Fixed speed competition on the configuration model with infinite variance degrees: unequal speeds

Enrico Baroni*   Remco van der Hofstad*   Júlia Komjáthy*

Abstract

We study competition of two spreading colors starting from single sources on the configuration model with i.i.d. degrees following a power-law distribution with exponent $\tau \in (2,3)$. In this model two colors spread with a fixed but not necessarily equal speed on the unweighted random graph. We show that if the speeds are not equal, then the faster color paints almost all vertices, while the slower color can paint only a random subpolynomial fraction of the vertices. We investigate the case when the speeds are equal and typical distances in a follow-up paper.

1 Introduction and results

1.1 The model and the main result

Let us consider the configuration model $\mathrm{CM}_n(\boldsymbol{d})$ on n vertices, where the degrees $D_v, v \in \{1, 2, \ldots, n\} := [n]$ are i.i.d. with a power-law tail distribution. That is, given the number of vertices n, to each vertex we assign a random number of half-edges drawn independently from a distribution F and the half-edges are then paired randomly to form edges. In case the total number of half-edges $\mathcal{L}_n := \sum_{v \in [n]} D_v$ is odd, then we drop one half-edge from D_n (see below for more details). We assume that

$$\frac{c_1}{x^{\tau-1}} \le 1 - F(x) = \mathbb{P}(D > x) \le \frac{C_1}{x^{\tau-1}}, \tag{1.1}$$

with $\tau \in (2,3)$, and all edges have weight 1. We assume $\mathbb{P}(D \ge 2) = 1$ guaranteeing that the graph has almost surely a unique connected component of size $n(1 - o(1))$ see e.g. [27, Theorem 10.1] or [35, 36].
We further denote the mass function of -1 plus the *size-biased version* of D by

$$f_j^* := \frac{(j+1)\mathbb{P}(D = j+1)}{\mathbb{E}[D]}, \ j \ge 0. \tag{1.2}$$

---

*Eindhoven University of Technology, The Netherlands. E-mail: e.baroni,rhofstad,j.komjathy@tue.nl.

We write $F^*(x)$ for the distribution function $F^*(x) = \sum_{0 < j \le x} f_j^*$.

Pick two vertices $\mathcal{R}_0$ (red source) and $\mathcal{B}_0$ (blue source) uniformly at random in $[n]$, and consider these as two sources of spreading infections. Each infection spreads *deterministically* on the graph: for color blue it takes λ time units to pass through an edge, while color red needs 1 unit of time for that. Without loss of generality we can assume that $\lambda > 1$. Each vertex is painted the color of the infection that reaches it first, keeps its color forever, and starts coloring the outgoing edges at the speed of its color. When the two colors reach a vertex at the same time, the vertex gets color red or blue with an *arbitrary adapted rule*, i.e. a rule that does not depend on the future. An example of such a rule is when the vertex is painted red or blue with probability $1/2$ each, independently of everything else. Another natural adapted rule is that a vertex, when the two colours arrive at it at the same time, is painted red or blue with probability proportional to the number of previously red and blue-colored neighbors of the vertex. Let $\mathcal{R}_t := \mathcal{R}_t(n)$ and $\mathcal{B}_t := \mathcal{B}_t(n)$ denote the set of red and blue vertices occupied up to time t, respectively and let R_t, B_t denote the sizes of these sets. In particular, let us denote by $R_\infty := R_\infty(n), B_\infty := B_\infty(n)$ the number of vertices eventually occupied by red and blue, respectively. We emphasise that the randomness in this model is only arising from the structure or topology of the graph and the uniform choice of the source vertices for the two colors; once these are settled, the dynamics is completely deterministic.

Roughly speaking, the first result of this paper, Theorem 1.2 below, tells us that in the quenched setting, i.e., for almost all realizations of the graph $\mathrm{CM}_n(\boldsymbol{d})$ and for almost all initial vertices $\mathcal{R}_0, \mathcal{B}_0$, the faster color wins, that is, it gets $n - o(n)$ many vertices. Furthermore, the number of vertices the slower color paints is subpolynomial in n. More precisely, blue paints whp $\exp\{(\log n)^{2/(\lambda+1)} G_n(Y_r^{(n)}, Y_b^{(n)})\}$ many vertices, i.e., a stretched exponential in $\log n$ with exponent $2/(\lambda + 1) < 1$, and where the coefficient $G_n(Y_r^{(n)}, Y_b^{(n)})$ is a random function that depends on n, λ, τ, and two random variables $Y_r^{(n)}$ and $Y_b^{(n)}$, that can intuitively be interpreted as some measure of 'how good' the neighbourhoods of the source vertices are: the faster the local neighbourhoods grow, the larger these variables are.[1] Moreover, $G_n(Y_r^{(n)}, Y_b^{(n)})$ does not converge: it has an oscillatory part that exhibits 'log log-periodicity'. We can 'filter out' these oscillations by introducing another function $H_n(Y_r^{(n)}, Y_b^{(n)})$ that oscillates and is tight, and then $G_n(Y_r^{(n)}, Y_b^{(n)})/H_n(Y_r^{(n)}, Y_b^{(n)})$ is a deterministic function of $Y_r^{(n)}, Y_b^{(n)}$ that converges, see the statement of Theorem 1.2 below.

The other result, Theorem 1.4, shows that the degree of the maximal-degree vertex that blue ever occupies obeys asymptotic behaviour similar to blue's total number, with a strictly smaller coefficient in the exponent, and the same $\log\log$-periodicity. This phenomenon is due to integer part issues coming from the fact that the edge weights are concentrated on a lattice.

Notation

We write $[n]$ for the set of integers $\{1, 2, \ldots, n\}$. We use the same notation and add a superscript $(r), (b)$ to corresponding random variables, sets or other quantities belonging to the red and blue processes, respectively. We write $E(\mathrm{CM}_n(\boldsymbol{d}))$ for the set of edges. For any set of vertices $S \subset [n]$, we write $N(S)$ for the set of their neighbors, i.e.,

$$N(S) = \{y \in [n] : \exists x \in S, (x, y) \in E(\mathrm{CM}_n(\boldsymbol{d}))\}. \tag{1.3}$$

---

[1] More precisely, these variables appear as the double exponential growth rate function of the local neighborhood of the two starting vertices.

For any event A, $\mathbb{P}_n(A) := \mathbb{P}(A | D_1, D_2, \ldots, D_n)$. As usual, we write i.i.d. for independent and identically distributed, lhs and rhs for left-hand side and right-hand side. We write $\lfloor x \rfloor, \lceil x \rceil$ for the lower and upper integer part of $x \in \mathbb{R}$, and $\{x\}$ for the fractional part of $x \in \mathbb{R}$. Slightly misusing the notation, we use curly brackets around set elements, events and exponents as well. We say that a sequence of events $\mathcal{E}_n$ occurs with high probability (whp) when $\lim_{n \to \infty} \mathbb{P}(\mathcal{E}_n) = 1$. In this paper, constants are typically denoted by c in lower and C in upper bounds (with possible indices), and their precise values might change even along lines. Typically, all the whp-events hold whp under the event $\{\mathcal{L}_n \in [1/2 \mathbb{E}[D]n, 2\mathbb{E}[D]n]\}$.

Results

To be able to state the main theorem precisely, let us define the following random variables:

Definition 1.1. Let $Z_k^{(r)}, Z_k^{(b)}$ denote the number of individuals in the kth generation of two independent copies of a Galton-Watson process described as follows: the size of the first generation has distribution F satisfying (1.1), and all the further generations have offspring distribution F^* from (1.2). Then, for a fixed but small $\rho > 0$ let us define

$$Y_r^{(n)} := (\tau - 2)^{t(n^\rho)} \log(Z_{t(n^\rho)}^{(r)}), \quad Y_b^{(n)} := (\tau - 2)^{\lfloor t(n^\rho)/\lambda \rfloor} \log(Z_{\lfloor t(n^\rho)/\lambda \rfloor}^{(b)}), \qquad (1.4)$$

where $t(n^\rho) = \inf_k \{Z_k^{(r)} \geq n^\rho\}$. Let us further introduce

$$Y_r := \lim_{k \to \infty} (\tau - 2)^k \log(Z_k^{(r)}), \quad Y_b := \lim_{k \to \infty} (\tau - 2)^k \log(Z_k^{(b)}). \qquad (1.5)$$

In Section 2 below we will see that these quantities are well defined and that $(Y_r^{(n)}, Y_b^{(n)}) \xrightarrow{d} (Y_r, Y_b)$ from (1.5) as $n \to \infty$. With this notation in mind, we have the following theorem:

Theorem 1.2. Fix $\lambda > 1$. Then, $\lim_{n \to \infty} R_\infty / n = 1$ whp. Further, there exists a bounded and strictly positive random function $H_n(Y_r^{(n)}, Y_b^{(n)})$ such that as $n \to \infty$

$$\frac{\log(B_\infty)}{(\log n)^{\frac{2}{\lambda+1}} H_n(Y_r^{(n)}, Y_b^{(n)})} \xrightarrow{d} \left(\frac{Y_b^\lambda}{Y_r} \right)^{\frac{1}{\lambda+1}}. \qquad (1.6)$$

We identify $H_n(Y_r^{(n)}, Y_b^{(n)})$ in (7.17) as a deterministic, oscillating (non-convergent) function of $\tau, \lambda, n, Y_r^{(n)}, Y_b^{(n)}$. $H_n(Y_r^{(n)}, Y_b^{(n)})$ has the uniform (non-tight) bounds

$$\left(\frac{(\tau - 2)}{(\tau - 1)^2} \frac{(3 - \tau)}{1 - (\tau - 2)^\lambda} \right)^{\frac{1}{\lambda+1}} < H_n(Y_r^{(n)}, Y_b^{(n)}) < (\tau - 2)^{-2} \cdot \frac{4 - \tau}{1 - (\tau - 2)^\lambda}.$$

Remark 1.3. We also give accompanying tight bounds on $H_n(Y_r^{(n)}, Y_b^{(n)})$, see (7.18).

Let us denote the degree of the maximal degree vertex eventually occupied by blue

$$D_{\max}^{(b,n)}(\infty) := \max_{i \in \mathcal{B}_\infty} D_i. \qquad (1.7)$$

As a side result of the proof of Theorem 1.2, we get the following theorem:

Theorem 1.4. Fix $\lambda > 1$. There exists a bounded and strictly positive random function $H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})$ defined below in (6.10), such that as $n \to \infty$

$$\frac{\log D_{\max}^{(b,n)}(\infty)}{(\log n)^{\frac{2}{\lambda+1}} H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})} \xrightarrow{d} \left(\frac{Y_b^\lambda}{Y_r} \right)^{\frac{1}{\lambda+1}} \qquad (1.8)$$

Further, $H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})$ is stochastically dominated by $H_n(Y_r^{(n)}, Y_b^{(n)})$ whp, and

$$\left(\frac{(\tau-2)^{2+\lambda}}{(\tau-1)^2} \right)^{\frac{1}{\lambda+1}} \leq H_n^{\max}(Y_r^{(n)}, Y_b^{(n)}) \leq \left(\frac{\tau-2}{4} \right)^{\frac{1}{\lambda+1}}. \tag{1.9}$$

Remark 1.5. We emphasise that these results are valid for any adapted decision rule when the two colors jump at the same time to a vertex. In case λ is irrational, clearly, this rule will never be used. If λ is rational, then the normalisation random variables $H_n(Y_r^{(n)}, Y_b^{(n)})$, $H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})$ depend on the rule – they are slightly different if the rule is so that these vertices are always painted red, from the case when there is a positive chance that these vertices are painted blue, but the upper and lower bounds on $H_n(Y_r^{(n)}, Y_b^{(n)})$, $H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})$ remain the same. On the other hand, when $\lambda = 1$, this rule will play an important role in the outcome. This case is more complex and can be found in the follow-up paper [29].

Remark 1.6 (More than two colors)**.** If there are a finite number of colors with edge passage-times $1 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_k$, then the statements of Theorems 1.2 and 1.4 stay valid for each λ_i, $2 \leq i \leq k$, with limit variables $(Y_i^{\lambda_i}/Y_1)^{1/(\lambda_i+1)}$ on the right hand side of (1.6), where Y_i are i.i.d. copies of Y. The reason for this is that with high probability each slower color only meets the fastest color and never meets the other slow ones. That is, the clusters of slower colors are separated from each other by the cluster of the fastest color.

1.2 Related work and discussion

First we give a (non-complete) overview of the literature on competition on different graph models. Then we mention some more applied results.

In a seminal paper [24] Häggström and Pemantle introduced competition on the grid $\mathbb{Z}^d$. The model is called the two-type Richardson's model, and it describes the dynamics of two (red and blue) infections with single source vertices $v_0, v_1 \in \mathbb{Z}^d$ that compete to conquer the grid $\mathbb{Z}^d$. In this continuous-time model, a vertex of $\mathbb{Z}^d$ gets a given color with rate proportional to the number of infected neighbours of that color; then, once a vertex is infected, it keeps its color forever. Note that the evolution of a single color without the presence of the other color has independent exponential passage times across edges, and a vertex gets infected at the time that equals the minimal length path from the source to the vertex. Hence, a single color process is often called *first passage percolation* in the literature. Multiple colours then lead *competing first passage percolation*.

For two colors, we have two possible evolution scenarios: in the first, one of the growing clusters completely blocks the growth of the other color – by surrounding it – and then it infects all the remaining healthy vertices. In the second scenario the two clusters continue to grow unboundedly forever: this is called *coexistence*. The important question is: does coexistence occur with positive probability? Häggström and Pemantle [24] proved that this is the case for the $\mathbb{Z}^2$ grid with i.i.d. exponential passage times. Later this result was extended by Garet and Marchand [21] for $\mathbb{Z}^d$, $d \geq 2$ for a vast class of passage time distributions under mild hypothesis. For further literature on the Richardson model see [14, 15, 22, 25, 26].

Recently, a noticeable scientific interest arose in understanding the structure of large but finite networks and the behaviour of spreading processes on these networks. Typically, results on these topics are called first passage percolation, see e.g. [6, 7, 8]. It is then natural to ask what happens when one considers competition of multiple spreading processes on these networks. When studying competitive spreading, one might also gain a more detailed understanding of the structure of these graphs.

The idea of competitive spreading on finite random graph sequences raises several questions. First and foremost, due to the finite size of the graphs the main questions about these models must be rephrased, since infinite growth can never happen. Thus, the definition of coexistence had to be modified in this setting. Consider two competing colors on a sequence of random graphs: is there an asymptotic coexistence of the two colors? That is, is it possible that both colors paint a positive proportion of vertices with positive probability, as the size of the graph tends to infinity? If this is not the case then can we determine the number of eventually occupied vertices for both colors in terms of the size of the graphs? What happens if we modify the passage dynamics so that the two infections have different rates of growth λ_1 and λ_2? Here we give a (non-complete) overview of the existing literature on these topics for different random graph models.

Antunovic, Dekel, Mossel and Peres [2] give a detailed analysis of competition on random regular graphs (degree at least 3) on n vertices with i.i.d. exponential edge weights. They analyse the number of eventually occupied vertices by both colors as a function of the speeds λ_1, λ_2 and of the initial number of infected vertices, that might even grow with n. They show that asymptotically almost surely the color with higher rate occupies $n - o(n)$ vertices and the slower color paints approximately n^β vertices for some deterministic function $\beta(\lambda_1, \lambda_2)$. Their result include asymptotic coexistence for equal speeds $\lambda_1 = \lambda_2$ for infections starting from single sources.

Next, van der Hofstad and Deijfen [16] investigate competition with exponential spreading times on the configuration model with i.i.d. degrees arising from a power-law distribution with exponent $\tau \in (2, 3)$. They prove that even if the speeds are not equal, the 'winner' color is *random*, i.e. the color with slower rate can still take most of the graph. Moreover, the winning color paints all but a finite number of vertices. The randomness of the 'winner' color comes from the fact that the underlying Markov branching process explodes in finite time, and the slower color has a positive chance to explode earlier than the faster color.

A slightly different, discrete time competition model is analysed by Antunović, Mossel and Rácz in [3]. There, the underlying random graph is the growing linear preferential attachment model, and vertices pick their color upon entering the network randomly from the colors of the vertices they attach to. The probability of picking a color is a (possibly linear) function of the number of neighbors with the given color, called the *coloring function*. The authors analyse coexistence of colors in terms of the properties of the coloring function. Note that in this case the graph has power law $\tau = 3$. The proofs are based on a comparison to Pólya urns.

Finally, this paper considers competition on the configuration model with i.i.d. power-law degrees with exponent $\tau \in (2, 3)$, but with *deterministic* unit edge-weights. Theorem 1.2 shows that the fact that the edge weights have a support separated from zero entirely changes the picture observed in [16]: when the speeds are unequal, the faster color always paints $n - o(n)$ vertices, and the slower color can paint only subpolynomial many vertices.

If the speeds are equal, then the phenomena is richer ([29]): as a side result of the analysis of the $\lambda = 1$ case, we obtain precise distributional limits of the second order terms in typical distances in the graph, as in [28]. Further, for the adapted tie-breaking rule of flipping an independent coin every time with some fixed positive probability, we show there that the following phenomena occurs: when the random variables $Y_r^{(n)}, Y_b^{(n)}$ are far enough from each other, i.e., one of the source vertices has a much faster growing local neighborhood than the other, then, the one with the better neighborhood occupies almost every vertex in the graph, while the loser color can paint polynomially many vertices with a random exponent that is less than 1. However, when the two variables are close to each other (within a factor of $\tau - 2$ of each other), then there is coexistence

and both colors can occupy asymptotically linearly many vertices. Due to the length of the analysis of this case and to highlight the richer phenomena that comes with it, we decided to put the equal-speed case in a subsequent paper, see [29].

From the more applied perspective, competition on networks is present in many aspects of our life. To start with an example, in marketing, companies compete for customers who are connected via their acquaintance network, and they provide word-of-mouth recommendations and opinions about the services of the different companies, see [19, 20]. For economic studies on the importance of word-of-mouth, see e.g. [4, 11, 18]. Recently, 'word-of-mouth' recommendations happen also on large scale in different social online media such as Facebook and Twitter. For a survey on how online feedback mechanisms differ from original word-of-mouth recommendations and what challenges they pose, see [17]. The paper [33] analyses recommendation-based viral marketing on social media, where they use viral marketing also to identify communities of online networks. For recent economic studies of the importance of word-of-mouth recommendations, see e.g. [12, 32].

In epidemiology, viruses and bacterial infections spread through populations. In this setting, competition can happen among different strains of a pathogen, see e.g. [34] for a study under what conditions coexistence can occur. In the physics community, [1, 31, 37] study the effect of the underlying network on co-existence of competing viruses.

The epidemiological analogies have been further exploited by [38], where they study a variation of susceptible-infectious-susceptible epidemic spread, in which two epidemics are immune to each other, and the authors show that one of them completely takes over (similarly as in [16]). Then, [5] studies how partial immunity can cause coexistence in the previous model.

Discussion and open problems

The analysis of competition on the configuration model is far from complete. One can for instance ask about different spreading dynamics (edge lengths) and different power-law exponents. Further, one can ask what happens if the colors have entirely different passage time distributions (e.g. one is explosive and the other is not), or what happens if one of the colours have a main advantage by starting from one or many initial vertices of very high degree. These can correspond to e.g. competition advantage (early birds) of different products on the network or to different marketing strategies. Here we list some conjectures for uniformly picked single vertex sources of infections on $\mathrm{CM}_n(\boldsymbol{d})$ with i.i.d. power law degrees of distribution D with exponent τ. We further assume that the passage times through edges can be represented as i.i.d. random variables, from distribution I_r, I_b for red and blue, respectively. For a more thorough list of open problems, see [29].

Competition problems for infinite variance degree graphs, $\tau \in (2, 3)$:

A. If the spreading dynamics are so that the underlying branching processes defined by D, I_r and D, I_b are both explosive, then we conjecture that there is *never* coexistence and either of the two colors can win. This is one of our ongoing research projects.

B. If the underlying branching process for one color has explosive spreading while the other one has not, than we conjecture that the explosive one always wins. This case could be even relaxed to show that this is still true whenever the non-explosive color starts from some function $f(n) = o(n^\varepsilon)$ many half-edges in total, for any $\varepsilon > 0$.

C. If both underlying branching processes are non-explosive, and further assume $I_r \overset{d}{=} \lambda I_b$, then we guess that there is no coexistence if $\lambda \neq 1$ (the fastest color wins). We suspect that the number of vertices the 'loser' color paints depends sensitively on the weight distribution. The outcome in the $\lambda = 1$ case might sensitively depend on the weight distribution.

Competition problems for finite variance degree graphs, $\tau > 3$:

D. We suspect that if the transmission times I_r, I_b both have continuous distribution, and the branching process approximations of them have different Malthusian parameters, then there is no coexistence, and the number of vertices painted by the slower color is n^β for some $\beta \in (0,1)$. When the Malthusian parameters agree, we suspect that there is asymptotic co-existence.

Competition problems for the critical power-law exponent, $\tau = 3$:

E. In this case $\mathbb{P}(D > x) = L(x)/x^2$, with $L(x)$ a slowly varying function at infinity. We suspect that $L(x)$ and the transmission distributions I_r, I_b jointly determine into which category among A, C, D above the spreading of the colours belongs to.

1.3 Overview of the proof and structure of the paper

The heuristic idea of the proof is as follows: we can start growing the two clusters simultaneously. The growth has six phases, each corresponding to a section below, described as follows:

(i) *Branching process phase.*
   At first, whp, the two colored clusters do not meet and the growth of both clusters is characterised by the growth rate of the branching process (BP) to which they can be coupled. This we call the *branching process phase*. The length of this phase is of order $\log \log n / |\log(\tau - 2)| + O(1)$. Then, the faster color (red) reaches the area where the coupling fails to remain valid: R_t reaches size n^ϱ for some $\varrho > 0$.

(ii) *Mountain climbing phase.*
   At this point, we start making use of the structure of high-degree vertices in the graph: due to high connectivity, the subgraph formed by high-degree vertices can be represented as a 'mountain' where the height function is linear in the $\log \log$-degree. Level sets of this mountain represent vertices with degree of the same order of magnitude, with the maximal degree in the graph at the top of the mountain. We partition this mountain into *layers* – that is, constant length intervals on a $\log \log$-scale – and we show that every vertex in a given layer has at least one neighbour in one layer higher. As a result, we show the existence of a path for red through these layers of vertices of higher and higher degree such that the path reaches some vertex with degree larger than $n^{(\tau-2)/(\tau-1)}$ at the end. This we call the *mountain climbing phase*. The climbing phase lasts only finitely many steps, but the constants turns out to be important, so we perform a rather careful analysis. We denote the total time of the branching process phase and the climbing phase for red by T_r.

(iii) *Crossing the peak of the mountain.*
   We handle how the color red goes through the peak of the 'degree-mountain' very carefully. Vertices of degree much larger than $\sqrt{n}$ form a subgraph that is a complete graph, hence it takes only one step to paint all the very high degree vertices, but the degree of vertices to which the faster color arrives at the end of this single step is delicately depending on the initial random growth rates of the branching processes and their integer and fractional parts.

(iv) *Red avalanche from the peak.*
   After crossing the mountain, red starts sloping down to layers of vertices of smaller and smaller degree. Since it is still true that each vertex in a layer is connected to at least one vertex in one layer higher, this means that in each additional step, red paints *all* the vertices in one layer lower. We call this the *avalanche-phase of red*. (One can imagine this as red being a very careless climber who – after crossing the

peak of a mountain – steps in the snow with a bucket of red paint and starts a huge painted avalanche.)

(v) *At the collision time.*
Now we turn our attention to the blue climber who does essentially the same as red except that it is slower: after getting out of its local neighbourhood corresponding to the branching process, blue starts its mountain climbing phase as well. Since it is slower, whp it will only reach some low layer of the degree-mountain when red starts its avalanche. With this picture in hand, we can identify the maximal degree vertex eventually painted blue - this is the vertex in the highest layer blue can still reach. The idea of the proof is to determine the value ℓ such that during the total time $T_r + \ell$, blue has climbed up to the same layer as the red avalanche has sloped down to. Since red occupies *every vertex* in a layer it reaches, it will necessarily bump into blue, who whp reaches only some vertices in that layer. This determines the time when red starts successfully blocking blue.

(vi) *Competing with the avalanche.*
After the meeting time $T_r + \ell$, blue cannot go higher up on the mountain since red already occupies every vertex having degree higher than the maximal degree of blue. Note that at this time most of the graph is still not reached by any color: we need to estimate the number of vertices that blue can still reach before the red avalanche closes up around the blue cluster. This is done in two steps: heuristically, every vertex that is close enough to a blue half-edge occupied at or before $T_r + \ell$ has a high chance to become blue later. Hence, first we calculate the size of the 'optional cluster of blue', i.e., we calculate the size of the k-neighborhood of blue half-edges via path counting methods. The size of the optional cluster is convergent if $k \to \infty$: due to the presence of the red avalanche, the degrees in the blue paths get more and more restricted and finally the red avalanche reaches constant order vertices and then the procedure stops.

It can still happen that some vertices in the optional cluster of blue are occupied by red simply because they are 'accidentally' also close to some red vertex. Thus, in the second step we estimate the size of the intersection between the optional cluster of blue and the red cluster. The two steps together provide a matching upper and lower bound for the number of vertices that blue occupies after the intersection. This phase has a non-negligible impact on the order of magnitude of vertices painted blue since the constant H_n in the denominator of (1.6) is influenced by this last phase.

2 The branching process phase

First we describe the exploration process of the local neighbourhood of a given vertex in order to relate it to a branching process.

The configuration model $\mathrm{CM}_n(\boldsymbol{d})$ (introduced in [10], for more see [9, 27]) on n vertices with i.i.d. degree distribution D can be briefly described as follows: for each vertex $i \in [n]$ we assign an i.i.d. random variable $D_i \sim D$, and attach D_i half-edges to that vertex. If the total degree $\mathcal{L}_n = \sum_{i=1}^n D_i$ is odd, then we add an extra half-edge to the vertex n. Then we number the half-edges in an arbitrary way from 1 to $\mathcal{L}_n$, and start pairing them uniformly at random, i.e. we pick an arbitrary unpaired half-edge and pair it to a uniformly chosen other unpaired half-edge to form an edge. Once paired, we remove them from the set of unpaired half-edges and continue the procedure until all half-edges are paired. We call the resulting multi-graph $\mathrm{CM}_n(\boldsymbol{d})$. Since the choice of the half-edge to be paired is arbitrary, we can start from any set of vertices, and explore their cluster

simultaneously with the construction of the graph. We call this procedure the *exploration process*, which is a version of a Breadth First Search Algorithm on the random graph $\mathrm{CM}_n(\boldsymbol{d})$. We describe the exploration process in more detail for the case when the initial set is a single uniformly chosen vertex $v \in [n]$, and relate it to a corresponding branching process as follows.

