Successive Minimum Spanning Trees

Svante Janson
Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
http://www.math.uu.se/svante-janson/
svante.janson@math.uu.se

Gregory B. Sorkin
Department of Mathematics, The London School of Economics and Political Science, Houghton Street, London WC2A 2AE, England
http://personal.lse.ac.uk/sorkin/
g.b.sorkin@lse.ac.uk

Abstract

In a complete graph K_n with edge weights drawn independently from a uniform distribution $U(0,1)$ (or alternatively an exponential distribution Exp(1)), let T_1 be the MST (the spanning tree of minimum weight) and let T_k be the MST after deletion of the edges of all previous trees T_i, $i < k$. We show that each tree's weight $w(T_k)$ converges in probability to a constant γ_k with $2k - 2\sqrt{k} < \gamma_k < 2k + 2\sqrt{k}$, and we conjecture that $\gamma_k = 2k - 1 + o(1)$. The problem is distinct from that of Frieze and Johansson [6], finding k MSTs of combined minimum weight, and the combined cost for two trees in their problem is, asymptotically, strictly smaller than our $\gamma_1 + \gamma_2$.

Our results also hold (and mostly are derived) in a multigraph model where edge weights for each vertex pair follow a Poisson process; here we additionally have $\mathbb{E}(w(T_k)) \to \gamma_k$. Thinking of an edge of weight w as arriving at time $t = nw$, Kruskal's algorithm defines forests $F_k(t)$, each initially empty and eventually equal to T_k, with each arriving edge added to the first $F_k(t)$ where it does not create a cycle. Using tools of inhomogeneous random graphs we obtain structural results including that $C_1(F_k(t))/n$, the fraction of vertices in the largest component of $F_k(t)$, converges in probability to a function $\rho_k(t)$, uniformly for all t, and that a giant component appears in $F_k(t)$ at a time $t = \sigma_k$. We conjecture that the functions ρ_k tend to time translations of a single function, $\rho_k(2k + x) \to \rho_\infty(x)$ as $k \to \infty$, uniformly in $x \in \mathbb{R}$.

Simulations and numerical computations give estimated values of γ_k for small k, and support the conjectures stated above.

2012 ACM Subject Classification Mathematics of computing → Random graphs; Mathematics of computing → Paths and connectivity problems; Mathematics of computing → Combinatorial optimization; Mathematics of computing → Matroids and greedoids

Keywords and phrases miminum spanning tree, second-cheapest structure, inhomogeneous random graph, optimization in random structures, discrete probability, multi-type branching process, functional fixed point, robust optimization, Kruskal's algorithm

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2019.60

Category RANDOM

Related Version A full version of the paper is available at https://arxiv.org/abs/1906.01533.

Funding The work was partly supported by the Knut and Alice Wallenberg Foundation.

Acknowledgements We thank Oliver Riordan for helpful comments which simplified our proof, and Balázs Mezei for assistance with Julia programming.

1 Introduction

1.1 Problem definition and main results

Consider the complete graph K_n with edge costs that are i.i.d. random variables, with a uniform distribution $U(0,1)$ or, alternatively, an exponential distribution Exp(1). A well-known problem is to find the minimum (cost) spanning tree T_1, and its cost or "weight" $w(T_1)$. A famous result by Frieze [7] shows that as $n \to \infty$, $w(T_1)$ converges in probability to $\zeta(3)$, in both the uniform and exponential cases.

Suppose now that we want a second spanning tree T_2, edge-disjoint from the first, and that we do this in a greedy fashion by first finding the minimum spanning tree T_1, and then the minimum spanning tree T_2 using only the remaining edges. (I.e., T_2 is the minimum spanning tree in $K_n \setminus T_1$, meaning the graph with edge set $E(K_n) \setminus E(T_1)$.) We then continue and define T_3 as the minimum spanning tree in $K_n \setminus (T_1 \cup T_2)$, and so on. The main purpose of the present paper is to show that the costs $w(T_2)$, $w(T_3)$, ... also converge in probability to some constants.

