Increasing paths in random temporal graphs

Nicolas Broutin [*†] Nina Kamčev [‡] Gábor Lugosi [§¶‖**]

June 20, 2023

Abstract

We consider random temporal graphs, a version of the classical Erdős–Rényi random graph $G(n, p)$ where additionally, each edge has a distinct random time stamp, and connectivity is constrained to sequences of edges with increasing time stamps. We study the asymptotics for the distances in such graphs, mostly in the regime of interest where np is of order $\log n$. We establish the first order asymptotics for the lengths of increasing paths: the lengths of the shortest and longest paths between typical vertices, the maxima of these lengths from a given vertex, as well as the maxima between any two vertices; this covers the (temporal) diameter.

1 Introduction

1.1 Motivation and model

A *temporal graph* $G = (V, E, \pi)$ is a finite simple graph together with an ordering $\pi : E \to \{1, 2, \ldots, |E|\}$ of the edges. The ordering may be interpreted as time stamps on the edges. An edge e precedes another edge f if $\pi(e) < \pi(f)$. Temporal graphs naturally model time-dependent propagation processes, for instance social interactions or infection processes.

The specific model we study is that of *random simple temporal graphs*, which was explicitly addressed by Casteigts, Raskin, Renken, and Zamaraev [7], as well as Becker, Casteigts, Crescenzi, Kodric, Renken, Raskin, and Zamaraev [4]. However, the same model had already been considered by Angel, Ferber, Sudakov, and Tassion [2]. A number of authors have studied related random temporal graph models – we refer the reader to [7,

[*]LPSM, Sorbonne Université, 4 Place Jussieu, 75005 Paris

[†]Institut Universitaire de France (IUF)

[‡]Department of Mathematics, Faculty of Science, University of Zagreb, Croatia. Research supported by the European Union's Horizon 2020 research and innovation programme [MSCA GA No 101038085].

[§]Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain

[¶]ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain

[‖]Barcelona Graduate School of Economics

[**]Research supported by Ayudas Fundación BBVA a Proyectos de Investigación Científica 2021 and the Spanish Ministry of Economy and Competitiveness, Grant PGC2018-101643-B-I00 and FEDER, EU

Section 6] for a thorough literature review. Specifically, reachability in similar models which allow repeated edge labels was studied since the 1970s by Boyd and Steele [5], Frieze and Grimmett [9], Haigh [10], Mocquard et al. [15, 16], Moon [17], van Ditmarsch et al. [19]. We return to these related models in Section 1.3, since our methods extend to them in a straightforward way.

In the *random simple temporal graph model*, G is an Erdős–Rényi random graph on n vertices and edge probability $p \in [0,1]$ and the time stamps are generated by a uniform random permutation on the edges. Such a random graph may be conveniently generated by assigning independent, identically distributed random *labels* to each edge of the complete graph on n vertices. In most of this article, the label assigned to an unoriented edge $\{i,j\}$ is an exponential random variable $W_{i,j}$ (of parameter 1), where $1 \le i < j \le n$. An edge is kept if and only if $W_{i,j} \le \tau$ where $\tau = -\log(1-p)$. Note that $\mathbb{P}\{W_{i,j} \le \tau\} = p$ and therefore the graph obtained by keeping only edges with $W_{i,j} \le \tau$ indeed is an Erdős-Rényi random graph[1] $G \sim \mathcal{G}(n,p)$. Hence, $V = [n]$ and $E = \{\{i,j\} : W_{i,j} \le \tau\}$. An edge $\{i,j\}$ *precedes* the edge $\{i',j'\}$ if an only if $W_{i,j} \le W_{i',j'}$. We denote such a random graph generated by the labels $W = (W_{i,j})$ by $G_p(W)$.

Our main objects of interest in temporal graphs are *increasing* paths. A path $(v_1,\dots,v_k)$ is *increasing* if the edge labels of the path are increasing, that is, $W_{v_\ell,v_{\ell+1}} \le W_{v_{\ell+1},v_{\ell+2}}$ for all $\ell \in \{1,\dots,k-2\}$. The *length* of a path $(v_1,\dots,v_k)$ is $k-1$. We call v_1 and v_k the *starting point and endpoint* of the path, respectively.

We study longest and shortest increasing paths. For a pair of vertices $i,j \in [n]$, we may define $\ell(i,j)$ and $L(i,j)$ as the minimum and maximum length of any increasing path from i to j. We refer to $\ell(i,j)$ as the *distance* from i to j when the 'direction' is clear from the context. In most of the paper we focus on the regime $np = c\log n$ for a constant c. The main reasons for singling out this range are that it is where the phase transition occurs, and where the path lengths exhibit an intriguing behaviour. Indeed, as shown below, for smaller values of p, there are no increasing paths between two typical vertices. When p is much larger, the longest path in the entire graph has roughly the same length as the longest path between two typical vertices. However, an interesting phenomenon in the regime $np \sim c\log n$ is that the longest path in the entire graph is significantly longer than the longest path from a typical vertex. This poses some nontrivial challenges as it is difficult to use constructive methods for finding the longest path (such as branching processes or coupling arguments.)

The phase transition occurring around $p = c\log n/n$ was first studied by Casteigts et al. [7].[2] Specifically, they studied the following properties: a typical pair of vertices is *connected* (by an increasing path), a typical vertex can reach all other vertices, and any pair of vertices is *connected*. They found that the threshold probabilities for these three properties are $\log n/n$, $2\log n/n$ and $3\log n/n$ respectively. In Theorem 3, we strengthen their results by finding paths of (asymptotically) minimal length throughout this probability range, addressing [6, Conjecture 5] in a strong form. The 'longest shortest' increasing path,

[1] Note that $-\log(1-p) = p + O(p^2)$, and none of our results are sensitive to such lower-order changes of p.

[2] We usually write $\log n/n$ for $(\log n)/n$; we believe that it should not cause any confusion.

denoted $\max_{i,j\in[n]} \ell(i,j)$, can be interpreted as the *diameter* of our temporal graph; such shortest paths also differ significantly between typical vertex pairs and 'most distant' vertex pairs, in contrast to other models such as the Erdős-Rényi random graph.

The authors of [7] also noticed an intriguing connection between their phase-transition results and Janson's celebrated results on percolation in weighted graphs [11]. We unveil an explicit connection between G_p and a random recursive tree, which is also related to minimum-weight paths in Janson's model [1]. Thus, our methods shed some light on the parallels between paths in these two random-graph models.

1.2 Main results

In this section we present the main findings of this paper. The subjects of our study are the following random variables.

(a) $L(1,2)$ and $\ell(1,2)$, that is, the lengths of the longest and shortest increasing paths between two fixed vertices (say vertex 1 and 2);

(b) $\max_{j\in\{2,\ldots,n\}} L(1,j)$ and $\max_{j\in\{2,\ldots,n\}} \ell(1,j)$, the maximum length of the longest and shortest increasing paths starting at a fixed vertex;

(c) $\max_{i,j\in[n]} L(i,j)$ and $\max_{i,j\in[n]} \ell(i,j)$, that is, the maximal length of the longest and shortest increasing paths in G.

Angel, Ferber, Sudakov, and Tassion [2] showed that if $p = o(n)$ such that $pn/\log n \to \infty$, then $\max_{i,j\in[n]} L(i,j) \sim_p enp$, answering a question of Lavrov and Loh [14]. Naturally, they asked about the corresponding estimate for $np = \Theta(\log n)$. When $pn/\log n \to \infty$, the upper bound $\max_{i,j\in[n]} L(i,j) \le (1+o(1))enp$ follows from a first-moment argument (see Section 2.1), and the same computation suggests that in this range, the length of the longest path between a *typical* pair of vertices (denoted $L(1,2)$) is also asymptotically epn. We confirm this intuition with a quick observation. More importantly, the situation changes when $p = c\log n/n$, and gaps appear between the typical behavior of $L(1,2)$, $\max_{j\in[n]} L(1,j)$ and $\max_{i<j\in[n]} L(i,j)$. Consequently, we use three different techniques for determining the three quantities, as detailed in Section 1.4. Note that for $c < 1$, with high probability, there is no increasing path between vertex 1 and vertex 2 (which was also shown by [7]), while we still have $\max_{j\in\{2,\ldots,n\}} L(1,j) \sim_p ec\log n$.

In order to formulate the main results, we need to introduce a few constants, depending on the parameter c. For any $c > 0$, define

$$\alpha(c) := \inf\{x > 0 : x\log(x/ec) = 1\} \tag{1.1}$$

and, for $c > 1$,

$$\beta(c) := \sup\{x > 0 : x\log(x/ec) = -1\}, \tag{1.2}$$
$$\gamma(c) := \inf\{x > 0 : x\log(x/ec) = -1\}\,. \tag{1.3}$$

Since $x \log(x/ec)$ is a strictly convex function of x, the equations $x \log(x/ec) = 1$ and $x \log(x/ec) = -1$ have at most two solutions for any $c > 0$. The real numbers $\beta(c)$ and $\gamma(c)$ are the two solutions of the latter equation. When $c = 1$, there is only one solution and $\beta(1) = \gamma(1) = 1$. For $c < 1$ there is no solution, while for $c > 1$ there are two distinct solutions. On the other hand, the equation $x \log(x/ec) = 1$ has a unique solution for all $c > 0$ which we denote by $\alpha(c)$. Observe also that as $c \to \infty$, $\alpha(c)/c \to e$ and $\beta(c)/c \to e$. An interesting example value is $\alpha(1) \approx 3.5911$.

The following result characterises the first-order asymptotics for the lengths of longest increasing paths in the regime $np = c \log n$. For a sequence of events (E_n) we say that E_n holds with high probability if $\mathbb{P}\{E_n\} \to 1$. We write $X_n \sim_p a_n$ to denote that $X_n/a_n \to 1$ in probability.

Theorem 1 (Longest paths)**.** *Suppose that $p = c \log n/n$ for some $c > 0$. Then*

(i) if $c \in (0, 1)$, there is no increasing path between 1 and 2 with high probability, and if $c \geq 1$, $L(1, 2) \sim_p \beta(c) \log n$;
(ii) for all $c > 0$, $\max_{j\in\{2,...,n\}} L(1, j) \sim_p ec \log n$.
(iii) for all $c > 0$, $\max_{i,j\in[n]} L(i, j) \sim_p \alpha(c) \log n$.

We also observe that the result of Angel et al. [2] can be strengthened. Namely, when $np/ \log n \to \infty$, even the longest path between two fixed vertices has length about enp. Note that this is consistent with the fact that $\beta(c) \to e$ as $c \to \infty$.

Proposition 2 (Longest paths for $np \gg \log n$)**.** *Assume that $p = o(1)$, and $pn/ \log n \to \infty$, then $L(1, 2) \sim_p enp$.*

We also find the asymptotics for the lengths of shortest paths. In particular, this confirms Conjecture 5 from [6] (which is an extended version of [7]).

Theorem 3 (Shortest paths)**.** *For $c \geq 1$, let $\gamma(c)$ be the smallest $x > 0$ such that $x \log(x/ec) + 1 = 0$, as defined in (1.3). Let $p = c \log n/n$. Then*

(i) for all $c > 1$, $\ell(1, 2) \sim_p \gamma(c) \log n$;
(ii) for all $c > 2$, $\max_{i\in[n]} \ell(1, i) \sim_p \gamma(c - 1) \log n$;
(iii) for all $c > 3$, $\max_{i,j\in[n]} \ell(i, j) \sim_p \gamma(c - 2) \log n$.