In each step of the exploration process, each vertex belongs to exactly one of three sets: it can be active ($\mathcal{A}$), explored ($\mathcal{E}$) or unexplored ($\mathcal{U}$). Initially $\mathcal{E}_0 = \varnothing$ and all vertices except v are in $\mathcal{U}_0$. We start setting the status of the initial vertex v to active: $\mathcal{A}_0 = \{v\}$, and we write $\mathcal{A}_i$ for the set of active vertices after the ith step of the exploration. In each step we pick a vertex v_{i+1} from $\mathcal{A}_i$ (we do this first-in-first-out way, i.e., we keep track of when a vertex enters the set $\mathcal{A}$) and do three things: remove v_{i+1} from $\mathcal{A}_i$; add it to the explored vertices $\mathcal{E}_i$; and put all its unexplored neighbors in the active set of vertices, i.e.,

$$\mathcal{A}_{i+1} := \mathcal{A}_i \setminus \{v_{i+1}\} \cup \{N(v_{i+1}) \cap \mathcal{U}_i\},$$
$$\mathcal{U}_{i+1} := \mathcal{U}_i \setminus N(v_{i+1}),$$

where $N(v_{i+1})$ denotes the neighbors of v_{i+1} in $\mathrm{CM}_n(\boldsymbol{d})$. The explored vertices form the sequence $\mathcal{E}_i = \{v = v_1, v_2, \ldots, v_i\}$.

Let D_i^f stand for the *forward-degree* of the vertex v_i in the exploration process, so that $D_i^f = |N(v_i) \cap \mathcal{U}_i|$. We aim to determine the distribution of D_i^f. For this we note that in the construction of the random graph $\mathrm{CM}_n(\boldsymbol{d})$, an *arbitrary* half-edge is chosen and paired to a uniformly chosen unpaired half-edge. Hence, we can do the construction of the graph together with the exploration process. Further, the probability of picking a half-edge which is belonging to a vertex with degree $j + 1$ is proportional to $(j+1)\mathbb{P}(D = j + 1)$, and as long as the size of the neighbourhood is small the probability that a vertex is connected to some vertex explored earlier vanishes. Hence, we get the size-biased distribution (1.2) as a natural candidate for the forward degrees of the vertices v_i in the exploration process. More precisely, we have the following result:

Proposition 2.1. [6, Proposition 4.7] *There exists $0 < \rho < 1$ such that the random vector $\left(D_i^f\right)_{i=2}^{n^\rho}$ of forward degrees can be coupled to an independent sequence of random variables $\left(D_i^\star\right)_{i=2}^{n^\rho}$ with probability mass function given in (1.2) and $\left(D_i^f\right)_{i=2}^{n^\rho} = \left(D_i^\star\right)_{i=2}^{n^\rho}$ whp.*

Proof. See [6, Proposition 4.5] and the proof of Proposition 4.7 in [6, Appendix A.2]. □

In our case, we have two source vertices red and blue with different spreading speed, thus, we need a slight modification of this proposition. Namely, we need that a similar coupling remains valid for two exploration processes from two uniformly chosen vertices up to the time when the red (first) color reaches size n^ρ. Let us temporarily denote the number of vertices occupied by blue (the other) color by this time by $h(n, \rho)$. This coupling is similar to [6, Proposition 4.8], but we state it for the reader's convenience:

Lemma 2.2. *Fix $\lambda \geq 1$. Let $t(n^\rho) := \inf\{t : R_t \geq n^\rho \text{ or } B_t \geq n^\rho\}$. Then there exists a $\rho > 0$ such that $\mathcal{R}_{t(n^\rho)} \cap \mathcal{B}_{t(n^\rho)} = \varnothing$ whp, and the forward degrees in both the red and the blue process can be coupled to i.i.d. sequences $\left(D_i^{\star(r)}\right)_{i=2}^{n^\rho}$ and $\left(D_j^{\star(b)}\right)_{j=2}^{h(n,\rho)}$ from distribution $F^\star$, where $h(n, \rho)$ is the random number of vertices reached by blue up to time $t(n^\rho)$. (The same statement holds true for $\lambda = 1$, but in this case the color to reach n^ρ vertices first is random.)*

Proof of Lemma 2.2. Consider the following version of the exploration process: start with the red source and do breath-first exploration, but one vertex at a time within the generation, giving the indices of $D_i^{f(r)}$ of the forward degrees. Once we finish a

generation, we move on the the next one. It is only important that in total we discover no more than n^ρ vertices when we stop the exploration. The last generation completely explored this way is random. First assume that $\lambda > 1$: consider the coupling given by Proposition 2.1 with forward degrees $\left(D_i^{\star(r)}\right)_{i=2}^{n^\rho}$ for the red cluster. After this, drop all the other active vertices from $\mathcal{A}_{t(n^\rho)}$ and re-start the exploration process with only vertex $\mathcal{B}_0$ being active. Since it takes time λ to cover an edge for blue, up to time $t(n^\rho)$ blue reaches all the vertices which have graph distance at most $\lfloor t(n^\rho)/\lambda \rfloor$ from the source vertex $\mathcal{B}_0$. Thus, continue the exploration process from the blue source up to finishing generation $\lfloor t(n^\rho)/\lambda \rfloor$. Since $\lambda > 1$, the total number of vertices found by this second phase has smaller order than n^ρ, so that the coupling still remains valid. Moreover, since each of the clusters have only at most n^ρ many vertices, with high probability they do not meet each other. Further, when $\lambda = 1$, the proof is the same, the first cluster to reach n^ρ vertices takes the role of red, and the other one takes the role of blue. □

An immediate consequence of Proposition 2.1 and Lemma 2.2 is that locally we can consider the growth of $\mathcal{R}_t$ and $\mathcal{B}_t$ as independent branching processes $(Z_k)_{k>0}$ with offspring distribution F^* for the second and further generations, and with offspring distribution given by F for the first generation.

Let us now investigate the growth of these branching processes. Since $\tau \in (2,3)$, the offspring distribution of this branching process has infinite mean for every individual in the second and later generations. To understand the behavior of this BP, we first look at what happens in a BP where all the degrees are distributed as F^*, including the first generation.

The following theorem by Davies [13] describes the growth rate of such a branching process:

Theorem 2.3 (Branching process with infinite mean [13])**.** *Let $\widetilde{Z}_k$ denote the k-th generation of a branching process with offspring distribution given by the distribution function F^*. Suppose there exists an $x_0 > 0$ and a function $x \mapsto \kappa(x)$ on $\mathbb{R}^+$ that satisfies the following conditions:*

(i) *$\kappa(x)$ is non-negative and non-increasing,*

(ii) *$x^{\kappa(x)}$ is non decreasing,*

(iii) *$\int_0^\infty \kappa\left(e^{e^x}\right) \mathrm{d}x < \infty$.*

Let us assume that for some $\tau \in (2,3)$, the tail of the offspring distribution satisfies that, for all $x \geq x_0$,

$$x^{-(\tau-2)-\kappa(x)} \leq 1 - F^*(x) \leq x^{-(\tau-2)+\kappa(x)}. \tag{2.1}$$

Then $(\tau - 2)^k \log(\widetilde{Z}_k \vee 1)$ converges almost surely to a random variable $\widetilde{Y}$. Further, the variable $\widetilde{Y}$ has exponential tails.

To be able to apply this theorem to our setting, we need to show that the distribution function F^* satisfies the condition (2.1). This is clearly the case since using the elementary re-arrangement of weights

$$1 - F^*(x) = \sum_{j=x+1}^\infty \frac{(j+1)\mathbb{P}(D=j)}{\mathbb{E}[D]} = \frac{1}{\mathbb{E}[D]}\left((x+2)[1-F(x+1)] + \sum_{j=x+2}^\infty [1-F(j)]\right),$$

combined with the bounds in (1.1) and elementary estimates immediately yields that there exist constants $0 < c_1^* \leq C_1^* < \infty$, such that for x large enough

$$x^{-(\tau-2)}c_1^* \leq 1 - F^*(x) \leq x^{-(\tau-2)}C_1^*. \tag{2.2}$$

Since $\mathbb{P}(D \geq 2) = \mathbb{P}(B \geq 1)$, these branching processes cannot die out, i.e., we can write $\log \widetilde{Z}_k$ instead of $\log(\widetilde{Z}_k \vee 1)$ and apply Davies' theorem to obtain the a.s. convergence of $\widetilde{Y}_k = (\tau - 2)^k \log(\widetilde{Z}_k)$ to a random variable $\widetilde{Y}$. Recall that the degree of the first vertex in the exploration process is distributed as F not F^*, hence we denote by Z_k the corresponding BP and call it the *delayed branching process*. The next lemma identifies the distribution of the limit of the properly scaled delayed branching process. We also identify the limit random variable Y in terms of $\widetilde{Y}$.

Lemma 2.4. *Let Y be the limiting random variable $\lim_{k \to \infty} (\tau - 2)^k \log Z_k$ of the delayed BP. Then Y satisfies the distributional identity*

$$Y \overset{d}{\equiv} (\tau - 2) \max_{1 \leq i \leq D} \widetilde{Y}^{(i)}, \tag{2.3}$$

where $\widetilde{Y}^{(i)}$ are i.i.d. copies of the limiting random variable of the original non-delayed BP. Further,

$$\widetilde{Y} \overset{d}{\equiv} (\tau - 2) \max_{1 \leq i < D^*} \widetilde{Y}^{(i)}.$$

Remark 2.5. An elementary calculation using (2.3) shows that Y also has exponential tails: we know from [13] that $\widetilde{Y}$ satisfies $\lim_{x \to \infty} \log(\mathbb{P}(\widetilde{Y} > x))/x = 1$. Then

$$\mathbb{P}(Y > x) = 1 - \mathbb{P}(\max_{1 \leq i \leq D} \widetilde{Y}_i < x/(\tau - 2)) = 1 - \mathbb{E}\left[\left(1 - \mathbb{P}(\widetilde{Y}_i > x/(\tau - 2))\right)^D\right].$$

Using that $D \geq 1$ and that $(1 - x)^D \geq 1 - Dx$, we obtain the following lower and upper bounds:

$$\mathbb{P}(\widetilde{Y}_i > x/(\tau - 2)) \leq \mathbb{P}(Y > x) \leq \mathbb{E}[D]\mathbb{P}(\widetilde{Y}_i > x/(\tau - 2)).$$

Taking logarithm an dividing by x on both sides yields that $\lim_{x \to \infty} \log(\mathbb{P}(Y > x))/x = 1/(\tau - 2)$.

Proof of Lemma 2.4. Since the number of offsprings in the first generation is distributed as D, by the branching property the subtrees starting from the first generation up to level k are distributed as $\widetilde{Z}_{k-1}$ and are independent of each other. Thus, for every $k \geq 1$,

$$Z_k \overset{d}{\equiv} \sum_{i=1}^{D} \widetilde{Z}_{k-1}^{(i)}, \tag{2.4}$$

where $\widetilde{Z}_{k-1}^{(i)}$ are i.i.d. copies of $\widetilde{Z}_{k-1}$. Hence

$$Y = \lim_{k \to \infty} (\tau - 2)^k \log(Z_k) = \lim_{k \to \infty} (\tau - 2)^k \log\left(\sum_{i=1}^{D} \widetilde{Z}_{k-1}^{(i)}\right). \tag{2.5}$$

We can bound the right hand side from both sides:

$$(\tau - 2)^k \log\left(\max_{i=1,\ldots,D} \widetilde{Z}_{k-1}^{(i)}\right) \leq (\tau - 2)^k \log\left(\sum_{i=1}^{D} Z_{k-1}^{(i)}\right) \leq (\tau - 2)^k \log\left(D \cdot \max_{i=1,\ldots,D} \widetilde{Z}_{k-1}^{(i)}\right). \tag{2.6}$$

Clearly $(\tau - 2)^k \log D \overset{\mathbb{P}}{\to} 0$, and by monotonicity we can exchange $\log$ and $\max$ and use Theorem 2.3 for the convergence of $(\tau - 2)^{k-1} \log(\widetilde{Z}_{k-1}^{(i)})$. Thus combining (2.5) with (2.6) yields

$$Y = \lim_{k \to \infty} \max_{i=1,\ldots,D} (\tau - 2)\widetilde{Y}_{k-1}^{(i)}. \tag{2.7}$$

Exchanging the limit with the maximum finishes the proof. The second statement of the lemma can be proved analogously. □

3 Mountain-climbing phase

In this section we describe the mountain-climbing phase. From now on we will concentrate on the growth of the *red* (the faster) cluster, but the very same methods will later be used for blue as well. Thus, in this section we neglect the superscript (r), and temporarily every quantity is belonging to the red cluster. We denote the set of red vertices at time t by $\mathcal{R}_t$ and its size by R_t. Since Proposition 2.1 only guarantees the coupling as long as the total number of explored vertices by red is at most n^ϱ for some $\varrho > 0$, let us first set some $\varrho' < \varrho(\tau - 2)^2$ and define

$$t(n^{\varrho'}) = \inf\{k : Z_k \geq n^{\varrho'}\}.$$

Note that by Lemma 2.2, and the fact that the total size of earlier generations are whp negligible compared to the last generation, $t(n^\rho) = T(n^{\varrho'})$ whp. Recall Definition 1.1, i.e.,

$$Y_r^{(n)} := (\tau - 2)^{t(n^{\varrho'})} \log Z_{t(n^{\varrho'})}. \tag{3.1}$$

Note that $t(n^{\varrho'})$ and thus $Y_r^{(n)}$ depends on n. Then, an easy calculation yields that with $\{x\} = x - \lfloor x \rfloor$,

$$t(n^{\varrho'}) = \frac{\log(\varrho'/Y_r^{(n)}) + \log \log n}{|\log(\tau - 2)|} + 1 - a_n, \tag{3.2}$$

where

$$a_n = \left\{ \frac{\log(\varrho'/Y_r^{(n)}) + \log \log n}{|\log(\tau - 2)|} \right\}. \tag{3.3}$$

Note that $1 - a_n$ is there to make the expression on the rhs of $t(n^{\varrho'})$ equal to its upper integer part. Due to this effect, the last generation has a bit more vertices than $n^{\varrho'}$, so let us introduce the notation ϱ'' for the random exponent of the overshoot

$$Z_{t(n^{\varrho'})} = n^{\varrho'(\tau - 2)^{a_n - 1}} := n^{\varrho''}, \tag{3.4}$$

We get this expression by rearranging (3.1) and using the value $t(n^\rho)$ from (3.2). The property $\varrho' < \varrho(\tau - 2)^2$ guarantees that the coupling is still valid, i.e. we can also couple the *degrees* of vertices in the $t(n^{\varrho'})$th generation of the branching process to i.i.d. size biased degrees.

After time $t(n^{\varrho'})$, we stop the coupling and focus on the graph: we start decomposing the graph to the following nested sets of vertices, that we call *layers*:

$$\Gamma_i := \{v : D_v > u_i\}, \tag{3.5}$$

where u_i is defined recursively by

$$u_{i+1} = \left(\frac{u_i}{C \log n} \right)^{1/(\tau - 2)}, \quad u_0 := \left(\frac{n^{\varrho''}}{C \log n} \right)^{1/(\tau - 2)} \tag{3.6}$$

for a large enough constant $C > 0$. We will see below that e.g. $C = 8/c_1$ is sufficient, where c_1 is from (1.1). It is not hard to see that

$$u_i = n^{\varrho''(\tau - 2)^{-(i+1)}} (C \log n)^{-e_i} \quad \text{with} \quad e_i = \frac{1}{3 - \tau} \left(\left(\frac{1}{\tau - 2} \right)^{i+1} - 1 \right). \tag{3.7}$$

Note that since $(\tau - 2)^{-1} > 1$, u_i is growing, hence $\Gamma_0 \supset \Gamma_1 \supset \Gamma_2 \supset \ldots$. First we need to show that $Z_{t(n^{\varrho'})}$ has a nonempty intersection with the initial layer Γ_0, and then we will build a path through the layers. The following lemma is a general lemma about the maximum of i.i.d. power-law random variables. It guarantees that $\mathcal{R}_{t(n^{\varrho'})} \cap \Gamma_0 \neq \varnothing$, and will also be repeatedly used to determine the maximum degree in a set of vertices:

Lemma 3.1. *Let X_i, $i = 1, \ldots, m$ be i.i.d. random variables with power-law tail exponent α, i.e. the distribution function of X_i satisfies (1.1) with $\tau - 1$ replaced by any $\alpha > 0$. Then for $K > 0$ and any $n > 1$,*

$$\mathbb{P}\left(\max_{i=1,\ldots,m} X_i < \left(\frac{m}{K \log n} \right)^{1/\alpha} \right) \leq \frac{1}{n^{c_1 K}}, \tag{3.8}$$

where c_1 arises from (1.1).

Proof. For the second statement, due to the independence the left hand side can be rewritten as follows:

$$\mathbb{P}\left(X_1 < \left(\frac{m}{K \log n} \right)^{1/\alpha} \right)^m \leq \left(1 - \frac{c_1 K \log n}{m} \right)^m \leq \exp\{-c_1 K \log n\},$$

where we used that $\mathbb{P}(X_1 < x) \leq 1 - c_1/x^\alpha$. $\qquad\square$

Note that the distribution F^* satisfies the condition of the lemma with $\alpha = \tau - 2$, see (2.2). So, we can apply this lemma (specially (3.8)) in the following setting: the i.i.d. variables X_i are the forward degrees $(B_i)_{i=1,\ldots,Z_{t(n\varrho')}} \sim F^*$ in the last generation of the branching process, thus $m := Z_{t(n\varrho')} = n^{\varrho''}$ and $\alpha = \tau - 2$. Note that the bound we get when applying (3.8) states that whp there is at least one vertex with degree at least u_0 (defined in (3.6)). Hence, we get that $\Gamma_0 \cap \mathcal{R}_{t(n\varrho')} \neq \varnothing$ whp.

We will repeatedly use concentration of binomial random variables of the following form

Lemma 3.2 (Concentration of binomial random variable). *Let X be a binomial random variable with parameters n, p_n. Then*

$$\begin{aligned} \mathbb{P}(X \geq 2\mathbb{E}[X]) &\leq \exp\{-\mathbb{E}[X]/8\}, \\ \mathbb{P}(X \leq \mathbb{E}[X]/2) &\leq \exp\{-\mathbb{E}[X]/8\}. \end{aligned} \tag{3.9}$$

Proof. Follows from standard estimates, see e.g. [27, Theorem 2.19] or [23]. $\qquad\square$

In what follows, we will build a path from $\Gamma_0 \cap \mathcal{R}_{t(n\varrho')}$ to the highest-degree vertices through successive layers Γ_i. The following lemma guarantees the existence of such a path. Recall that $N(S)$ stands for the neighbors of the set S in $\mathrm{CM}_n(\boldsymbol{d})$.

Lemma 3.3. *With u_i and Γ_i defined as in (3.7) and (3.5), for every $v \in \Gamma_i$, whp there is a vertex $w \in \Gamma_{i+1}$, such that $(v, w) \in E(\mathrm{CM}_n(\boldsymbol{d}))$. Equivalently,*

$$\Gamma_i \subset N(\Gamma_{i+1}) \quad whp. \tag{3.10}$$

Furthermore, the previous statement can be applied repeatedly *to build a path from Γ_0 to Γ_i as long as $u_i = o(n^{1/(\tau-1)})$, which is*

$$i < -\frac{\log((\tau-1)\varrho'')}{|\log(\tau-2)|}. \tag{3.11}$$

Proof. Let us denote the total number of half-edges in Γ_i by $\mathcal{S}_i$. Then, since the degrees are i.i.d., we have $|\Gamma_{i+1}| \sim \mathrm{Bin}(n, 1 - F(u_{i+1}))$, and each vertex $w \in \Gamma_{i+1}$ has degree at least u_{i+1}. Thus by Lemma 3.2,

$$\mathbb{P}\left(\mathcal{S}_{i+1} < \frac{u_{i+1} n \left[1 - F(u_{i+1}) \right]}{2} \right) \leq \exp\left\{ -\frac{n \left[1 - F(u_{i+1}) \right]}{8} \right\}, \tag{3.12}$$
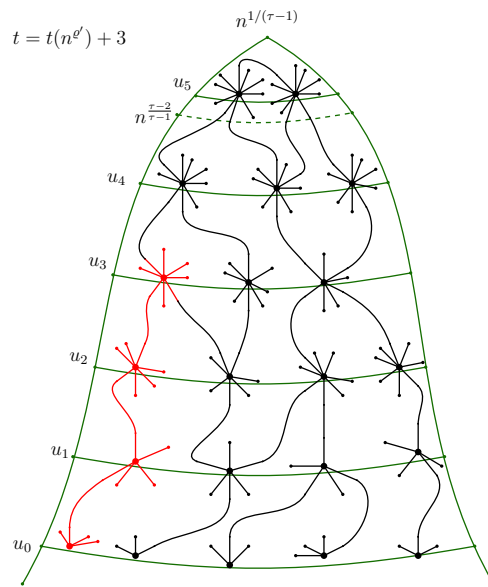
Figure 1: An illustration of the layers and the mountain climbing phase at time $t(n^{\varrho'}) + 3$. Disclaimer: the degrees on the picture are only an illustration.

Recall that $\mathcal{L}_n$ denotes the total number of half-edges in the graph. On the event $\{\mathcal{S}_{i+1} > u_{i+1}n\left[1 - F(u_{i+1})\right]/2\} \cap \{\mathcal{L}_n < 2\mathbb{E}[D]n\}$, the probability that there is a vertex $v \in \Gamma_i$ not connected to Γ_{i+1} can be bounded from above by

$$\mathbb{P}_n\left(\exists v \in \Gamma_i, v \nleftrightarrow \Gamma_{i+1}\right) \leq |\Gamma_i| \left(1 - \frac{\mathcal{S}_{i+1}}{\mathcal{L}_n}\right)^{u_i/2}$$
$$\leq n \exp\left\{-\frac{u_i u_{i+1} n\left[1 - F(u_{i+1})\right]}{8n\mathbb{E}[D]}\right\}, \tag{3.13}$$

where we recall that $\mathbb{P}_n(\cdot) := \mathbb{P}(\cdot | D_1, \ldots, D_n)$. We have used that $|\Gamma_i| < n$, and the estimate S_{i+1} in (3.12). The factor $1/2$ in the exponent $u_i/2$ comes from the worst-case scenario estimate when we connect all the first $u_i/2$ half-edges back to v. Similar calculations (with indices of u_i and u_{i+1} exchanged) are worked out in more detail in [27, Volume II., Chapter 5]. Then, using the defining recursion (3.6), it is easy to see that

$$u_i u_{i+1} \left[1 - F(u_{i+1})\right] = C \log n (u_{i+1})^{\tau-1} \left[1 - F(u_{i+1})\right] \geq C c_1 \log n,$$

that is, the error term in (3.13) is bounded by

$$\varepsilon_i := \exp\left\{\left(1 - \frac{C c_1}{8\mathbb{E}[D]}\right) \log n\right\} + \exp\{-n\left[1 - F(u_{i+1})\right]/8\}.$$

The first assertion of the lemma follows if ε_i small: the first term is small when picking C large enough. For the second term we need $n\left[1 - F(u_{i+1})\right] > O(1)$, which exactly translates to the condition $u_i = o(n^{1/(\tau-1)})$ and to (3.11) using (3.7). Note that as long as (3.11) is satisfied, even $\sum_{j=1}^i \varepsilon_j = o(n^{-1})$. This means that we can apply the lemma consecutively for the layers (Γ_i)s and build a path $(v_0, v_1, \ldots, v_i)$ such that for all $0 \leq j \leq i$ we have $v_j \in \Gamma_j$ and $(v_j, v_{j+1}) \in E(\mathrm{CM}_n(\boldsymbol{d}))$ whp, as long as i satisfies (3.11). This finishes the proof of the second statement of the lemma. $\qquad\square$

With Lemma 3.3 in hand we can determine how long it takes to climb up through the layers Γ_i to the highest-degree vertices. Lemma 3.1 with $X_i = D_i \sim F$, $\alpha = \tau - 1$ shows

that the maximal degree in $\mathrm{CM}_n(\boldsymbol{d})$ is of order $n^{1/(\tau-1)}$. We write i^* for the last index when Γ_i is whp nonempty, i.e.,

$$i^* := \inf\{i : u_i \le n^{1/(\tau-1)} < u_{i+1}\}. \tag{3.14}$$

An easy calculation using (3.7) shows that

$$i^* = -1 + \frac{-\log((\tau-1)\varrho'')}{|\log(\tau-2)|} - b_n, \quad \text{with} \quad b_n = \left\{ \frac{-\log((\tau-1)\varrho'')}{|\log(\tau-2)|} \right\}. \tag{3.15}$$

Note that i^* satisfies (3.11), thus all the error terms up to this point stay small. Using the value of the overshoot exponent ϱ'' in (3.4) and then the value a_n in (3.3), plus the fact that $\{x - 1 + \{y\}\} = \{x+y\}$, we get that

$$b_n = \left\{ \frac{-\log((\tau-1)\varrho')}{|\log(\tau-2)|} + a_n - 1 \right\} = \left\{ \frac{-\log((\tau-1)Y_r^{(n)}) + \log\log n}{|\log(\tau-2)|} \right\}. \tag{3.16}$$

From (3.7) one can easily calculate that

$$u_{i^*} = n^{\frac{(\tau-2)^{b_n}}{\tau-1}} (C\log n)^{-e_{i^*}}, \quad \text{with}$$
$$e_{i^*} = \frac{1}{3-\tau}\left(\frac{(\tau-2)^{b_n}}{(\tau-1)\varrho''} - 1 \right) \le \frac{1}{(3-\tau)}\left(\frac{1}{(\tau-1)\varrho''} - 1 \right). \tag{3.17}$$

We will repeatedly need the total time to reach the top, so let us introduce the notation

$$T_r := t(n^{\varrho'}) + i^* = \frac{\log\log n - \log\left((\tau-1)Y_r^{(n)}\right)}{|\log(\tau-2)|} - 1 - b_n, \tag{3.18}$$

which only depends on ϱ' via the approximating $Y_r^{(n)}$, and b_n is exactly the fractional part of the expression on the rhs of T_r. Since also $Y_r^{(n)} \to Y_r$ irrespective of the choice of ϱ', this establishes that the choice of ϱ' is not relevant in the proof.