▶ **Theorem 1.** *For each $k \geqslant 1$, there exists a constant γ_k such that, as $n \to \infty$, $w(T_k) \xrightarrow{\mathrm{p}} \gamma_k$ (for both uniform and exponential cost distributions).*

The result extends easily to other distributions of the edge costs (see full version for details), but we consider in this paper only the uniform and exponential cases.

A minor technical problem is that T_2 and subsequent trees do not always exist; it may happen that T_1 is a star and then $K_n \setminus T_1$ is disconnected. This happens only with a small probability, and w.h.p. (with high probability, i.e., with probability $1 - o(1)$ as $n \to \infty$) T_k is defined for every fixed k; see the full version for details. However, in the main part of the paper we avoid this problem completely by modifying the model: we assume that we have a multigraph, which we denote by K_n^∞, with an infinite number of copies of each edge in K_n, and that each edge's copies' costs are given by the points in a Poisson process with intensity 1 on $[0,\infty)$. (The Poisson processes for different edges are, of course, independent.) Note that when finding T_1, we only care about the cheapest copy of each edge, and its cost has an Exp(1) distribution, so the problem for T_1 is the same as the original one. However, on K_n^∞ we never run out of edges and we can define T_k for all integers $k = 1, 2, 3, \dots$. Asymptotically, the three models are equivalent (see full version for details), and Theorem 1 holds for any of the models. In particular:

▶ **Theorem 2.** *For each $k \geqslant 1$, as $n \to \infty$, $w(T_k) \xrightarrow{\mathrm{p}} \gamma_k$ also for the multigraph model with Poisson process costs.*

Frieze [7] also proved that the expectation $\mathbb{E}\, w(T_1)$ converges to $\zeta(3)$. For the multigraph model just described, this too extends.

▶ **Theorem 3.** *For the Poisson multigraph model, $\mathbb{E}\, w(T_k) \to \gamma_k$ for each $k \geqslant 1$ as $n \to \infty$.*

1.2 Motivations

Frieze and Johansson [6] recently considered a related problem, where instead of choosing spanning trees $T_1, T_2, \dots$ greedily one by one, they choose k edge-disjoint spanning trees with minimum total cost. It is easy to see, by small examples, that selecting k spanning trees greedily one by one does not always give a set of k edge-disjoint spanning trees with minimum cost, so the problems are different.

We show in Theorem 19 that, at least for $k = 2$, the two problems also asymptotically have different answers, in the sense that the limiting values of the minimum cost – which exist for both problems – are different. (Also, as discussed in Section 3.1, we improve on the upper bound from [6, Section 3] on the cost of the net cheapest k trees, since our upper bound (3.1) on the cost of the first k trees is smaller.)

Both our question and that of Frieze and Johansson [6] are natural, both seem generally relevant to questions of robust network design, and both have mathematically interesting answers.

Another reason for interest in T_2 comes from the field of algorithmic mechanism design. Imagine that each edge of $G = K_n$ is owned by a different "agent"; the agent owning edge e values it at $w(e)$, an amount known only to them. We, an "auctioneer", want to buy a spanning tree, at low cost. One "mechanism" for doing so is a sealed-bid auction where each agent posts a price $w'(e)$ for their edge, and we buy the tree that is cheapest according to these prices. Here, agents will naturally inflate their prices, posting prices $w'(e) > w(e)$.

One alternative is a VCG (Vickrey–Clarke–Groves) auction, a generalization of a single-item second-price auction. Here, we again buy the tree that is cheapest according to the posted prices w', but for each edge e purchased, we pay an amount that is a function of w'_{-e}, i.e., of all posted prices *except* that of e; for details see for example [16, Chapter 9]. This means that varying $w'(e)$ affects only whether edge e is purchased, not how much is paid for it if it is, and results in the mechanism being *truthful*: it is in each agent's selfish interest to set $w'(e) = w(e)$. Thus, the tree purchased is simply T_1, the tree cheapest according to the values w. However, the amount paid for it is more than $w(T_1)$, as the mechanism ensures the amount paid for each edge e purchased is at least $w(e)$ and typically more. A central question is the extent of this overpayment, measured by the "frugality ratio" of the VCG cost V (or that of any mechanism) to some benchmark.