Specifically, taking $c = 1 + \varepsilon$, $c = 2 + \varepsilon$ or $c = 3 + \varepsilon$, we are able to reprove the *phase transition* established in Becker et al. [4], but also attaining the minimal possible distance (asymptotic to $\log n$) between the considered vertices.

For $c < 1$, a typical pair of vertices is not connected by any increasing path. Still, we can enquire about the set of increasing paths starting from a typical vertex. This is crucial for proving Theorem 1 (ii), but it is also of independent interest as we utilise a connection between the increasing paths from vertex 1 and a random recursive tree. Specifically, we will cite the typical height of a random recursive tree, but many other properties can be transferred to random temporal graphs.

Let $B_\ell(v)$ be the set of vertices from vertex v by increasing paths in G consisting of at most ℓ edges (including the vertex v), and note that $B_n(v)$ is the set of all vertices reachable from v. Occasionally we abbreviate $B_\ell = B_\ell(1)$.

Theorem 4 (Reachability from vertex 1). *Let $p = c \log n/n$ with $c \leq 1$ and $\varepsilon > 0$.*

(i) With probability $1 - n^{-\Omega(\varepsilon^2)}$, the random temporal graph G_p contains an increasing path from 1 of length at least $ec \log n(1 - \varepsilon)$.

(ii) With high probability, $|B_n(1)| \geq e^{pn(1-\varepsilon)}$, and for most vertices in $v \in B_n(1)$, we have $\ell(1, v) \leq pn(1 + \varepsilon)$.

We remark that (ii) can also be used to derive an elementary proof of Theorem 3 i for $c = 1 + \varepsilon$. (This is due to the fact that typical longest and shortest paths from 1 to v both have length asymptotic to $\log n$ in this regime, so we do not have to *look for* atypical paths.) It also yields an alternative proof for the phase transition studied in [7], which we will comment on in Section 4.

First-moment arguments also show that the exponent in (ii) is optimal, since the number of vertices reachable from 1 is at most $e^{pn(1+\varepsilon/2)}$.

1.3 An alternative model

As mentioned above, related models with repeated edges have been studied since the 1970s. For instance, the following model, sometimes referred to as a *random gossipping protocol*, is considered in [5, 10, 15–17, 19]. For a given m, let $(e_1, \dots, e_m)$ be a sequence of independent uniformly random edges of the complete graph (so repetitions are allowed), and let H_m be a random temporal graph with ordered edges $e_1, \dots, e_m$. Increasing paths in H_m are defined as before. It is not difficult to see that all our results translate to the model H_m, with $m \sim pn^2/2$.

Indeed, from the graph H_m, one can construct a simple temporal graph $\widetilde{H}_m$ by discarding the repeated edges, and with high probability, the simple graph $\widetilde{H}_m$ contains the random simple temporal graph $G_p(W)$ with $p = 2m/n^2(1 - o(1))$ (with a natural definition of containment). Hence, any increasing path in $G_p(W)$ is also contained in $\widetilde{H}_m$ and hence in H_m. Conversely, all of our first-moment arguments for non-existence of increasing paths can be adapted to the model H_m.

1.4 Techniques and plan of the paper

In Section 2, we present the first moment bounds for the longest and shortest paths. These prove the upper bound on the longest paths in Theorem 1 (i)–(iii) as well as the lower bound on the shortest typical path in Theorem 3, part (i).

In order to prove Theorem 1 (ii) for $c < 1$, we show by coupling that we may embed a *random recursive tree* of size almost n in the graph. A random recursive tree is by now a well-understood object, obtained by repeatedly attaching a leaf to a random vertex of the existing tree. The height of such a tree of size m is known to be asymptotic to $e \log m$ [8, 18]. This construcive method is also useful for Theorem 4. For $c \geq 1$, we partition the *probability range* into a bounded number of intervals of length at most $\log n/n$, and show that we may *compose* the paths found in these intervals. Note that $\max_j L(1, j) \sim_p epn$ is linear in p, so the *path composition* strategy indeed yields optimal results. These results are proved in Section 4.

Section 3 is devoted to the lower bounds for the maximum length of an increasing path (i.e., part (iii) of Theorem 1). The intuition for $\max_{i,j} L(i,j) \sim_p \alpha(c)\log n$ can also be derived from the random-recursive-tree argument, at least for $c \leq 1$. Namely, $\alpha(c)\log n$ is the maximal height over n independently sampled random recursive trees, which in our graph correspond to exploration processes initiated from the n vertices. Certainly, this heuristic cannot be used to prove that $\max_{i,j} L(i,j) \sim_p \alpha(c)\log n$, since the exploration processes from the different vertices are correlated.

Instead, we resort to the second moment method. However, the second moment of the *obvious* random variable – the number of paths of length roughly $\alpha \log n$ – is simply too large. Therefore, we need to restrict the set of paths we are counting in a way that deterministically avoids certain undesired intersecting pairs of paths, and still has a sufficiently large expectation. This utilises some of the ideas and tools from Addario-Berry et al. [1].

The proof of the upper bound on $\ell(1,2)$, that is, the length of the shortest path between two typical vertices is found in Section 5. It is based on a branching process argument. Roughly speaking, we use a branching process to find paths from vertex 1 which are (atypically) short. We show that such short paths can reach almost all other vertices. The same argument also proves that there exists an increasing path between 1 and 2 whose length is close to $\beta(c)\log n$, thereby proving the lower bound of Theorem 1 (i).

Finally, let us give some intuition regarding the asymptotics for $\max\{\ell(1,i) : i \in [n]\}$ and $\max\{\ell(i,j) : i,j \in [n]\}$. For $\max\{\ell(1,i) : i \in [n]\}$, note that typically, $G_p(W)$ contains a vertex i with no incident edges whose labels lie in the interval $(0,(1-\varepsilon)\log n/n)$. Thus all increasing paths starting at i will use only edges with labels in an interval of length roughly $(c-1)\log n/n$, which suggests that the shortest path from 1 should be of length $\gamma(c-1)\log n$ (recalling the definition of γ in (1.3)). This lower bound on $\max\{\ell(1,i) : i \in [n]\}$ is proved in Section 5. This intuition turns out to give the correct estimate, and the upper bound is proved using the fact that any vertex *attaches* (in a single step) to a *short* increasing path with edge labels in a restricted probability range $(\log n/n, c\log n/n)$. Similar arguments yield the asymptotics for $\max\{\ell(i,j) : i,j \in [n]\}$.

2 First moment bounds: longest and shortest paths

In this section, we present the first moments bounds for the lengths of longest and shortest paths. They are based on elementary counting arguments.

2.1 First moment bound: longest paths

When $np/\log n \to \infty$, the upper bound for $\max\{L(i,j) : 1 \leq i,j \leq n\}$ is established in [2], which also implies the upper bounds for $\max\{L(1,i) : 1 \leq i \leq n\}$ and $L(1,2)$ as well. Thus we focus on the range where $np \sim c\log n$ for some $c > 0$ and prove the upper bounds in Theorem 1.

Proof of the upper bounds for Theorem 1. For $k \in [n-1]$, let X_k be the number of increasing paths of length k in G. Similarly, let Y_k denote the number of increasing paths of length k starting at vertex 1, and let Z_k be the number of increasing paths of length k from vertex 1 to vertex 2. Clearly,

$$\begin{aligned}
\mathbb{E}X_k &= \binom{n}{k+1}(k+1)!\frac{p^k}{k!} \ , \\
\mathbb{E}Y_k &= \binom{n-1}{k}k!\frac{p^k}{k!} \ , \\
\mathbb{E}Z_k &= \binom{n-2}{k-1}(k-1)!\frac{p^k}{k!} \ .
\end{aligned} \tag{2.1}$$

To prove part (iii), observe that, if there is an increasing path of length ℓ, then there is also an increasing path of length s, for any $1 \le s \le \ell$. As a consequence, by Markov's inequality,

$$\mathbb{P}\Big(\max_{i,j\in[n]} L(i,j) \ge \ell\Big) = \mathbb{P}(X_\ell \ge 1) \le \mathbb{E}[X_\ell] \ . \tag{2.2}$$

Therefore, in order to establish an upper bound for $\max\{L(i,j) : i,j \in [n]\}$, it suffices to obtain an upper bound on $\mathbb{E}[X_k]$. By Stirling's formula, for any natural number $k \ge 0$,

$$\mathbb{E}[X_k] \le n\frac{(np)^k}{k!} \le n\left(\frac{nep}{k}\right)^k \ . \tag{2.3}$$

Setting $\ell = \lceil a \log n \rceil$ for some $a > \alpha(c)$, (2.3) implies that, for any $\epsilon > 0$ and all n large enough,

$$\mathbb{E}[X_\ell] \le n\left(\frac{ec(1+\epsilon)}{a}\right)^\ell \le n^{1-a\log(a/(ec(1+\epsilon)))} \to 0 \ ,$$

provided $\epsilon > 0$ is small enough: indeed, by definition of $\alpha(c)$ we have $1 - a\log(a/(ec) < 0$, and by continuity the same holds for the exponent in the right-hand side above for all $\epsilon > 0$ small enough. (We remark that $a > ec$.)

For part (ii), focusing now on $\max\{L(1,i) : i \in [n]\}$, we have a relation similar to the one in (2.2):

$$\mathbb{P}\Big(\max_{i\in[n]} L(1,i) \ge \ell\Big) \le \mathbb{P}(Y_\ell \ge 1) \le \mathbb{E}[Y_\ell] \, .$$

However, for any $p \in [0,1]$ and $\ell = \lceil (1+\epsilon)enp \rceil$,

$$\mathbb{E}[Y_\ell] \le \left(\frac{nep}{\ell}\right)^\ell \le (1+\epsilon)^{-\ell} \le e^{-\epsilon enp} \, ,$$

which tends to zero as soon as $np \to \infty$.

To prove part (i), we finally consider $L(1,2)$. Suppose first that $c < 1$. Let k be an arbitrary positive integer, and define $\beta_n = k/\log n$. Then

$$\mathbb{E}Z_k \le \frac{1}{n}\left(\frac{nep}{k}\right)^k = \frac{1}{n}\left(\frac{ce}{\beta_n}\right)^{\beta_n \log n} = n^{-\beta_n \log(\beta_n/(ec))-1} \le n^{c-1} \ , \tag{2.4}$$

since $\inf\{x\log(x/(ec)) - 1 : x > 0\} = c - 1$. Hence, for $c < 1$, we have

$$\begin{aligned}\mathbb{P}\{L(1,2) \geq 1\} &\leq \mathbb{P}(Z_k > 0 \text{ for some } k \geq 1) \\ &\leq \mathbb{P}(Z_k > 0 \text{ for some } k \in \{1, \ldots, 3c\log n\}) \\ &\quad + \mathbb{P}(Y_k > 0 \text{ for some } k \geq 3c\log n)\ .\end{aligned}$$

We have already shown that the second term on the right-hand side converges to zero (since $e < 3$), and by (2.4), the first term is at most $(3c\log n)n^{c-1} \to 0$. Hence, we conclude that $\mathbb{P}\{L(1,2) = 0\} \to 1$ whenever $p \leq c\log n$ for some $c < 1$.

