, V_{z_1} = y_1, \dots, V_{z_l} = y_l, H \cap (\text{Core}_t \cup [t^a]) = \emptyset) \\ & \leq \left(\frac{\alpha(t)}{2(\log t)^{1+\varepsilon}}\right)^l \mathbb{P}(\widehat{U} = H, V_{z_1} \notin H, V_{z_2} \notin H, \dots, V_{z_l} \notin H). \end{aligned} \quad (6.30)$$

As in (6.17), we can write

$$H = \{v_s\}_{s \in \mathcal{W}} \quad \text{and} \quad y_1 = v_{\bar{s}_1}, \dots, y_l = v_{\bar{s}_l} \quad \text{for some } \bar{s}_1, \dots, \bar{s}_l \in \mathcal{W}.$$

To obtain a factorizable event, we must specify the incoming neighborhoods  $\mathcal{N}_{v_s} = N_{v_s}$  for all  $s \in \mathcal{W}$ , which must be compatible with  $H$  and satisfy the constraint (6.19). If we write

$$z_1 = (x_1, j_1), \dots, z_l = (x_l, j_l), \quad \text{for some } x_1, \dots, x_l \in \mathcal{W}, j_1, \dots, j_l \in [m],$$

then we also impose the constraint that obviously generalizes (6.21), namely

$$(v_{x_1}, j_1) \in N_{v_{\bar{s}_1}}, \dots, (v_{x_l}, j_l) \in N_{v_{\bar{s}_l}}.$$

The analogue of (6.22) then becomes

$$\begin{aligned} & \mathbb{P}(\mathcal{N}(v_s) = N_{v_s} \text{ for every } s \in \mathcal{W}) \\ & \leq \left(\frac{\alpha(t)}{2(\log t)^{1+\varepsilon}}\right)^l \mathbb{P}(\mathcal{N}(v_s) = N_{v_s} \text{ for every } s \in \mathcal{W} \setminus \{\bar{s}_1, \dots, \bar{s}_l\}, \\ & \quad \mathcal{N}(v_{\bar{s}_i}) = N_{v_{\bar{s}_i}} \setminus \{(v_{x_i}, j_i)\} \text{ for every } i = 1, \dots, l). \end{aligned} \quad (6.31)$$

When we define  $N$  as in (6.23), the probability in the left-hand side of (6.31) can be factorized in a product of terms of two different types, which are given precisely by (6.24) and (6.25). In order to obtain the probability in the right-hand side of (6.31), we have to divide the left-hand side by a product of factors analogous to (6.26) and (6.27). More precisely, (6.26) becomes

$$\prod_{i=1}^l \left(\frac{D_{v_{x_i},j_i-1}(v_{\bar{s}_i}) + \delta}{c_{v_{x_i},j_i}}\right) \left(1 - \frac{D_{v_{x_i},j_i-1}(H) + |H \cap [v_{x_i}-1]| \delta}{c_{v_{x_i},j_i}}\right)^{-1}, \quad (6.32)$$

while (6.27) becomes

$$\prod_{i=1}^l \frac{D_t(v_{\bar{s}_i}) + \delta}{D_{v_{x_i},j_i-1}(v_{\bar{s}_i}) + \delta}.$$

We define  $\beta$  accordingly, namely we take the product for  $i = 1, \dots, l$  of (6.28) with  $x, j, \bar{s}$  replaced respectively by  $x_i, j_i, \bar{s}_i$ . Then it is easy to show that

$$\beta \leq \left( \frac{\alpha(t)}{2(\log t)^{1+\varepsilon}} \right)^l,$$

arguing as in the case  $l = 1$ . This completes the proof of (6.31).  $\square$

**6.2. Proof of Statement 2.6.** The next step is to prove that the boundaries of the  $k_t^+$ -exploration graphs are at most at distance

$$h_t = \lceil B \log \log \log t + C \rceil \quad (6.33)$$

from  $\text{Core}_t$ , where  $B, C$  are constants to be chosen later on. Similarly to the proof in Section 5.2, we consider a  $k_t^+$ -exploration graph, and we enumerate the vertices on the boundary as  $x_1, \dots, x_N$ , where  $N \geq s(m, l)m^{k_t^+}$  from Lemma 6.6 and  $l$  is chosen as in Lemma 6.5. We next define what it means to have a success:

**Definition 6.8** (Success). *Consider the vertices  $x_1, \dots, x_N$  on the boundary of a  $k_t^+$ -exploration graph. We say that  $x_i$  is a success when the distance between  $x_i$  and  $\text{Core}_t$  is at most  $2h_t$ .*

The next lemma is similar to Lemma 5.7 (but only deals with vertices in  $[t/2]$ ):

**Lemma 6.9** (Probability of success). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Consider  $v \in [t/2] \setminus \text{Core}_t$  and its  $k_t^+$ -exploration graph. Then there exists a constant  $\eta > 0$  such that*

$$\mathbb{P}(S_{x_1} \mid \text{PA}_{t/2}) \geq \eta, \quad (6.34)$$

and for all  $j = 2, \dots, N$ ,

$$\mathbb{P}(S_{x_1} \mid \text{PA}_{t/2}, S_{x_1}^c, \dots, S_{x_{j-1}}^c) \geq \eta. \quad (6.35)$$

The aim is to define a sequence of vertices  $w_0, \dots, w_h$  that connects a vertex  $x_i$  on the boundary with  $\text{Core}_t$ . In order to do this, we need some preliminary results. We start with the crucial definition of a  $t$ -connector:

**Definition 6.10** ( $t$ -connector). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$ . Consider two subsets  $A, B \subseteq [t/2]$ , with  $A \cap B = \emptyset$ . We say that a vertex  $j \in [t] \setminus [t/2]$  is a  $t$ -connector for  $A$  and  $B$  if at least one of the edges incident to  $j$  is attached to a vertex in  $A$  and at least one is attached to a vertex in  $B$ .*

The notion of  $t$ -connector is useful, because, unlike in the configuration model, in the preferential attachment model typically two high-degree vertices are not directly connected. From the definition of the preferential attachment model, it is clear that the older vertices have with high probability large degree, and the younger vertices have lower degree. When we add a new vertex, this is typically attached to vertices with large degrees. This means that, with high probability, two vertices with high degree can be connected by a young vertex, which is the  $t$ -connector.

A further important reason for the usefulness of  $t$ -connectors is that we have effectively *decoupled* the preferential attachment model at time  $t/2$  and what happens in between times  $t/2$  and  $t$ . When the sets  $A$  and  $B$  are appropriately chosen, then each vertex will be a  $t$ -connector with reasonable probability, and the events that distinct vertices are  $t$ -connectors are close to being independent. Thus, we can use comparisons to binomial random variables to investigate the existence of  $t$ -connectors. In order to make this work, we need to identify the structure of  $\text{PA}_{t/2}$  and show that it has sufficiently many vertices of large degree, and we need to show that  $t$ -connectors are likely to exist. We start with the latter.

In more detail, we will use  $t$ -connectors to generate the sequence of vertices  $w_1, \dots, w_h$  between the boundary of a  $k_n^+$ -exploration graph and the  $\text{Core}_t$ , in the sense that we use a  $t$ -connector to link the vertex  $w_i$  to the vertex  $w_{i+1}$ . (This is why we define a vertex  $x_i$  to be a success if its distance from  $\text{Core}_t$  is at most  $2h_t$ , instead of  $h_t$ .) We rely on a result implying the existence of  $t$ -connectors between sets of high total degree:

**Lemma 6.11** (Existence of  $t$ -connectors). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . There exists a constant  $\mu > 0$  such that, for every  $A \subseteq [t/2]$ , and  $i \in [t/2] \setminus A$ ,*

$$\mathbb{P}\left(\#\{j \in [t] \setminus [t/2] : j \text{ is a } t\text{-connector for } i \text{ and } A \mid \text{PA}_{t/2}\} \leq \exp\left(-\frac{\mu D_A(t/2) D_i(t/2)}{t}\right)\right), \quad (6.36)$$

where  $D_A(t/2) = \sum_{v \in A} D_v(t/2)$  is the total degree of  $A$  at time  $t/2$ .

*Proof.* The proof of this lemma is present in the proof of [DHH10, Proposition 3.2].  $\square$

**Remark 6.12.** Notice that this bound depends on the fact that the number of possible  $t$ -connectors is of order  $t$ .