The question applies of course to problems other than MSTs, including the purchase of a cheapest path between two given points in a graph, or of a basis in a bridgeless matroid. In any of these contexts, let us continue to use T_1 for the cheapest structure and T_2 for the cheapest structure disjoint from T_1. The cost $w(T_1)$ is not a useful benchmark because $V/w(T_1)$ is unbounded in even the simplest examples (such as buying one of two identical items).

Instead, Talwar [17] and Archer and Tardos [1] propose $w(T_2)$ as the benchmark. (An often-equivalent benchmark, based on a Nash equilibrium, is given by [14] and [16, Chapter 13].) [17] shows that for any bridgeless matroid, $V/w(T_2) \leqslant 1$, and, focusing on the worst case over all weights w, this bound is achieved by some weights (namely weights 0 on T_1, 1 on T_2, and infinity elsewhere). By contrast, for paths the ratio is unbounded. The interpretation, based on worst-case weights, is that this frugality ratio is 1 for amenable problems like MSTs and other matroids, and larger for other problems.

In our setting of an MST in K_n with random weights, though, the frugality ratio is naturally less than its maximum of 1. Specifically, [4] and [11] show that the VCG cost is typically $2w(T_1)$, which by [7] is $2\zeta(3) \doteq 2.4041$. We show here that $w(T_2)$ is typically γ_2, which by Remark 21 is at least 2.9683, making the frugality typically at most 0.8099. (We estimate non-rigorously that γ_2 is about 3.09 – see Table 1 – in which case the frugality ratio is typically about 0.78.) Specifically, this holds w.h.p. for n large, and also holds for the ratio between the expected VCG cost and the expected cost $w(T_2)$.

1.3 Further results, structural properties, and conjectures

It is well known that the minimum spanning tree (with any given costs, obtained randomly or deterministically) can be found by *Kruskal's algorithm* [15], which processes the edges in order of increasing cost and keeps those that join two different components in the forest obtained so far. (I.e., it keeps each edge that does not form a cycle together with previously chosen edges.) As in many other previous papers on the random minimum spanning tree problem, from [7] on, our proofs are based on analyzing the behavior of this algorithm.

Rescale weight as time, thinking of an edge of weight w as arriving at time $t = nw$. Kruskal's algorithm allows us to construct all trees T_k simultaneously by growing forests $F_k(t)$, with $F_k(0)$ empty and $F_k(\infty) = T_k$: taking the edges of K_n (or K_n^∞) in order of time arrival (increasing cost), an edge is added to the first forest F_k where it does not create a cycle. We will also consider a sequence of graphs $G_k(t) \supseteq F_k(t)$, where when we add an edge to F_k we also add it to all the graphs $G_1, \dots, G_k$; see Section 2.2 for details.

The proof of Theorem 1 is based on a detailed structural characterization of the graphs $G_k(t)$, given by Theorem 9 (too detailed to set forth in full here in the introduction), relying heavily on the theory of inhomogeneous random graphs from [3] and related works. Where $C_1(G_k(t))$ denotes the number of vertices in the largest component of $G_k(t)$ (or equivalently of $F_k(t)$, as by construction they have the same components), Theorem 9 shows that $C_1(G_k(t))/n$ converges in probability to some function $\rho_k(t)$, uniformly for all times t. Moreover, each G_k has its own giant-component threshold: $\rho_k(t)$ is 0 until some time σ_k, and strictly positive thereafter.

The functions $\rho_k(t)$ are of central interest. For one thing, an edge is rejected from F_k, making it a candidate for F_{k+1}, precisely if its two endpoints are within the same component of F_k, and we show that this is essentially equivalent to the two endpoints both being within the largest component. This line of reasoning yields the constants γ_k explicitly, albeit not in a form that is easily evaluated. We are able, at least, to re-prove that $\gamma_1 = \zeta(3)$, as first shown in [7].

The functions ρ_k also appear to have a beautiful structure, tending to time-translated copies of a single universal function:

▶ **Conjecture 4.** *There exists a continuous increasing function $\rho_\infty(x) : (-\infty, \infty) \to [0, 1)$ such that $\rho_k(2k + x) \to \rho_\infty(x)$ as $k \to \infty$, uniformly in $x \in \mathbb{R}$.*

This suggests, though does not immediately imply, another conjecture.