Suppose now that $c \geq 1$, and set $\ell = \lceil (1+\epsilon)\beta(c)\log n \rceil$. By an analogous calculation we then have $\sum_{k\geq\ell} \mathbb{E}[Z_k] = \exp(-\Omega(\varepsilon np))$, which completes the proof of the upper bounds in Theorem 1. ■

2.2 First moment bound: shortest path

We will now prove a lower bound on $\ell(1,2)$. The remaining statements of Theorem 3 are proved in Section 5.

Proof of the lower bound for Theorem 3 i. that is, we will consider the length of the *shortest* increasing path between two typical vertices, denoted $\ell(1,2)$. We have already seen that when $p = c\log n/n$ for some $c < 1$, with high probability, there are no increasing paths starting at vertex 1 and ending at vertex 2 (Theorem 1 (i)).

Recall that Z_k denotes the number of increasing paths between vertices 1 and 2. For any given $\ell \in \mathbb{N}$,

$$\mathbb{P}(\ell(1,2) \leq \ell) = \mathbb{P}\left(\sum_{k=1}^{\ell} Z_k \geq 1\right) \leq \sum_{k=1}^{\ell} \mathbb{E}Z_k\,.$$

It follows from the expression in (2.1) and Stirling's formula, that

$$\mathbb{P}(\ell(1,2) \leq \ell) \leq \sum_{k=1}^{\ell} \frac{1}{n}\left(\frac{nep}{k}\right)^k = \frac{1}{n}\sum_{k=1}^{\ell}\left(\frac{ce\log n}{k}\right)^k\ .$$

Let $\epsilon > 0$ be arbitrary. If $\ell \leq (1-\epsilon)\gamma(c)\log n$, then each term of the sum in the right-hand side above may be bounded by

$$\left(\frac{ce\log n}{k}\right)^k \leq \left(\frac{ce}{(1-\epsilon)\gamma(c)}\right)^{(1-\epsilon)\gamma(c)\log n} = n^{(1-\epsilon)\gamma(c)\log\frac{ce}{(1-\epsilon)\gamma(c)}}\ .$$

(To see this, note that the left-hand side is an increasing function of k for $k \leq \gamma(c)\log n$.) Since the function $x\log(ec/x)$ is strictly convex and increasing for $x \leq \gamma(c)$, there exists a positive number $\phi(\epsilon)$ such that $(1-\epsilon)\gamma(c)\log\frac{ce}{(1-\epsilon)\gamma(c)} \leq \gamma(c)\log\frac{ce}{\gamma(c)} - \phi(\epsilon) = 1 - \phi(\epsilon)$. Therefore, we have that

$$\mathbb{P}(\ell(1,2) \leq (1-\epsilon)\gamma(c)\log n) \leq \gamma(c)n^{-\phi(\epsilon)}\log n \to 0\ ,$$

proving that, for every $\epsilon > 0$, with high probability, $\ell(1,2) \geq (1-\epsilon)\gamma(c)\log n$. ■

3 Longest increasing path: the lower bound

The calculation from the previous section shows that the expected number of increasing paths of length $\alpha(c)\log n(1-o(1))$ tends to infinity, and a natural strategy for proving that such paths are indeed likely to exist is the second moment method. However, it turns out that for a fixed path P, the expected number of increasing paths Q intersecting P in, say, one short segment, is too large, so we have to exclude such paths deterministically, which is the aim of the following subsection.

3.1 Paths with few intersections

We will now restrict our attention to a collection of paths of length k which does not contain any pairs of paths (P,Q) intersecting in *undesired* ways. For instance, let us give some intuition of how a long segment of $Q \setminus P$ can be avoided. Given a path P of length $k \sim \alpha(c)\log n$, the edge labels on it increase significantly more slowly than on a typical path. Hence, if we can enforce the property that all the vertices on P are *typical*, in the sense that the paths leaving them increase and decrease at the rate at most $1/(en)$, then this will ensure that all the paths *leaving* P are in fact shorter than the corresponding segment of P. The majority of the work in this section consists in showing that restricting our collection of paths does not significantly decrease the expected number of such paths.

Throughout the section, we work with the model $G_p(W)$ whose labels are i.i.d. exponential random variables $(W_e : e \in K_n)$. Recall that by definition of $\alpha = \alpha(c)$, we have $\alpha\log(\alpha/c) - \alpha - 1 = 0$. Given $\varepsilon > 0$ and $p = c\log n/n$, we fix $k = k_{c,\varepsilon} = \lfloor (1-\varepsilon)\alpha(c)\log n \rfloor$. We say that a labelled path $P = (v_1, \dots, v_{k+1})$ with edge labels $w_1 \le w_2 \le \cdots \le w_k$ is *C-legal* if for $i = 1, \dots, k$

$$\left| \frac{k}{w_k} \cdot w_i - i \right| \le C\sqrt{2i\log\log i}\ , \qquad \text{and} \tag{3.1}$$

$$\left| \frac{k}{w_k} \cdot (w_k - w_{k-i+1}) - i \right| \le C\sqrt{2i\log\log i}\ . \tag{3.2}$$

This definition is identical to the one in Addario-Berry et al. [1], and the fact that for some C, a path of length k is C-legal with positive probability follow from the argument of that paper. In [1] it was also shown that if P is C-legal, then for any $1 \le i < j \le k$,

$$\left| \frac{k}{w_k}(w_j - w_i) - (j-i) \right| \le 4C\sqrt{k\log\log k}\ ,$$

or, more conveniently,

$$w_j - w_i \in \frac{w_k}{k}(j-i)\left(1 \pm 4(j-i)^{-1}C\sqrt{k\log\log k}\right)\ . \tag{3.3}$$

For convenience, we set w_0=0, so that (3.3) also holds for $i = 0$.

For the vertex v_i and $w > 0$, let $Inc(v_i, w)$ be the set of increasing paths from v_i with labels in $[w_{i-1}, w_{i-1} + w]$. If $w = -\log(1-p) - w_{i-1}$ (the maximal possible value), we simply write $Inc(v_i)$. Similarly, let $Dec(v_i, w)$ be the set of decreasing paths from v_i with labels in $[w_i - w, w_i]$. A vertex v_i is called (i,β)-*typical* if the following hold.

(B1) For $w \le \log^{2/3} n/n$, any path in $Inc(v_i, w) \cup Dec(v_i, w)$ has length at most $2en \log^{2/3} n$, and

(B2) for $w \ge \log^{2/3} n/(2n)$, any path in $Inc(v_i, w) \cup Dec(v_i, w)$ has length at most $(1+\beta)ewn$.

P is called *β-typical* if each vertex v_s on P is (s, β)-typical.

We show that any increasing path P of length k is C-legal and β-typical (for appropriate $C, \beta > 0$) with positive probability. More formally, we first expose the ordering of the labels on P, and condition on the event that P is increasing. Then we expose the labels themselves, and bound the probability that P is C-legal. This follows directly from the arguments of [1]. Finally, we expose the paths starting at $v_1, \dots, v_{k+1}$ (disjoint from P), which are independent of the edges of P.

Lemma 5. *Let $p = c \log n/n$ and let P be a path of length $k \ge (\alpha(c) + ec) \log n/2$. First, for some C, P is C-legal with probability at least $3/4$. Second, for any $\beta > 0$, conditionally on being increasing and C-legal, P is β-typical with probability at least $1 - e^{-\Omega(\beta \log^{2/3} n)}$.*

Proof. The first statement follows from [1]. Namely, let $T_1, \dots, T_k \sim \mathrm{Exp}(1)$ be iid random variables. Then, by the memoryless property of the exponential distribution, if we condition on the event that P is increasing, then the i-th label w_i on P has the same distribution as $T_1 + \cdots + T_i$. Hence, by Corollary 12 from [1], the probability that P is C-legal is at least $3/4$.

For the second statement, fix a vertex v_i on P and $w \in (0, p)$. Note that for any edge f, and $x, w < 1/2$,

$$\mathbb{P}\left[W_f \in [x, x+w] = e^{-x} - e^{-x-w}\right] = w(1 + O(x+w)) \ .$$

Hence, the number of paths in $Inc(v_i, w)$ of length k has the same distribution as the variable Y_k defined in Section 2, after replacing p by $w(1+O(\log n/n))$. The same holds for $Dec(v_i, w)$. For (B1), let $w = \log^{2/3} n/n$, and recalling (2.1), the expected number of paths in $\mathrm{Inc}(v_i, w) \cup \mathrm{Dec}(v_i, w)$ of length at least $2ewn$ is at most

$$\left(\frac{ewn(1+O(w))}{2ewn}\right)^{2ewn} \le 2^{-\log^{2/3} n} \ .$$

For (B2), note that it suffices to impose the condition with integer values for wn (and β replaced by $\beta/2$), because the bound for other values follows. Hence, recalling (2.1), the expected number of paths at v_i violating (B2) is at most $e^{-\Omega(\beta wn)} = e^{-\Omega(\beta \log^{2/3} n)}$. Summing over all values of wn and v_i (with $O(\log^2 n)$ many options), we have that the expected number of paths violating (B1) or (B2) is $e^{-\Omega(\beta \log^{2/3} n)}$, so the probability that P is not β-typical is at most $e^{-\Omega(\beta \log^{2/3} n)}$. ■

Let $\Lambda(P)$ be the collection of paths Q of length k intersecting P in at most $2\alpha + 2$ connected components with $|P \setminus Q| \ge \log^{3/4} n$. We show that $\Lambda(P)$ contains no other C-legal and β-typical paths, which will be crucial for controlling the second moment of the number of desired paths. The main idea is similar to that in [1]: for any path $Q \in \Lambda(P))$, the labels on $Q \setminus P$ are increasing at the rate at most $1/(en)$, which is strictly higher than the rate of increase on P, namely $\frac{c}{\alpha(c)n}$.

Lemma 6. *Let n be sufficiently large and $k \geq (\alpha - \varepsilon)\log n$ for some $\varepsilon > 0$. Let P be a C-legal β-typical path of length k for $\beta = o(1)$ and an arbitrary constant C. The collection $\Lambda(P)$ contains no other paths which are C-legal and β-typical.*

Proof. Let $P = v_1, \ldots, v_{k+1}$ be an increasing path with edge labels $w_1, \ldots, w_k$. We define the weight of the path by $w(P) = w_k$, that is, the largest weight on the edges of the path. It is sufficient to show that if $Q \in \Lambda(P)$, then $w(Q) > w(P)$ (i.e., P is in some sense locally optimal). Indeed, then P is also in $\Lambda(Q)$ by symmetry, and if Q is C legal and β-typical, then $w(P) > w(Q)$, which is a contradiction.

For $Q \in \Lambda(P)$, let $S_1, S_2, \ldots, S_{j+1}$ be the (maximal) segments of $P \setminus Q$, and let $S'_1, \ldots, S'_{j+1}$ be the corresponding segments of $Q \setminus P$, with $j \leq 2\alpha + 2$. For convenience, define $w_0 = 0$ and $w_{k+1} = w_k$. Assuming $w(Q) \leq w(P)$, we show that

$$\sum_{t=1}^{j+1} \left(|S_t| - |S'_t|\right) > 0 \ , \tag{3.4}$$

so Q cannot have length k.