A last preliminary result that we need is a technical one, which plays the role of Lemma 5.9 for the configuration model and shows that at time  $t/2$  there are sufficiently many vertices of high degree, uniformly over a wide range of what ‘large’ could mean:

**Lemma 6.13** (Tail of degree distribution). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model, with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Then, for all  $\zeta > 0$  there exists a constant  $c = c(\zeta)$  such that, for all  $1 \leq x \leq (\log t)^q$ , for any  $q > 0$ , and uniformly in  $t$ ,*

$$P_{\geq x}(t) = \frac{1}{t} \sum_{v \in [t]} \mathbf{1}_{\{D_v(t) \geq x\}} \geq cx^{-(\tau-1+\zeta)}. \quad (6.37)$$

*Proof.* The degree distribution sequence  $(p_k)_{k \in \mathbb{N}}$  in (1.12) satisfies a power law with exponent  $\tau \in (2, 3)$ . As a consequence, for all  $\zeta > 0$  there exists a constant  $\bar{c} = \bar{c}(\zeta)$  such that

$$p_{\geq x} := \sum_{k \geq x} p_k \geq \bar{c}x^{-(\tau-1+\zeta)}. \quad (6.38)$$

We now use a concentration result on the empirical degree distribution (for details, see [Hof17, Theorem 8.2]), which assures us that there exists a second constant  $C > 0$  such that, with high probability, for every  $x \in \mathbb{N}$ ,

$$|P_{\geq x} - p_{\geq x}| \leq C \sqrt{\frac{\log t}{t}}. \quad (6.39)$$

Fix now  $\zeta > 0$ , then from this last bound we can immediately write, for a suitable constant  $\bar{c}$  as in (6.38),

$$P_{\geq x} \geq p_{\geq x} - C \sqrt{\frac{\log t}{t}} \geq \bar{c}x^{-(\tau-1+\zeta)} - C \sqrt{\frac{\log t}{t}} \geq \frac{\bar{c}}{2}x^{-(\tau-1+\zeta)}, \quad (6.40)$$

if and only if

$$C \sqrt{\frac{\log t}{t}} = o(x^{-(\tau-1+\zeta)}). \quad (6.41)$$

This is clearly true for  $x \leq (\log t)^q$ , for any positive  $q$ . Taking  $c = \bar{c}/2$  completes the proof.  $\square$

With the above tools, we are now ready to complete the proof of Lemma 6.9:

*Proof of Lemma 6.9.* As in the proof of Proposition 5.7, we define the super-exponentially growing sequence  $g_\ell$  as in (5.17), where  $\gamma > 0$  is chosen small enough, as well as  $\zeta > 0$ , so that (5.16) holds. The constants  $B$  and  $C$  in the definition (6.33) of  $h_t$  are fixed as prescribed below (5.17).

We will define a sequence of vertices  $w_0, \dots, w_h$  such that, for  $i = 1, \dots, h$ ,  $D_{w_i}(t) \geq g_i$  and  $w_{i-1}$  is connected to  $w_i$ . For this, we define, for  $i = 1, \dots, h-1$ ,

$$H_i = \left\{ u \in [t] : D_u(t/2) \geq g_i \right\} \subseteq [t/2], \quad (6.42)$$

so that we aim for  $w_i \in H_i$ .

We define the vertices recursively, and start with  $w_0 = x_1$ . Then, we consider  $t$ -connectors between  $w_0$  and  $H_1$ , and denote by  $w_1$  the vertex in  $H_1$  with minimal degree among the ones that are connected to  $w_0$  by a  $t$ -connector. Recursively, consider  $t$ -connectors between  $w_i$  and  $H_{i+1}$ ,

and denote by  $w_{i+1}$  the vertex in  $H_{i+1}$  with minimal degree among the ones that are connected to  $w_i$  by a  $t$ -connector. Recall (5.18) to see that  $g_{h_t} \geq (\log t)^\sigma$ , where  $h_t$  is defined in (6.33). The distance between  $w_0$  and  $\text{Core}_t$  is at most  $2h_t = 2\lceil B \log \log \log t + C \rceil$ . If we denote the event that there exists a  $t$  connector between  $w_{i-1}$  and  $H_i$  by  $\{w_{i-1} \sim H_i\}$ , then we will bound from below

$$\mathbb{P}(S_{x_1} \mid \text{PA}_{t/2}) \geq \mathbb{E} \left[ \prod_{i=1}^{h_t} \mathbb{1}_{\{w_{i-1} \sim H_i\}} \mid \text{PA}_{t/2} \right]. \quad (6.43)$$