4 Crossing the peak of the mountain

Next we investigate what happens when the path through the layers reaches the highest degree vertices. We have just seen that the exponent of n in u_{i^*} is $\frac{(\tau-2)^{b_n}}{\tau-1} \in \left(\frac{\tau-2}{\tau-1}, \frac{1}{\tau-1} \right)$. Recall that the maximum degree in the graph has exponent $\frac{1}{\tau-1}$ whp, i.e. $\Gamma_{i^*+1} = \varnothing$ whp, meaning the path can not jump 'up' one more step. On the other hand, we can make use of the following lemma from [27, Volume II., Chapter 5]:

Lemma 4.1. *Consider two sets of vertices A and B. If for the number of half-edges $\mathcal{S}_A = o(n)$ and $\mathcal{S}_B$ satisfy*

$$\frac{\mathcal{S}_A \mathcal{S}_B}{n} > h(n),$$

for some function $h(n)$, then conditioned on the degree sequence with $\mathcal{L}_n \le 2\mathbb{E}[D]n$, the probability that the two sets are not directly connected can be bounded from above by

$$\mathbb{P}_n(A \nleftrightarrow B) < e^{-\frac{h(n)}{4\mathbb{E}[D]}}.$$

Proof. When pairing the half-edges coming out from A, the probability that the i-th one paired is not directly connected to a half-edge in B is $(1 - \mathcal{S}_B/(\mathcal{L}_n - 2i - 1))$. Thus,

$$\mathbb{P}_n(A \nleftrightarrow B) = \prod_{i=0}^{\lceil \mathcal{S}_A/2 \rceil} \left(1 - \frac{\mathcal{S}_B}{\mathcal{L}_n - 2i - 1} \right) \le \exp\left\{ -\frac{\mathcal{S}_A \mathcal{S}_B}{2(\mathcal{L}_n - 2\mathcal{S}_A)} \right\}$$
$$\le \exp\left\{ -\frac{h(n)}{4\mathbb{E}[D]} \right\}.$$

The product only goes until $\mathcal{S}_A/2 - 1$, since in the worst case scenario the first $\lfloor \mathcal{S}_A/2 \rfloor$ half-edges are all paired back to another half-edge in A, thus the last $\lfloor \mathcal{S}_A/2 \rfloor$ half-edges are not used anymore. In both cases, we can pair at least $\lceil \mathcal{S}_A/2 \rceil$ many half-edges. $\square$

Let us introduce

$$\alpha := 1 - \frac{(\tau-2)^{b_n}}{\tau - 1}, \quad \beta := 1 + \frac{1}{(3-\tau)}\left(\frac{1}{(\tau-1)\varrho''} - 1\right), \tag{4.1}$$

and

$$\widetilde{u}_1 := (C \log n)n/u_{i^*} = n^\alpha (C \log n)^\beta, \tag{4.2}$$

and the following layer:

$$\widetilde{\Gamma}_1 := \{v \in \mathrm{CM}_n(\boldsymbol{d}), D_v > \widetilde{u}_1\}. \tag{4.3}$$

Recall the definition of the T_r from (3.18), that denotes the total time for the red color to reach the last layer Γ_{i^*}. The next lemma helps us describe how the process goes through the highest-degree vertices:

Lemma 4.2. *All the vertices in $\widetilde{\Gamma}_1$ are occupied by red at time $T_r + 1$, i.e.,*

$$\widetilde{\Gamma}_1 \subset \mathcal{R}_{T_r+1} \quad \text{whp.} \tag{4.4}$$

Proof. By Lemma 3.3, there is a red path up to Γ_{i^*}, and hence, red occupies some vertices in layer Γ_{i^*} at time T_r. Hence, $\mathcal{R}_{T_r} \cap \Gamma_{i^*} \neq \varnothing$, and we have at least one vertex v_{i^*} in $\mathcal{R}_{T_r}$ for which the degree is at least u_{i^*}, see (3.17). We claim that this vertex is whp connected to *every* vertex in $\widetilde{\Gamma}_1$. To see this, let us set $A := \{v_{i^*}\}$ and $B := \{w\}$ with any $w \in \widetilde{\Gamma}_1$, that is, any single vertex in $\widetilde{\Gamma}_1$ with degree at least $(C \log n)n/u_{i^*}$. Then apply Lemma 4.1 with this setting to see that v_{i^*} is whp connected to w. Further, note that $\mathcal{S}_A \mathcal{S}_B/n = C \log n$ by the definition of $\widetilde{u}_1$. Hence, using the error bound in Lemma 4.1 and a union bound,

$$\mathbb{P}_n(\exists w \in \widetilde{\Gamma}_1, (v_{i^*}, w) \notin E(\mathrm{CM}_n(\boldsymbol{d}))|\widetilde{\Gamma}_1) \leq |\widetilde{\Gamma}_1|\frac{1}{n^{C/4\mathbb{E}[D]}}. \tag{4.5}$$

Clearly $|\widetilde{\Gamma}_1| < n$: picking a large enough C, we see that the error probability tends to zero. Calculating $C \log n \cdot n/u_{i^*}$ yields the formula for $\widetilde{u}_1$. $\square$

It is important to note that vertices with degree larger than $\widetilde{u}_1$ do whp exist in $\mathrm{CM}_n(\boldsymbol{d})$ by Lemma 3.1. Moreover, i^* is the first index when we can apply Lemma 4.1, since for all smaller values $i < i^*$, there are whp no vertices with degree at least n/u_i by Lemma 3.1.

This completes the *crossing the peak of the mountain* phase.

5 Red avalanche from the peak and the blue climber

Using the value $\widetilde{u}_1$ in (4.2), let us again recursively define

$$\widetilde{u}_{\ell+1} = C \log n \cdot (\widetilde{u}_\ell)^{\tau-2}. \tag{5.1}$$

and also the increasing sequence of sets

$$\widetilde{\Gamma}_\ell := \{v : D_v > \widetilde{u}_\ell\},$$

i.e., now $\widetilde{\Gamma}_1 \subset \widetilde{\Gamma}_2 \subset \ldots$ holds. Since (5.1) is the very same as the recursion in (3.6) with indices exchanged, we can apply Lemma 3.3 to $\left(\widetilde{\Gamma}_\ell\right)_{\ell \geq 1}$, now yielding that for any $\varepsilon > 0$, for all $\ell < (1-\varepsilon)\log\log n/|\log(\tau-2)|$

$$\widetilde{\Gamma}_{\ell+1} \subset N(\widetilde{\Gamma}_\ell) \quad \text{and} \quad \widetilde{\Gamma}_\ell \subset \mathcal{R}_{T_r+\ell} \quad \text{whp.} \tag{5.2}$$

This means that in the 'sloping down' phase, whp red occupies *all* vertices in $\widetilde{\Gamma}_\ell$ at time $T_r + \ell$. Solving the recursion (5.1) yields that

$$\widetilde{u}_\ell = n^{\alpha(\tau-2)^{\ell-1}}(C \log n)^{\beta(\tau-2)^{\ell-1}+\frac{1}{3-\tau}\left(1-(\tau-2)^{\ell-1}\right)}, \tag{5.3}$$

where α and β were defined in (4.1). Note that the exponent of $C \log n$ stays bounded even when $\ell \to \infty$. Hence this procedure can be continued even to reach lower degree vertices, for every fixed $\varepsilon > 0$ up until $\ell < (1-\varepsilon) \log \log n / |\log(\tau-2)|$.

In what follows, we determine the point where red and blue meet. More precisely, we calculate the value ℓ such that during the time $T_r + \ell$, the maximum degree vertex in the cluster of blue is of the same order as $\widetilde{u}_\ell$. Since at time $T_r + \ell$, red occupies whp almost every vertex with degree at least $\widetilde{u}_\ell$, the growing cluster of blue bumps into the occupied vertices and cannot spread to higher-degree vertices anymore.

The following proposition about the maximal degree of blue is our main building block for the proof of Theorem 1.4:

Proposition 5.1. *Let us denote by $D_{\max}^{(b,n)}(t)$ the forward degree in $\mathrm{CM}_n(\boldsymbol{d})$ of the maximal degree vertex in the blue cluster at time t. Then, whp at time $T_r + t$ and for any real $1 \le t \le O^\diamond(1)$,*

$$D_{\max}^{(b,n)}(T_r + t) = \exp\left\{Y_b^{(n)}\left(\frac{1}{\tau-2}\right)^{\left\lfloor \frac{T_r+t}{\lambda} \right\rfloor+1}(1+o_{\mathbb{P}}(1))\right\}, \tag{5.4}$$

as long as t is so that the quantity on the rhs is less than $\widetilde{u}_{[t]}$, and where $Y_b^{(n)}$ is defined in Definition 1.5.

Before the proof we need some important definitions that will be used also outside the proof. Similarly as in (3.5), let us define:

$$\begin{aligned}
\widehat{u}_0^{(b)} &:= (Z_{\lfloor t(ne')/\lambda \rfloor}^{(b)} \cdot C \log n)^{1/(\tau-2)}, & u_0^{(b)} &:= (Z_{\lfloor t(ne')/\lambda \rfloor}^{(b)}/C \log n)^{1/(\tau-2)} \\
\widehat{u}_{i+1}^{(b)} &:= (\widehat{u}_i^{(b)} \cdot C \log n)^{1/(\tau-2)}, & u_{i+1}^{(b)} &:= \left(u_i^{(b)}/C \log n\right)^{1/(\tau-2)} \\
\widehat{\Gamma}_i^{(b)} &:= \{v \in \mathrm{CM}_n(\boldsymbol{d}) : d_v \ge \widehat{u}_i^{(b)}\}, & \Gamma_i^{(b)} &:= \{v \in \mathrm{CM}_n(\boldsymbol{d}) : d_v \ge u_i^{(b)}\}.
\end{aligned} \tag{5.5}$$

Note that $\Gamma_i^{(b)}$ grows exactly as Γ_i while $\widehat{\Gamma}_i^{(b)}$ grows faster: there is always an extra $(C \log n)^2$ factor causing an initial 'gap' of order $(\log n)^2$ between $u_0^{(b)}, \widehat{u}_0^{(b)}$ and 'opening up' as i gets larger.

Further, let us say that a quantity $Q \asymp O^\diamond(x)$ for $x \in R^+$, if Q satisfies

$$Q = x \frac{\log \log n}{|\log(\tau-2)|} + O_{\mathbb{P}}(1). \tag{5.6}$$

We will see below in (6.5) that blue cannot make more jumps than $O^\diamond\left(\frac{\lambda-1}{\lambda(\lambda+1)}\right)$ in its climbing phase. In order to show Proposition 5.1, we need a lower and an upper bound on the maximal degree in each step. The next lemma handles the upper bound, but first some definitions.

We say that a sequence of vertices and half-edges $(\pi_0, s_0, t_1, \pi_1, s_1, t_2, \ldots, t_k, \pi_k)$ forms a path in $\mathrm{CM}_n(\boldsymbol{d})$, if for all $0 < i \le k$, the half edges s_{i-1}, t_i form an edge between π_{i-1}, π_i. Let us denote the vertices in a path starting from a half-edge in $Z_{\lfloor t(ne')/\lambda \rfloor}^{(b)}$ by $\pi_0, \pi_1, \ldots$. We say that a path is *good* if $\deg(\pi_i) \le \widehat{u}_i^{(b)}$ holds for every i. Otherwise we call it *bad*. We decompose the set of bad paths in terms of where they turn bad, i.e. we say that a bad path is belonging to $\mathcal{B}ad\mathcal{P}_k$ if it turns bad at the kth step:

$$\begin{aligned}
\mathcal{B}ad\mathcal{P}_k := \{ & (\pi_0, s_0, t_1, \pi_1, s_1 \ldots, t_k, \pi_k) \text{ is a path}, \\
& \pi_0 \in \mathcal{B}_{t(ne')}, \ \deg(\pi_i) \le \widehat{u}_i^{(b)} \ \forall i \le k-1, \ \deg(\pi_k) \ge \widehat{u}_k^{(b)} \}.
\end{aligned}$$

Let us write $\mathcal{F}_{t(n^{\varrho'})}$ for the σ-algebra generated by constructing the configuration model using the exploration process as described before, until time $t(n^{\varrho'})$, i.e., reaching *the vertices* in the last $(\lfloor t(n^{\varrho'})/\lambda \rfloor)$th generation of the blue BP, (but not their degrees). The following lemma tells us that the probability of having a bad path is tending to zero, conditioned on $\mathcal{F}_{t(n^{\varrho'})}$.

Lemma 5.2. *Fix* $0 < x \leq (\lambda - 1)/\lambda(\lambda + 1)$. *Then for any* $k_0 \leq O^{\diamond}(x)$, *the following bound on the probability of having any bad paths holds:*

$$\mathbb{P}(\exists k \leq k_0 : \mathcal{B}ad\mathcal{P}_k \neq \varnothing | \mathcal{F}_{t(n^{\varrho'})}) \leq \frac{2 \log \log n}{C \log n}. \tag{5.7}$$

Proof. The proof uses path counting methods that we describe in the appendix. Hence we put the proof there. □

Proof of Proposition 5.1. Since the method for the lower bound is very much the same as for red, plus we will need a more detailed analysis of this process below in Lemma 5.4, we just sketch the proof (read further to the proof of Lemma 5.4 for more details). First, Lemma 2.1 ensures that we can couple both the blue and the red cluster to their BP approximation until time $t(n^{\varrho'})$ given in (3.3). Since it takes $\lambda > 1$ unit of time to cover an edge for blue, the number of generations covered by the branching process approximation $Z^{(b)}$ of blue is $\lfloor t(n^{\varrho'})/\lambda \rfloor$. The size of the last generation in the blue BP is thus

$$Z^{(b)}_{\lfloor t(n^{\varrho'})/\lambda \rfloor} = \exp \left\{ Y_b^{(n)} \left(\frac{1}{\tau - 2} \right)^{\lfloor t(n^{\varrho'})/\lambda \rfloor} \right\}. \tag{5.8}$$

We start applying the method in the *Mountain climbing* phase for blue from this point on. With the same technique as we used to show that $\mathcal{R}_{t(n^{\varrho})} \cap \Gamma_0 \neq \varnothing$ using Lemma 3.1, we define $u_0^{(b)}$ and a corresponding layer $\Gamma_0^{(b)}$ in (5.5) and can show that $\mathcal{B}_{t(n^{\varrho'})} \cap \Gamma_0^{(b)} \neq \varnothing$. Then, we define the nested sequence of layers $(\Gamma_i^{(b)})_{i \geq 0}$ with $u_i^{(b)}$-s such that at time $\lambda(\lfloor t(n^{\varrho'})/\lambda \rfloor + i)$, blue occupies at least 1 vertex in $\Gamma_i^{(b)}$ by Lemma 3.3. Note that from $\Gamma_i^{(b)}$ to $\Gamma_{i+1}^{(b)}$, the exponent of $1/(\tau - 2)$ on the right hand side of (5.8) is increased by one. Further, there is an extra $+1$ in the exponent for the initial maximization of the degrees in $u_0^{(b)}$ similarly as in (3.6).

The total number of layers $\Gamma_i^{(b)}$ jumped by blue at time $T_r + t$ is then $\lfloor (T_r + t)/\lambda \rfloor - \lfloor t(n^{\varrho'})/\lambda \rfloor$, that, combined with (5.8), yields formula (5.4).

We still need to check that the term arising from $C \log n$ in the definition of $u_i^{(b)}$'s can be put in a $(1 + o_{\mathbb{P}}(1))$ factor in the exponent. For this, write $Z^{(b)}_{\lfloor t(n^{\varrho'})/\lambda \rfloor} := m$, then

$$u_i^{(b)} = m^{(\tau - 2)^{-(i+1)}} (C \log n)^{(1 - (\tau - 2)^{-(i+1)})/(3 - \tau)} \tag{5.9}$$

and the last layer before time $T_r + t$ is reached after climbing $i = \lfloor (T_r + t)/\lambda \rfloor - \lfloor t(n^{\varrho'})/\lambda \rfloor$ many $\Gamma_i^{(b)}$ layers, so by (3.2) and (3.18) we calculate

$$\lfloor (T_r + t)/\lambda \rfloor - \lfloor t(n^{\varrho'})/\lambda \rfloor = (i^* + t)/\lambda(1 + o(1)) \tag{5.10}$$

Thus, if $t \leq O^{\diamond}(1)$, when taking the logarithm, then the term corresponding to $(C \log n)^{(\tau - 2)^{-i}/(3 - \tau)}$ in $\log(u_i^{(b)})$ at time $T_r + t$ is of order $(\tau - 2)^{-t/\lambda} = o((\tau - 2)^{-(T_r + t)/\lambda})$ Hence, these terms vanish when taking out $(\tau - 2)^{-\lfloor (T_r + t)/\lambda \rfloor}$ in the statement of the lemma. We will see below in (6.5) that in fact the procedure stops at $t = O^{\diamond}(\frac{\lambda - 1}{\lambda(\lambda + 1)})$ since after that red will block the growth of blue entirely.

For the upper bound, according to Lemma 5.2, whp $\{\mathcal{B}ad\mathcal{P}_k = \varnothing \ \forall k \leq k_0\}$, and on this event the maximal degree of blue at time $\lambda \lfloor t(n^{\varrho'})/\lambda \rfloor + \lambda i$ is at most $\widehat{u}_i^{(b)}$. Since the exponent of $C \log n$ in $\widehat{u}_i^{(b)}$ is exactly (-1) times the exponent of $C \log n$ in $u_i^{(b)}$, these

terms can also be put in the $(1 + o_{\mathbb{P}}(1))$ factor by the same argument as for the lower bound. $\qquad\square$

We will later need more information than the maximal degree of blue, namely, we also need an upper bound on *how many* vertices blue occupies in each layer. For this, first we show that the probability that blue goes above $\widehat{u}_i^{(b)}$ at time $\lambda \lfloor t(n^{\varrho'})/\lambda \rfloor + \lambda i$ is small, then we estimate the number of vertices blue paints in each layer based on this bound. We carry these out in a claim and a lemma.

Let us denote the total number of half-edges attached to vertices with degree larger than y_n by $\mathcal{E}_{\geq y_n}$. Then, we have the following tail bound for $\mathcal{E}_{\geq y_n}$:

Claim 5.3. *Let $\mathcal{E}_{\geq y_n}$ denote the total number of half-edges incident to vertices with degree at least y_n, and $V_{\geq y_n}$ the total number of vertices of degree at least y_n, for a sequence $y = y_n$. Then, for any $\omega(n) \to \infty$, and a large enough constant $C < \infty$, a small enough constant $0 < c$ and for some constant $0 < c_2 < \infty$,*

$$
\begin{aligned}
\mathbb{P}(\mathcal{E}_{\geq y_n} \geq C\omega(n) \cdot n \cdot y_n^{2-\tau}) &\leq c_2/\omega(n), \\
\mathbb{P}(V_{\geq y_n} \geq C \cdot n \cdot y_n^{1-\tau}) &\leq \exp\{-c_2 \cdot n y_n^{1-\tau}\}, \\
\mathbb{P}(V_{\geq y_n} \leq c \cdot n \cdot y_n^{1-\tau}) &\leq \exp\{-c_2 \cdot n y_n^{1-\tau}\}
\end{aligned}
\tag{5.11}
$$

Proof. Note that

$$
\mathcal{E}_{\geq y_n} = \sum_{i=1}^{n} D_i \mathbb{1}_{\{D_i \geq y_n\}} \leq \sum_{i=1}^{n} \sum_{k=1}^{\infty} y_n 2^k \mathbb{1}_{\{2^{k-1} y_n < D_i \leq 2^k y_n\}}.
$$

Exchanging sums we can write

$$
\mathcal{E}_{\geq y_n} \leq \sum_{k=1}^{\infty} y_n 2^k X_k^{(n)},
$$

where the marginals of the random variables on the rhs are $X_k^{(n)} \overset{d}{\leq} \mathrm{Bin}(n, C_1(y_n 2^{k-1})^{1-\tau})$, where C_1 is from (1.1). Calculating the expected value and using Markov's inequality yields (5.11). The proof for $V_{\geq y_n}$ is easier and directly follows from the fact that $V_{\geq y_n} \sim \mathrm{Bin}(n, 1 - F(y_n))$ and usual concentration of Binomial random variables, see Lemma 3.2. $\qquad\square$

Let us denote the set and number of blue vertices in the ith layer $\Gamma_i^{(b)}$ right at the time when blue reaches it by

$$
\mathcal{A}_i := \mathcal{B}_{\lambda \lfloor t(n^{\varrho'})/\lambda \rfloor + \lambda i} \cap \Gamma_i^{(b)}, \qquad A_i := |\mathcal{A}_i|. \tag{5.12}
$$

Lemma 5.4. *Let $k_0 \leq O^{\diamond}((\lambda - 1)/\lambda(\lambda + 1))$. On the event $\{\mathcal{B}ad\mathcal{P}_k = \varnothing \ \forall k \leq k_0\}$, whp for all $i \leq O^{\diamond}(\frac{\lambda-1}{\lambda(\lambda+1)})$,*

$$
A_i \leq \exp\left\{ \log(C \log n) \cdot \frac{2(\tau-2)^{-i}}{(3-\tau)^2} \right\}. \tag{5.13}
$$

Hence, for $x \leq (\lambda - 1)/\lambda(\lambda + 1)$, for some constant K_2, whp

$$
\log A_{O^{\diamond}(x)} \leq K_2 (\log n)^x \log \log n \cdot (1 + o(1)). \tag{5.14}
$$

Proof. First, Lemma 5.2 guarantees that $\{\mathcal{B}ad\mathcal{P}_k = \varnothing \ \forall k \leq k_0\}$ holds whp, and on this event $\widehat{u}_i^{(b)}$ serves as an upper bound on the maximal degree of blue at time $\lambda \lfloor t(n^{\rho'})/\lambda \rfloor + \lambda i$. So, we can give a recursive upper bound on the number of vertices reached by blue in a given layer $\Gamma_i^{(b)}$ by using $u_i^{(b)}$ as a lower and $\widehat{u}_i^{(b)}$ as an upper bound on the degrees.

Let us condition on the number of blue vertices A_i in layer $\Gamma_i^{(b)}$. Then, we have at most A_i half-edges in $\Gamma_i^{(b)} \cap \mathcal{B}_{\lambda \lfloor t(n^{\varrho'})/\lambda \rfloor + \lambda i}$, with degree at most $\widehat{u}_i^{(b)}$, hence we get the stochastic domination

$$A_{i+1} = |N(\mathcal{A}_i) \cap \Gamma_{i+1}^{(b)}| \overset{d}{\leq} \mathsf{Bin}\left(A_i \widehat{u}_i^{(b)}, \frac{\mathcal{E}_{\geq u_{i+1}^{(b)}}}{\mathcal{L}_n (1 + o(1))}\right). \tag{5.15}$$

Then, we can use Claim 5.3 with $y_n = u_{i+1}^{(b)}$ to bound $\mathcal{E}_{\geq u_{i+1}^{(b)}}$. Hence, on the event that $\mathbb{E}[D] n/2 < \mathcal{L}_n < 2\mathbb{E}[D] n$ holds, by Lemma 3.2 applied on the binomial variable in (5.15),

$$\mathbb{P}\left(A_{i+1} > \frac{8C_1}{\mathbb{E}[D](2 - \tau)} A_i \widehat{u}_i^{(b)} (u_{i+1}^{(b)})^{2-\tau} \Big| A_i\right) \leq \exp\left\{-A_i u_i^{(b)} (u_{i+1}^{(b)})^{2-\tau} \frac{C_1}{4c_1 \mathbb{E}[D]}\right\}. \tag{5.16}$$

Thus, with the error probability in the previous display, whp

$$A_{i+1} \leq A_i \widehat{u}_i^{(b)} (u_{i+1}^{(b)})^{2-\tau} \frac{8C_1}{\mathbb{E}[D](2 - \tau)} < A_i \frac{\widehat{u}_i^{(b)}}{u_i^{(b)}} C' \log n \tag{5.17}$$

with $C' := 8C_1 C/(\mathbb{E}[D](2 - \tau))$, where the last inequality follows from using (5.5). As a result, we get the recursion $A_{i+1} \leq A_i C' \log n \cdot \widehat{u}_i^{(b)}/u_i^{(b)}$. The bound in (5.13) is nothing but

$$A_i \leq A_0 \prod_{j=0}^{i-1} \left(C' \log n \cdot \widehat{u}_j^{(b)}/u_j^{(b)}\right). \tag{5.18}$$

Solving the recursions for $u_i^{(b)}$ and $\widehat{u}_i^{(b)}$ in (5.5) (or, see (5.9)) we get that

$$\widehat{u}_i^{(b)}/u_i^{(b)} = (C \log n)^{2\left((\tau-2)^{-(i+1)} - 1\right)/(3-\tau)}. \tag{5.19}$$

Initially $A_0 \leq 2C \log n$ whp. This can be seen as follows: by the coupling of the exploration process to the branching process in Section 2, the last generation has size $Z_{\lfloor t(n^{\varrho'})/\lambda \rfloor}^{(b)}$, and the degrees are i.i.d. of distribution $D^\star$. Hence, the number of vertices in this last generation that have degree at least $u_0^{(b)}$ has distribution $A_0 \sim \mathsf{Bin}(Z_{t(n^{\varrho'})}^{(b)}, \mathbb{P}(D^\star > u_0^{(b)}))$. Note that by the choice of $u_0^{(b)}$, $\mathbb{E}[A_0] \leq CC_1 \log n$, and the Lemma 3.2 implies that $A_0 < 2CC_1 \log n$ holds with probability at least $\exp\{CC_1 \log n/8\}$, which is small when C is large enough.

Using (5.19) and evaluating (5.18) finishes the proof of (5.13). We dropped some negative terms in the exponent in (5.13). If we set $i \leq O^\diamond(x) = x \log \log n/|\log(\tau - 2)| + O_\mathbb{P}(1)$, then by picking a large enough C, the error terms are $o(n^{-A_i})$ in (5.16). Thus we can also iterate the argument up to time $O^\diamond(x)$ to see that at time $\lambda \lfloor t(n^{\varrho'})/\lambda \rfloor + \lambda O^\diamond(x)$, in $\Gamma_{O^\diamond(x)}^{(b)}$, the number of vertices blue occupies is bounded by the right hand side of (5.13). □

6 At the collision time - the maximal degree of blue

6.1 The maximum degree of blue

In this section we analyse how red and blue collide and prove Theorem 1.4, i.e., we determine the degree of the maximum degree vertex that blue ever occupies. There are two different processes running at time $T_r + \ell$: the red process is in its *avalanche* phase and occupies every vertex that has degree higher than $\widetilde{u}_\ell$, while the slower blue process is still in its *mountain-climbing* phase and keeps increasing its maximal degree (but does not occupy all vertices of this degree). To obtain a good approximation when the two processes meet let us first neglect integer part issues and determine the collision

time of the two processes (the red avalanche and the blue climber) in continuous time.[2] Neglecting the terms containing $C \log n$, we can compare the maximal degree (5.4) to $\widetilde{u}_\ell$ in (5.3) and define t_c as the solution of the equation

$$\exp\left\{ Y_b^{(n)}(\tau-2)^{-\frac{T_r+t_c}{\lambda}-1} \right\} = \exp\left\{ \log n \cdot \alpha \cdot (\tau-2)^{t_c-1} \right\}. \tag{6.1}$$

Note that the left-hand side is approximately equal to the maximum degree $D_{\max}^{(b,n)}(T_r + \ell)$ of blue, while the right-hand side is the approximate value of $\widetilde{u}_\ell$. Thus t_c is the (non-integer valued) time left till the intersection of these two functions after time T_r.