For a segment $S = v_i, \ldots, v_{i+s}$, we define its *weight* to be $w(S) = w_{i+s} - w_{i-1}$ (the difference between the labels of the edges preceding and following S), and note that $w(P)$ is defined consistently since $w_0 = 0$ and $w_{k+1} = w_k$. For the segments S_t of weight at most $\log^{2/3} n/n$, condition (B1) implies

$$|S_t| - |S'_t| \geq -|S'_t| \geq -2e\log^{2/3} n \ ,$$

and note that the condition $w(Q) \leq w_k$ was used for bounding $|S'_t|$ in case S'_t is the final segment of $Q \setminus P$.

Now consider a segment $S \in \{S_1, \ldots, S_{j+1}\}$ of weight at least $\log^{2/3} n/n$. Using (3.3) and $w_k \leq \frac{c\log n}{n}$, the length of S is at least

$$\frac{w(S)k}{w_k}(1 - o(1)) \geq \frac{w(S)n(\alpha(c) - 2\varepsilon)}{c} \ .$$

Moreover, using (B2), the weight of the corresponding segment of Q is at most $enw(S)(1+\beta)$, so we obtain

$$|S| - |S'| \geq w(S)\left(\frac{k}{w_k} - en - 2\beta en\right) \geq w(S)n\left(\frac{\alpha - \varepsilon}{c} - e - 2\beta\right) \ . \tag{3.5}$$

For sufficiently small positive β and ε and the considered segments S, we have $|S| - |S'| \geq 0$.

Now, at least one segment, say S_{j+1}, has to have weight at least $\frac{c\log^{3/4} n}{4\alpha(2\alpha+3)n}$, since otherwise by (3.3),

$$|S_1| + \cdots + |S_{j+1}| \leq \frac{k}{w_k}\left(\frac{c\log^{3/4} n}{4\alpha n} + O(\sqrt{k\log\log k})\right) < \log^{3/4} n \ ,$$

which contradicts the assumption of the lemma. Hence, using (3.5), $\alpha > ec$, and taking $\varepsilon > 0$ and $\beta > 0$ sufficiently small,

$$|S_{j+1}| - |S'_{j+1}| = \Omega(\log^{3/4} n) \ .$$

We conclude that

$$|S_{j+1}| - |S'_{j+1}| + \sum_{t=1}^{j}(|S_t| - |S'_t|) \geq \Omega(\log^{3/4} n) - 2(2\alpha+2)e\log^{2/3} n > 0 \ ,$$

as required. ■

We remark that the previous proof indicates why the case where $Q \setminus P$ consists of segments of length $O(1)$ is hard to control by imposing local restrictions on P. Instead, this case is covered using the second moment method in Lemma 8 below.

3.2 The second moment bound

Let $\mathcal{S}_{k,i,j}$ denote the set of pairs (P,Q) of paths in K_n of length k such that $|P \cap Q|$ consists of i edges in j components, where the components are considered in the graph $P \cup Q$. Let $I(P)$ denote the event that P is an increasing path in $G_p(W)$. For *most* values of i and j, we prove bounds on $\Delta_{k,i,j} := \sum_{(P,Q)\in\mathcal{S}_{k,i,j}} \mathbb{P}[I(P) \wedge I(Q)]$ which are sufficient for an application of Chebyshev's inequality; the remaining cases have been handled in the previous subsection.

We use the following lemma from [1]. (The upper bound can actually be strengthened for our setting, but this strengthening is not needed.)

Lemma 7 (Lemma 8 in [1])**.** *For every $n,i,j,k \in \mathbb{N}$, we have $|\mathcal{S}_{k,i,j}| \leq n^{2k+2-i-j}(2k^3)^j$.*

Recall that by definition of $\alpha = \alpha(c)$, we have $\alpha\log(\alpha/c) - \alpha - 1 = 0$. Recalling that $\alpha(1) \approx 3.591$, we have the general upper bound

$$\alpha(c) \leq \max(3.6, 3.6c) \ .$$

Denote

$$\mu_k = \frac{n^{k+1}p^k}{k!} \ . \tag{3.6}$$

The following lemma is applied when c is a constant, although it is stated with slightly weaker assumptions.

Lemma 8. *Let $p = c\log n/n$ with $c \leq n^{1/15}$, and $j \leq i < k \leq \alpha(c)\log n$.*

(i) If $j \geq 2\alpha+2$, and n is sufficiently large $\frac{1}{\mu_k^2}\Delta_{k,i,j} \leq n^{-1/4}$.

(ii) If $j \geq 2$ and $k - i \leq \varphi k$, where $\varphi\log(1/\varphi) < 1/(10ec)$, then $\frac{1}{\mu_k^2}\Delta_{k,i,j} \leq n^{-1/4}$.

(iii) If $k \leq (1-\varepsilon)\alpha\log n$ and $k-i \leq \varepsilon k/100$ for some $\varepsilon = \varepsilon(n) > 0$, then $\frac{1}{\mu_k^2}\Delta_{k,i,1} \leq n^{-\Omega(\varepsilon)}$.

Proof. For $P, Q \in \mathcal{S}_{k,i,j}$, the probability of $I(Q)$ given $I(P)$ is at most $\frac{p^{k-i}}{(k-i)!}$, since the $k-i$ edges in $Q \setminus P$ have to have increasing labels along Q. Hence, by Lemma 7,

$$\frac{1}{\mu_k^2}\Delta_{k,i,j} \le \frac{1}{\mu_k^2} \cdot \frac{n^{2k+2-i-j}(2k^3)^j p^{2k-i}}{k!(k-i)!} = \frac{k!}{(np)^i(k-i)!} \cdot \left(\frac{2k^3}{n}\right)^j \le \frac{k! n^{-3j/4}}{(np)^i(k-i)!} . \tag{3.7}$$

For $j \ge 2\alpha(c) + 2$, we now have

$$\frac{1}{\mu_k^2}\Delta_{k,i,j} \le \left(\frac{k}{np}\right)^i n^{-3j/4} \le \left(\frac{\alpha}{c}\right)^{\alpha \log n} n^{-3j/4} = n^{\alpha+1} n^{-3j/4} \le n^{-j/4} ,$$

which proves (i).

For (ii), we consider two cases. First assume that $j \ge 2$ and $i \ge i_0 = (1-\varphi)k$ with $\varphi < 1/50$ such that $\varphi \log(1/\varphi) < 1/(10ec) \le 1/(5\alpha)$. We start by bounding $\frac{k!}{(k-i)!}$. By Stirling's formula, for all n,

$$n \log n - n \le \log n! \le n \log n - n + \frac{1}{2}\log n + 10 .$$

Thus, we have

$$\begin{aligned}\log \frac{k!}{k^i(k-i)!} &\le k \log k - k + O(\log k) - i \log k - (k-i)\log(k-i) + k - i \\ &= (k-i)\log\frac{k}{k-i} - i + O(\log k) \le -k\varphi \log \varphi - i + O(\log k) \le \frac{k}{4\alpha} - i .\end{aligned}$$

Hence $\frac{k!}{(k-i)!} \le \left(\frac{k}{i}\right)^i n^{1/4}$, and substituting this into (3.7), we get that for $i \ge (1-\varphi)k$ and $j \ge 2$,

$$\frac{1}{\mu_k^2}\Delta_{k,i,j} \le \left(\frac{k}{enp}\right)^i n^{1/4-3j/4} \le e^{\log n} n^{1/4-3j/4} \le n^{-1/4} ,$$

as required for (ii).

Finally, we prove (iii), that is, the case when $j = 1$. For this case, we need a more precise estimate for the probability that both P and Q are increasing paths. Fix $P = v_1, \dots, v_{k+1}$ and $Q = w_1, \dots, w_{k+1}$, and let $s_P \in [1, k-i]$ be the number of edges in the segment of $P \setminus Q$ containing v_1, and similar for $Q \setminus P$; thus the segment of $P \setminus Q$ containing v_{k+1} has length $k - i - s_P$, and note that s_P may be $k-i$. The probability that P and Q are both increasing (given that their labels are at most p) is exactly

$$\begin{aligned}f(i, s_P, s_Q) &:= \left(\binom{2k-i}{i}\binom{2k-2i}{s_P+s_Q} i!\, s_P!\, s_Q!(k-i-s_P)!(k-i-s_Q)!\right)^{-1} \\ &= \frac{(2k-2i-s_P-s_Q)!\,(s_P+s_Q)!}{(2k-i)!\, s_P!\, s_Q!(k-i-s_P)!(k-i-s_Q)!} ,\end{aligned}$$

which is seen by assigning random labels from $[2k-i]$ to the edges of $P \cup Q$; the first two terms are the probability that $P \cap Q$ and the initial segments get the correct set of labels, and the remaining five terms are the probability of correctly ordering each of the five segments. Note that

$$\frac{1}{\mu_k^2}\Delta_{k,i,1} \le 2n^{-1}(np)^{-i}(k!)^2 \sum_{0\le s_Q\le s_P} f(k, s_P, s_Q)\ , \tag{3.8}$$

and we now proceed to proving an upper bound for the right-hand side by establishing that $f(i, s_P, s_Q)$ is maximised when $s_P = s_Q = k-i$. First, assume that $s_P \ge s_Q + 1$ (noting that in this case $k-i-s_Q \ge 1$), and we show that $f(i, s_P, s_Q)/f(i, s_P, s_Q+1) \le 1$. Indeed, we have

$$\frac{f(i, s_P, s_Q)}{f(i, s_P, s_Q+1)} = \frac{(s_Q+1)(2k-2i-s_P-s_Q)}{(k-i-s_Q)(s_P+s_Q+1)} \le \frac{(2k-2i-2s_Q-1)(s_Q+1)}{(k-i-s_Q)(2s_Q+2)} \le 1\ .$$

Hence $f(i, s_P, s_Q)$ is maximised when $s_P = s_Q$. Second, assume that $s_P \ge k-i-s_P$ (without loss of generality), and we show that $f(i, s_P, s_P)$ is maximised when $s_P = k-i$. We have

$$\begin{aligned}\frac{f(i, s_P, s_P)}{f(i, s_P+1, s_P+1)} &= \frac{(2k-2i-2s_P)(2k-2i-2s_P-1)(s_P+1)^2}{(k-i-s_P)^2(2s_P+2)(2s_P+1)}\\ &= \left(2-\frac{1}{k-i-s_P}\right)\left(2-\frac{1}{s_P+1}\right)^{-1}\ ,\end{aligned}$$

which is at most 1 since $k-i-s_P \le s_P+1$.

Finally, we bound $f(i, k-i, k-i)$, which corresponds to the case when P and Q intersect in the final i edges. Substituting $a = k-i$, we have

$$\begin{aligned}\log((k!)^2 f(k-a, a, a)) &= \log\frac{(k!)^2(2a)!)}{(k+a)!(a!)^2}\\ &\le 2k\log k - 2k + O(\log k) + 2a\log(2a) - 2a - (k+a)\log(k+a) + k + a - 2a\log a + 2a\\ &\le (k-a)\log k - k + a + 2a\log 2 + O(\log k) = (k-a)\log(k/e) + 2a\log 2 + O(\log k)\ .\end{aligned}$$

Substituting this bound into (3.8), recalling that the term $f(k-a, a, a)$ is the largest term, taking $a \le k\varepsilon/100$ and using the definition of α,

$$\frac{1}{\mu_k}\Delta_{k,k-a,1} \le n^{-1}2^{2a+O(\log\log n)}\left(\frac{k}{enp}\right)^k \le n^{-1}\left(\frac{\alpha e^{-\varepsilon/2}}{ec}\right)^{\alpha\log n} = n^{-\Omega(\varepsilon)}\ ,$$

which ends the proof. ∎

3.3 Finding a long path

Now we are prepared to complete the proof of part (iii) of Theorem 1. The upper bound is shown in Section 2.1, so it remains to prove the existence of a long path. This follows from the next lemma.