In Lemma 6.11, the bound on the probability that a vertex  $j \in [t] \setminus [t/2]$  is a  $t$ -connector between two subsets of  $[t]$  is independent of the fact that the other vertices are  $t$ -connectors or not. This means that, with  $\mathcal{F}_i$  the  $\sigma$ -field generated by the path formed by  $w_0, \dots, w_i$  and their respective  $t$ -connectors,

$$\mathbb{E} \left[ \mathbb{1}_{\{w_{i-1} \sim H_i\}} \mid \text{PA}_{t/2}, \mathcal{F}_{i-1} \right] \geq 1 - e^{-\mu D_{w_{i-1}}(t/2) D_{H_i}(t/2)/t}, \quad (6.44)$$

where  $D_{H_i}(t) = \sum_{u \in H_i} D_u(t/2)$ . This means that

$$\mathbb{E} \left[ \prod_{i=1}^{h_t} \mathbb{1}_{\{w_{i-1} \sim H_i\}} \mid \text{PA}_{t/2} \right] \geq \prod_{i=1}^{h_t} \left( 1 - e^{-\mu D_{w_{i-1}}(t/2) D_{H_i}(t/2)/t} \right). \quad (6.45)$$

We have to bound every term in the product. Using Lemma 6.13, for  $i = 1$ ,

$$1 - e^{-\mu D_{w_0}(t/2) D_{H_1}(t/2)/t} \geq 1 - e^{-\mu D_{w_0}(t/2) g_1 P_{\geq g_1}(t/2)}, \quad (6.46)$$

while, for  $i = 2, \dots, h-1$

$$1 - e^{-\mu D_{w_{i-1}}(t/2) D_{H_i}(t/2)/t} \geq 1 - e^{-\mu g_{i-1} g_i P_{\geq g_i}(t/2)}. \quad (6.47)$$

Applying (6.37) and recalling (5.25)–(5.26), the result is

$$\begin{aligned} \mathbb{P}(S_{x_1} \mid \text{PA}_t) &\geq \left( 1 - e^{-\mu D_{w_0}(t/2) g_1 P_{\geq g_1}(t/2)} \right) \prod_{i=2}^{h_t} \left( 1 - e^{-\mu g_{i-1} g_i P_{\geq g_i}(t/2)} \right) \\ &\geq \left( 1 - e^{-\mu m g_1 P_{\geq g_1}(t/2)} \right) \prod_{i=2}^{\infty} \left( 1 - e^{-\tilde{c}(g_i)^\xi} \right), \end{aligned} \quad (6.48)$$

for some constant  $\tilde{c}$ . Since  $h_t = \lceil B \log \log \log t + C \rceil$ , and

$$P_{\geq g_1}(t/2) \rightarrow \sum_{k \geq g_1} p_k > 0 \quad (6.49)$$

with high probability as  $t \rightarrow \infty$ , we can find a constant  $\eta$  such that

$$\left( 1 - e^{-\eta m g_1 P_{\geq g_1}(t/2)} \right) \prod_{i=2}^{h_t} \left( 1 - e^{-\tilde{c}(g_i)^\xi} \right) > \eta > 0, \quad (6.50)$$

which proves (6.34).

To prove (6.35), we observe that all the lower bounds that we have used on the probability of existence of  $t$ -connectors only depend on the existence of sufficiently many potential  $t$ -connectors. Thus, it suffices to prove that, on the event  $S_{x_1}^c \cap \dots \cap S_{x_{j-1}}^c$ , we have not used too many vertices as  $t$ -connectors. On this event, we have used at most  $h_t \cdot (j-1)$  vertices as  $t$ -connectors, which is  $o(t)$ . Thus, this means that, when we bound the probability of  $S_{x_j}$ , we still have  $t - h_t \cdot (j-1)$  possible  $t$ -connectors, where  $j$  is at most  $(\log t)^{1+\varepsilon}$ . Thus, with the same notation as before,

$$\mathbb{E} \left[ \mathbb{1}_{\{w_{i-1} \sim H_i\}} \mid \text{PA}_{t/2}, S_{x_1}^c, \dots, S_{x_{j-1}}^c \right] \geq 1 - e^{-\mu D_{w_{i-1}}(t/2) D_{H_i}(t/2)/t}, \quad (6.51)$$

so that we can proceed as we did for  $S_{x_1}$ . We omit further details.  $\square$