▶ **Conjecture 5.** *For some δ, as $k \to \infty$, $\gamma_k = 2k + \delta + o(1)$.*

If this conjecture holds, then necessarily $\delta \in [-1, 0]$, see Remark 17.

A variety of computational results are given in Section 5. They are supportive of Conjecture 4 and a stronger version of Conjecture 5 where we take $\delta = -1$:

▶ **Conjecture 6.** *As $k \to \infty$, $\gamma_k = 2k - 1 + o(1)$.*

Although we cannot prove these conjectures, some bounds on γ_k are obtained in Section 3 by a more elementary analysis of the sequence of forests F_k. In particular, Theorem 12 and Corollary 13 lead to the following, implying that $\gamma_k \sim 2k$ as $k \to \infty$.

▶ **Corollary 7.** *For every $k \geqslant 1$,*

$$2k - 2k^{1/2} < \gamma_k < 2k + 2k^{1/2}. \tag{1.1}$$

▶ Remark 8. For the minimum spanning tree T_1, various further results are known, including refined estimates for the expectation of the cost $w(T_1)$ [5], a normal limit law [9], and asymptotics for the variance [9, 13, 18]. It seems challenging to show corresponding results for T_2 or later trees. ◀

1.4 Notes on this extended abstract

A full version of this work can be found as [12]. The present extended abstract omits most proofs as well as many further results. However, Sections 2 and 3 here are reasonably complete. We will say a few words in Section 2.5 on the approach to proving Theorem 9, but the technicalities are substantial.

2 Model and main structural results

2.1 Some notation

We use $:=$ as defining its left-hand side, and $\stackrel{\text{def}}{=}$ as a reminder that equality of the two sides is by definition. We write $\doteq$ for numerical approximate equality, and $\approx$ for approximate equality in an asymptotic sense (details given where used).

If x and y are real numbers, then $x \vee y := \max(x, y)$ and $x \wedge y := \min(x, y)$. Furthermore, $x_+ := x \vee 0$. These operators bind most strongly, e.g., $t - \tau(i) \vee \tau(j)$ means $t - (\tau(i) \vee \tau(j))$.

We use "increasing" and "decreasing" in their weak senses; for example, a function f is increasing if $f(x) \leqslant f(y)$ whenever $x \leqslant y$.

Unspecified limits are as $n \to \infty$. As said above, w.h.p. means with probability $1 - o(1)$. Convergence in probability is denoted $\xrightarrow{\text{p}}$. Furthermore, if X_n are random variables and a_n are positive constants, $X_n = o_{\text{p}}(a_n)$ means, as usual, $X_n/a_n \xrightarrow{\text{p}} 0$; this is also equivalent to: for every $\varepsilon > 0$, w.h.p. $|X_n| < \varepsilon a_n$.

Graph means, in general, multigraph. (It is usually clear from the context whether we consider a multigraph or simple graph.) If G is a multigraph, then $\dot{G}$ denotes the simple graph obtained by merging parallel edges and deleting loops. (Loops do not appear in the present paper.) The number of vertices in a graph G is denoted by $|G|$, and the number of edges by $e(G)$.

For a graph G, let $\mathcal{C}_1(G)$, $\mathcal{C}_2(G)$, ... be the largest component, the second largest component, and so on, using any rule to break ties. (If there are less than k components, we define $\mathcal{C}_k(G) = \emptyset$.) Furthermore, let $C_i(G) := |\mathcal{C}_i(G)|$; thus $C_1(G)$ is the the number of vertices in the largest component, and so on. We generally regard components of a graph G as sets of vertices.

2.2 Model

We elaborate the multigraph model in the introduction.