Lemma 9. *Let $p = c\log n/n$ and let $\varepsilon = \varepsilon(n)$ be such that $\varepsilon/\log^{-1/4} n \to \infty$. If $k = k_{c,\varepsilon} = (\alpha(c) - \varepsilon)\log n$, then, with high probability, $G_p(W)$ contains an increasing path of length at least k.*

Proof. Let C be a sufficiently large constant so that Lemma 5 holds and let $\beta = (\log n)^{-1/4}$. Let $\mathcal{P}$ be the collection of paths of length $k_{c,\varepsilon}$ which are C-legal and β-typical. It suffices to show that $|\mathcal{P}| > 0$ with high probability. To this end, we use the second-moment method.

A fixed path $P = v_1, \ldots, v_{k+1}$ in K_n is increasing in $G_p(W)$ with probability $\frac{p^k}{k!}$, and conditioned on this event, the probability that P is in $\mathcal{P}$ is at least $\frac{1}{2}$, by Lemma 5. Hence,

$$\mathbb{E}|\mathcal{P}| \geq \frac{1}{2}\mathbb{E}|X_k| = \frac{1}{2}\binom{n}{k+1}(k+1)p^k \geq \frac{n^{k+1}p^k}{4k!} = \frac{1}{4}\mu_k \ ,$$

recalling the definition of μ_k in (3.6). It follows that

$$\mathbb{E}|\mathcal{P}| \geq \frac{1}{2}\log n^{1-(\alpha-\varepsilon)\log((\alpha-\varepsilon)/ec)-o(1)} = n^{\Omega(\varepsilon)} \ .$$

For the second moment, Lemma 6 implies that $\mathcal{P}$ (deterministically) contains no two paths P and Q intersecting in at most $2\alpha+2$ components with $|P\cap Q| \leq k-\log^{3/4} n$. So, for a given number of components j, define $I_j = \{j, \ldots, k\}$ if $j \geq 2\alpha+2$, and $I_j = \{k-\lfloor\log^{3/4} n\rfloor, \ldots, k\}$ otherwise. Thus, we have

$$\frac{1}{\mathbb{E}[|\mathcal{P}|^2]}\mathrm{Var}(|\mathcal{P}|) \leq \frac{4}{\mu_k^2}\sum_{j=1}^{k}\sum_{i\in I_j}\Delta_{k,i,j} \ .$$

Since $\log^{3/4} n \ll \varepsilon\log n$, Lemma 8 implies that $\Delta_{k,i,j}/\mu_k^2 = n^{-\Omega(\varepsilon)}$ for all j and $i \in I_j$, and therefore $\mathrm{Var}(|\mathcal{P}|) = o(\mathbb{E}[|\mathcal{P}|^2])$. By Chebyshev's inequality, $|\mathcal{P}| \geq 1$ with high probability, as desired. ■

3.4 Proof of Proposition 2

The asymptotics for $\max\{L(i,j) : 1 \leq i,j \leq n\}$ is a result of Angel et al. [2]. Our contribution is to prove that the same actually holds for $L(1,2)$; the claim for $\max\{L(1,i) : 1 \leq i \leq n\}$ is then a straightforward consequence. Partition the interval $[0,p)$ into three pieces $[0,p_1) \cup [p_1,p_2) \cup [p_2,p)$, where $np_1 = 2\log n$ and $np_2 = np - 2\log n$. Let W_1, W_2, W_3 be the collections of edge weights falling in the respective intervals $[0, -\log(1-p_1))$, $[-\log(1-p_1), -\log(1-p_2))$, and $[-\log(1-p_2), -\log(1-p))$. This decomposes $G(W)$ into the union of three disjoint random simple temporal graphs $G(W_1), G(W_2)$, and $G(W_3)$. Note that the concatenation of three monotone paths one from $G(W_1)$, one from $G(W_2)$, and one from $G(W_3)$ (such that the endpoint of the first path is the starting point of the second and the endpoint of the second path is the starting point of the third) is a monotone path in $G(W)$.

Let $i^\star$ and $j^\star$ be the extremities of the longest increasing path in $G(W_2)$. As shown in [2], this path has length $(e-o(1))(p_2-p_1)n$, with high probability. By part (i) of Theorem 1, there exists an increasing path in $G(W_1)$ from vertex 1 to $i^\star$ and another one in $G(W_3)$ from $j^\star$ to vertex 2. As a consequence, with high probability, there exists an increasing path connecting 1 to 2, whose length is at least $(e - o(1))(p_2 - p_1)n = enp - o(np)$.

4 Reachability from a single vertex

The main result of this section is the missing part of the proof of (ii) of Theorem 1. The upper bound for the length of the longest monotone path with vertex 1 as a starting point is shown in Section 2.1. In order to prove the corresponding lower bound, we show how to construct an increasing path from vertex 1 of length $epn(1 - o(1))$, when $p = c \log n/n$ for some $c > 0$. This is done by analysing an exploration process on increasing paths from vertex 1. By doing so, we are able to a answer some further questions. Namely, how many vertices can be reached from a specified vertex, and in how many steps? It turns out that for $p = c \log n/n$ with $c \leq 1$, with high probability, vertex 1 can reach $e^{pn(1-o(1))} = n^{c-o(1)}$ vertices, and most of these vertices can be reached in $pn(1+\varepsilon) = (1+\varepsilon)c \log n$ steps. The former property was also shown by Casteigts et al. [7].

We study these questions by constructing a tree in $G_p(W)$ rooted at vertex 1, consisting of increasing paths from 1. The resulting random tree is distributed as a uniform random recursive tree on $e^{np(1-o(1))}$ vertices. It is well known (see Devroye [8], Pittel [18]) that, with high probability, such a tree has height $enp(1 - o(1))$, which gives an increasing path in $G_p(W)$. For $p = \frac{c \log n}{n}$ with $c > 1$, we may compose paths constructed in roughly $\lfloor 1/c \rfloor$ disjoint *layers*.

Recall that $G_p(W)$ is the random graph generated using exponentially distributed labels W. Moreover, $B_\ell(v)$ is the set of vertices reachable from vertex v by increasing paths in G consisting of at most ℓ edges (including the vertex v), and note that $B_n(v)$ is the set of all vertices reachable from v. Moreover, we have $B_\ell(1) \leq \sum_{k \leq \ell} Y_k$. where Y_k is the number of increasing paths of length k starting at vertex 1. Occasionally we abbreviate $B_\ell = B_\ell(1)$.

4.1 The upper bound

Before proving Theorem 4, we show some upper bounds for the number of vertices reachable from vertex 1, using the first-moment method. The first one states that part (ii) of Theorem 4 is approximately optimal, and the second bound implies that most vertices in $B_n(1)$ are *not too far* from vertex 1.

Proposition 10. *For $p = p_n \in (0, 0.1)$ and $\varepsilon \in (0, 1)$, if n is sufficiently large, then*

$$\mathbb{E}|B_n(1)| \leq e^{np+np^2} \qquad \text{and} \qquad \mathbb{E}\left[\sum_{k>(1+\varepsilon)np} Y_k\right] \leq 2e^{np(1-\varepsilon^2/4)} . \tag{4.1}$$

Proof. Firstly, note that for any k

$$\mathbb{E}Y_k = \binom{n-1}{k} p^k \leq (1-p)^{-n+1} \sum_{k=1}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} , \tag{4.2}$$

where the last inequality follows from $(1-p)^{-k} \geq 1$. The terms $\binom{n-1}{k} p^k (1-p)^{n-1-k}$ are just point probabilities of the appropriate binomial distribution.

Hence, $\mathbb{E}|B_n(1)| \leq (1-p)^{-n+1} \leq e^{np+np^2}$. For the second inequality, using a standard binomial tail bound (see, e.g., [13, Corollary 2.3]), we have

$$\mathbb{E}\left[\sum_{k>(1+\varepsilon)np} Y_k\right] \leq e^{np+np^2}\mathbb{P}(\mathrm{Bin}(n-1,p) > (1+\varepsilon)pn) \leq 2e^{np-\frac{\varepsilon^2}{4}np} \ ,$$

as required. ■

4.2 Embedding a random recursive tree

In this section, we construct a coupling of $G_p(W)$ with a uniform random recursive tree on r vertices, denoted by T_r (where r will be set to $e^{np(1-o(1))}$). To control the labels in $G_p(W)$, we relate them to the following exponential random variables. For natural numbers n_* and i, let Z'_i be mutually independent random variables with $Z'_i \sim \mathrm{Exp}(in_*)$ (so that $\mathbb{E}Z'_i = 1/(in_*)$). First we state a concentration bound that is needed for the coupling. The proof can be found at the end of this subsection.

Lemma 11. *For $n_*, i \in \mathbb{N}$, let Z'_i be as above. For $r \geq \log n_*$ and $\varepsilon > 0$,*

$$\mathbb{P}\left(\sum_{i=1}^{r} Z'_i > \frac{(1+\varepsilon)\log r}{n_*}\right) = e^{-\Omega(\varepsilon^2 \log r)} \ .$$

Let $A_r(\varepsilon)$ be the event that $\sum_{i=1}^r Z'_i \leq \frac{(1+\varepsilon)\log r}{n_*}$. We set $n_* = \lfloor n(1-2/\log n)\rfloor$. Moreover, let E be the event that for $p = 2\log n/n$, the graph $G_p(W)$ has maximum degree at most $10\log n$. We established that for $r \sim n^c$, $\mathbb{P}(A_r(\varepsilon)) = 1 - n^{-\Omega(\varepsilon^2)}$, and Chernoff bounds imply that $\mathbb{P}(E) \geq 1 - n^{-\Omega(1)}$.

Lemma 12. *Let $p = c\log n/n$ with $c \leq 1$, $n_* = \lfloor n(1-2/\log n)\rfloor$, $\varepsilon > 0$ and $r = \lfloor e^{np(1-\varepsilon)}\rfloor$. For all sufficiently large n, there is a coupling between the random variables $(Z'_i)_{i\in\mathbb{N}}$, the graph $G_p(W)$ and the random recursive tree T_r such that, assuming $A_r(\varepsilon/2)$ and E, there is a tree $\widetilde{T}_r$ rooted at vertex 1 which consists of increasing paths from 1 and is isomorphic to T_r.*

Proof. In order to construct the desired coupling, we expose the edge labels and build a tree distributed as T_r according to the following procedure. Let $m = \lceil n/\log n\rceil$, so that $n_* = n - 2m$.

Let $v_1 = 1$, and $V(\widetilde{T}_1) = \{v_1\}$. Let $S_1(1)$ be an arbitrary set of $n-m$ edges incident to v_1. Let $(v_1, v_2) \in S_1(1)$ be the edge of minimal label Z_1. Note that Z_1 is the minimum of $n-m$ exponential random variables with parameter 1, so $Z_1 \sim \mathrm{Exp}(n-m)$. We may couple (or jointly sample) Z_1 with Z'_1 so that $Z_1 \leq Z'_1$. Now let $S_2(1) = S_1(1) \setminus \{(v_1, v_2)\}$, and note that $S_2(1)$ is a set of $n-m-1$ edges incident to v_1 of label at least Z_1. Let $S_2(2)$

be an arbitrary set of $n-m-1$ edges incident to v_2 whose label is at least Z_1, which exists assuming the event E. Let $\widetilde{T}_2$ be the tree with a single edge $\{v_1, v_2\}$.