We are now ready to identify the distance between the vertices outside the core and the core:

**Proposition 6.14** (Distance between periphery and  $\text{Core}_t$ ). *Let  $(\text{PA}_t)_{t \geq 1}$  be a preferential attachment model with  $m \geq 2$  and  $\delta \in (-m, 0)$ . Then, with high probability and for all  $v \in [t] \setminus \text{Core}_t$ ,*

$$\text{dist}_{\text{PA}_t}(v, \text{Core}_t) \leq k_t^+ + 2h_t. \quad (6.52)$$

*Proof.* We start by analyzing  $v \in [t/2]$ . By Lemma 6.3, with high probability there exists  $a \in (0, 1]$  such that  $[t^a] \subseteq \text{Core}_t$ . Consider  $l > 1/a$ , and fix a vertex  $v \in [t/2]$ . Then, by Lemma 6.5 and with high probability, the  $k_t^+$ -exploration graph starting from  $v$  has at most  $l$  collisions before hitting  $\text{Core}_t$ . By Lemma 6.6 and with high probability, the number of vertices on the boundary of the  $k_t^+$ -exploration graph is at least  $N = s(m, l)(\log t)^{1+\varepsilon}$ . It remains to bound the probability that none of the  $N$  vertices on the boundary is a success, meaning that it does not reach  $\text{Core}_t$  in at most  $2h_t = 2\lceil B \log \log t + C \rceil$  steps.

By Lemma 6.9,

$$\mathbb{P}(S_{x_1}^c \cap \dots \cap S_{x_N}^c \mid \text{PA}_{t/2}) \leq (1 - \eta)^N = o(1/t), \quad (6.53)$$

thanks to the bound  $N \geq s(m, l)(\log t)^{1+\varepsilon}$ . This means that the probability that there exists a vertex  $v \in [t/2]$  such that its  $k_n^+$ -exploration graph is at distance more than  $A \log \log \log t$  from  $\text{Core}_t$  is  $o(1)$ . This proves the statement for all  $v \in [t/2]$ .

Next, consider a vertex  $v \in [t] \setminus [t/2]$ . Lemma 6.5 implies that the probability that there exists a vertex  $v \in [t] \setminus [t/2]$  such that its  $k_t^+$ -exploration graph contains more than one collision before hitting  $\text{Core}_t \cup [t/2]$  is  $o(1)$ . As before, the number of vertices on the boundary of a  $k_t^+$ -exploration graph starting at  $v \in [t] \setminus [t/2]$  is at least  $N \geq s(m, 1)m^{k_t^+} = s(m, 1)(\log t)^{1+\varepsilon}$ . We denote these vertices by  $x_1, \dots, x_N$ . We aim to show that, with high probability,

$$\Delta_N = \sum_{i=1}^N \mathbb{1}_{(x_i \in [t/2])} \geq N/4. \quad (6.54)$$

For every  $i = 1, \dots, N$ , there exists a unique vertex  $y_i$  such that  $y_i$  is in the  $k_t^+$ -exploration graph and it is attached to  $x_i$ . Obviously, if  $y_i \in [t/2]$  then also  $x_i \in [t/2]$ , since  $x_i$  has to be older than  $y_i$ . If  $y_i \notin [t/2]$ , then

$$\mathbb{P}(x_i \in [t/2] \mid \text{PA}_{y_i-1}) = \mathbb{P}(y_i \rightarrow [t/2] \mid \text{PA}_{y_i-1}) \geq \frac{1}{2}, \quad (6.55)$$

and this bound does not depend on the attaching of the edges of the other vertices  $\{y_j : j \neq i\}$ .