We will soon see that neglecting the integer part does have an influence on the highest degree vertex blue can occupy. To get a more precise picture, we should compare which color is first and second to jump after $T_r + t_c$, since if it is blue, it can still increase its exponent. So, let us introduce the time of the last jump of red and blue before time $T_r + t_c$:

$$r_-^* := \lfloor t_c \rfloor, \qquad b_-^* := \lambda \cdot \left\lfloor \frac{T_r + t_c}{\lambda} \right\rfloor. \tag{6.2}$$

In words, red jumps at times $T_r + r_-^*, T_r + r_-^* + 1, \ldots$ while blue jumps at times $b_-^*, b_-^* + \lambda, \ldots$ and t_c satisfies $T_r + r_-^* \le t_c < T_r + r_-^* + 1$ and $b_-^* \le t_c < b_-^* + \lambda$. We need to determine who jumps first after time $T_r + t_c$, (that is, $T_r + r_-^* + 1 < b_-^* + \lambda$ or the other way round), so let us also introduce the remaining times till the next jump *after* the intersection for both colors:

$$J_r := r_-^* + 1 - t_c = 1 - \{t_c\}, \quad J_b := b_-^* + \lambda - (T_r + t_c) = \lambda\left(1 - \left\{\frac{T_r + t_c}{\lambda}\right\}\right). \tag{6.3}$$

J_r and J_b stands for the *additional time* needed for red and blue till their next jump after time $T_r + t_c$.

Remark 6.1. Note that given the values $Y_b^{(n)}, T_r$ and α, with each additional jump, red decreases the exponent of $1/(\tau - 2)$ by 1 and blue increases its exponent by 1. (Here we again neglect the terms including $C \log n$.) Thus, for red, when plotting the exponents of $1/(\tau - 2)$ of $\log \widetilde{u}_\ell / \log n$ one gets a line of slope -1, starting from time $T_r + 1$ from the value α. The exponent of $1/(\tau - 2)$ in $\log u_i^{(b)} / \log n$ in the cluster of blue is a line of slope $1/\lambda$, since it increases by one with every additional λ time units, see (5.4). These lines can be seen in Fig. 2, 3, 4, 5 and Fig. 6. The definition of t_c is the intersection time of these two lines.

Intuitively, the final exponent of the maximal degree of blue depends on two things: which color jumps first after the intersection time $T_r + t_c$ and how large the difference $d(t_c)$ is between the exponents of $1/(\tau - 2)$ in the $\log(\text{degree})/\log n$ of red and blue before time $T_r + t_c$. Since with each jump the exponent of $1/(\tau - 2)$ of the jumping color is changed by one, it is crucial whether this difference is less than or larger than 1. Let us temporarily postpone the calculations and believe that this difference is

$$d(t_c) := \{t_c\} + \left\{\frac{T_r + t_c}{\lambda}\right\}. \tag{6.4}$$

---

[2]The simplified picture is this: red occupies smaller and smaller degree vertices (jumps down), blue occupies some higher and higher degree vertices (jumps upwards), the question is, when do they meet? The problem is, that red jumps at unit time intervals while blue jumps at times that are multiples of λ. Further, the layers of red $\widetilde{\Gamma}_\ell$-s are not the same as the layers of blue, the $\Gamma_i^{(b)}$-s, that is, these layers might and will be shifted from each other. We still would like to determine the last possible up-jump of blue and the degrees that it reaches. It turned out that (probably) the easiest way to do this is to first solve where the two processes would meet if their jumps would be 'continuous', i.e., they would increase their degrees in a continuous fashion, (forgetting about the unit/multiple of λ issue and setting time to be continuous). This continuous 'intersection time' is denoted by t_c. Of course, this 'intersection' is never realised on the graph itself, where we do have to take into account that vertices are occupied only at times that are multiples of 1 or λ.
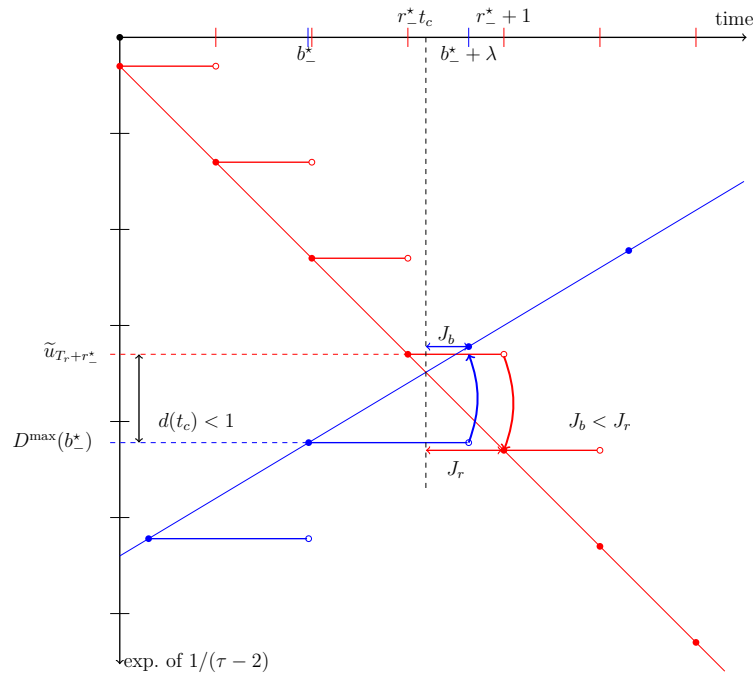
Figure 2: In this picture, the horizontal axes corresponds to time, the vertical axes to the exponent of $1/(\tau - 2)$ for $\log(\text{degree})/\log n$. The red and blue right-continuous stepfunctions indicate the exponent of $1/(\tau - 2)$ of the degrees red and blue occupies: red decreases this exponent by 1 at every unit time interval while blue increases it at every λ-interval. The two slanted lines are the continuous approximations of these stepfunctions as in (6.1), and their intersection gives t_c. In this particular picture, $d(t_c)$ is smaller than 1 and $J_b < J_r$. Thus, blue jumps first after the intersection and can occupy higher degree vertices up to $\widetilde{u}_{T_r+r^*}$. After that, red jumps and closes the gap, so the maximal degree of blue in this case is $\widetilde{u}_{T_r+r^*}$.

We will later analyse this difference in detail around equation (6.8). Since $d(t_c)$ is the sum of two fractional parts, it is at most 2. Recall also that J_r, J_b stands for the time till the next jump of red and blue after time $T_r + t_c$, respectively (see (6.3)). With these notations in mind, there are five cases:

(B1) $J_b < J_r$ and $d(t_c) < 1$. Blue jumps first after the intersection and occupies some vertices up to $\widetilde{\Gamma}_{r^*}$, i.e. blue can increase the exponent by a factor $(\tau - 2)^{-d(t_c)}$. (Vertices with higher degree than that are already red). See Fig 2.

(B2) $J_b < J_r$ and $d(t_c) > 1$. Blue jumps first after the intersection and occupies some vertices one layer higher, namely the total exponent of $1/(\tau - 2)$ in (5.4) reached by blue is $\lfloor \frac{T_r+t_c}{\lambda} \rfloor + 1$. However, since $1 < \lambda$, the next jump after this must be a red jump, hence red occupies every vertex with higher degree than this value. See Fig. 4.

(R1) $J_r < J_b$ and $d(t_c) < 1$. Red jumps first after the intersection, and occupies every not-yet blue vertex down to $\widetilde{\Gamma}_{r^*+1}$, which means that blue cannot increase its exponent anymore. Thus the exponent of $1/(\tau - 2)$ in (5.4) of the maximal degree reached by blue is $\lfloor \frac{T_r+t_c}{\lambda} \rfloor$. See Fig. 3.
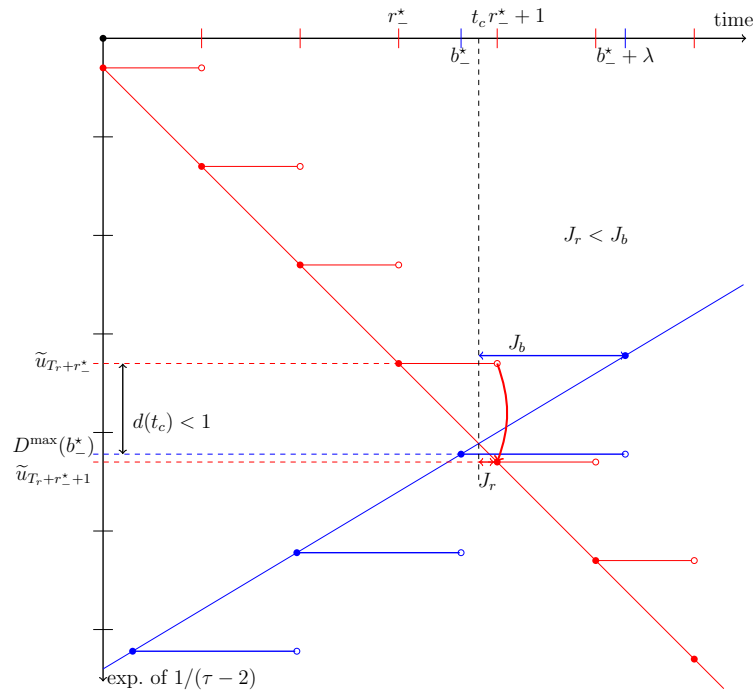
Figure 3: In this picture, the distance between the two exponents, $d(t_c)$ is smaller than 1 and $J_r < J_b$. Thus, red jumps first after the intersection and occupies every vertex of degree at least $\widetilde{u}_{T_r+r^*_+1}$ that is already less than $D^{\max}(b^\star_-)$. So the maximal degree of blue is $D^{\max}(b^\star_-)$.

(R2) $J_r < J_b < J_r + 1$ and $d(t_c) > 1$. Red can make only one jump after the intersection and occupies every vertex in $\widetilde{\Gamma}_{r^*_-+1}$, while blue jumps after this and can reach some vertices with degree up to $\widetilde{\Gamma}_{r^*_-+1}$ with its next jump. Thus the maximal degree of blue in this case is determined by $\widetilde{\Gamma}_{r^*_-+1}$, see Fig. 5.

(R3) $J_r + 1 < J_b$ and $d(t_c) > 1$. Red can make at least two consecutive jumps after the intersection and occupies every not-yet occupied vertex in $\widetilde{\Gamma}_{r^*_-+2}$, which means that blue can not increase its exponent. The exponent of $1/(\tau-2)$ in (5.4) of the maximal degree reached by blue is again $\lfloor \frac{T_r+t_c}{\lambda} \rfloor$, see Fig. 6.

Note that above we only handle the cases when $J_b \neq J_r$: this can be ensured by restricting λ to be irrational. If $\lambda = p/q$, $p, q \in \mathbb{N}$ is rational with p and q co-primes, then every vertex that is qt away from the blue source and pt away from the red source for arbitrary $t \in \mathbb{N}$ might be occupied at the same (i.e, at time pt). In this case, the color of such a vertex is chosen with probability $1/2$ independently of everything else. For the meeting time of the red avalanche and blue climber, a rational λ implies cases when $J_b = J_r$ or $J_b = J_r + 1$, i.e. the two processes jump at the same time after t_c. Here we list what happens in these cases, to be able to merge them in the cases above. We assume here that the adapted rule is so that there is a strictly positive probability (i.e. bounded away from 0 for all times) that a vertex becomes blue upon co-occupation.

(BR1) $J_b = J_r$ and $d(t_c) < 1$. Since there are lots of vertices just slightly smaller than $\widetilde{u}_{r^*_-}$, blue whp occupies some vertices up to that point, i.e. blue can increase the exponent by a factor $(\tau-2)^{-d(t_c)}$ again. This case can be merged into Case B1.
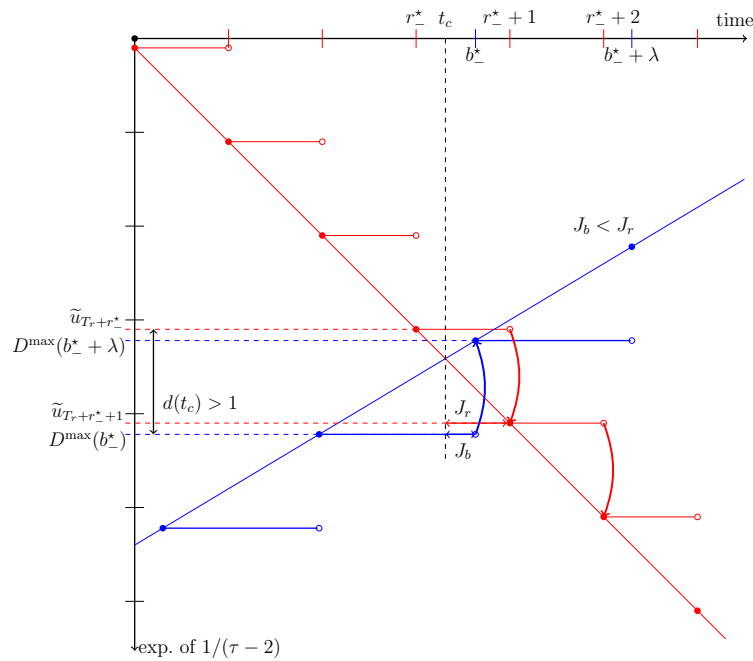
Figure 4: In this picture, the distance between the two exponents, $d(t_c)$ is larger than 1 and $J_b < J_r$. Thus, blue jumps first after the intersection and can occupy higher degree vertices up to $D^{\max}(b^\star + \lambda)$. After that, red jumps and closes the gap so the maximal degree vertex that blue can occupy has degree $D^{\max}(b^\star + \lambda)$.

(BR2) $J_b = J_r$ and $d(t_c) > 1$. In this case, blue can occupy some of the vertices up to one $\Gamma_i^{(b)}$ higher. This case can be merged into Case B2.

(BR3) $J_b = J_r + 1$ and $d(t_c) > 1$. In this case, red jumps first and occupies all the vertices down to $\Gamma_{r_-^*+1}$, and then the two processes jump together, so blue can occupy some vertices right below that. This case can be merged into Case R2.

Remark 6.2. If the adapted rule is so that the probability that a vertex is going to be red with probability one upon co-occupation, then Case BR1 merges into Case R1, case BR2 merges into Case B2, and Case BR3 merges into Case R3. We see that the adapted rule only influences the place where the strict and non-strict inequality signs appear inside the indicators in $f(d(t_c), J_r, J_b)$ in (6.9) below. Hence, the main result still holds true with a slightly different $f(d(t_c), J_r, J_b)$. For other adapted rules, the function f can be determined similarly.

Now we formalize these heuristics by finishing the proof of Theorem 1.4. An elementary calculation is to solve (6.1) yielding

$$
\begin{aligned}
t_c &= \frac{\lambda}{\lambda+1} \frac{\log\log n + \log(\alpha/Y_b^{(n)})}{|\log(\tau-2)|} - \frac{T_r}{\lambda+1} \\
&= \frac{\lambda-1}{\lambda+1} \frac{\log\log n}{|\log(\tau-2)|} + \frac{\log\left(\alpha^\lambda(\tau-1)Y_r^{(n)}/\left(Y_b^{(n)}\right)^\lambda\right)}{(\lambda+1)|\log(\tau-2)|} + \frac{1+b_n^{(r)}}{\lambda+1};
\end{aligned}
$$

$$
\frac{T_r + t_c}{\lambda} = \frac{1}{\lambda+1} \frac{2\log\log n - \log(Y_r^{(n)}Y_b^{(n)}(\tau-1)/\alpha)}{|\log(\tau-2)|} - \frac{1+b_n^{(r)}}{\lambda+1}.
$$

(6.5)

Hence, the red avalanche right before the intersection occupies every vertex with degree larger than $\widetilde{u}_{r^*}$, where $r_-^* = [t_c] = t_c - \{t_c\}$. Combining the formula of t_c with $\widetilde{u}_\ell$ in (5.3),
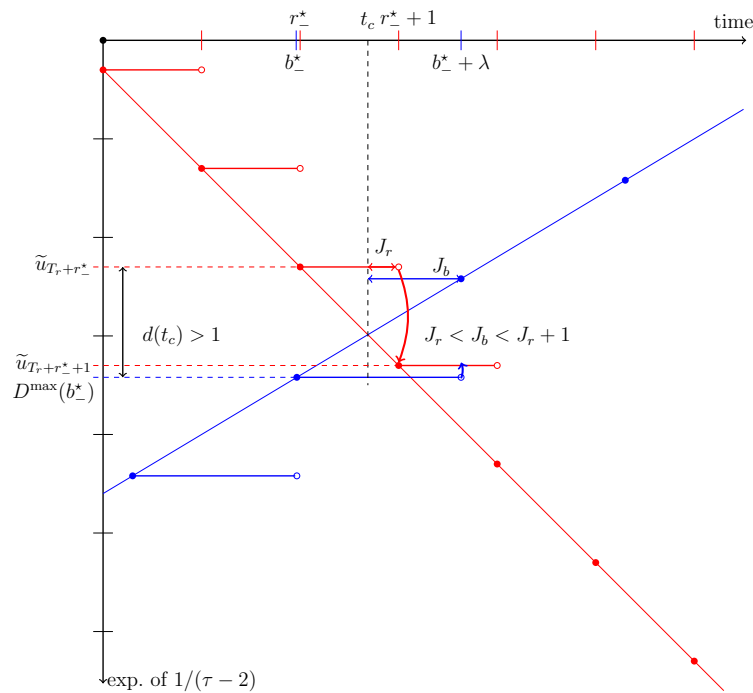
Figure 5: In this picture, the distance between the two exponents, $d(t_c)$ is larger than 1 and $J_r < J_b < J_r + 1$. Thus, red jumps first after t_c, occupies every vertex of degree at least $\widetilde{u}_{T_r+r_+^\star 1}$. Then, blue jumps and can occupy higher degree vertices up to $\widetilde{u}_{T_r+r_+^\star 1}$. After that, red jumps and closes the gap so the maximal degree vertex that blue can occupy has degree $\widetilde{u}_{T_r+r_+^\star 1}$.

we get

$$\log\left(\widetilde{u}_{r_-^\star}\right) = (\log n)^{\frac{2}{\lambda+1}} \left(\frac{\left(Y_b^{(n)}\right)^\lambda}{Y_r^{(n)}} \frac{\alpha}{\tau-1}\right)^{\frac{1}{\lambda+1}} (\tau-2)^{\frac{1+b_n^{(r)}}{\lambda+1}-1-\{t_c\}}(1+o(1)). \qquad (6.6)$$

On the other hand, since the last jump of blue before time $T_r + t_c$ is at time $\lambda[(T_r + t_c)/\lambda]$, blue could do $(T_r+t_c)/\lambda - \{(T_r+t_c)/\lambda\}$ many up-jumps, hence right before the intersection, blue occupies some vertices that satisfy

$$\log\left(D_{\max}^{(b,n)}(T_r + t_c)\right) = (\log n)^{\frac{2}{\lambda+1}} \left(\frac{\left(Y_b^{(n)}\right)^\lambda}{Y_r^{(n)}} \frac{\alpha}{\tau-1}\right)^{\frac{1}{\lambda+1}} (\tau-2)^{\frac{1+b_n^{(r)}}{\lambda+1}-1+\{\frac{T_r+t_c}{\lambda}\}}(1+o(1)), \qquad (6.7)$$

where we have used (6.5) for $(T_r + t_c)/\lambda$ combined with (5.4) at time $\lfloor (T_r + t_c)/\lambda \rfloor$.

Note that the formulas (6.6) and (6.7) only differ in the exponents of $1/(\tau - 2)$, and this difference is exactly $d(t_c)$, introduces in defined in (6.4). More precisely,

$$\frac{\log\left(\frac{\log\left(\widetilde{u}_{r_-^\star}\right)}{\log\left(D_{\max}^{(b,n)}(T_r+t_c)\right)}\right)}{|\log(\tau-2)|} = \{t_c\} + \left\{\frac{T_r+t_c}{\lambda}\right\} + o(1) := d(t_c) + o(1). \qquad (6.8)$$

Recall from (6.3) that the remaining time to the next jump for red and blue after the intersection at time $T_r + t_c$ is denoted by J_r and J_b, respectively.
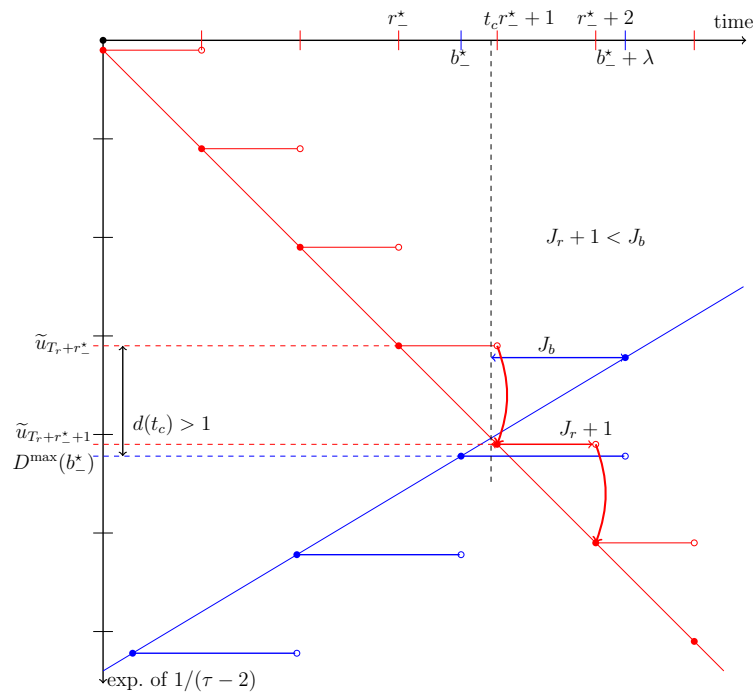
Figure 6: In this picture, the distance between the two exponents, $d(t_c)$ is larger than 1 and $J_r + 1 < J_b$. Thus, red jumps twice after the intersection and occupies every vertex of degree at least $\widetilde{u}_{T_r + r_+^\star 2}$ that is already less than $D^{\max}(b_-^\star)$. So the maximal degree of blue is $D^{\max}(b_-^\star)$.

Since (6.7) is the exponent of the maximal degree vertex that blue occupies before the intersection, to determine the maximal degree of blue, we need to investigate whether blue can jump once more before the red avalanche reaches lower degrees than (6.7). If yes, then blue can gain an additional factor to the rhs of (6.7).

Obviously, if $d(t_c) < 1$, then even though blue jumps first, it cannot increase its exponent by a whole factor $(\tau - 2)^{-1}$, since vertices with degree larger than $(D^{(b,n)}_{\max}(T_r + t_c))^{(\tau-2)^{-1}}$ are already all red. It is not hard to see that blue in this case will occupy some vertices 'right below' $u_{[t_c]}$ (that is, say, higher than $\widetilde{u}_{[t_c]}/(C \log n)$), hence blue in this case can increase its exponent by a factor of $(\tau - 2)^{-d(t_c)}$.

This case illustrates that the additional factor that we need to add to the rhs of (6.7) depends on two things: (1) which color jumps first (and possibly second) after the intersection and (2) whether $d(t_c) > 1$ or not. There are five cases, described above (after formula (6.4)). As a result, the gain in the exponent for blue can be summarized by multiplying (6.7) by the following function containing indicators for these five cases (the order is Case B1, R1, B2, R2, R3 here, and the cases where λ rational are also included):

$$
\begin{aligned}
f\left(d(t_c), J_r, J_b\right) &:= \mathbb{1}_{\{d(t_c)<1\}} \left(\mathbb{1}_{\{J_b \le J_r\}} (\tau - 2)^{-d(t_c)} + \mathbb{1}_{\{J_r < J_b\}} \right) \\
&+ \mathbb{1}_{\{d(t_c)>1\}} \left(\mathbb{1}_{\{J_b \le J_r\}} (\tau - 2)^{-1} + \mathbb{1}_{\{J_r < J_b \le J_r + 1\}} (\tau - 2)^{1-d(t_c)} + \mathbb{1}_{\{J_r + 1 < J_b\}} \right).
\end{aligned}
\tag{6.9}
$$

Note that every expression in this formula, i.e., $d(t_c), J_r, J_b$ are simple functions of the (random) fractional parts $\{t_c\}$, $\left\{ \frac{T_c + t_c}{\lambda} \right\}$ and $b_n^{(r)}$. Further, $\{t_c\}$, $\left\{ \frac{T_c + t_c}{\lambda} \right\}$ and $b_n^{(r)}$ depend only on n and also on $Y_r^{(n)}, Y_b^{(n)}$. Hence knowing these values determines f uniquely.

Recall the value α from (4.1), combine (6.7) with the additional factor $f\left(d(t_c), J_r, J_b\right)$,

so that we can introduce the 'oscillation-filtering' random variable

$$H_n^{\max}(Y_r^{(n)}, Y_b^{(n)}) := \left(\frac{(\tau - 1) - (\tau - 2)^{b_n^{(r)}}}{(\tau - 1)^2} \right)^{\frac{1}{\lambda + 1}} (\tau - 2)^{\frac{-\lambda + b_n^{(r)}}{\lambda + 1} + \left\{ \frac{T_r + t_c}{\lambda} \right\}} f(d(t_c), J_r, J_b), \quad (6.10)$$

which oscillates with n and is random, but depends on the same randomness as $Y_r^{(n)}, Y_b^{(n)}$, i.e., they are defined on the same probability space.

At this point we have shown that

$$\log D_{\max}^{(b,n)}(\infty) = (\log n)^{\frac{2}{\lambda+1}} \frac{\left(Y_b^{(n)} \right)^\lambda}{Y_r^{(n)}} H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})(1 + o_{\mathbb{P}}(1)). \qquad (6.11)$$

To obtain the statement of Theorem 1.4, we have to check the conditions of Lemma 5.1. For this, note that the last layer of blue is reached after climbing $i_{\max} = \lfloor (T_r + t_c)/\lambda \rfloor - \lfloor t(n^{\varrho'})/\lambda \rfloor + \mathbb{1}_{\{\text{Case B2}\}}$ many $\Gamma_i^{(b)}$ layers, and by (6.5) and (3.3) we can calculate

$$i_{\max} := \lfloor (T_r + t_c)/\lambda \rfloor - \lfloor t(n^{\varrho'})/\lambda \rfloor + \mathbb{1}_{\{\text{Case B2}\}} = \frac{\lambda - 1}{\lambda(\lambda + 1)} \frac{\log \log n}{|\log(\tau - 2)|} + O_{\mathbb{P}}(1). \quad (6.12)$$

Thus, $i_{\max} = O^\diamond(\frac{\lambda - 1}{\lambda(\lambda+1)})$, hence the conditions of the lemma hold. Thus, we get the first statement of Theorem 1.4 by noting that $(Y_r^{(n)}, Y_b^{(n)}) \xrightarrow{d} (Y_r, Y_b)$.