In general, for $i \geq 2$ our inductive hypothesis is that we constructed $\widetilde{T}_i$, its final vertex is a leaf v_i, and the only exposed edge incident to v_i has label $\sum_{k=1}^{i-1} Z_k$. Moreover, for $j \in \{1, \ldots, i\}$, there is a set $S_i(j)$ consisting of $n-m-i+1$ edges incident to v_j, whose other endpoint is not in $V(\widetilde{T}_i)$. The labels of the edges in $S_i(j)$ have not been exposed, but they are known to be at least $\sum_{k=1}^{i-1} Z_k$.

To construct $\widetilde{T}_{i+1}$, expose the edge $e_i \in \cup_{j=1}^{i} S_i(j)$ with a minimal label, and let its endpoint outside of $V(\widetilde{T}_i)$ be v_{i+1}. Denote the label of e_i by $Z_1 + Z_2 + \cdots + Z_i$. Crucially, v_{i+1} is equally likely to be attached to any of the vertices $v_1, \ldots, v_i$, which is why $\widetilde{T}_{i+1}$ is distributed as T_{i+1}.

Moreover, we claim that $Z_i \sim \mathrm{Exp}(i(n-m-i+1))$. Indeed, using the memoryless property of the exponential distribution, for each edge $e \in \cup_{j=1}^{i} S_i(j)$, $W_e - \sum_{k=1}^{i-1} Z_k$ is distributed as $\mathrm{Exp}(1)$, and the minimum of those $i(n-m-i+1)$ variables has the distribution $\mathrm{Exp}(i(n-m+i))$. Since $i \leq r \leq m$, we have $n-m-i+1 \geq n-2m = n_*$, so we may couple Z_i with Z_i' so that $Z_i \leq Z_i'$.

It remains to construct the sets $S_{i+1}(j)$. Let $F = F_{i+1}$ be the set of edges incident to v_{i+1} whose labels are at most $Z_1 + \cdots + Z_i \leq \frac{(1+\varepsilon/2)\log r}{n_*} \leq \frac{2\log n}{n}$. Assuming the event E, we have $|F| \leq 10 \log n < m/2$. Hence there exists a set of $n-m-i$ edges incident to v_{i+1} which are not incident to $v_1, \ldots, v_i$ and do not lie in F; denote such a set by $S_{i+1}(i+1)$. For $j \leq i$, note that the labels of all edges in $S_i(j)$ are at least $Z_1 + \cdots + Z_i$, and that we have only exposed the label of e_i. Thus we can let $S_{i+1}(j) \subset S_i(j)$ be an arbitrary set of $n-m-i$ edges.

The procedure is continued until $i = r = \lfloor e^{np(1-\varepsilon)} \rfloor$. To verify that all the edges of $\widetilde{T}_r$ have labels at most $-\log(1-p)$, note that, conditionally on the event $A_r(\varepsilon/2)$, for sufficiently large n,

$$\sum_{i=1}^{r} Z_i \leq \sum_{i=1}^{r} Z_i' \leq \frac{(1+\varepsilon/2)\log r}{n_*} \leq \frac{(1+\varepsilon)(1-\varepsilon)np}{n(1-(3\log n)^{-1})} \leq -\log(1-p).$$

This completes the proof. ∎

Lemma 12 is crucial in the proof of Theorem 4 that uses the typical properties of random recursive trees and the fact that the events E and $A_r(\varepsilon)$ occur with high probability.

Proof of Theorem 4. We generate $G_p(W)$ using exponential labels W. Let $r = \lfloor e^{np(1-\varepsilon/2)} \rfloor = \lfloor n^{c(1-\varepsilon/2)} \rfloor$. Since the events $A_r(\varepsilon/4)$ and E occur with probability $1 - n^{-\Omega(\varepsilon^2)}$, we may assume that $G_p(W)$ contains a tree $\widetilde{T}_r$ rooted at vertex 1 which consists of increasing paths from 1 and is distributed as a uniform random recursive tree on r vertices. Note that this event implies

$$|B_n(1)| \geq n^{c(1-\varepsilon/2)} \ . \tag{4.3}$$

Part (i) of the theorem follows from the fact that, with probability $1 - n^{-\Omega(\varepsilon)}$, $\widetilde{T}_r$ contains a path of length at least $(1-\varepsilon/2)e\log r \geq (1-2\varepsilon)ec\log n$ [Corollary 1.3]af13. (Earlier proofs without explicit bounds on the failure probability can be found in [8, 18].)

In order to prove part (ii) of the theorem, note that Proposition 10 and Markov's inequality imply that, with high probability, for $\ell = (1+10\sqrt{\varepsilon})c\log n$, there are at most $n^{c(1-\varepsilon/4)}$ paths starting at vertex 1 of length at least ℓ. Hence, using (4.3), most vertices in $B_n(1)$ are at distance at most ℓ from vertex 1. ∎

For $c = 1+\varepsilon$, Theorem 4 (ii) implies that $B_{\log n}(1) = n^{1-o(1)}$. However, simply taking $\varepsilon \log n$ additional steps, each using an interval of labels of length, say $10/n$, shows that in fact $B_{(1+\varepsilon)\log n} = (1-o(1))n$, which implies that $\ell(1,2) \sim \log n$ with high probability. Using arguments which will be presented in Section 5.4, one can reprove the 1-2-3 phase transition from [7] with path lengths at most $(1+o(1))\log n$.

It remains to prove Lemma 11, which gives a lower bound for the probability of $A_r(\varepsilon)$. It follows easily from the following result due to Janson [12].

Lemma 13. *Let $X_i \sim \mathrm{Exp}(a_i)$ be mutually independent random variables. Define*

$$\mu = \mathbb{E}\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} \frac{1}{a_i} \quad \textit{and} \quad a_* = \min_{i\in[r]} a_i \ .$$

For $\varepsilon > 0$, we have

$$\mathbb{P}\left(\sum_{i\in[r]} X_i \geq (1+\varepsilon)\mu\right) \leq (1+\varepsilon)^{-1} e^{-a_*\mu(\varepsilon - \log(1+\varepsilon))} \ .$$

Proof of lemma 11. Apply Lemma 13 with $a_i = in_*$ and $a_* = n_*$. Note that in our case,

$$\mu = \sum_{i\in[r]} \frac{1}{in_*} = \frac{\log r + O(1)}{n_*} \ .$$

Using $\varepsilon - \log(1+\varepsilon) = \Omega(\varepsilon^2)$, we obtain

$$\mathbb{P}\left(\sum_{i\in[r]} U_i' \geq (1+\varepsilon)\frac{\log r}{n_*}\right) \leq \mathbb{P}\left(\sum_{i\in[r]} U_i' \geq (1+\varepsilon/2)\mu\right) = e^{-\Omega(\log r \varepsilon^2)} \ ,$$

as required. ∎

4.3 Composing long paths

We close this section by proving part (ii) of Theorem 1. In fact, the following theorem allows much larger values of p than the $p = c\log n/n$ considered in Theorem 1. On the other hand, note that for $p \gg \log n/n$, the statement follows from Proposition 2.

Our proof uses Theorem 4 to construct a path of length $epn(1-o(1))$ for any p of order $\log n/n$.

Theorem 14. *Let $a \geq 1$ and let $\varepsilon > 0$. If $p \leq \frac{(\log n)^a}{n}$, then, with high probability, G_p contains an increasing path of length at least $(1-\varepsilon)epn$.*

Proof. For $p = c' \log n/n$ with $c' \le 1$, the statement follows from Theorem 4. Let us write the probability p as $p = Bc \log n/n$ for some integer $B = B(n) = \log n^{(O(1))}$ and $c \in (1/2, 1)$.

We sample $G_{p'}$ with $p' < p$ in *layers* as follows. Let $L_1, L_2, \dots, L_B$ be mutually independent random temporal graphs on $[n]$ with probability $q = c \log n/n$ sampled as follows: in L_i, each edge e is present with probability q, and subject to that, it gets a uniformly random label $U_e(i) \in [(i-1)q, iq)$; otherwise, we write $U_e(i) = -1$. Clearly each L_i has the claimed distribution.

Informally, we form $G_{p'}$ as a union of $L_1, \dots, L_B$, and if an edge belongs to two layers L_i and L_j, it gets its label from $\max(i, j)$. In other words, (G, M) is a labelled graph, where G is the union of $L_1, \dots, L_B$, and the label M_e of the edge e is

$$M_e = \max(U_e(1), U_e(2), \dots, U_e(B)) \ .$$

The ordered graph induced by M on G is distributed as $G_{p'}$ with $1 - p' = (1-q)^B > 1 - qB = 1 - p$. This follows from the fact that the labels M on the graph G are still i.i.d. We refer to the edges of G as *i-single edges* if they occur in only one of the layers $L_1 \cup \dots \cup L_i$.

We inductively build an increasing path in G as follows. P_1 is a path in L_1 from 1 to some vertex v_2 of length $eqn(1-\varepsilon)$, which exists with high probability by Theorem 4.

Assume that P_i is an increasing path of length at least $ieqn(1-\varepsilon)$ in $L_1 \cup \dots \cup L_i$ from 1 to v_i consisting of i-single edges. Now we expose $L_{i+1} \cap P_i$. We assume that no edge of L_{i+1} is on the path P_i, as this occurs with probability

$$(1-q)^{|P_i|} \ge 1 - q|P_i| \ge 1 - \frac{\log n}{n} \cdot epn \ge 1 - n^{-1/2}.$$

Finally, let $K = K_i = [n] \backslash V(P_i) \cup \{v_i\}$, so $K \ge n(1-ep)$. By Theorem 4 (i), with probability at least $1 - n^{-\Omega(\varepsilon^2)}$, $L_{i+1}[K]$ has an increasing path of length at least $(1-\varepsilon/2)eq|K| \ge (1-\varepsilon)eqn$. Denote the other endpoint of this path by v_{i+1}, and let P_{i+1} be the obtained increasing path in $L_1 \cup \dots \cup L_{i+1}$. The failure probability in step i is at most $n^{-\Omega(\varepsilon^2)}$.

Thus we may continue the process until $i = B$, obtaining the desired path P_B. Recalling that $B = (\log n)^{O(1)}$, the failure probability is at most $Bn^{-\Omega(\varepsilon^2)} = o(1)$. ∎

5 Shortest increasing paths

In this section, we assume that $c > 1$. Let $\psi(x) = x \log(ce/x)$. Recall that the equation $\psi(x) = 1$ has two distinct solutions $\gamma(c) < \beta(c)$. Moreover, $\psi(x) > 1$ for all $x \in (\gamma(c), \beta(c))$. We prove that for any $x \in (\gamma(c), \beta(c))$ with high probability, there exists an increasing path containing $(1 + o(1))x \log n$ edges. This will simultaneously prove the upper bound for $\ell(1, 2)$ (part (i) of Theorem 3) and the lower bound for $L(1, 2)$ (part (i) of Theorem 1).

5.1 The general strategy

In this section, we generate the random temporal graph using independent uniform edge labels: each edge (i, j) is assigned a label $U_{ij} \sim \mathrm{Unif}[0, 1]$, and an edge is kept if $U_{ij} \le p$.