This means that we obtain the stochastic domination

$$\Delta_N \geq \sum_{i=1}^N \mathbb{1}_{(x_i \in [t/2])} \succeq \text{Bin}(N, \frac{1}{2}), \quad (6.56)$$

where we write that  $X \succeq Y$  when the random variable  $X$  is stochastically larger than  $Y$ . By concentration properties of the binomial,  $\text{Bin}(N, \frac{1}{2}) \geq N/4$  with probability at least

$$1 - e^{-N/4} = 1 - e^{-s(m, 1)(\log t)^{1+\varepsilon}/4} = 1 - o(1/t). \quad (6.57)$$

Thus, the probability that none of the vertices on the boundary intersected with  $[t/2]$  is a success is bounded by

$$\mathbb{P}(S_{x_1}^c \cap \dots \cap S_{x_{\Delta_N}}^c \mid \text{PA}_{t/2}) \leq (1 - \eta)^{N/4} + o(1/t) = o(1/t). \quad (6.58)$$

We conclude that the probability that there exists a vertex in  $[t] \setminus [t/2]$  such that it is at distance more than  $k_t^+ + 2h_t$  from  $\text{Core}_t$  is  $o(1)$ .  $\square$

This completes the proof of Statement 2.6, and thus of Theorem 1.5.  $\square$

## ACKNOWLEDGEMENTS

We thank the referees for their detailed comments, which improved the paper considerably.

The work of FC was partially supported by the ERC Advanced Grant 267356 VARIS and by the PRIN 20155PAWZB “Large Scale Random Structures”. The work of AG was partially supported by University of Milano Bicocca through an EXTRA scholarship that sponsored his visit to Eindhoven University of Technology to complete his master project in March-May 2014. The work of AG and RvdH is supported by the Netherlands Organisation for Scientific Research (NWO) through the Gravitation NETWORKS grant 024.002.003. The work of RvdH is also supported by NWO through VICI grant 639.033.806.

## REFERENCES

- [Blo09] Mindaugas Bloznelis. A note on loglog distances in a power law random intersection graph. *Preprint arXiv.org:0911.5127 [math.PR]*, 2009.
- [Bol01] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.
- [Bol81] Béla Bollobás. The diameter of random graphs. *Transactions of the American Mathematical Society*, 267(1):41–52, 1981.
- [BR04] Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.
- [BRST01] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18(3):279–290, 2001.
- [CGH16] Francesco Caravenna, Alessandro Garavaglia and Remco van der Hofstad. Diameter in ultra-small scale-free random graphs: Extended version. *arXiv.org: 1605.02714 [math.PR]*, 2016.
- [CL01] Fan Chung and Linyuan Lu. The diameter of sparse random graphs. *Advances in Applied Mathematics*, 26(4):257 – 279, 2001.
- [CL02] Fan Chung and Linyuan Lu. The average distances in random graphs with given expected degrees. *Internet Mathematics*, 1:15879–15882, 2002.
- [DHH10] Sander Dommers, Remco van der Hofstad, and Gerard Hooghiemstra. Diameters in preferential attachment models. *J. Stat. Phys.*, 139(1):72–107, 2010.
- [DMM12] Steffen Dereich, Christian Mönch, and Peter Mörters. Typical distances in ultrasmall random networks. *Adv. in Appl. Probab.*, 44(2):583–601, 2012.
- [EHH08] Henri van den Esker, Remco van der Hofstad, and Gerard Hooghiemstra. Universality for the distance in finite variance random graphs. *Journal of Statistical Physics*, 133(1):169–202, Oct 2008.
- [FR07] Daniel Fernholz and Vijaya Ramachandran. The diameter of sparse random graphs. *Random Structures Algorithms*, 31(4):482–516, 2007.
- [Hof17] Remco van der Hofstad. *Random Graphs and Complex Networks, Volume 1*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.
- [Hof18+] Remco van der Hofstad. *Random graphs and complex networks, Volume 2*. 2018+. Book in preparation.
- [HHVM05] Remco van der Hofstad, Gerard Hooghiemstra, and Piet Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures Algorithms*, 27(1):76–123, 2005.
- [HZH07a] Remco van der Hofstad, Gerard Hooghiemstra, and Dmitri Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Probab.*, 12:no. 25, 703–766, 2007.
- [HZH07b] Remco van der Hofstad, Gerard Hooghiemstra, and Dmitri Znamenski. A phase transition for the diameter of the configuration model. *Internet Math.*, 4(1):113–128, 2007.
- [HK17] Remco van der Hofstad and Júlia Komjáthy. When is a scale-free graph ultra-small? *Journal of Statistical Physics*, Sep 2017.
- [Jan09] Svante Janson. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, 18(1–2):205–225, 2009.
- [RW10] Oliver Riordan and Nick Wormald. The diameter of sparse random graphs. *Combinatorics, Probability and Computing*, 19(5-6):835–926, 2010.