By maximizing and minimizing the constants and the fractional parts in (6.10) and in the indicators in $f(d(t_c), J_r, J_c)$ we get the bounds in (1.9). This finishes the proof of Theorem 1.4.

6.2 Number of maximum degree vertices

With the last up-jump of blue, blue occupies some vertices of degree of order $D_{\max}^{(b,n)}(\infty)$. In this section we investigate *how many* maximum degree vertices are reached by blue. We show that in some cases (namely, Cases B1, R2) the number of these vertices is so large that it corresponds to an additional factor for the total number of half-edges in maximum degree vertices of blue.

More precisely, let us denote the set of outgoing half-edges from these maximal degree vertices by $\mathcal{M}_n^{(b)}$, and its size by $M_n^{(b)}$. Later we will determine how many vertices blue can occupy after this phase, and to be able to count that we need to know how many half-edges are in the highest layer of blue.

Lemma 6.3. *For $M_n^{(b)}$, the number of outgoing half-edges from the set of maximal degree vertices, i.e. the sum of the forward degrees reached by blue for which (1.8) holds, we have*

$$\frac{\log M_n^{(b)}}{(\log n)^{\frac{2}{\lambda+1}} H_n^{\text{half-edge}}(Y_r^{(n)}, Y_b^{(n)})} \xrightarrow{\mathbb{P}} \left(\frac{Y_b^\lambda}{Y_r} \right)^{1/(\lambda+1)},$$

where $H_n^{\text{half-edge}}(Y_r^{(n)}, Y_b^{(n)}) \geq H_n^{\max}(Y_r^{(n)}, Y_b^{(n)})$ is a bounded random variable given below in formula (6.17).

Proof. Recall that A_i denotes the number of vertices blue occupies in layer $\Gamma_i^{(b)}$ upon reaching it, see (5.12). In the cases where blue finishes its last jump at a certain layer $\Gamma_i^{(b)}$, that is, in Case R1 (Fig **??** and Case R3 (Fig **??**) and also in Case B2 (Fig **??**) the statement is a direct consequence of Lemma 5.4, since blue is stuck with its maximal degree at a given layer $\Gamma_{i_{\max}}^{(b)}$, and hence $M_n^{(b)} = A_{i_{\max}} D_{\max}^{(b,n)}(\infty)(1 + o(1))$. Taking logarithm we get

$$\log M_n^{(b)} \leq \log D_{\max}^{(b,n)}(\infty)(1 + o(1)) + \log A_{i_{\max}}. \qquad (6.13)$$

By (6.12), $i_{\max} = O^{\diamond}(\frac{\lambda-1}{\lambda(\lambda+1)})$ in Lemma 5.4, so we can use the bound in (5.13) with $x = (\lambda - 1)/\lambda(\lambda + 1)$. Hence, the last term in (6.13) disappears when we divide by $(\log n)^{2/(\lambda+1)}$.

We are left with handling the cases where the last jump of blue is not a full layer, i.e., Cases B1 and R2. In these cases, after reaching layer $\Gamma^{(b)}_{i_{\max}}$, blue still jumps up, but not a full layer: due to the presence of red the forward degrees are truncated at $\widetilde{u}_{r^*}$ in Case B1 and at $\widetilde{u}_{r^*+1}$ in Case R2.

First, we apply Lemma 5.4 to see that $\log A_{i_{\max}}$ in the last 'full' layer $\Gamma^{(b)}_{i_{\max}}$ is small. Let us recall the notation $u^{(b)}_{i_{\max}} = D^{(b,n)}_{\max}(T_r + t_c)$, and introduce the extra factor of the log(degrees) reached at the last up-jump of blue by

$$\kappa := (\tau - 2)^{-d(t_c)\mathbb{1}\{B1\} + (1-d(t_c))\mathbb{1}\{R2\}}.$$

Then we introduce a new layer

$$\Gamma^{\diamond} := \left\{ v \in \mathrm{CM}_n(\boldsymbol{d}) : d_v \geq \frac{(u^{(b)}_{i_{\max}})^{\kappa}}{(\log n)^{1/(\tau-2)}} \right\},$$

and we denote the number of half-edges in this set by $\mathcal{E}_{\kappa}$.

By Lemma 5.4, whp blue does not reach higher degrees than $\widehat{u}^{(b)}_{i_{\max}}$ at time $i_{\max}$. Recall that there are $A_{i_{\max}}$ many blue vertices in layer $\Gamma^{(b)}_{i_{\max}}$. Hence, the total number of blue half-edges in this layer is at most $A_{i_{\max}} \widehat{u}^{(b)}_{i_{\max}}$. Thus, the number of vertices in $\Gamma^{\diamond}$ to which blue is connected is dominated by

$$\mathcal{B} \cap \Gamma^{\diamond} \overset{d}{\leq} \mathrm{Bin}\left(A_{i_{\max}} \widehat{u}^{(b)}_{i_{\max}}, \frac{\mathcal{E}_{\kappa}}{\mathcal{L}_n(1 + o(1))} \right). \tag{6.14}$$

Using Claim 5.3, $\mathcal{E}_{\kappa} \leq (u^{(b)}_{i_{\max}})^{-\kappa(\tau-2)}$ whp. Thus, conditioned on $A_{i_{\max}}$, the expected value of the Binomial variable in (6.14) is bounded above by

$$\frac{2C_1^2}{c_1} A_{i_{\max}} \widehat{u}^{(b)}_{i_{\max}} (u^{(b)}_{i_{\max}})^{-\kappa(\tau-2)} = \frac{2C_1^2}{c_1} A_{i_{\max}} \frac{\widehat{u}^{(b)}_{i_{\max}}}{u^{(b)}_{i_{\max}}} (u^{(b)}_{i_{\max}})^{1-\kappa(\tau-2)}.$$

Since red occupies every vertex with degree larger than $(u^{(b)}_{i_{\max}})^{\kappa}$, the previous formula bounds the number of vertices with degree in the interval $[(u^{(b)}_{i_{\max}})^{\kappa}/C \log n, (u^{(b)}_{i_{\max}})^{\kappa})$. Thus, the total number of half-edges going out from maximal degree vertices can be bounded by

$$\frac{2C_1^2}{c_1} A_{i_{\max}} \frac{\widehat{u}^{(b)}_{i_{\max}}}{u^{(b)}_{i_{\max}}} (u^{(b)}_{i_{\max}})^{1+\kappa(3-\tau)}. \tag{6.15}$$

Since $i_{\max} = O^{\diamond}(\frac{\lambda-1}{\lambda(\lambda+1)})$, we can use (5.13) and the calculations in the proof of Lemma 5.4 to see that

$$A_{i_{\max}} \frac{\widehat{u}^{(b)}_{i_{\max}}}{u^{(b)}_{i_{\max}}} \leq (A_{O^{\diamond}(\frac{\lambda-1}{\lambda(\lambda+1)})})^2$$

is still small, i.e., it disappears when taking logarithm and dividing by $(\log n)^{2/(\lambda+1)}$. Hence, the main contribution comes from $(u^{(b)}_{i_{\max}})^{1+\kappa(3-\tau)} = (D^{(b,n)}_{\max}(T_r + t_c))^{1+\kappa(3-\tau)}$.

Hence, in Cases B1 and R2, blue can get more half-edges than of order $D^{(b,n)}_{\max}(\infty)$. To get the total number of half-edges at the last up-jump, we need to modify the function $f(d(t_c), J_r, J_b)$. An elementary rearranging of the indicators of the cases and the constants shows that the extra factor needed for (6.7) to get $M^{(b)}_n$ is

$$\begin{aligned} g(d(t_c), J_r, J_b) := {} &\mathbb{1}_{\{d(t_c)<1\}}\left(\mathbb{1}_{\{J_b \leq J_r\}}(1 + (3-\tau)(\tau-2)^{-d(t_c)} + \mathbb{1}_{\{J_r < J_b\}}) \right) \\ &+ \mathbb{1}_{\{d(t_c)>1\}}\left(\mathbb{1}_{\{J_b \leq J_r\}}(\tau-2)^{-1} + \mathbb{1}_{\{J_r < J_b \leq J_r+1\}}(1 + (3-\tau)(\tau-2)^{1-d(t_c)}) \right) \\ &+ \mathbb{1}_{\{d(t_c)>1\}}\mathbb{1}_{\{J_r+1 < J_b\}}. \end{aligned} \tag{6.16}$$

Then the normalizing constant for $M_n^{(b)}$ is given by

$$H_n^{\text{half-edge}}(Y_r^{(n)}, Y_b^{(n)}) := \left(\frac{(\tau-1) - (\tau-2)^{b_n^{(r)}}}{(\tau-1)^2} \right)^{\frac{1}{\lambda+1}} (\tau-2)^{-\frac{\lambda+b_n^{(r)}}{\lambda+1} + \left\{ \frac{T_r+t_c}{\lambda} \right\}} g\left(d(t_c), J_r, J_b \right).$$

(6.17)

This finishes the proof of Lemma 6.3. □

Before moving on to the next section, let us introduce the time when the maximal degree is reached, which is nothing else but the time of the last possible up-jump of blue, i.e.,

$$
\begin{aligned}
t_b &:= \lambda \left(\left\lfloor \frac{T_r + t_c}{\lambda} \right\rfloor + \mathbb{1}_{\{J_b < J_r\}} + \mathbb{1}_{\{J_r < J_b < J_r+1\}} \mathbb{1}_{\{d(t_c) > 1\}} \right) \\
&= T_r + t_c + \lambda \left(\mathbb{1}_{\mathcal{E}} - \left\{ \frac{T_r + t_c}{\lambda} \right\} \right) \\
&= \frac{\lambda}{\lambda+1} \frac{2 \log \log n - \log(Y_r^{(n)} Y_b^{(n)}(\tau-1)/\alpha)}{|\log(\tau-2)|} - \frac{1 + b_n^{(r)}}{\lambda+1} + \lambda \left(\mathbb{1}_{\mathcal{E}} - \left\{ \frac{T_r + t_c}{\lambda} \right\} \right),
\end{aligned}
$$

(6.18)

where $\mathcal{E}$ stands for the event that blue has an additional up-jump after time t_c, i.e. Case B1, B2 or R2 happens.

7 Path counting methods for blue

By time t_b, only $o(n)$ vertices are reached by red and blue together – most of the vertices are still not colored. Thus, it still remains to determine how many vertices blue can reach after time t_b. We do this via giving matching upper and lower bounds on how many vertices blue occupies in this last phase.

For the upper bound, the idea is that we count the size of the local neighborhood of the half edges that are just occupied at time t_b. Since the red avalanche continues to be in its avalanche phase and occupies all vertices of smaller and smaller degrees as time passes, the spreading of blue is more and more restricted, so this local neighborhood is quite small. We call this the *optional cluster* of blue. Since its size is random, we give a concentration result on its size, i.e., we give a concentrated upper bound on what blue can get.

For the lower bound, we estimate how much the red color might 'bite out' of this optional cluster. This can happen since even a constant degree vertex might by chance be close to both colors. We show that this intersection of the clusters is negligible compared to the size of the optional cluster.

We start describing the first step – the optional cluster of blue – in more detail. At time t_b, the half-edges in the set $\mathcal{M}_n^{(b)}$ start their own *exploration clusters*, i.e., an exploration process from the half-edge to not-yet occupied vertices. At time $t_b + \lambda j$, we color every vertex v, whose distance is exactly j from some half-edge h in $\mathcal{M}_n^{(b)}$, and the degrees of vertices on the path from h to v are less than what red occupies at that moment, blue. That is, the degree of the jth vertex on the path must be less than $\widetilde{u}_{\lfloor t_b - \lambda i + \lambda j - T_r \rfloor}$. We do this via estimating the number of paths with degree restrictions from $\mathcal{M}_n^{(b)}$ and call this the *optional cluster of blue*, denote the set by $\mathcal{O}_{\max}$ and its size by $O_{\max}$. Corollary 7.2 below determines its asymptotic behavior.

On the other hand, not just the half-edges in $\mathcal{M}_n^{(b)}$ can gain extra blue vertices: from half-edges in $\mathcal{A}_{i_{\max}-z}$, $z = 0, 1, 2 \ldots$ the explorations start a bit earlier (at time $t_b - \lambda z$) towards small degree vertices. Let us denote the vertices reached via half-edges from layer $\mathcal{A}_{i_{\max}-z} \setminus \mathcal{A}_{i_{\max}-z+1}$ by $\mathcal{O}_{-z}, z \geq 0$. At time $t_b - \lambda z + \lambda j$, we color a vertex v blue if its distance is exactly j from a half-edge h in $\mathcal{A}_{i_{\max}-z}$, and the degrees of vertices on the

path from h to v are less than $u^{(b)}_{i_{\max}-z+j}$ and also what red occupies at that moment, i.e., the degree of the jth vertex on the path must be less than $\min\{u^{(b)}_{i_{\max}-z+j}, \widetilde{u}_{\lfloor t_b - \lambda z + \lambda j - T_r \rfloor}\}$. This extra truncation is needed since we want to avoid *double counting*, that is, we do not want to count vertices explored from $\mathcal{A}_{i_{\max}-z}$ towards $\mathcal{A}_{i_{\max}-z+1}$, hence the additional restriction. We show that the total number of optional blue vertices in lower layers, $\sum_{z \geq 0} \mathrm{O}_{-z}$ with these additional explorations is at most the same order as $\mathrm{O}_{\max}$ in Lemma 7.3.

For the lower bound of what blue can occupy after time t_b, note that not every vertex in $\mathcal{O}_{\max}$ will be occupied by blue: red can still bite out some parts of these vertices by simply randomly being close to some parts of the blue cluster. We estimate the number of vertices in the intersection of $\mathcal{O}_{\max}$ and red, and then subtracting the gained estimate from the lower bound on $\mathrm{O}_{\max}$ gives a lower bound on what blue occupies from the graph after t_b, see Lemma 7.4. Now we turn to the calculations.

We introduce the expected truncated degree of a vertex that is distance j away from the set $\mathcal{M}^{(b)}_n$

$$\nu_j := \mathbb{E}\left[D^\star \mathbb{1}_{\{D^\star < \widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}\}}\right] = \frac{1}{\mathbb{E}[D]} \int_0^{\widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}} [1 - F^*(x)]\mathrm{d}x. \tag{7.1}$$

Then, by (1.1),

$$\frac{c_1}{\mathbb{E}[D]} \left(\widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}\right)^{3-\tau} \leq \nu_j \leq \frac{C_1}{\mathbb{E}[D]} \left(\widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}\right)^{3-\tau}. \tag{7.2}$$

Let us also define

$$\kappa_j := \frac{1}{\mathbb{E}[D]} \mathbb{E}[D(D-1)(D-2)\mathbb{1}_{\{D < \widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}\}}]$$

Then again by (1.1),

$$\frac{c_1}{\mathbb{E}[D]} \left(\widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}\right)^{4-\tau} \leq \kappa_j \leq \frac{C_1}{\mathbb{E}[D]} \left(\widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}\right)^{4-\tau}.$$

Let us call a path of length k from $\mathcal{M}^{(b)}_n$ with vertices $(\pi_j)_{j \leq k}$ *good* if $\pi_j \leq \widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}$, and *good-directed* if $u_{\lfloor t_b + \lambda j + 1 - T_r \rfloor} \leq \pi_j \leq \widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor}$. Similarly as before Lemma 5.2, let us introduce the σ-algebra $\mathcal{F}_{t_b}$, generated by the construction of the configuration model - i.e., the red and blue cluster - until time t_b, the last possible up-jump of blue, including now the half-edges in $\mathcal{M}^{(b)}_n$ (but not their pairs).

Lemma 7.1. *For $k \geq 0$, denote by $\mathrm{O}_{\max}(k), \mathrm{O}^d_{\max}(k)$ the number of vertices that are on good and good-directed paths of distance k away from $\mathcal{M}^{(b)}_n$, respectively. Then for every $\varepsilon > 0$, whp*

$$M^{(b)}_n \cdot \prod_{j=1}^k \nu_j (\tfrac{1}{2} - \varepsilon)^{2k} \leq \mathbb{E}[\mathrm{O}_{\max}(k) \mid \mathcal{F}_{t_b}] \leq M^{(b)}_n \cdot \prod_{j=1}^k \nu_j \cdot (1+\varepsilon)^{2k} \left(1 + O\left(\frac{k^2}{n}\right)\right), \tag{7.3}$$

and

$$M^{(b)}_n \cdot \prod_{j=1}^k (\nu_j - \nu_{j+1})(1-\varepsilon)^{2k} \leq \mathbb{E}[\mathrm{O}^d_{\max}(k) \mid \mathcal{F}_{t_b}] \leq M^{(b)}_n \cdot \prod_{j=1}^k (\nu_j - \nu_{j+1})(1+\varepsilon)^{2k} \cdot \left(1 + O\left(\frac{k^2}{n}\right)\right), \tag{7.4}$$

while for the variance of the latter:

$$\mathbf{Var}[\mathrm{O}^d_{\max}(k) | \mathcal{F}_{t_b}] \leq \mathbb{E}[\mathrm{O}^d_{\max}(k) | \mathcal{F}_{t_b}]$$
$$+ \overline{\mathbb{E}[\mathrm{O}_{\max}(k) | \mathcal{F}_{t_b}]}^2 \cdot \left(\frac{\nu_{k-1}}{(\nu_{k-1} - 1)} \frac{\kappa_1}{\nu_1^2}\left(\frac{1}{M^{(b)}_n} + \frac{2}{\mathcal{L}_n}\right) + \frac{\nu_{k-1}^2}{(\nu_{k-1}^2 - 1)^2} \frac{\kappa_1^2}{\nu_1^4} \frac{2}{M^{(b)}_n \mathcal{L}_n} + e_{k,n}\right), \tag{7.5}$$

where $\overline{\mathbb{E}[\mathrm{O}_{\max}^d(k)|\mathcal{F}_{t_b}]}$ means the upper bound on $\mathbb{E}[\mathrm{O}_{\max}^d(k)|\mathcal{F}_{t_b}]$ in (7.4), and the error term $e_{k,n}$ is

$$e_{k,n} = \left(\prod_{i=1}^{k} \frac{\mathcal{L}_n - 2i + 1}{\mathcal{L}_n - 2i - 2k + 1} - 1\right) \tag{7.6}$$
$$+ \left(1 + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2}\frac{1}{M_n^{(b)}}\right)\left(1 + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2}\frac{1}{c\mathcal{L}_n}\right)\frac{k}{\nu_{k-1} - 1}\left(e^{k^2 \kappa_1^2 \nu_{k-1}/(\nu_1^4 \mathcal{L}_n)} - 1\right).$$

The proof of this lemma uses path counting methods and is similar to that of [30, Lemma 5.1]. Similar techniques can also be found in [27, Section 10.4.2]. Since our case is slightly different than the cases handled there, we work out the details in Appendix A.

Now we state the immediate corollary of Lemma 7.1. Recall the definition of t_b from (6.18).

Corollary 7.2 (Chebyshev's inequality for blue vertices)**.** *Take $c_3 \leq \frac{2-\varepsilon}{\lambda+1}|\log(\tau - 2)|^{-1}$ and any $k \leq c_3 \log\log n$. Then, conditioned on the number of blue half-edges $M_n^{(b)}$ at time t_b, the number of vertices optionally occupied by blue up to time $t_b + \lambda k$ satisfies that, conditionally on $M_n^{(b)}$,*

$$\frac{\log(\mathrm{O}_{\max}(k))}{\log M_n^{(b)} + \sum_{i=1}^{k-1} \log \nu_i} \xrightarrow{\mathbb{P}} 1.$$

Proof. Let us write $\mathrm{O}_{\max}^{\mathrm{non\text{-}d}}(k)$ for paths that are good but not good-directed. We show that they have a negligible contribution, while $\mathrm{O}_{\max}^d(k)$ is well-concentrated. Let us write

$$J := \mathbb{P}\left(\left|\mathrm{O}_{\max}(k) - \mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]\right| \geq \frac{1}{2}\mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]\right)$$
$$\leq \mathbb{P}\left(\left|\mathrm{O}_{\max}^d(k) - \mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]\right| \geq \frac{1}{3}\mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]\right) + \mathbb{P}\left(\mathrm{O}_{\max}^{\mathrm{non\text{-}d}}(k) \geq \frac{1}{6}\mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]\right).$$

Now we can apply Chebyshev's inequality on the first term while Markov's inequality on the second term (both conditioned on $\mathcal{F}_{t_b}$), using Lemma 7.1:

$$J \leq \frac{9\mathbf{Var}[\mathrm{O}_{\max}^d(k)|\mathcal{F}_{t_b}]}{\mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]^2} + \frac{\mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}] - \mathbb{E}[\mathrm{O}_{\max}^d(k)|\mathcal{F}_{t_b}]}{6^{-1}\mathbb{E}[\mathrm{O}_{\max}(k)|\mathcal{F}_{t_b}]}$$
$$\leq \left(1 + O\left(\frac{k^2}{n}\right)\right)\left(\frac{1}{M_n^{(b)}}\frac{\kappa_1}{\nu_1^2}\frac{\nu_{k-1}}{(\nu_{k-1} - 1)} + \frac{2k^4 \nu_{k-1}}{\nu_{k-1} - 1}\frac{\kappa_1^2}{\nu_1^4 \mathcal{L}_n}\left(1 + O\left(\frac{1}{M_n^{(b)}}\frac{\kappa_1}{\nu_1^2}\right)\right)\right) \tag{7.7}$$
$$+ 6\left(1 - \prod_{j=1}^{k}\left(1 - \frac{\nu_{j+1}}{\nu_j}\right)\right)$$

The term containing $\kappa_1^2/\nu_1^4\mathcal{L}_n$ is coming from the Taylor expansion of the exponential factor in the formula for $e_{k,n}$. We only have to verify that the rhs of the previous display is tends to 0. For this we need $\kappa_1/(\nu_1^2 M_n^{(b)}) \to 0$ and also $\kappa_1^2/(\nu_1^4\mathcal{L}_n) \to 0$. For the first term, note that $M_n^{(b)} \geq D_{\max}^{(b,n)}(\infty)$, since it counts the number of half-edges with maximal degree $D_{\max}^{(b,n)}(\infty)$. Further, $\kappa_1/\nu_1^2 = \widetilde{u}_{\lfloor t_b+\lambda-T_r\rfloor} = o(D_{\max}^{(b,n)}(\infty))$, since it is not hard to see that at time $t_b + \lambda$, the degree above which red occupies everything (i.e., $\widetilde{u}_{\lfloor t_b+\lambda-T_r\rfloor}$) is already less than $D_{\max}^{(b,n)}(\infty)$, otherwise blue could have still increased its maximal degree at $t_b + \lambda$ by an extra jump. (Technically, this was the definition of t_b. Alternatively, compare the exact values of $D_{\max}^{(b,n)}(\infty)$ in (1.8) and (6.10), and compare it to that of $\widetilde{u}_{\lfloor t_b+\lambda-T_r\rfloor}$, which can be derived from (6.6) by adding the appropriate number of $(\tau - 2)$ factors in the exponent corresponding to the five different cases. This calculation is left to the reader.)

Similarly, the second term, $\kappa_1^2/(\nu_1^4\mathcal{L}_n) = \widetilde{u}_{\lfloor t_b+\lambda-T_r\rfloor}^{2(\tau-2)}/\mathcal{L}_n = \widetilde{u}_{\lfloor t_b+\lambda-T_r\rfloor+1}^2$ is less than of order $D_{\max}^{(b,n)}(\infty)^2/n$ and hence is small as long as $D_{\max}^{(b,n)}(\infty) = o(\sqrt{n})$. Note that this is the case by Theorem 1.4 since $\lambda > 1$.

Finally, we show that the last term in (7.7) is also small. Since $\lambda > 1$, $[t_b + \lambda(j+1) - T_r] \geq [t_b + \lambda j - T_r] + 1$, and $\nu_i \in (c_1, C_1) \times \widetilde{u}_{[t_b + \lambda i - T_r]}^{3-\tau}$ hence the last term is less than 6 times

$$\sum_{j=1}^{k} \frac{\nu_{j+1}}{\nu_j} \leq \sum_{\ell = [t_b + \lambda - T_r]}^{[t_b + \lambda k - T_r]} \frac{C_1}{c_1} \left(\frac{\widetilde{u}_{\ell+1}}{\widetilde{u}_\ell} \right)^{3-\tau} \leq \sum_{\ell = [t_b + \lambda - T_r]}^{[t_b + \lambda k - T_r]} \frac{C_1 (C \log n)^{3-\tau}}{c_1 \widetilde{u}_\ell},$$

where we have used the recursion $\widetilde{u}_{\ell+1} = C \log n \, \widetilde{u}_\ell^{2-\tau}$ in (5.1). Again, by the same recursion, for some large enough constant C', the sum on the rhs is at most

$$\frac{C_1 (C \log n)^{3-\tau}}{c_1} \frac{C'}{\widetilde{u}_{[t_b + \lambda k - T_r]}},$$

which is small as long as $\log \widetilde{u}_{[t_b + \lambda k - T_r]}$ is of larger order than $\log((C \log n)^{3-\tau})$. Note that this holds for an appropriate choice of k, since using (6.6) and the recursion for $\widetilde{u}_\ell$ again, $\log \widetilde{u}_{[t_b + \lambda k - T_r]}$ is of order

$$\log \widetilde{u}_{[t_b + \lambda k - T_r]} = O((\log n)^{2/(1+\lambda)} (\tau - 2)^k).$$

Note that if we now pick $k = o(\log \log n)$, then the exponent $(\log n)^{2/(1+\lambda)}$ stays unchanged and the expression is much larger order than $\log \log n$. □

Recall that $\mathcal{A}_i, A_i$ stands for the set and number of blue vertices in layer $\Gamma_i^{(b)}$ at the time when blue reaches the layer – at time $\lambda[t(n_\rho')/\lambda] + \lambda i$. Also recall that $i_{\max}$ stands for the index of the last $\Gamma_i^{(b)}$ layer ever reached by blue, see (6.12). Further, $\mathrm{O}_{-z}(k)$ is the number of vertices explored via a path of length k starting from a half-edge in $\mathcal{A}_{i_{\max}-z}$ that are not explored via a half-edge from $\mathcal{A}_{i_{\max}-z+1}$. Next we show that $\sum_{z \geq 0} \mathrm{O}_{-z}(k)$ is at most the same order of magnitude as $\mathrm{O}_{\max}(k)$:

Lemma 7.3. *With the notation introduced before,*

$$\log \left(\sum_{z \geq 0} \mathrm{O}_{-z}(k) \right) \leq \log(\mathrm{O}_{\max}(k))(1 + o_{\mathbb{P}}(1)).$$

Proof. Let us denote the number of half-edges in $\mathcal{A}_{i_{\max}-z}$ that are *not connected* directly to $\Gamma_{i_{\max}-z+1}^{(b)}$ by H_{-z}. From Lemma 5.4 we have a bound on the number of vertices A_i in layer $\Gamma_i^{(b)}$, and Lemma 5.2 says that the maximal degree in $\mathcal{A}_i$ is at most $\widehat{u}_i^{(b)}$ whp.