The resulting ordered graph is denoted by $G_p = G_p(U)$. Recall that $G_I = G_I(U)$ is the random temporal graph consisting of edges whose labels are in a given interval $I \subset [0,1]$.

The strategy consists of looking for increasing paths from vertex 1 along which the labels increase roughly as they should to have length $x \log n$ along the whole range. Similarly, we look for decreasing paths from vertex 2, again with the constraints that the labels decrease at a rate that ensures that, if extended for the whole range of labels, the length would be $x \log n$. We only conduct this search up to half the distance from each end, namely $\frac{1}{2}x \log n$, and show that with high probability the two sets of end points of the path must intersect. This is because, for $x \in (\gamma(c), \beta(c))$, with high probability the sets at distance $\frac{x}{2} \log n$ are of size at least $n^{1/2}$.

5.2 The branching process construction

We now describe the construction of the increasing paths from vertex 1 in $G_{[0,p/2]}$. In the following section, the construction of the decreasing paths in $G_{[p/2,p]}$ from 2 is done similarly and symmetrically.

Lemma 15. *Let $p = c \log n$ for some $c > 1$ and let $x \in (\gamma(c), \beta(c))$. With high probability, $G_{p/2}$ contains a tree T_1 which consists of increasing paths starting at vertex 1 and has at least $n^{1/2+\delta}$ leaves at distance $\frac{1}{2}x \log n \pm (\log n)^{1/2}$ from vertex 1 for some constant $\delta = \delta(x) > 0$.*

Proof. We fix constants $x \in (\gamma(c), \beta(c))$ and $A > 0$ to be chosen later. Let $m = \lfloor xA \rfloor$. We split the interval $[0, p/2]$ into $r = \lfloor (\log n)/(2A) \rfloor$ disjoint intervals $I_j = [jcA/n, (j+1)cA/n)$ for $0 \le j < r$. The first interval I_0 is used differently from the remaining ones $(I_1, \dots, I_{r-1})$: the intervals $I_1, \dots, I_{r-1}$ are used to construct a supercritical branching process that builds the desired path, while the first interval is used to ensure that we have enough starting points to achieve a low failure probability. The constructed paths will consist of $1 + (r-1)m$ edges, which is indeed $\frac{1}{2}x \log n \pm (\log n)^{1/2}$ for large n.

To build the branching process, we need to ensure independence and avoid collisions (this is avoidable for short paths, but not for long paths). Towards this objective, we now fix an arbitrary $\epsilon > 0$. We maintain a set of vertices which are not considered, whose size is at any time at most ϵn. Initially, we discard vertices arbitrarily to keep a *target set* of size $\lceil (1-\epsilon)n \rceil$, but later on, we first discard the vertices that have been used before, and then complete with arbitrary vertices. This guarantees that we always work with a target set of nodes of fixed size $\lceil (1-\epsilon)n \rceil$. This only runs into trouble if, at some point, we have discovered more than ϵn vertices, but we prove that this only ever happens with small probability.

In stage 0, we use the first interval $I_0 = [0, cA/n)$ to discover many nodes that will be used as the ancestors of branching processes. Let $\hat{\xi}$ denote the number of neighbours of vertex 1 in the target set of size $\lceil (1-\epsilon)n \rceil$. Then $\hat{\xi}$ is a binomial random variable with parameters $\lceil (1-\epsilon)n \rceil$ and cA/n, so that

$$\mathbb{E}[\hat{\xi}] = \lceil (1-\epsilon)n \rceil \cdot \frac{cA}{n} \ge (1-\epsilon)cA\,.$$

Let B_0 be the event that $\hat{\xi} \le cA/2$. Then, by a standard binomial tail bound, we have

$$\mathbb{P}(B_0) = \mathbb{P}(\mathrm{Bin}(\lceil (1-\epsilon)n \rceil, cA/n) \le cA/2) \le e^{-cA/10}\,, \tag{5.1}$$

for all $\epsilon > 0$ small enough and all n large enough. If it turns out that $\hat{\xi} > A$, we keep an arbitrary set $N(1)$ consisting of only $\lfloor A \rfloor$ vertices.

We now proceed with the stages $j = 1, \ldots, r-1$. We construct a branching process which will help us find the atypical paths in the graph. For each j, the j-th step of the branching process consists of m 'levels' in the graph G_{I_j}. The basic ingredient is the following. For a given vertex u and a given interval I_j of length cA/n, let $S_j(u)$ be the set of vertices v such that the graph G_{I_j} contains an increasing path of length m from u to v using only vertices from the current target set. Let ξ be the random variable describing $|S_j(u)|$. Then, for any fixed natural number $m \ge 1$, we have

$$\mathbb{E}[\xi] = \binom{\lfloor (1-\epsilon)n \rfloor}{m} \left(\frac{cA}{n}\right)^m \sim \left(\frac{(1-\epsilon)ceA}{m}\right)^m,$$

as $n \to \infty$. By choosing $A \in \mathbb{N}$ large enough (so that $m \ge (1-\varepsilon)xA$), we can ensure that

$$\mathbb{E}[\xi] \ge \left(\frac{(1-\epsilon)ce}{x}\right)^{(1-\epsilon)xA} = (1-\epsilon)^{(1-\epsilon)xA} e^{(1-\epsilon)A\psi(x)} \ge \left(e^{-2\epsilon x + \psi(x)}\right)^{(1-\epsilon)A}. \tag{5.2}$$

Recall that, since $c > 1$, the equation $\psi(x) = x\log(ec/x) = 1$ has two solutions $\gamma(c) < \beta(c)$, and that $\psi(x) > 1$ for $x \in (\gamma(c), \beta(c))$. In particular, by choosing $\epsilon > 0$ small enough, we may ensure that $\mathbb{E}[\xi] > 1$.

Later, we shall also need that $\mathbb{E}[\xi^2] < \infty$. We provide a proof for the sake of completeness.

Claim 16. *There is a universal constant $C \in \mathbb{R}$ (independent of n) such that $\mathbb{E}[\xi^2] \le C$.*

Proof. Observe that, by construction, ξ is stochastically dominated by the number of individuals in a branching process whose progeny distribution is binomial with parameters $(n, cA/n)$ (this is an upper bound on the number of simple paths of length m). As $n \to \infty$, $\mathrm{Bin}(n, cA/n)$ converges in distribution to a random variable with $\mathrm{Poisson}(cA)$ distribution. So, for all n large enough, and for all $i \ge 1$ we have $\mathbb{P}(\mathrm{Bin}(n, cA/n) = i) \le 2\mathbb{P}(\mathrm{Poisson}(cA) = i)$. Let f be the probability generating function of a $\mathrm{Poisson}(cA)$ random variable N, namely $f(s) = \mathbb{E}[s^N]$. Let f_n be the probability generating function of a $\mathrm{Binomial}(n, cA/n)$ random variable. Then $f_n(s) \le 2f(s)$ for all n large enough. It is standard that ξ is stochastically dominated by a random variable whose probability distribution is the m-th composite function $f_n \circ f_n \circ \cdots \circ f_n = f_n^{\circ m}$. However, by the previous arguments, $f_n^{\circ m}$ has a positive radius of convergence, independent of n. This implies that all moments of ξ are finite, and thus that $\mathbb{E}[\xi^2] < \infty$. ■

Fix a vertex $v_0 \in N(1)$. In a target set which avoids $N(1) \cup \{1\}$, we expose increasing paths from v_0 in G_{I_1} to discover the set $S_1^* = S_1(v_0)$. This step is used repeatedly to construct a branching process containing the nodes at distances mj, for $j = 1, 2, \ldots, r-1$ from v_0 (recalling that $m(r-1) \sim x\log n$). Then the same exploration process is run for each $v \in N(1)$ to *boost* the survival probability. To ensure independence and the global increasing property of the paths, the interval $[jcA/n, (j+1)cA/n]$ is used to construct the part of the path between distance mj and $m(j+1)$.

Let us describe the j-th step of the branching process. Suppose that we have discovered a set S^*_{j-1} of vertices at distance $(j-1)m$ from v_0. We then use breadth-first search in G_{I_j} for each $u \in S^*_{j-1}$ up to distance m, discarding the previously discovered vertices in the process. This yields the sets $S_j(u)$ for $u \in S^*_{j-1}$ whose sizes are distributed as ξ, and let

$$S^*_j(u) = \bigcup_{u \in S^*_{j-1}(u)} S_j(u) \; .$$

Note that the number of discarded vertices while exposing $S_j(u)$ is at most $m|S_j(u)|$. Let B_1 be the bad event that we discover more than $n/\log n$ vertices before reaching level r. Then, recalling that we only keep at most A ancestors from stage 0, Markov's inequality implies that

$$\mathbb{P}(B_1) \le \frac{\log n}{n} \sum_{i=0}^{r} A\mathbb{E}[\xi]^i \le n^{-1/4}\,, \tag{5.3}$$

for any n large enough (depending on ϵ and A).

We now claim that

$$\mathbb{P}(|S_{r-1}(v_0)| < n^{1/2+\delta}) \le q \tag{5.4}$$

for some constants $\delta > 0$, $q < 1$ and for large n. Although the branching process is only run to level $r-1$ in the graph, we may then complete it using independent copies of ξ. To this end, let $(Z_i)_{i\ge1}$ be the branching process defined from a single ancestor by $Z_1 = 1$, and $Z_{i+1} = \xi^n_1 + \xi^n_2 + \cdots + \xi^n_{Z_i}$, where $(\xi^n_j)_{n\ge0, j\ge1}$ are iid copies of the variable ξ. By the choice of $\epsilon > 0$, we know that $\mathbb{E}[\xi] > 1$, so that the process is supercritical. It follows that, denoting by $\mathcal{E}$ the event that there exists some $n \ge 0$ for which $Z_n = 0$, we have $\mathbb{P}(\mathcal{E}) = q' < 1$, and the process survives forever (and in particular to level $r-1$) with probability $1 - q' > 0$.

Furthermore, the non-negative martingale $Z_i/\mathbb{E}[\xi]^i$ converges almost surely to a limit W. Claim 16 implies that $\mathbb{E}(\xi^2) < \infty$. Hence, by the Kesten–Stigum Theorem (see, e.g., Athreya and Ney [3, Chapter I, Section 6, Theorem 2]), W is almost surely positive on the survival event $\mathcal{E}^c$, that is,

$$\mathbb{P}(W = 0 \mid \mathcal{E}^c) = 0 \; .$$

Since $r \to \infty$ as $n \to \infty$, $Z_r/\mathbb{E}(\xi)^r$ converges in distribution to W. Using (5.2), and recalling that $r = \lfloor (\log n)/(2A) \rfloor$,

$$\mathbb{E}(\xi)^{r-1} \ge \left(e^{-2\epsilon x + \psi(x)}\right)^{(1-\epsilon)A(r-1)} \ge n^{(1-2\epsilon)\psi(x)/2} \; . \tag{5.5}$$

Since $\psi(x) > 1$ because $x \in (\gamma(c), \beta(c))$, this can be made at least $n^{1/2+2\delta}$ for some small $\delta > 0$ by making $\epsilon > 0$ small enough once again. With this choice, and for any $\varepsilon' > 0$ and for sufficiently large n, have

$$\mathbb{P}(Z_{r-1} < n^{1/2+\delta} \mid \mathcal{E}^c) \le \mathbb{P}(W \le n^{-\delta} \mid \mathcal{E}^c) + \varepsilon' \le 2\varepsilon'$$

for all n large enough. The previous estimate, the fact that $|S_{r-1}(v_0)|$ is distributed as Z_{r-1}, and $\mathbb{P}(\mathcal{E}) \le q' < 1$ imply that (5.4), as claimed.