We consider (random) (multi)graphs on the vertex set $[n] := \{1, \dots, n\}$; we usually omit n from the notation. The graphs will depend on time, and are denoted by $G_k(t)$ and $F_k(t)$, where $k = 1, 2, 3, \dots$ and $t \in [0, \infty]$; they all start as empty at time $t = 0$ and grow as time increases. We will have $G_k(t) \supseteq G_{k+1}(t)$ and $F_k(t) \subseteq G_k(t)$ for all k and t. Furthermore, $F_k(t)$ will be a forest. As $t \to \infty$, $F_k(t)$ will eventually become a spanning tree, $F_k(\infty)$, which is the kth spanning tree T_k produced by the greedy algorithm in the introduction, operating on the multigraph $G_1(\infty)$.

Since the vertex set is fixed, we may when convenient identify the multigraphs with sets of edges. We begin by defining $G_1(t)$ by letting edges arrive as independent Poisson processes with rate $1/n$ for each pair $\{i, j\}$ of vertices; $G_1(t)$ consists of all edges that have arrived at

or before time t. (This scaling of time turns out to be natural and useful. In essence this is because what is relevant is the cheapest edges on each vertex, and these have expected cost $\Theta(1/n)$ and thus appear at expected time $\Theta(1)$.) We define the cost of an edge arriving at time t to be t/n, and note that in $G_1(\infty)$, the costs of the edges joining two vertices form a Poisson process with rate 1. Hence, $G_1(\infty)$ is the multigraph model defined in Section 1.

Thus, for any fixed $t \geqslant 0$, $G_1(t)$ is a multigraph where the number of edges between any two fixed vertices is $\mathrm{Po}(t/n)$, and these numbers are independent for different pairs of vertices. This is a natural multigraph version of the Erdős–Rényi graph $G(n,t)$. (The process $G_1(t)$, $t \geqslant 0$, is a continuous-time version of the multigraph process in e.g. [2] and [10, Section 1], ignoring loops.) Note that $\dot{G}_1(t)$, i.e., $G_1(t)$ with multiple edges merged, is simply the random graph $G(n,p)$ with $p = 1 - e^{-t/n}$.

Next, we let $F_1(t)$ be the subgraph of $G_1(t)$ consisting of every edge that has arrived at some time $s \leqslant t$ and at that time joined two different components of $G_1(s)$. Thus, this is a subforest of $G_1(t)$, as stated above, and it is precisely the forest constructed by Kruskal's algorithm (recalled in the introduction) operating on $G_1(\infty)$, at the time all edges with cost $\leqslant t/n$ have been considered. Hence, $F_1(\infty)$ is the minimum spanning tree T_1 of $G_1(\infty)$.

Let $G_2(t) := G_1(t) \setminus F_1(t)$, i.e., the subgraph of $G_1(t)$ consisting of all edges rejected from $F_1(t)$; in other words $G_2(t)$ consists of the edges that, when they arrive to $G_1(t)$, have their endpoints in the same component.

We continue recursively. $F_k(t)$ is the subforest of $G_k(t)$ consisting of all edges in $G_k(t)$ that, when they arrived at some time $s \leqslant t$, joined two different components in $G_k(s)$. And $G_{k+1}(t) := G_k(t) \setminus F_k(t)$, consisting of the edges rejected from $F_k(t)$.

Hence, the kth spanning tree T_k produced by Kruskal's algorithm equals $F_k(\infty)$, as asserted above.

Note that $F_k(t)$ is a spanning subforest of $G_k(t)$, in other words, the components of $F_k(t)$ (regarded as vertex sets) are the same as the components of $G_k(t)$; this will be used frequently below. Moreover, each edge in $G_{k+1}(t)$ has endpoints in the same component of $G_k(t)$; hence, each component of $G_{k+1}(t)$ is a subset of a component of $G_k(t)$. It follows that an edge arriving to $G_1(t)$ will be passed through $G_2(t), \dots, G_k(t)$ and to $G_{k+1}(t)$ (and possibly further) if and only if its endpoints belong to the same component of $G_k(t)$, and thus if and only if its endpoints belong to the same component of $F_k(t)$.

2.3 More notation

We say that a component $\mathcal{C}$ of a graph G is the *unique giant* of G if $|\mathcal{C}| > |\mathcal{C}'|$ for every other component $\mathcal{C}'$; if there is no such component (i.e., if the maximum size is tied), then we define the unique giant to be $\emptyset$.