First, let us describe the following construction of the blue cluster spreading through the layers $\Gamma_i^{(b)}$. After an extra time unit λ, A_{i+1} half-edges out of the at most $A_i \widehat{u}_i^{(b)}$ half-edges of blue are connected to half-edges in $\Gamma_{i+1}^{(b)}$, while the other half-edges are not. In the construction of $\mathrm{CM}_n(\boldsymbol{d})$ in Section 2, each half-edge is paired to a uniformly chosen other half-edge. The uniform distribution restricted to a set is still uniform on that set, thus we can think of this procedure by picking A_{i+1} many of the half-edges out of the at most $A_i \widehat{u}_i$ half-edges uniformly at random and connecting them to uniformly chosen half-edges in $\Gamma_{i+1}^{(b)}$.

The rest of the half-edges in $\Gamma_i^{(b)}$ are connected to lower degree vertices, i.e., we can simply pair these half-edges to lower degree vertices than $u_{i+1}^{(b)}$, and apply the path counting method similar as for $\mathcal{M}_n^{(b)}$ in Lemma 7.1, with the restriction that the degree of the j-th vertex on such a path must be less than the degree in $\Gamma_{i_{\max}-z+j}^{(b)}$ if $j \leq z$ and less than the degree where the red avalanche is at the current time when $j > z$, respectively. The restriction for $j \leq z$ is needed to avoid double counting.

Clearly, $H_{-z} \leq A_{i_{\max}-z} \widehat{u}_{i_{\max}-z}$. Then the degree truncation for this process at λj time unit later is at $u_{i_{\max}-z+j}^{(b)}$ if $j \leq z$ and $\widetilde{u}_{t_b + \lambda(j-z) - T_r}$ if $j > z$.

A simple modification of Lemma 7.1 gives the number of vertices found from these half-edges. Moreover, to show that vertices reached from $\mathcal{A}_{i_{\max}-z}$, for $z \geq 1$ are of less order than that reached via $\mathcal{M}_n^{(b)}$, we can use Markov's inequality:

Similarly as in (7.12),

$$\log O_{-z}(k) \leq \log \left(A_{i_{\max}-z} \widehat{u}_{i_{\max}-z} \right) + \sum_{j=1}^{z} \log((\widehat{u}_{i_{\max}-z+j})^{3-\tau}) + \sum_{j=1}^{k-z-1} \log \nu_j, \qquad (7.8)$$

where the exponent $3 - \tau$ comes from a similar calculation than that in (7.2). We claim that the maximum of this quantity is at $z = 0$.

Since $\log O_{\max}(k) = \log M_n^{(b)} + \sum_{j=1}^{k-1} \log \nu_j$, to prove that $O_{-z}(k) \leq O_{\max}(k - z)$, we need to show that the sum of the first two terms in (7.8) are less than $\log M_n^{(b)}$.

By the recursive definition of $\widehat{u}_i$ in (5.5), $\log \widehat{u}_{i_{\max}-z} = (\tau - 2)^z \log \widehat{u}_{i_{\max}}(1 + o(1))$. We can also use the fact from Lemma 5.4 that $A_i = o(\widehat{u}_i)$ for any $i \leq i_{\max}$. Hence

$$\log \left(A_{i_{\max}-z} \widehat{u}_{i_{\max}-z} \right) = (\tau - 2)^z \log(\widehat{u}_{i_{\max}})(1 + o(1)) \text{ whp} \qquad (7.9)$$

and the second term in (7.8) is $(1 + o(1))$ times

$$\sum_{j=-z+1}^{0} (3 - \tau) \log(\widehat{u}_{i_{\max}-z+j}) = \log(\widehat{u}_{i_{\max}})(3 - \tau) \sum_{j=0}^{z-1} (\tau - 2)^j = \log(\widehat{u}_{i_{\max}}) \left(1 - (\tau - 2)^z \right). \qquad (7.10)$$

We see that the sum of the right hand sides of (7.9) and (7.10) is exactly $\log(\widehat{u}_{i_{\max}})$. Thus, returning to (7.8),

$$\log O_{-z}(k) \leq \log \widehat{u}_{i_{\max}} + \sum_{j=1}^{k-z-1} \log \nu_j + o(1).$$

The right hand side is indeed maximal for $z = 0$, for which we have

$$\log O_{-0}(k) \leq \log \widehat{u}_{i_{\max}} + \sum_{j=1}^{k-1} \log \nu_j + o(1).$$

Compare this quantity to $\log O_{\max}(k)$ in Corollary 7.2. Since $\widehat{u}_{i_{\max}}(1 + o(1)) \leq D_{\max}^{(b,n)}(\infty) \leq M_n^{(b)}$, this finishes the proof of Lemma 7.3, since

$$\sum_{z=0}^{O(\log \log n)} O_{-z}(k) \leq O(\log \log n) \max_z O_{-z}(k),$$

and the $\log \log n$ factor becomes a negligible additive term when taking logarithm. $\qquad \square$

Having analysed the size of the optional cluster of blue, we are ready to finish the upper bound of Theorem 1.2 by combining the previous results.

Proof of the upper bound in Theorem 1.2. First, fix $k = k(n) \to \infty$ so that $k(n) = o(\log \log n)$. Then, Lemma 7.3 implies that the logarithm of the total number of vertices that blue paints in the last phase is at most $\log O_{\max}(k)(1 + o_{\mathbb{P}}(1))$. Corollary 7.2 says that the order of magnitude of $\log(O_{\max}(k)) = \log M_n^{(b)} + \sum_{j=1}^{k-1} \log \nu_j + o_{\mathbb{P}}(1)$, where $M_n^{(b)}$ is the number of blue half-edges in the highest layer that blue can reach. Further, Lemma 6.3 determines the order of magnitude of $\log M_n^{(b)}$, which is

$$\log M_n^{(b)} = \left((Y_b^{(n)})^\lambda / Y_r^{(n)} \right)^{1/(\lambda+1)} (\log n)^{\frac{2}{\lambda+1}} H_n^{\text{half-edge}}(Y_r^{(n)}, Y_b^{(n)})(1 + o_{\mathbb{P}}(1)) \qquad (7.11)$$

and hence converges in distribution to $(Y_b^\lambda / Y_r)^{1/(\lambda+1)}$ when divided by the second two factors.

Thus, to get the asymptotic of $\log(\mathrm{O}_{\max}(k))$, it remains to calculate $\sum_{j=1}^{k} \log \nu_j$ and compare it to the order of $\log M_n^{(b)}$. For this recall the definitions of ν_j in (7.1), t_b in (6.18), t_c in (6.5), $\widetilde{u}_\ell$ in (5.3), and the upper bound on ν_j in (7.2). With $\delta_j := \{t_b - T_r + \lambda j\}$,

$$\sum_{j=1}^{k} \log \nu_j \leq \sum_{j=1}^{k} \log \left(C_1 \left(\widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor} \right)^{3-\tau} \right) \tag{7.12}$$

$$\leq \sum_{j=1}^{k} \left\{ \left(\alpha \log n + \beta \log(C \log n) \right) (\tau - 2)^{-1 + t_b - T_r + \lambda j - \delta_j} (3 - \tau) \right\} + k \log(C_1 C \log n).$$

We rewrite $t_b - T_r = t_c + \lambda \left(\mathbb{1}_{\mathcal{E}} - \left\{ \frac{T_r + t_c}{\lambda} \right\} \right)$ in the exponent using (6.18), and then use formula (6.5) to see that $t_c = O^{\diamond}(\frac{\lambda - 1}{\lambda(\lambda + 1)})$. Hence the main order term in $(\tau - 2)^{t_c}$ is $(\log n)^{(\lambda - 1)/(\lambda(\lambda + 1))}$. This implies that the two smaller order terms $k \log(C_1 C \log n)$ and $\beta \log(C \log n)(\tau - 2)^{t_c}$ are $o(\log n (\tau - 2)^{t_c})$ and can be put in a $(1 + o_{\mathbb{P}}(1))$ factor of the main term. Using the exact value of t_c in (6.5) we obtain then

$$\sum_{j=1}^{k} \log \nu_j \leq (1 + o_{\mathbb{P}}(1)) \cdot (\log n)^{\frac{2}{\lambda + 1}} \left(\frac{\left(Y_b^{(n)} \right)^\lambda}{Y_r^{(n)}} \right)^{1/(\lambda + 1)}$$

$$\times \left(\frac{\alpha}{\tau - 1} \right)^{\frac{1}{\lambda + 1}} \left(\sum_{j=1}^{k} (\tau - 2)^{\lambda j - \delta_j} (3 - \tau) \right) (\tau - 2)^{\frac{-\lambda + b_n^{(r)}}{\lambda + 1} + \lambda \left(\mathbb{1}_{\mathcal{E}} - \left\{ \frac{T_r + t_c}{\lambda} \right\} \right)}. \tag{7.13}$$

Note that we have used the formula in (5.3) for $\widetilde{u}_\ell$, but this is valid only until $\ell < O^{\diamond}(1 - \varepsilon)$ steps, so $k = k(n)$ must satisfy $t_b - T_r + \lambda k \leq O^{\diamond}(1 - \varepsilon)$. Again, by (6.18), $t_b - T_r = O^{\diamond}(\frac{\lambda - 1}{\lambda(\lambda + 1)})$, hence any $k := k(n) \to \infty$ with $k(n) = o(\log \log n)$ is still a good choice.

Similarly as in (7.11), we get

$$\sum_{j=1}^{k} \log \nu_j \leq \left(\left(Y_b^{(n)} \right)^\lambda / Y_r^{(n)} \right)^{1/(\lambda + 1)} (\log n)^{\frac{2}{\lambda + 1}} C = H_n^{\mathrm{paths,k}}(Y_r^{(n)}, Y_b^{(n)}), \tag{7.14}$$

where we introduce

$$H_n^{\mathrm{paths,k}}(Y_r^{(n)}, Y_b^{(n)}) := \left(\frac{\tau - 1 - (\tau - 2)^{b_n^{(r)}}}{(\tau - 1)^2} \right)^{\frac{1}{\lambda + 1}} (\tau - 2)^{\frac{-\lambda + b_n^{(r)}}{\lambda + 1} + \lambda \left(\mathbb{1}_{\mathcal{E}} - \left\{ \frac{T_r + t_c}{\lambda} \right\} \right)}$$

$$\times (3 - \tau) \sum_{j=1}^{k} (\tau - 2)^{\lambda j - \delta_j}, \tag{7.15}$$

and where we used that $\alpha = 1 - (\tau - 2)^{b_n^{(r)}}/(\tau - 1)$, see (4.1), where b_n is replaced by $b_n^{(r)}$. Let

$$H_n^{\mathrm{paths}}(Y_r^{(n)}, Y_b^{(n)}) := \lim_{k \to \infty} H_n^{\mathrm{paths,k}}(Y_r^{(n)}, Y_b^{(n)}). \tag{7.16}$$

Note that in this formula, $\delta_j = \{t_b - T_r + \lambda j\} \in [0, 1)$, where both t_b, T_r are integers. This implies that if $\lambda = p$ for some $p \geq 2$ integer, then all $\delta_j \equiv 0$ and the last sum can be carried out and tends to $1/(1 - (\tau - 2)^\lambda)$. Otherwise, the sum is less than $1/(1 - (\tau - 2)^\lambda)$.

Now, recall again that $\log \mathrm{O}_{\max}(k) = \log M_n^{(b)} + \sum_{j=1}^{k} \nu_j + o_{\mathbb{P}}(1)$ by Corollary 7.2, and combine (7.14) with (7.11)

$$\frac{\log \mathrm{O}_{\max}(k)}{(\log n)^{\frac{2}{\lambda + 1}} \left(H_n^{\mathrm{half-edge}}(Y_r^{(n)}, Y_b^{(n)}) + H_n^{\mathrm{paths}}(Y_r^{(n)}, Y_b^{(n)}) \right)} \leq \left(\frac{\left(Y_b^{(n)} \right)^\lambda}{Y_r^{(n)}} \right)^{\frac{1}{\lambda + 1}} + o_{\mathbb{P}}(1).$$

Now we can finally use $Y_r^{(n)} \xrightarrow{d} Y_r$, $Y_b^{(n)} \xrightarrow{d} Y_b$ (by Theorem 2.3). Hence, the right hand side converges to $(Y_b^\lambda / Y_r)^{1/(\lambda+1)}$.

Let us denote

$$H_n(Y_r^{(n)}, Y_b^{(n)}) := H_n^{\text{half-edge}}(Y_r^{(n)}, Y_b^{(n)}) + H_n^{\text{paths}}(Y_r^{(n)}, Y_b^{(n)}). \tag{7.17}$$

Some elementary calculations – optimizing the fractional parts appearing in the exponents – yield the following bounds:

$$\left(\frac{(\tau-2)(3-\tau)}{(\tau-1)^2 \left(1 - (\tau-2)^\lambda\right)} \right)^{\frac{1}{\lambda+1}} \frac{\lambda+1}{\lambda^{\frac{\lambda}{\lambda+1}}} \leq H_n(Y_r^{(n)}, Y_b^{(n)}) \leq \frac{(\tau-2)^{-\frac{2\lambda+1}{\lambda+1}}}{4^{\frac{1}{\lambda+1}}} \left(1 + \frac{3-\tau}{1 - (\tau-2)^\lambda} \right). \tag{7.18}$$

The bounds given in Theorem 1.2 are the simplified versions of (7.18). $\qquad \square$

For the lower bound of Theorem 1.2 we need to show that most of the optional cluster of blue is actually going to be occupied by blue. For this, let us introduce the notation $\mathcal{O}(k) := \bigcup_{z \geq 0} \mathcal{O}_{-z}(k) \cup \mathcal{O}_{\max}(k)$, and set $\text{Opt}(k) := |\mathcal{O}(k)|$, where k stands for the length of the paths we are counting. The next lemma shows that essentially all the vertices in $\mathcal{O}(k)$ for some $k = k_n = o(\log\log n)$ will indeed be painted blue, i.e., red cannot accidentally bite out too much from this set.

Lemma 7.4. *Set $k = k(n) = o(\log\log n)$. The number of vertices in the intersection of $\mathcal{R}_{t_b + \lambda k}$ and $\mathcal{O}(k)$ is small, i.e.,*

$$|\mathcal{O}(k) \cap \mathcal{R}_{t_b + \lambda k}| = o_{\mathbb{P}}(\text{O}_{\max}(k)),$$

hence

$$|\mathcal{O}(k) \setminus (\mathcal{O}(k) \cap \mathcal{R}_{t_b + \lambda k})| = \text{O}_{\max}(k)(1 + o_{\mathbb{P}}(1)).$$

The proof of the lemma will follow from the following claim:

Claim 7.5. *Suppose $S \subset \text{CM}_n(\boldsymbol{d})$ is an arbitrary subset of vertices such that $|S| = o(n)$ but $|S| := s_n \to \infty$. Then with high probability, the total number half-edges that point out of S, $H(S)$ is at most $|S|^{(\tau-2)/(\tau-1)} n^{1/(\tau-1)}$. Hence, $H(S) = o(n)$ holds also.*

Proof. In the worst case scenario, S contains the $|S|$ highest degree-vertices. Order the degrees $D_1, D_2, \ldots, D_n$ in $\text{CM}_n(\boldsymbol{d})$ of vertices in non-increasing order: $D^{(1)} \geq D^{(2)} \geq \cdots \geq D^{(n)}$. Then with an arbitrarily chosen $K_n > 0$ the following bound always hold:

$$\sum_{v \in S} d_v \leq \sum_{j=1}^{|S|} D^{(j)} \mathbb{1}_{\{D^{(j)} < K_n\}} + \sum_{i=1}^{n} D_i \mathbb{1}_{\{D_i > K_n\}}. \tag{7.19}$$

By Lemma 5.3, the second term is whp at most $n K_n^{2-\tau}$, with the error probability $\exp\{-c n K_n^{1-\tau}\}$ being small as long as $K_n = o(n^{1/(\tau-1)})$. Hence the rhs of (7.19) is whp at most

$$\sum_{v \in S} d_v \leq |S| K_n + n K_n^{2-\tau}.$$

For the right hand side to be minimal we set $K_n := ((\tau-2)n/|S|)^{1/(\tau-1)}$, which is $o(n^{1/(\tau-1)})$ as long as $|S| \to \infty$ with n. With this choice of K_n,

$$\sum_{v \in S} d_v \leq |S|^{\frac{\tau-2}{\tau-1}} n^{\frac{1}{\tau-1}}.$$

Since the exponents sum up to 1, the rhs is always $o(n)$ if $|S| = o(n)$. $\qquad \square$

Proof of Lemma 7.4. Note that we can construct the configuration model by pairing the half-edges in an arbitrarily chosen order. This enables the joint construction of the graph and the spread of the red and blue cluster. Hence, we can assume that if a vertex is not yet colored, its half-edges are still free, and we do not have to take into account the effect that whole paths can be blocked away from one color by the other color by painting one or a few vertices only.

Fix the length of the blue exploration path k. For a set S of vertices, we denote by $H(S)$ the total number of half-edges that point out of the set S. As a lower bound, we can use the adapted rule that whenever red and blue arrives at a vertex at the same time, it is going to be red deterministically. We can further assume that if this is the case, i.e., there are simultaneous jumps of red and blue, then we always pair the red half-edges first, i.e., when pairing the blue half-edges at time $t_b + \lambda i$, we consider $\mathcal{R}_{t_b + \lambda i}$ as already determined.

Let us consider a path π ending in $\mathcal{O}(k)$ given by the sequence of half-edges and vertices $(\pi_0, s_0, t_1, \pi_1, s_1, \ldots, t_k, \pi_k)$, that is, s_i is the half-edge pointing out of vertex π_i that we pair to t_{i+1}, a half-edge belonging to vertex π_{i+1}. We call this path *thinned at step i* if the half-edge s_{i-1} is paired to the half-edge t_i where π_i is already red, i.e. $\pi_i \in \mathcal{R}_{t_b + \lambda i}$. We call a path *thinned* if it is thinned at some $i \le k$.

Clearly, each time we pair a blue half-edge at time $t_b + \lambda i$, it is with probability $H(\mathcal{R}_{t_b + \lambda i})/\mathcal{L}_n(1 + o(1))$ paired to a red half-edge. Let us denote $\sigma_k := \{\mathcal{R}_{t_b + \lambda i}\}_{i=1}^{k}$. Hence, the probability that a particular path ending in $\mathcal{O}(k)$ to be thinned can be bounded using a union bound

$$\mathbb{P}\left((\pi_0, s_0, t_1, \pi_1, s_1, \ldots, t_k, \pi_k) \in \mathcal{O}(k) \text{ thinned} \,\big|\, \sigma_k\right) \le \sum_{i=1}^{k} \frac{H(\mathcal{R}_{t_b + \lambda i})}{\mathcal{L}_n} =: p_{\text{th,k}},$$

as long as $k = k(n)$ is so that the quantity on the rhs is less than 1. Hence, for any function $\delta_{n,k}$ so that $\delta_{n,k} p_{\text{th,k}} < 1$, the proportion of vertices in $\mathcal{O}(k)$ that are thinned - denoted by $\mathcal{O}_{\text{th}}(k)$ - by Markov's inequality is at most

$$\mathbb{P}\left(\frac{|\mathcal{O}_{\text{th}}(k)|}{|\mathcal{O}(k)|} \ge \delta_{n,k} p_{\text{th,k}} \,\Big|\, \sigma_k, \mathcal{O}(k)\right) \le \frac{\mathbb{E}[\frac{1}{\mathcal{O}(k)} \sum_{\pi \in \mathcal{O}(k)} \mathbb{1}_{\{\pi \text{ thinned}\}} | \sigma_k, \mathcal{O}(k)]}{\delta_{n,k} p_{\text{th},k}} \le \frac{1}{\delta_{n,k}}. \tag{7.20}$$

Now, note that we are done with the lower bound if we can pick a $k = k(n) \to \infty$ and an $\delta_{n,k} \to \infty$ so that $\delta_{n,k} \to \infty$ and $\delta_{n,k} p_{\text{th,k}} < 1$.

For this, let us temporarily believe that $\widetilde{k} := \widetilde{k}(n) = \log\log\log n$ has the property that $\mathcal{R}_{t_b + \lambda \widetilde{k}} = o_{\mathbb{P}}(n)$. Then, let us write $\mathcal{R}_{t_b + \lambda \widetilde{k}} := O(n/\omega_{n,\widetilde{k}})$ where $\omega_{n,\widetilde{k}} \to \infty$ with $n \to \infty$. Set

$$k := k(n) = \min\{\log\log\log n, (\omega_{n,\widetilde{k}})^{\frac{\tau-2}{2(\tau-1)}}\}. \tag{7.21}$$

Clearly, $k \le \widetilde{k}$ holds, hence, by monotonicity we have $\mathcal{R}_{t_b + \lambda k} \le \mathcal{R}_{t_b + \lambda \widetilde{k}}$. Applying Claim 5.3 on each term in the sum,

$$p_{\text{th},k} = \sum_{i=1}^{k} \frac{H(\mathcal{R}_{t_b + \lambda i})}{\mathcal{L}_n} \le \frac{k \, n}{\mathcal{L}_n (\omega_{n,\widetilde{k}})^{(\tau-2)/(\tau-1)}}.$$

On the event $\{\mathcal{L}_n \in (1/2\mathbb{E}[D]n, 2\mathbb{E}[D]n)\}$, using (7.21),

$$p_{\text{th},k} \le \frac{2}{\mathbb{E}[D]} \left(\frac{1}{\omega_{n,\widetilde{k}}}\right)^{\frac{\tau-2}{2(\tau-1)}}.$$

This allows us to pick $\delta_{n,k} := (\omega_{n,\widetilde{k}})^{\frac{\tau-2}{4(\tau-1)}}$, and then $\delta_{n,k} p_{\text{th,k}} \to 0$ as well as $\delta_{n,k} \to \infty$ holds with $n \to \infty$. As a result, the rhs of (7.20) tends to zero, showing that whp, only a negligible fraction of the vertices in $\mathcal{O}(k)$ will be thinned.

We are left showing that with $\widetilde{k} = \widetilde{k}(n) = \log \log \log n$, we have $|\mathcal{R}_{t_b + \lambda k}| = o_\mathbb{P}(n)$. One way to see this is to use [28, Theorem 1.2] about typical distances: typical distances in the graph are $2 \log \log n / |\log(\tau - 2)|$ with bounded fluctuations around this value, while $t_b + \lambda \widetilde{k} < (1 + \varepsilon) 2\lambda/(\lambda + 1) \log \log n / |\log(\tau - 2)|$. Hence, the number of vertices at most $t_b + \lambda \widetilde{k}$ away from the uniformly chosen red source vertex must be $o(n)$.

To keep the paper self-contained, we provide another proof of this fact here. For this, note that $t_b + \lambda \widetilde{k}$, t_b defined in (6.18) is at most $\frac{(1+\varepsilon)\lambda}{\lambda+1} \frac{2 \log \log n}{|\log(\tau-2)|}$ for some $\varepsilon > 0$ whp. To estimate the expected size of the red cluster, we write

$$\mathbb{E}\left[|\mathcal{R}_{t_b + \lambda \widetilde{k}}|\right] \leq n \, \mathbb{P}\left(\mathcal{D}(\mathcal{R}_0, v) \leq t_b + \lambda k\right)$$
$$\leq n \, \mathbb{P}\left(\mathcal{D}(\mathcal{R}_0, v)/2 \leq \frac{(1+\varepsilon)\lambda}{\lambda+1} \frac{\log \log n}{|\log(\tau-2)|}\right)$$

Now, using that $\mathcal{D}(\mathcal{R}_0, v)/2$ has the distribution $t_b\big|_{\lambda=1}$, we can continue the bound as

$$\mathbb{E}\left[|\mathcal{R}_{t_b + \lambda \widetilde{k}}|/n\right] \leq \mathbb{P}\left(t_b\big|_{\lambda=1} \leq \frac{(1+\varepsilon)\lambda}{\lambda+1} \frac{\log \log n}{|\log(\tau-2)|}\right)$$
$$\leq \mathbb{P}\left(\log(Y_r^{(n)} Y_b^{(n)}) \geq \frac{1 - 2\varepsilon\lambda}{\lambda+1} \log \log n\right).$$

In the last line, we used (6.18) with $\lambda = 1$ and put bounded terms there in the $(1 - 2\varepsilon\lambda)$ factor on the rhs inside the probability sign. Further, note that the random variables $Y_r^{(n)} Y_b^{(n)}$ have asymptotically exponential tails. Hence, the probability is tending to zero as $n \to \infty$. This ensures that most vertices are further away from the source of the red infection than $t_b + \lambda k$ and hence $\mathcal{R}_{t_b + \lambda \widetilde{k}} = o(n)$. This finishes the proof of the lemma. $\square$

Proof of the lower bound in Theorem 1.2. First, note that the time T_r for red to reach the top of the mountain was a lower bound, i.e., we have shown the existence of a path that reaches the top in time T_r whp in Lemma 3.3. Clearly, if red reaches the top earlier, then there is less time for blue to increase its degree, hence, it will occupy fewer vertices. Fortunately, an adaptation of Lemma 5.2 for red instead of blue shows that this cannot happen. That is, one can define the sequence $\widehat{u}_i^{(r)}$ by the recursion

$$u_0^{(r)} := (n^{\varrho''} C \log n)^{1/(\tau-2)}, \qquad u_i^{(r)} := (u_0^{(r)} C \log n)^{1/(\tau-2)},$$

and then exchange every superscript (b) to (r) in the definition of $\mathcal{B}ad\mathcal{P}_k$ (see right before Lemma 5.2). Applying Lemma 5.2 yields that with high probability, red cannot jump a layer ahead, and hence the time to reach the top remains as defined in T_r.

Next, everything from this point on was a concentrated estimate, hence, we only need to check what happens in the last phase, how many vertices blue can actually get from its optional cluster.