We have established that the constructed branching process from a single vertex $v_0 \in N(1)$ reaches $n^{1/2+\delta}$ vertices with positive probability. Now we run the same exploration process to define a set $S_{r-1}(v)$ for each $v \in N(1)$; using (5.3), we only discard $o(n)$ vertices in total. The probability that $|S_{r-1}(v)| \ge n^{1/2+\delta}$ for some $v \in N(1)$ is at least $1-q^A$, which can be made arbitrarily close to 1 by choosing a sufficiently large constant A. ∎

5.3 The construction of atypical paths: Proof of part (i) of Theorem 1 and part (i) of Theorem 3

Now we are prepared to complete the proof of part (i) of Theorem 1 and part (i) of Theorem 3.

As before, let $x \in (\beta(c), \gamma(c))$, and let $p = (c \log n)/n + 1/n$. It suffices to prove that there is a path of length $x \log n(1 + o(1))$ in G_p (as opposed to $G_{c \log n/n}$). Construct the tree T_1 in $G_{[0,c\log n/(2n)]}$ using Lemma 15. Similarly, construct the tree T_2 consisting of decreasing paths from vertex 2 with labels in $[p - c\log n/(2n), p]$. Note that T_2 can be taken to be vertex-disjoint from T_1, since $|V(T_1)| = o(n)$. Let L_1 and L_2 denote the sets of leaves of T_1 and T_2. The number of vertex pairs in $L_1 \times L_2$ is at least $n^{1+\delta}$, and the labels of these pairs have not been exposed while constructing T_1 and T_2. Hence, the probability that there is an edge $e \in L_1 \times L_2$ whose label is in the 'middle interval' $(c\log n/(2n), c\log n/(2n)+1/n)$ is at least $1 - O\left(e^{-n^\delta}\right)$. This edge e completes an increasing path from 1 to 2 of length $(1 + o(1))x \log n$, as required.

5.4 From typical to worst case shortest paths: Proof of parts (ii) and (iii) of Theorem 3

It remains to prove parts (ii) and (iii) of Theorem 3. We do this by relating the 'worst-case' shortest path to typical shortest paths in a modified graph (where only the range of edge labels is changed). Indeed, in the graph, there are always atypical vertices that do not have any edge in a given range of labels of length around $(\log n)/n$. Getting in or out of these vertices already 'costs' a significant portion of the interval of labels, and thus the typical length of the shortest path between such vertices is much longer than between two typical vertices. This explains the constant $\gamma(c-1)$ and $\gamma(c-2)$, accounting for such a fixed cost at one or both ends, respectively. We prove that this is essentially the worst situation.

We first establish the negative result, that is, a lower bound on $\ell(1, j)$ and $\ell(i, j)$. Let $I_1 = [0, (1-\varepsilon)\log n/n]$, $I_3 = [(c-1+\varepsilon)\log n/n, c\log n/n]$ and $I_2 = [0, c\log n/n] \setminus (I_1 \cup I_3)$. As in the previous section, we sample the graph G_p as the union of independent copies of G_{I_1}, G_{I_2} and G_{I_3}. If an edge e is in more than one of the three graphs, it *chooses* its minimal label for G_p, and we ignore the lower-order change in edge probability caused by this sampling. Suppose that $c \ge 2$. Let N be the number of vertices which are isolated in G_{I_3}. Then, with high probability $N \ge 1$. Indeed, since we only require to avoid an interval of length $(1-\varepsilon)\log n/n$, we have $\mathbb{E}[N] = n(1-(1-\varepsilon)\log n/n)^{n-1} \sim n^{\varepsilon}$, while a similar argument yields $\mathbb{E}[N^2] \sim n^{2\varepsilon}$. Chebyshev's Inequality then implies that $\mathbb{P}(N>0) \to 1$. Let $j^\star$ be such a vertex. By part (i) of Theorem 3, with high probability, the shortest path between 1 and $j^\star$ in $G_{I_1} \cup G_{I_2}$ has length at least $(\gamma(c-1+\varepsilon)+o(1))\log n$. By continuity of $\gamma(\cdot)$ and since $\varepsilon > 0$ was arbitrary, it follows that, for any $\varepsilon' > 0$, with high probability $\max\{\ell(1,i) : 1 \le i \le n\} \ge (\gamma(c) - \varepsilon')\log n$.

The argument is easily adapted to find two vertices, one which is isolated in G_{I_1}, and another one which is isolated in G_{I_3}. This shows that for $c \ge 3$, $\max\{\ell(i,j) : 1 \le i, j \le n\} \ge (\gamma(c-2)-\varepsilon')\log n$ with high probability.

It thus remains to prove the claimed upper bounds on $\ell(1,j)$ and $\ell(i,j)$. For this, we prove that for $c > 2$, with high probability, all pairs of vertices are connected by a path whose length is close to $\gamma(c-2)\log n$. Let $\varepsilon > 0$ be such that $c > 2+2\varepsilon$. We split the range $[0,p]$ into $I_1 = [0,p_1)$, $I_2 = [p_1,p_2)$ and $I_3 = [p_3,p]$, with $np_1 = (1+\varepsilon)\log n$ and $np_2 = (c-1-\varepsilon)\log n$. For an interval $I \subseteq [0,1]$, let $B_k(u,I)$ denote the set of vertices that $u \in [n]$ can reach using an increasing path of length to k in G_I. Let $k_2 = \lfloor(\gamma(c-2-2\varepsilon)+\varepsilon)\log n\rfloor$. Then, by exchangeability of the vertices,

$$\begin{aligned}\mathbb{P}\left(2 \in B_{k_2}(1,I_2)\right) &\le (1-\varepsilon/2)\mathbb{P}\left(|B_{k_2}(1,I_2)| < (1-\varepsilon/2)n\right) + \mathbb{P}\left(|B_{k_2}(1,I_2)| \ge (1-\varepsilon/2)n\right) \\ &= 1-(\varepsilon/2)\cdot\mathbb{P}\left(B_{k_2}(1,I_2) < (1-\varepsilon/2)n\right) .\end{aligned}$$

Part (i) of Theorem 3 implies that the left-hand side above tends to 1, so that $|B_{k_2}(1,I_2)| \ge (1-\varepsilon/2)n$ with high probability. Now, if there is an edge with label in I_3 between $i \in B_k(1,I_2)$ and some vertex j, then $j \in B_{k_2+1}(1, I_2 \cup I_3)$. Furthermore, if j is not already in $B_{k_2}(1,I_2)$, and this set turns out to be of size at least $(1-\varepsilon/2)n$, then there are at least $(1-\varepsilon/2)n - \Delta$ potential edges whose label might be in I_3, where Δ denotes the maximal degree of the entire graph. As a consequence, using the union bound,

$$\begin{aligned}\mathbb{P}(B_{k_2+1}(1,I_2\cup I_3) \ne [n]) &\le \mathbb{P}(|B_{k_2}(1,I_2)| < (1-\varepsilon/2)n) + \mathbb{P}(4\Delta > \varepsilon n) + n(1-|I_3|)^{(1-3\varepsilon/4)n} \\ &\le o(1) + n\exp(-(1-3\varepsilon/4)(1+\varepsilon)\log n) = o(1)\,,\end{aligned}$$

as $n \to \infty$; here we used the classical fact that the maximum degree of $\mathcal{G}(n, c\log n/n)$ is with high probability smaller than $\varepsilon n/4$ for any $\varepsilon > 0$. Together with the continuity of $\gamma(\cdot)$, this completes the proof of part (ii). Now, just as above, this shows that the set $S = \{u \in [n] : B_{k_2+1}(u, I_2 \cup I_3) = [n]\}$ has cardinality at least $(1-\varepsilon/2)n$ with high probability. From there, an argument similar as the one we have just used, but using an extra edge with label in I_1 (instead of I_3) shows that with high probability, the set $S' = \{u \in [n] : B_{k_2+2}(u, I_1 \cup I_2 \cup I_3) = [n]\}$ is actually $[n]$ with high probability. This concludes the proof of part (iii).

References

[1] Addario-Berry, L., Broutin, N., and Lugosi, G. (2010). The longest minimum-weight path in a complete graph. *Combin. Probab. Comput.*, 19(1):1–19.

[2] Angel, O., Ferber, A., Sudakov, B., and Tassion, V. (2020). Long monotone trails in random edge-labellings of random graphs. *Combinatorics, Probability and Computing*, 29(1):22–30.

[3] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes*. Springer, Berlin.

[4] Becker, R., Casteigts, A., Crescenzi, P., Kodric, B., Renken, M., Raskin, M., and Zamaraev, V. (2022). Giant components in random temporal graphs. *arXiv preprint arXiv:2205.14888*.

[5] Boyd, D. W. and Steele, J. M. (1979). Random exchanges of information. *J. Appl. Probab.*, 16(3):657–661.

[6] Casteigts, A., Raskin, M., Renken, M., and Zamaraev, V. (2020). Sharp thresholds in random simple temporal graphs. *arXiv preprint arXiv:2011.03738*.

[7] Casteigts, A., Raskin, M., Renken, M., and Zamaraev, V. (2022). Sharp thresholds in random simple temporal graphs. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–326.

[8] Devroye, L. (1987). Branching processes in the analysis of the heights of trees. *Acta Informatica*, 24:277–298.

[9] Frieze, A. M. and Grimmett, G. R. (1985). The shortest-path problem for graphs with random arc-lengths. *Discrete Appl. Math.*, 10(1):57–77.

[10] Haigh, J. (1981). Random exchanges of information. *J. Appl. Probab.*, 18(3):743–746.

[11] Janson, S. (1999). One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combin. Probab. Comput.*, 8(4):347–361.

[12] Janson, S. (2018). Tail bounds for sums of geometric and exponential variables. *Statist. Probab. Lett.*, 135:1–6.

[13] Janson, S., Łuczak, T., and Rucinski, A. (2000). *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York.

[14] Lavrov, M. and Loh, P.-S. (2016). Increasing Hamiltonian paths in random edge orderings. *Random Structures Algorithms*, 48(3):588–611.

[15] Mocquard, Y., Sericola, B., and Anceaume, E. (2020). Probabilistic analysis of rumor-spreading time. *INFORMS J. Comput.*, 32(1):172–181.

[16] Mocquard, Y., Sericola, B., Robert, S., and Anceaume, E. (2016). Analysis of the propagation time of a rumour in large-scale distributed systems. In *2016 IEEE 15th International Symposium on Network Computing and Applications (NCA)*, pages 264–271.

[17] Moon, J. W. (1972). Random exchanges of information. *Nieuw Arch. Wisk. (3)*, 20:246–249.

[18] Pittel, B. (1994). Note on the height of random recursive trees and m-ary search trees. *Random Structures and Algorithms*, 5:337–347.

[19] van Ditmarsch, H., Kokkinis, I., and Stockmarr, A. (2017). Reachability and expectation in gossiping. In *PRIMA 2017: Principles and Practice of Multi-Agent Systems: 20th International Conference, Nice, France, October 30–November 3, 2017, Proceedings*, pages 93–109. Springer.