Using Lemma 7.4, we see that the log-size of the blue cluster at time $t_b + \lambda k$ is whp $\log O(k) = \log O_{\max}(k)(1 + o_\mathbb{P}(1))$. Note also that by Corollary 7.2, $\log O_{\max}(k)$ is concentrated and is equal to $\log M_n^{(b)} + \sum_{j=1}^k \log \nu_j + o_\mathbb{P}(1)$ by Lemma 6.3. Hence, it only remains to give a lower bound on $\sum_{j=1}^k \log \nu_j$. For this, note that the lower bound on ν_j is the same as the upper bound, with a factor C_1 replaced by c_1. This factor becomes an additive term when taking the logarithm, and hence contributing only inside the $o_\mathbb{P}(1)$ factor. Hence,

$$\frac{\log O_{\max}(k)}{(\log n)^{\frac{2}{\lambda+1}} \left(H_n^{\text{half-edge}}(Y_r^{(n)}, Y_b^{(n)}) + H_n^{\text{paths}}(Y_r^{(n)}, Y_b^{(n)})\right)} \geq \left(\frac{(Y_b^{(n)})^\lambda}{Y_r^{(n)}}\right)^{\frac{1}{\lambda+1}} + o_\mathbb{P}(1).$$

The right hand side converges to $(Y_b^\lambda / Y_r)^{1/(\lambda+1)}$. Combining this with the upper bound completes the proof of the Theorem 1.2. $\square$

A Path counting methods for restricted paths

For the proof of Lemmas 5.2 and 7.1 we need a preliminary result considering the behavior of the degree sequence.

Proposition A.1 (Error probabilities for degrees in restricted sets)**.** *for $y_n \geq 0$, define the following events $\mathcal{G}_2(y_n, \varepsilon), \mathcal{G}_3(y_n, \varepsilon)$ concerning the empirical second and third factorial moments of the restricted degrees*

$$\mathcal{G}_2(y_n, \varepsilon) := \left\{ \left| \sum_{i=1}^{n} D_i(D_i - 1)\mathbb{1}\{D_i \leq y_n\} - n\mathbb{E}[D(D-1)\mathbb{1}_{D \leq y_n}] \right| \geq \varepsilon n \mathbb{E}[D(D-1)\mathbb{1}_{D \leq y_n}] \right\},$$

$$\mathcal{G}_3(y_n, \varepsilon) := \left\{ \left| \sum_{i=1}^{n} D_i(D_i - 1)(D_i - 2)\mathbb{1}\{D_i \leq y_n\} - n\mathbb{E}[D(D-1)(D-2)\mathbb{1}_{D \leq y_n}] \right| \right.$$

$$\left. \geq \varepsilon n \mathbb{E}[D(D-1)(D-2)\mathbb{1}_{D \leq y_n}] \right\}.$$

Then, for $j = 1, 2$, some $C > 0$ that is independent of y_n, n

$$\mathbb{P}(\mathcal{G}_j(y_n, \varepsilon)) \leq C y_n^{\tau - 1}/n,$$

Further, for the total number of half-edges $\mathcal{L}_n$, we have the following large deviation lower bound for some constant $c(\varepsilon) > 0$:

$$\mathbb{P}\left(\sum_{i=1}^{n} D_i \leq (1-\varepsilon)n\mathbb{E}[D] \right) \leq \exp\{-nc(\varepsilon)\}.$$

Proof. The proof of the first statement follows from Chebishev's inequality and the fact that the degrees are i.i.d. from distribution F, satisfying (1.1):

$$\mathbb{P}(\mathcal{G}_2(y_n, \varepsilon)) = \frac{\operatorname{Var}\left[D(D-1)\mathbb{1}\{D \leq y_n\} \right]}{\varepsilon^2 n \mathbb{E}[D(D-1)\mathbb{1}_{D \leq y_n}]^2} \leq \frac{C y_n^{5-\tau}}{n\varepsilon^2 y_n^{2(3-\tau)}} = C\frac{y_n^{\tau-1}}{n}.$$

For $\mathcal{G}_3(y_n, \varepsilon)$, the proof is analogous: on the rhs the exponent of y_n in the numerator is $7 - \tau$ and $2(4 - \tau)$ in the denominator. The proof of the second statement uses the usual Chernoff bound. For any $t > 0$,

$$\mathbb{P}\left(\mathrm{e}^{-t\sum_{i=1}^{n} D_i} \geq \mathrm{e}^{-t(1-\varepsilon)n\mathbb{E}[D]} \right) \leq \mathrm{e}^{t(1-\varepsilon)n\mathbb{E}[D]} \mathbb{E}\left[\mathrm{e}^{-tD} \right]^n$$

From (1.1) it follows that $\mathbb{E}\left[\mathrm{e}^{-tD} \right] = 1 - t\mathbb{E}[D] + O(t^{\tau-1})$, hence, the rhs can be bounded from above by

$$\exp\{-n(\varepsilon t\mathbb{E}(D) + O(t^{\tau-1}))\}.$$

Setting the value of t and modifying the constant to swallow the error term finishes the proof. $\qquad\square$

Proof of Lemma 5.2. Here we follow the notation of [27, Section 10.4.2] as much as we can. We will use union bound and Markov's inequality to bound the probability of the existence of bad paths:

$$\mathbb{P}_n(\exists k \leq k_0, \mathcal{B}ad\mathcal{P}_k \neq \varnothing | \mathcal{F}_{t(n^{\varrho'})}) \leq \sum_{k=1}^{k_0} \mathbb{E}_n[|\mathcal{B}ad\mathcal{P}_k| | \mathcal{F}_{t(n^{\varrho'})}]. \tag{A.1}$$

First we give an upper bound on the expected number of bad paths conditioned on the degree sequence, so let us fix the degrees first and write d_v for the degree of the vertex v. A (directed) path of length k from vertex a to some vertex π_k can be described as

$$\{(\pi_0, s_0), (\pi_1, t_1, s_1), \dots, (\pi_{k-1}, t_{k-1}, s_{k-1}), (\pi_k, t_k)\}, \tag{A.2}$$

where $\pi_i \in [n]$ is the i-th mid-vertex along the path, $s_i \in [d_{\pi_i}]$ denotes the label of the outgoing and $t_i \in [d_{\pi_i}]$ the label of the incoming half-edge of π_i. Recall that we call a path *good* if $\deg(\pi_i) \le \widehat{u}_i^{(b)}$ for all $0 \le i \le k$, and $\mathcal{B}ad\mathcal{P}_k$ is a subset of bad paths with $\pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}$, $d_{\pi_i} \le \widehat{u}_i^{(b)}$ for all $0 \le i \le k-1$ but $d_{\pi_k} > \widehat{u}_k^{(b)}$.

Since the number of half-edges out of vertex π_i is d_{π_i}, there are many possible paths via the vertices $(\pi_i)_{i=0}^k$. Thus, the expected number of paths through fixed vertices $\pi_0, \ldots, \pi_k$ equals the probability that a given path in (A.2) is present in $\mathrm{CM}_n(\boldsymbol{d})$ multiplied by the combinatorial factor of picking the possible half-edges for the paths, i.e.,

$$\prod_{i=1}^k \frac{1}{\mathcal{L}_n^* - 2i + 1} \cdot d_{\pi_0} \left(\prod_{i=1}^{k-1} d_{\pi_i}(d_{\pi_i} - 1) \right) d_{\pi_k}, \tag{A.3}$$

where $\mathcal{L}_n^*$ is the number of free half-edges when the procedure starts, i.e., at time $t(n^{\rho'})$. Thus, the expected number of all self-avoiding *bad paths in* $\mathcal{B}ad\mathcal{P}_k$ equals

$$\mathbb{E}_n[|\mathcal{B}ad\mathcal{P}_k| | \mathcal{F}_{t(n^{\varrho'})}] = \sum_{\pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}} d_{\pi_0} \left(\prod_{i=1}^k \frac{\mathcal{L}_n^*}{\mathcal{L}_n^* - 2i + 1} \right) \sum_{\substack{\pi_1, \ldots, \pi_{k-1} \\ \forall i \; \pi_i \le \widehat{u}_i^{(b)}}}^* \left(\prod_{i=1}^{k-1} \frac{d_{\pi_i}(d_{\pi_i} - 1)}{\mathcal{L}_n^*} \right) \left(\sum_{\substack{\pi_k \\ d_{\pi_k} \ge \widehat{u}_k^{(b)}}} \frac{d_{\pi_k}}{\mathcal{L}_n^*} \right), \tag{A.4}$$

where $\sum^*$ means that we sum over distinct vertices. Allowing non-distinct vertices, we get the upper bound

$$\mathbb{E}_n[|\mathcal{B}ad\mathcal{P}_k| | \mathcal{F}_{t(n^{\varrho'})}] \le \sum_{\pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}} d_{\pi_0} e^{\frac{k^2}{\mathcal{L}_n^*}} \prod_{i=1}^{k-1} \left(\sum_{d_{\pi_i} \le \widehat{u}_i^{(b)}} \frac{d_{\pi_i}(d_{\pi_i} - 1)}{\mathcal{L}_n^*} \right) \left(\sum_{\substack{\pi_k \\ d_{\pi_k} \ge \widehat{u}_k^{(b)}}} \frac{d_{\pi_k}}{\mathcal{L}_n^*} \right), \tag{A.5}$$

where the factor $e^{\frac{k^2}{\mathcal{L}_n^*}}$ is a bound on the term $\prod_{i=1}^k \mathcal{L}_n^*/(\mathcal{L}_n^* - 2i + 1)$ above. Note that $\mathcal{L}_n^*$ is $\mathcal{L}_n$ minus the already used or active half-edges at time $t(n^{\varrho'})$, and these are the ones in or incident to the red or blue Branching Process, and as a result, their total number is no more than of order $n^{\varrho'/(\tau-2)}$. Hence, $\mathcal{L}_n^* = \mathcal{L}_n - O(n^{\varrho'/(\tau-2)})$. Hence, when we set $k_0 = O(\log \log n)$, the factor $\exp\{\frac{k^2}{\mathcal{L}_n^*}\} = 1 + o(1)$ and stays close to 1. Further, when we switch from $\mathcal{L}_n^*$ to $\mathcal{L}_n$ in the denominators, this causes an error factor $(\mathcal{L}_n/\mathcal{L}_n^*)^k$, which is at most $(1 + O(n^{\rho'-1}))^k \le \exp\{kn^{\varrho'/(\tau-2)-1}\}$. Since $k = O(\log \log n)$ and $\rho'/(\tau-2) < 1$, this factor also tends to 1 with n and can be bounded from above by a constant.

By the Law of Large Numbers, $\mathcal{L}_n/n \to \mathbb{E}[D]$, hence the i-th factor on the right hand side is close to

$$\frac{1}{\mathbb{E}[D]} \mathbb{E}\left[D(D-1)\mathbb{1}_{\left\{ D \le \widehat{u}_i^{(b)} \right\}} \right] \le C_3 (\widehat{u}_i^{(b)})^{3-\tau}, \tag{A.6}$$

while the last factor in (A.5) is close to

$$\frac{1}{\mathbb{E}[D]} \mathbb{E}\left[D\mathbb{1}_{\left\{ D \ge \widehat{u}_k^{(b)} \right\}} \right] \le C_3 (\widehat{u}_k^{(b)})^{2-\tau} \tag{A.7}$$

by the tail behavior (1.1) of the distribution function of D, for some constant C_3. Recall that π_0 is a vertex that belongs to the last generation of the branching process approximation phase. By the coupling to the BP, the degrees in this generation are i.i.d. from distribution F^* in (1.2), hence for some constants $C', C_2 > 0$ we expect

$$\sum_{\pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}} d_{\pi_0} \le (C' \log n Z_{[t(n^{\varrho'})/\lambda]})^{1/(\tau-2)} \le C_2 \widehat{u}_0^{(b)}, \tag{A.8}$$

by the definition of $\widehat{u}_0^{(b)}$ in (5.5). (Note that only the vertices are included in $\mathcal{F}_{t(n^{\varrho'})}$ not their degrees, i.e., it makes sense to take conditional expectation of this event later w.r.t.

$\mathcal{F}_{t(n^{\varrho'})}$.) Making the statements in (A.6), (A.7) and (A.8) rigorous, the complement of the following event[3]

$$\mathcal{E}_n := \left\{ \forall i < k \sum_{d_{\pi_i} \leq \widehat{u}_i^{(b)}} \frac{d_{\pi_i}(d_{\pi_i} - 1)}{\mathcal{L}_n^*} < \omega(n)(\widehat{u}_i^{(b)})^{3-\tau} \text{ and } \sum_{\pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}} d_{\pi_0} < \omega(n)\widehat{u}_0^{(b)} \right.$$

$$\left. \text{and} \sum_{\substack{\pi_k \\ d_{\pi_k} \geq \widehat{u}_k^{(b)}}} \frac{d_{\pi_k}}{\mathcal{L}_n^*} \leq \omega(n)(\widehat{u}_k^{(b)})^{2-\tau} \right\} \tag{A.9}$$

can be estimated by Markov inequality and union bound for any function $\omega(n) \to \infty$, yielding

$$\mathbb{P}(\mathcal{E}_n | \mathcal{F}_{t(n^{\varrho'})}) \geq 1 - \frac{k+1}{\omega(n)}.$$

Further note that, with $m := Z_{[t(n^{\varrho'})/\lambda]}$ and the definition of $\widehat{u}_0^{(b)}$ again, by Lemma 3.1,

$$\mathbb{P}(\exists \pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}, d_{\pi_0} > \widehat{u}_0^{(b)} | \mathcal{F}_{t(n^{\varrho'})}) \leq \frac{1}{n^{c_1 C}}. \tag{A.10}$$

Then on $\mathcal{E}_n$, we can write

$$\mathbb{E}_n[|\mathcal{B}ad\mathcal{P}_k| | \mathcal{F}_{t(n^{\varrho'})}, \mathcal{E}_n] \leq \omega(n)^2 \widehat{u}_0^{(b)}(\widehat{u}_k^{(b)})^{2-\tau} \prod_{i=1}^{k-1} \left(\omega(n)(\widehat{u}_i^{(b)})^{3-\tau} \right). \tag{A.11}$$

The recursion for $\widehat{u}_i^{(b)}$ in (5.5) gives

$$\widehat{u}_i^{(b)} = (\widehat{u}_0^{(b)})^{(\tau-2)^{-i}} (C \log n)^{((\tau-2)^{-i} - 1)/(3-\tau)}.$$

Using this in (A.11), after a somewhat longish but elementary calculation, the powers of $(\tau - 2)^{-1}$ cancel in the exponent of $\widehat{u}_0^{(b)}$ and $C \log n$, and the formula simplifies to

$$\mathbb{E}_n[|\mathcal{B}ad\mathcal{P}_k| | \mathcal{F}_{t(n^{\varrho'})}] \leq \omega(n)(\omega(n)/C \log n)^k.$$

This estimate and (A.10) together implies that the union bound in (A.1) leads to

$$\mathbb{P}(\exists k \leq k_0, \mathcal{B}ad\mathcal{P}_k \neq \varnothing) \leq \omega(n) \sum_{k=1}^{\infty} (\omega(n)/C \log n)^k + \mathbb{P}(\exists \pi_0 \in Z_{[t(n^{\varrho'})/\lambda]}, d_{\pi_0} > \widehat{u}_0^{(b)}) \leq \frac{2\omega(n)^2}{C \log n}.$$

Setting e.g. $\omega(n) := \sqrt{\log \log n}$ completes the proof of Lemma 5.2. $\qquad \square$

Proof of Lemma 7.1. Similarly as in the proof of Lemma 5.2, first we give upper and lower bound on the expected number of k-length paths starting from half-edges in the set $\mathcal{M}_n^{(b)}$ conditioned on the degree sequence and on $\mathcal{F}_{t_b}$. Then, we relate the number of paths to the total number of vertices found in $\mathrm{O}_{\max}(k)$ or $\mathrm{O}_{\max}^{\mathrm{d}}(k)$.

Let us fix the degrees first and write d_v for the degree of the vertex v. Let us also introduce the nested sets $\Lambda_1 \supset \Lambda_2 \supset \ldots$-s by

$$\Lambda_i = \left\{ v \in [n] : d_v \leq \widetilde{u}_{\lfloor t_b + \lambda i - T_r \rfloor} \right\}.$$

Similarly as in the previous lemma, a (directed) path of length k from vertex $a = \pi_0$ to $b = \pi_k$ can be described as

$$\{(\pi_0, s_0), (\pi_1, t_1, s_1), \ldots, (\pi_{k-1}, t_{k-1}, s_{k-1}), (\pi_k, t_k)\}. \tag{A.12}$$

---

[3]note that in the definition of $\mathcal{E}_n$, the first intersection could be switched to a sharper estimate $\bigcap_{1 \leq i \leq k} \mathcal{G}(\widehat{u}_i^{(b)}, \varepsilon)$, on the cost of having larger error terms, (order $1/\varepsilon$). The extra factor $(1 + \varepsilon)^k$ or $\omega(n)^k$ is swallowed either way at the end, when dividing by $\log n$.

We call a path now *good* if $\pi_i \in \Lambda_i$ and *good-directed* if $\pi_i \in \Lambda_i \setminus \Lambda_{i+1}$ for all $1 \le i \le k$.

We write $N_k(a,b)$, $N_k^{\mathrm{d}}(a,b)$ for the number of self-avoiding good paths and good-directed paths going from vertex a to b, respectively, and $\mathcal{L}_n^*$ for the total number of half-edges present in $\mathrm{CM}_n(\boldsymbol{d})$ at time t_b. Similarly as in (A.4), the expected value of all self-avoiding *good paths* equals

$$\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}] = \frac{d_a d_b}{\mathcal{L}_n^* - 2k + 1}\left(\prod_{i=1}^{k-1}\frac{\mathcal{L}_n}{\mathcal{L}_n^* - 2i + 1}\right) \cdot \sum_{\substack{\pi_1,\ldots,\pi_{k-1} \\ \forall i \ \pi_i \in \Lambda_i}}^* \left(\prod_{i=1}^{k-1}\frac{d_{\pi_i}(d_{\pi_i}-1)}{\mathcal{L}_n}\right) \quad \text{(A.13)}$$

where $\sum^*$ means that we sum over distinct vertices. Now clearly we have the upper bound

$$\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}] \le d_a \frac{d_b}{\mathcal{L}_n}\left(\frac{\mathcal{L}_n}{\mathcal{L}_n^*}\right)^k \cdot \mathrm{e}^{\frac{k^2}{\mathcal{L}_n^*}} \cdot \prod_{i=1}^{k-1}\left(\sum_{\pi_i \in \Lambda_i}\frac{d_{\pi_i}(d_{\pi_i}-1)}{\mathcal{L}_n}\right) := \overline{n}_k(a,b). \quad \text{(A.14)}$$

Note that $\mathcal{L}_n/\mathcal{L}_n^* \to 1$ by the argument in Lemma 7.4. Similarly as in the proof of Lemma 5.2, we can have an estimate on $\mathcal{L}_n^*$ by calculating the number of half-edges that are non-free by time t_b: this includes (1) the sizes of the two branching processes, $Z_{t(n^\varrho)} + Z^{(b)}_{\lfloor (t(n^{\varrho'})/\lambda \rfloor} = O(n^{\varrho'/(\tau-2)})$, (2) the total number of half-edges in the last layer red occupies, i.e., the half-edges in $\widetilde{\Gamma}_{t_b}$, (3) vertices in different layers of blue, denoted by A_i-s, which has a total order at most the same as $M_n^{(b)} = O(\exp\{(\log n)^{2/\lambda+1}C_n\})$. Note that the worst case is given by (2), since the total number of half-edges with degree at least $\widetilde{u}_{t_b}$ can be calculated by Claim 5.3 and is of order $n\widetilde{u}_{t_b}^{2-\tau}$. As a result the factor $(\mathcal{L}_n/\mathcal{L}_n^*)^k$ is at most of order

$$(1 - \widetilde{u}_{t_b}^{2-\tau})^k \le \exp\{k\widetilde{u}_{t_b}^{2-\tau}\} \to 1,$$

since $\widetilde{u}_{t_b} = O(\exp\{(\log n)^{2/\lambda+1}C\})$ while $k = O(\log\log n)$. Let us define the event

$$\mathcal{E}_n' := \bigcap_{1 \le i \le k-1}\mathcal{G}_2(\widetilde{u}_{\lfloor t_b + \lambda i - T_r\rfloor}, \varepsilon)^C \cap \mathcal{G}_3(\widetilde{u}_{\lfloor t_b + \lambda i - T_r\rfloor}, \varepsilon)^C \bigcap \{\mathcal{L}_n \ge n\mathbb{E}[D]/(1+\varepsilon)\}. \quad \text{(A.15)}$$

Proposition A.1 yields that

$$\mathbb{P}(\mathcal{E}_n') \ge 1 - C\sum_{1 \le j \le k-1}(\widetilde{u}_{\lfloor t_b + \lambda i - T_r\rfloor})^{\tau-1}/n - \exp\{-nc(\varepsilon)\} \to 1,$$

since the terms $\widetilde{u}_{\lfloor t_b + \lambda i - T_r\rfloor}$ decrease double-exponentially and the dominating term is $O(\exp\{(\log n)^{2/\lambda+1}\})$. As a result, on $\mathcal{E}_n'$,

$$\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}, \mathcal{E}_n'] \le d_a \frac{d_b}{\mathcal{L}_n}\left(\prod_{i=1}^{k-1}(1+\varepsilon)^2\nu_i\right) \cdot \mathrm{e}^{\frac{k^2}{\mathcal{L}_n^*}} =: \overline{\mathbb{E}[N_k(a,b)|\mathcal{F}_{t_b}]} \quad \text{(A.16)}$$

where ν_i is from (7.1). By contracting all the vertices belonging to the set $\mathcal{M}_n^{(b)}$, we have $d_a = M_n^{(b)}$ and by letting b be the contraction of all the vertices with degree less than K for some arbitrary constant $K \ge 2$, we have $d_b/\mathcal{L}_n^* \le 1$ is of constant order again. Note that the total number of explored vertices on paths of length k is bounded from below and from above by

$$N_k(a,b) \le \mathrm{O}_{\max}(k) \le \sum_{i=1}^k N_i(a,b). \quad \text{(A.17)}$$

Noting that ν_i grows super-exponentially, $\sum_{i=1}^k \overline{\mathbb{E}_n[N_i(a,b)|\mathcal{F}_{t_b}]} = \overline{\mathbb{E}[N_k(a,b)|\mathcal{F}_{t_b}]}(1+o(1))$. This finishes the proof of the upper bound in (7.3).

We can get a lower bound on (A.13) if in the sum over distinct vertices, we leave out the i highest degree vertices $V_i := \{v_{i,1}^{\max}, \ldots, v_{i,i}^{\max}\}$ from each set Λ_i. That is, we have

$$\mathbb{E}_n[N_k(a, b)|\mathcal{F}_{t_b}] \geq d_a \frac{d_b}{\mathcal{L}_n} \prod_{i=1}^{k-1} \left(\sum_{\substack{\pi_i \in \Lambda_i \\ \pi_i \notin V_i}} \frac{d_{\pi_i}(d_{\pi_i} - 1)}{\mathcal{L}_n} \right) := \underline{n}_k(a, b). \tag{A.18}$$

Note that since we leave out only finitely many vertices, the ith sum within the product still converges to ν_i. More precisely, on $\mathcal{E}'_n \cap \{\mathcal{L}_n \in n(\mathbb{E}[D]/2, 2\mathbb{E}[D])\}$, each factor is at least $(1 - \varepsilon)\nu_i/2$.[4] Again contracting all the vertices belonging to the set $\mathcal{M}_n^{(b)}$, we have $d_a = M_n^{(b)}$ and by letting b be the contraction of all the vertices with degree less than K for some arbitrary constant $2 \leq K$, we have $d_b/\mathcal{L}_n$ is of constant order again. Combining with the lower bound in (A.17) finishes the proof of the lower bound in (7.3).

The proof of the bounds (7.4) for *good-directed* paths are analogous, but now one has to use the restricted sets

$$\Lambda_j^{\mathrm{d}} := \Lambda_j \setminus \Lambda_{j+1} = \left\{ v \in [n] : \widetilde{u}_{\lfloor t_b + \lambda(j+1) - T_r \rfloor} < d_v \leq \widetilde{u}_{\lfloor t_b + \lambda j - T_r \rfloor} \right\}.$$

Next we prove the variance formula for $O_{\max}^{\mathrm{d}}(k)$ following more or less the lines of [27, Section 10.4.2 and 9.4]. Note that the major difference between the proof of [27, Proposition 9.17] and our case is that here we have the extra restriction $\pi_i, \rho_i \in \Lambda_i^{\mathrm{d}}$, and the Λ_i^{d} sets are disjoint.

First write $N_k^{\mathrm{d}}(a, b)$ as the sum of indicators that a given good-directed path is present, and write $|\pi \cap \rho|$ for the number of edges the two paths share. Then we have the variance formula

$$\mathbf{Var}_n[N_k^{\mathrm{d}}(a, b)|\mathcal{F}_{t_b}] = \sum_{\ell=0}^{k} \sum_{\substack{\pi, \rho \\ |\pi \cap \rho| = \ell}} [\mathbb{P}_n(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b}) -$$

$$\mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})\mathbb{P}_n(\rho \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})].$$

Consider first the inner sum for $\ell = 0$, i.e. when the two path have disjoint edge-sets. Since at the time of pairing the ith half-edge, there are $\mathcal{L}_n^* - 2i + 1$ free half-edges to pick from, the probability that both π, ρ are present is exactly $\prod_{i=1}^{2k}(\mathcal{L}_n^* - 2i + 1)^{-1}$. On the other hand, the square of the probability that a path present is $\mathbb{P}(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d}))^2 = \prod_{i=1}^{k}(\mathcal{L}_n^* - 2i + 1)^{-2}$. Comparing the two, we get for $\ell = 0$, the inner sum is

$$\sum_{\substack{\pi, \rho \\ |\pi \cap \rho| = 0}} [\mathbb{P}_n(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b}) - \mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})\mathbb{P}_n(\rho \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})]$$

$$= \sum_{\substack{\pi, \rho \\ |\pi \cap \rho| = 0}} \left(\prod_{i=1}^{k} \frac{\mathcal{L}_n^* - 2i + 1}{\mathcal{L}_n^* - 2i - 2k + 1} - 1 \right) \mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})\mathbb{P}_n(\rho \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b}), \tag{A.19}$$

which, combined with the earlier observation about $\mathcal{L}_n^* = \mathcal{L}_n - O(n \exp\{(\log n)^{2/(\lambda+1)}(2 - \tau)\})$, gives rise to the first term in the error term $e_{k,n}$ in (7.6).

For $\ell = k$, that is, the two paths are identical, $\mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b}) - \mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})^2 \leq \mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b})$, hence the inner sum can be bounded by the inequality

$$\sum_{\substack{\pi, \rho \\ |\pi \cap \rho| = k}} \mathbb{P}_n(\pi \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b}) \leq \mathbb{E}_n[N_k^{\mathrm{d}}(a, b)|\mathcal{F}_{t_b}], \tag{A.20}$$

---

[4]Note that we have this weaker lower bound since we do not have upper large deviations for $\mathcal{L}_n$, i.e. we do not have a bound for $\mathbb{P}(\mathcal{L}_n \geq (1 + \varepsilon)\mathbb{E}[D]n)$. This is because D only has a finite first moment but infinite variance, and also the moment generating function does not exist for positive values.

explaining the first term on the right hand side of (7.5).

Now we are left with handling the cases $1 \le \ell \le k-1$. Note that in these cases we have to evaluate over all possible overlaps between the paths π, ρ. For this, note that the restriction that $\pi_i, \rho_i \in \Lambda_i^{\mathrm{d}}$ and Λ_i^{d} are disjoint sets implies that for each i there are only two cases: either $\pi_i = \rho_i$ or $\pi_i \ne \rho_i$, but in both cases they are disjoint from all $\pi_j, \rho_j, j \ne i$. We will merge these cases into shapes. Let us call an *excursion* of length s a connected component of $\pi \setminus \rho$, that is, a consecutive sequence of edges where the two paths are not the same. Formally, for some i, $(\pi_i, \pi_{i+1}) \ne (\rho_i, \rho_{i+1}), \ldots, (\pi_{i+s-1}, \pi_{i+s}) \ne (\rho_{i+s-1}, \rho_{i+s})$, is an excursion if it is started and ended by the common edges $(\pi_{i-1}, \pi_i) = (\rho_{i-1}, \rho_i)$ and $(\pi_{i+s}, \pi_{i+s+1}) = (\rho_{i+s}, \rho_{i+s+1})$ unless $i-1 = -1$ or $i+s+1 = k+1$, in which cases there is no edge before/after the excursion, respectively. Due to the property that Λ_i^{d} are disjoint, note that there are exactly the same number of edges on the π part of an excursion as on the ρ part of the excursion.

Let us denote by m the number of excursions, and again, we denote by $\ell := |\pi \cap \rho| = \ell$ the total number of shared edges. For a fix m, there can be $m-1, m$ or $m+1$ many segments of $\pi \cap \rho$, depending on whether none of, only one of, or both a, b are part of an excursion. Let us thus introduce the indicators $\delta_a = 1, \delta_b = 1$ if vertex a, b are parts of an excursion.

We can now define the class of shapes called $\mathrm{Shape}_{m,\ell}$ corresponding to pairs of paths for which $|\pi \cap \rho| = \ell$ and $\pi \setminus \rho$ consists of m excursions. That is, ρ has m edge-disjoint excursions from π, and between two consecutive excursions there is at least one edge in $\pi \cap \rho$. Note that the number of excursions m is thus at most $\ell + 1$. Also note that each shape in $\mathrm{Shape}_{m,\ell}$ can be uniquely characterised by a sequence of numbers of the form

$$(\delta_a, \delta_b, (\ell_1, \ell_2, \ldots, \ell_{m+1-\delta_a-\delta_b}), (e_1, e_2 \ldots, e_m)), \tag{A.21}$$

where δ_a, δ_b are indicators as before; ℓ_i is the number of edges in the ith connected component of $\pi \cap \rho$, with the property that $\sum_{i=1}^{m+1-\delta_a-\delta_b} \ell_i = \ell$; and e_i is the number of edges on the ith excursion with the property that $\sum_{i=1}^{m} e_i = k - \ell$. See Figure for examples.

The contribution of paths with $1 \le \ell \le k-1$ to the variance $\mathbf{Var}[N_k^{\mathrm{d}}(a,b)]$ can be bounded from above as follows:

$$\sum_{\substack{\pi, \rho \\ 1 \le |\pi \cap \rho| \le k-2}} \mathbb{P}_n(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d}) | \mathcal{F}_{t_b}) \le \sum_{\ell=1}^{k-1} \sum_{m=1}^{\ell+1} \sum_{\sigma \in \mathrm{Shape}_{m,\ell}} \sum_{\substack{\pi, \rho \\ \mathrm{Shape}(\pi,\rho)=\sigma}} \mathbb{P}_n(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d}) | \mathcal{F}_{t_b}).$$
$$\tag{A.22}$$

Note that if $|\pi \cap \rho| = \ell$ is fixed, then there are exactly $2k - \ell$ different edges in $\pi \cup \rho$, so that with fixed vertices and fixed half-edges,

$$\mathbb{P}_n(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d}) | \mathcal{F}_{t_b}) = \prod_{i=1}^{2k-\ell} \frac{1}{\mathcal{L}_n^* - 2i + 1}. \tag{A.23}$$

If we now fix only the vertices, but not the half-edges, then we have to multiply this with a combinatorial factor similar to that in (A.3) counting the number of possible variations of half-edges for fixed vertices $(\pi_i, \rho_i)_{1 \le i \le k}$. Recall again that $\delta_a = \mathbb{1}\{a \in \text{ first excursion of } \pi \setminus \rho\}$ and $\delta_b = \mathbb{1}\{b \in \text{ last excursion of } \pi \setminus \rho\}$. Let us write $d_\sigma(v)$ for the number of half-edges of v used in the union of paths π, ρ of shape σ, and in text we write degree_σ for this degree. At the end of every excursion we have degree_σ-3 vertices, while on the excursions and inside segments of $\pi \cap \rho$ we have degree_σ-2 vertices. Thus the combinatorial factor to pick half-edges, once fixing the vertices along the path (but
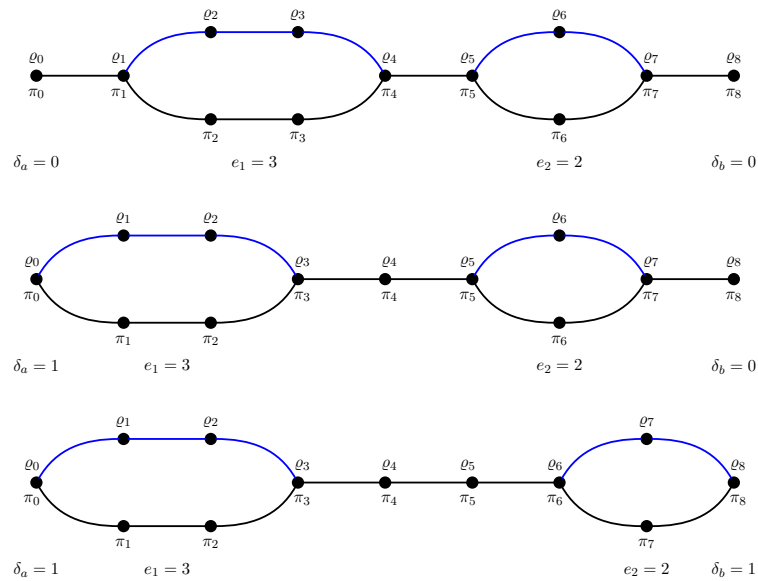
Figure 7: Paths of length 8 belonging to $\mathrm{Shape}_{2,3}$: $m = 2$ indicates that there are two excursions, $\ell = 3$ means that the two paths share 3 edges in total. On the first picture, the excursion do not start at the ends of the path, hence $\delta_a = \delta_b = 0$, on the second picture, $\delta_a = 1, \delta_b = 0$, while on the third picture both excursions start at the ends, hence $\delta_a = \delta_b = 1$. Note that in all cases, the number of degree three vertices is $2m - \delta_a - \delta_b$, and the shared edges form $m + 1 - \delta_a - \delta_b$ many connected components.

not the half-edges) is at most

$$d_a(d_a - 1)^{\delta_a} d_b(d_b - 1)^{\delta_b} \prod_{\substack{s: \\ d_\sigma(\pi_s)=3}} d_{\pi_{i_s}}(d_{\pi_{i_s}} - 1)(d_{\pi_{i_s}} - 2) \prod_{\substack{t: \\ d_\sigma(\pi_t)=2}} d_{v_t}(d_{v_t} - 1) \prod_{\substack{u: \\ d_\sigma(\rho_u)=2 \\ \rho_u \cap \pi = \varnothing}} d_{v_u}(d_{v_u} - 1). \quad \text{(A.24)}$$

Thus, establishing the contribution of a given fixed shape σ in (A.22), we should sum the product of (A.23) and (A.24) over all possible configurations $(\pi_i, \rho_i)_{i=1,\dots,k}$ forming the shape σ. Note that the number of factors in (A.23) equals the number of edges while the number of factors in (A.24) equals the number of vertices in σ. Since the number of edges minus the number of vertices without counting a and b is at least $m + 1$ (equality if all the vertices on excursions are different – see Remark A.2 below when this is not true), we obtain

$$\sum_{\substack{\pi, \rho \\ \mathrm{Shape}(\pi,\rho)=\sigma}} \mathbb{P}(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d}) | \mathcal{F}_{t_b}) \leq \frac{1}{\mathcal{L}_n^{m+1}} \prod_{i=1}^{2k-\ell} \frac{\mathcal{L}_n}{\mathcal{L}_n^* - 2i + 1} \cdot d_a(d_a - 1)^{\delta_a} d_b(d_b - 1)^{\delta_b}$$

$$\times \prod_{\substack{s: \\ d_\sigma(\pi_s)=3}} \left(\sum_{i \in \Lambda_{i_s}^{\mathrm{d}}}^* \frac{d_i(d_i - 1)(d_i - 2)}{\mathcal{L}_n} \right) \prod_{\substack{t: \\ d_\sigma(\pi_t)=2}} \left(\sum_{i \in \Lambda_{i_t}^{\mathrm{d}}}^* \frac{d_i(d_i - 1)}{\mathcal{L}_n} \right) \prod_{\substack{u: \\ d_\sigma(\rho_u)=2 \\ \rho_u \cap \pi = \varnothing}} \left(\sum_{i \in \Lambda_{i_u}^{\mathrm{d}}}^* \frac{d_i(d_i - 1)}{\mathcal{L}_n} \right).$$

$$\text{(A.25)}$$

Recall the event $\mathcal{E}_n'$ from (A.15). We put $\mathcal{G}_3$ in there exactly since now we can use it again: on $\mathcal{E}_n'$, the sums in the previous display can be bounded by $(1 + \varepsilon)^2 \kappa_{i_s}$ and

$(1+\varepsilon)^2 \nu_{i_t}, (1+\varepsilon)^2 \nu_{i_u}$, respectively. Thus, we get that on $\mathcal{E}'_n$ the rhs of (A.25) is at most

$$\frac{1}{\mathcal{L}_n^{m-1}} \prod_{i=1}^{2k-\ell} \frac{\mathcal{L}_n}{\mathcal{L}_n^* - 2i + 1} \cdot d_a^{1+\delta_a} \frac{d_b^{1+\delta_b}}{\mathcal{L}_n^2} \prod_{\substack{s: \\ d_\sigma(\pi_s)=3}} \kappa_{i_s} \prod_{\substack{t: \\ d_\sigma(\pi_t)=2}} \nu_{i_t} \prod_{\substack{u: \\ d_\sigma(\rho_u)=2 \\ \rho_u \cap \pi = \varnothing}} \nu_{i_u} (1+\varepsilon)^{2k}. \qquad \text{(A.26)}$$

Remark A.2. Note that even though the *edges* of an excursion are different by definition, the *vertices* in the excursions might still coincide, i.e. $\pi_i = \rho_i$ can happen even if π_i, ρ_i are sitting in the middle of an excursion. In this case, instead of having the factor $d_{\pi_i}(d_{\pi_i} - 1)d_{\rho_i}(d_{\rho_i} - 1)$ in (A.24) we have $d_{\pi_i}(d_{\pi_i} - 1)(d_{\pi_i} - 2)(d_{\pi_i} - 3)$. Since in (A.25), every original vertex gets a factor $\mathcal{L}_n^{-1}$ as a normalisation, when $\pi_i = \rho_i$ is merged within an excursion, there is a factor of the form

$$\sum_{\pi_i \in \Lambda_i^{\mathrm{d}}} d_{\pi_i}(d_{\pi_i} - 1)(d_{\pi_i} - 2)(d_{\pi_i} - 3)/\mathcal{L}_n^2$$

in (A.25). Similarly as in the proof of Proposition A.1, it can be shown that this expression is less than $C\widetilde{u}_{f(i)}^{5-\tau}/n$ where $f(i) = \lfloor t_b + \lambda i - T_r \rfloor$. If the vertices π_i, ρ_i are not merged, we have a factor $\nu_i^2 = \widetilde{u}_{f(i)}^{2(3-\tau)}$ in (A.25) instead. Note that

$$\widetilde{u}_{f(i)}^{5-\tau}/n = o(\widetilde{u}_{f(i)}^{2(3-\tau)})$$

as long as $\widetilde{u}_{f(i)} = o(n^{1/(\tau-1)})$. Since the maximal degree in the graph is of this order, this is always the case. That is, it in not worth merging vertices on excursions. We can continue analysing formula (A.26).

Now we identify the indices i_s, i_t, i_u, using the restrictions $\pi_i, \rho_i \in \Lambda_i^{\mathrm{d}}$. The crucial observation is the following: follow the indices $\pi_1, \pi_2, \ldots, \pi_{k-1}$ and $\rho_1, \rho_2, \ldots, \rho_{k-1}$ along the two paths. If for some i, the vertices $\pi_i \neq \rho_i$ are degree$_\sigma$-2 vertices on an excursion, then the corresponding ν_i^2 appears in the product in (A.26). If $\pi_i = \rho_i$ is a degree$_\sigma$-3 vertex, then we have a factor κ_i *replacing* ν_i^2 in the product. If $\pi_i = \rho_i$ is a degree$_\sigma$-2 vertex in $\pi \cap \rho$, then we only have a factor ν_i in the product (instead of ν_i^2) in (A.26).

Thus, dividing (A.26) by $\prod_{i=1}^{2k-2} \nu_i^2$ yields that for each degree$_\sigma$-3 vertex we have a factor κ_i/ν_i^2 and for each coinciding degree$_\sigma$-2 vertex we have a $1/\nu_i \leq 1/\nu_{k-1}$ in the product. Elementary calculation shows that

$$\kappa_i/\nu_i^2 = \left(\widetilde{u}_{\lfloor t_b + \lambda i - T_r \rfloor}\right)^{\tau-2} \qquad \text{(A.27)}$$

and since the sequence $\widetilde{u}_i$ is decreasing in i, we get

$$\frac{\kappa_i}{\nu_i^2} \leq \frac{\kappa_1}{\nu_1^2}.$$

Thus, we can bound the contribution of every degree$_\sigma$-3 vertex by a factor κ_1/ν_1^2, and every coinciding degree$_\sigma$-2 vertex by a factor $1/\nu_{k-1}$. Since there are $m - \delta_a - \delta_b$ many degree$_\sigma$-3 vertices and $\sum_{j=1}^{m+1-\delta_a-\delta_b} (\ell_i - 1) = \ell - m - 1 + \delta_a + \delta_b$ many coinciding degree$_\sigma$-2 vertices, if we take out $\overline{\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}]}^2$ as in (A.16) from (A.26), we are left with the following upper bound

$$\sum_{\substack{\pi, \rho \\ \mathrm{Shape}(\pi,\rho)=\sigma}} \mathbb{P}_n(\pi, \rho \subseteq \mathrm{CM}_n(\boldsymbol{d})|\mathcal{F}_{t_b}) \leq \overline{\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}]}^2 \frac{1}{d_a^{1-\delta_a} d_b^{1-\delta_b}} \left(\frac{\kappa_1}{\nu_1^2}\right)^{2m-\delta_a-\delta_b} \frac{1}{\nu_{k-1}^{\ell-m-1+\delta_a+\delta_b} \mathcal{L}_n^{m-1}}$$

$$\leq \overline{\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}]}^2 \cdot \frac{1}{\mathcal{L}_n^{m-1}} \left(\frac{\kappa_1^2 \nu_{k-1}}{\nu_1^4}\right)^{m-1} \frac{1}{\nu_{k-1}^\ell} \left(\frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_a}\right)^{1-\delta_a} \left(\frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_b}\right)^{1-\delta_b}.$$
$$\text{(A.28)}$$

Recall that we have to sum this formula over all the shapes in $\text{Shape}_{m,\ell}$, and over $m \geq 1$ and ℓ as in (A.22).

When $m = 1$, the number of $\sigma \in \text{Shape}_{1,\ell}$ is $\ell - 1$ if $\delta_a + \delta_b = 0$, and 1 if $\delta_a + \delta_b = 1$, and 0 if $\delta_a = \delta_b = 1$. Thus, fixing $m = 1$ and summing in δ_a, δ_b, ℓ gives at most

$$\left(\frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_a} + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_b} \right) \cdot \sum_{\ell=1}^{\infty} \frac{1}{\nu_{k-1}^\ell} + \frac{\kappa_1^2 \nu_{k-1}^2}{\nu_1^4 d_a d_b} \sum_{\ell=1}^{\infty} \frac{\ell - 1}{\nu_{k-1}^\ell}$$

Thus the contribution from $m = 1$ can be bounded from above by

$$\overline{\mathbb{E}_n[N_k(a,b)|\mathcal{F}_{t_b}]}^2 \left[\frac{\kappa_1 \nu_{k-1}}{\nu_1^2 (\nu_{k-1} - 1)} \left(\frac{1}{d_a} + \frac{1}{d_b} \right) + \frac{\kappa_1^2 \nu_{k-1}^2}{\nu_1^4 (\nu_{k-1} - 1)^2} \frac{1}{d_a d_b} \right]. \tag{A.29}$$

We are left with counting the contribution of shapes with $m \geq 2$. We can bound the number of shapes in $\text{Shape}_{m,l}$ (for fixed δ_a, δ_b) similarly as in [27, Lemma 9.18]

$$|\{\sigma : \sigma \in \text{Shape}_{m,l}\}| \leq \binom{k - \ell - 1}{m - 1} \binom{\ell - 1}{m - \delta_a - \delta_b}. \tag{A.30}$$

To explain this formula, recall the description of the shape in (A.21). Note that since excursions are separated by at least one common edge, we have to pick $\ell_1, \ell_2, \ell_{m+1-\delta_a-\delta_b}$ in (A.21) so that for all i, $\ell_i \geq 1$ and $\sum_i \ell_i = \ell$. This can be done in $\binom{\ell-1}{m-\delta_a-\delta_b}$ many ways. Then, we also have to pick the length of excursions so that $\sum_{i=1}^m e_i = k - \ell$ and each $e_i \geq 1$. This can be done in $\binom{k-\ell-1}{m-1}$ many ways.

For $m \geq 2$, the factors in (A.30) can be bounded by

$$\binom{k - \ell - 1}{m - 1} \leq \frac{k^{m-1}}{(m-1)!}, \qquad \binom{\ell - 1}{m - \delta_a - \delta_b} \leq \frac{\ell^{m - \delta_a - \delta_b}}{(m - \delta_a - \delta_b)!} \leq k^m,$$

since $\ell \leq k$. So, the total number of shapes in $\text{Shape}_{m,\ell}$ can be bounded by

$$|\{\sigma : \sigma \in \text{Shape}_{m,\ell}\}| \leq k \frac{k^{2(m-1)}}{(m-1)!}, \tag{A.31}$$

which is independent of ℓ. Using this bound in (A.28) when summing over the number of shapes, summing over $\delta_a, \delta_b \in \{0, 1\}$, then over $\ell \geq 1$, we get that the contribution of $m \geq 2$ terms can be bounded from above by $\overline{\mathbb{E}[N_k(a,b)]}^2$ times

$$\left(1 + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_a} \right) \left(1 + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_b} \right) \frac{k}{\nu_{k-1} - 1} \sum_{m=2}^{\infty} \left(\frac{\kappa_1^2 \nu_{k-1}}{\nu_1^4 \mathcal{L}_n} \right)^{m-1} \frac{k^{2(m-1)}}{(m-1)!}$$

$$= \left(1 + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_a} \right) \left(1 + \frac{\kappa_1 \nu_{k-1}}{\nu_1^2 d_b} \right) \frac{k}{\nu_{k-1} - 1} \left(e^{k^2 \kappa_1^2 \nu_{k-1}/(\nu_1^4 \mathcal{L}_n)} - 1 \right). \tag{A.32}$$

Now let us set $d_a := M_n^{(b)}$ and in d_b we collect all the vertices that are less then ν_k: these contain vertices with constant degree (say all the degrees smaller than K for $F(K) = 1/2$). This implies that $d_b \geq \mathcal{L}_n/2$ whp. Combining the contribution for $\ell = 0$ in (A.19), $\ell = k$ in (A.20), and then $m = 1$ in (A.29) and finally $m \geq 2$ in (A.32) yields (7.5). $\qquad \square$

References

[1] Yong-Yeol Ahn, Hawoong Jeong, Naoki Masuda, and Jae Dong Noh, *Epidemic dynamics of two species of interacting particles on scale-free networks*, Physical Review E **74** (2006), no. 6, 066113.

[2] Tonći Antunović, Yael Dekel, Elchanan Mossell, and Yuval Peres, *Competing first passage percolation on random regular graphs*, arXiv:1109.2575 [math.PR], 2011.

[3] Tonći Antunović, Elchanan Mossel, and Miklós Z. Rácz, *Coexistence in preferential attachment networks.*, arXiv: 1307.2893 [math.PR], 2013.

[4] Johan Arndt, *Role of product-related conversations in the diffusion of a new product*, Journal of marketing Research (1967), 291–295.

[5] Alex Beutel, B Aditya Prakash, Roni Rosenfeld, and Christos Faloutsos, *Interacting viruses in networks: can both survive?*, Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2012, pp. 426–434.

[6] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra, *First passage percolation on random graphs with finite mean degrees*, Ann. Appl. Probab. **20(5)** (2010), 1907–1965.

[7] S. Bhamidi, Remco van der Hofstad, and G. Hooghiemstra, *First passage percolation on the Erdős-Rényi random graph*, Combinatorics, Probability and Computing **20** (2011), 683–707.

[8] Shankar Bhamidi, Remco van der Hofstad, and Gerard Hooghiemstra, *Universality for first passage percolation on sparse random graphs*, arXiv:1210.6839, 2014.

[9] B. Bollobás, *Random graphs*, Cambridge University Press, 2001.

[10] Béla Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European Journal of Combinatorics **1** (1980), no. 4, 311 – 316.

[11] Francis A Buttle, *Word of mouth: understanding and managing referral marketing*, Journal of strategic marketing **6** (1998), no. 3, 241–254.

[12] Judith A Chevalier and Dina Mayzlin, *The effect of word of mouth on sales: Online book reviews*, Journal of marketing research **43** (2006), no. 3, 345–354.

[13] P. L. Davies, *The simple branching process: a note on convergence when the mean is infinite*, J. Appl. Probab. **15** (1978), no. 3, 466–480. MR-0483050

[14] Maria Deijfen and Olle Häggström, *The initial configuration is irrelevant for the possibility of mutual unbounded growth in the two-type Richardson model*, Combin. Probab. Comput. **15** (2006), no. 3, 345–353. MR-2216472

[15] _____, *The two-type Richardson model with unbounded initial configurations*, Ann. Appl. Probab. **17** (2007), no. 5-6, 1639–1656. MR-2358637

[16] Maria Deijfen and Remco van der Hofstad, *The winner takes it all*, arXiv:1306.6467 [math.PR], 2013.

[17] Chrysanthos Dellarocas, *The digitization of word of mouth: Promise and challenges of online feedback mechanisms*, Manage. Sci. **49** (2003), no. 10, 1407–1424.

[18] Ernest Dichter, {*How word-of-mouth advertising works*}, Harvard business review **44** (1966), no. 6, 147–160.

[19] Drew Fudenberg and Abhijit Banerjee, *Word of mouth learning*, Games and Economic Behavior **46** (2004), 1–22.

[20] Drew Fudenberg and Glenn Ellison, *Word of mouth communication and social learning*, Quarterly Journal of Economics **110** (1995), 93–126.

[21] Olivier Garet and Régine Marchand, *Coexistence in two-type first-passage percolation models*, Ann. Appl. Probab. **15** (2005), no. 1A, 298–330. MR-2115045

[22] _____, *First-passage competition with different speeds: positive density for both species is impossible*, Electron. J. Probab. **13** (2008), no. 70, 2118–2159. MR-2461538

[23] Torben Hagerup and Christine Rüb, *A guided tour of Chernoff bounds*, Information Processing Letters **33** (1990), no. 6, 305 – 308.

[24] Olle Häggström and Robin Pemantle, *First passage percolation and a model for competing spatial growth*, J. Appl. Probab. **35** (1998), no. 3, 683–692. MR-1659548

[25] Christopher Hoffman, *Coexistence for Richardson type competing spatial growth models*, Ann. Appl. Probab. **15** (2005), no. 1B, 739–747. MR-2114988

[26] _____, *Geodesics in first passage percolation*, Ann. Appl. Probab. **18** (2008), no. 5, 1944–1969. MR-2462555

[27] Remco van der Hofstad, *Random graphs and complex networks*, Springer, 2010, book in preparation.

[28] Remco van der Hofstad, Gerard Hooghiemstra, and Dmitri Znamenski, *Distances in random graphs with finite mean and infinite variance degrees*, Electron. J. Probab. **12** (2007), no. 25, 703–766. MR-2318408

[29] Remco van der Hoftad and Júlia Komjáthy, *Fixed speed competition on the configuration model with infinite variance degrees: equal speeds*, arXiv:1503.09046 [math.PR], 2015.

[30] Svante Janson, *Susceptibility of random graphs with given vertex degrees*, J. Comb. **1** (2010), no. 3-4, 357–387. MR-2799217

[31] Brian Karrer and MEJ Newman, *Competing epidemics on complex networks*, Physical Review E **84** (2011), no. 3, 036106.

[32] Viswanathan Kumar, J Andrew Petersen, and Robert P Leone, *How valuable is word of mouth?*, Harvard Business Review **85** (2007), no. 10, 139.

[33] Jure Leskovec, Lada A Adamic, and Bernardo A Huberman, *The dynamics of viral marketing*, ACM Transactions on the Web (TWEB) **1** (2007), no. 1, 5.

[34] Marc Lipsitch, Caroline Colijn, Ted Cohen, William P Hanage, and Christophe Fraser, *No coexistence for free: neutral null models for multistrain pathogens*, Epidemics **1** (2009), no. 1, 2–13.

[35] Michael Molloy and Bruce Reed, *A critical point for random graphs with a given degree sequence*, Random Structures and Algorithms **6** (1995), no. 2-3, 161–180.

[36] _____, *The size of the giant component of a random graph with a given degree sequence*, Comb. Probab. Comput. **7** (1998), no. 3, 295–305.

[37] Mark EJ Newman, *Threshold effects for two pathogens spreading on a network*, Physical review letters **95** (2005), no. 10, 108701.

[38] B Aditya Prakash, Alex Beutel, Roni Rosenfeld, and Christos Faloutsos, *Winner takes all: competing viruses or ideas on fair-play networks*, Proceedings of the 21st international conference on World Wide Web, ACM, 2012, pp. 1037–1046.