We say that a component $\mathcal{C}$ of $F_k(t)$ is the *permanent giant* of $F_k(t)$ (or of $G_k(t)$) if it is the unique giant of $F_k(t)$ and, furthermore, it is a subset of the unique giant of $F_k(u)$ for every $u > t$; if there is no such component then the permanent giant is defined to be $\emptyset$.

Let $\mathfrak{C}_k(t)$ denote the permanent giant of $F_k(t)$. Note that the permanent giant either is empty or the largest component; thus $|\mathfrak{C}_k(t)|$ is either 0 or $C_1(F_k(t)) = C_1(G_k(t))$. Note also that the permanent giant $\mathfrak{C}_k(t)$ is an increasing function of t: $\mathfrak{C}_k(t) \subseteq \mathfrak{C}_k(u)$ if $t \leqslant u$. Furthermore, for sufficiently large t (viz. t such that $G_k(t)$ is connected, and thus $F_k(t)$ is the spanning tree T_k), $\mathfrak{C}_k(t) = \mathfrak{C}_k(\infty) = [n]$.

2.4 A structure theorem

The basis of our proof of Theorems 1 and 2 is the following theorem on the structure of the components of $G_k(t)$. Recall that $F_k(t)$ has the same components as $G_k(t)$, so the theorem applies as well to $F_k(t)$.

For $k = 1$, the theorem collects various known results for $G(n, p)$. Our proof includes this case too, making the proof more self-contained.

▶ **Theorem 9.** *With the definitions above, the following hold for every fixed $k \geqslant 1$ as $n \to \infty$.*

(i) *There exists a continuous increasing function $\rho_k : [0, \infty) \to [0, 1)$ such that*

$$C_1(G_k(t))/n \xrightarrow{\mathrm{p}} \rho_k(t), \tag{2.1}$$

uniformly in $t \in [0, \infty)$; in other words, for any $\varepsilon > 0$, w.h.p., for all $t \geqslant 0$,

$$\rho_k(t) - \varepsilon \leqslant C_1(G_k(t))/n \leqslant \rho_k(t) + \varepsilon. \tag{2.2}$$

(ii) $\sup_{t \geqslant 0} C_2(G_k(t))/n \xrightarrow{\mathrm{p}} 0$.

(iii) *There exists a threshold $\sigma_k > 0$ such that $\rho_k(t) = 0$ for $t \leqslant \sigma_k$, but $\rho_k(t) > 0$ for $t > \sigma_k$. Furthermore, ρ_k is strictly increasing on $[\sigma_k, \infty)$.*

(iv) *There exist constants $b_k, B_k > 0$ such that*

$$\rho_k(t) \geqslant 1 - B_k e^{-b_k t}, \qquad t \geqslant 0. \tag{2.3}$$

In particular, $\rho_k(t) \to 1$ as $t \to \infty$.

(v) *If $t > \sigma_k$, then w.h.p. $G_k(t)$ has a non-empty permanent giant. Hence, for every $t \geqslant 0$,*

$$|\mathfrak{C}_k(t)|/n \xrightarrow{\mathrm{p}} \rho_k(t). \tag{2.4}$$

We note also a formula for the number of edges in $G_k(t)$, and two simple inequalities relating different k.

▶ **Theorem 10.** *For each fixed $k \geqslant 1$ and uniformly for t in any finite interval $[0, T]$,*

$$e(G_k(t))/n \xrightarrow{\mathrm{p}} \frac{1}{2} \int_0^t \rho_{k-1}(s)^2 \, \mathrm{d}s. \tag{2.5}$$

▶ **Theorem 11.** *$\rho_k(t) \leqslant \rho_{k-1}(t)$ for every $t \geqslant 0$, with strict inequality when $\rho_{k-1}(t) > 0$ (equivalently, when $t > \sigma_{k-1}$). Furthermore,*

$$\sigma_k \geqslant \sigma_{k-1} + 1. \tag{2.6}$$

Inequality (2.6) is weak in that we conjecture that as $k \to \infty$, $\sigma_k = \sigma_{k-1} + 2 + o(1)$.

2.5 The proof approach

Proofs of the results in this section are by induction on k, relying heavily on the theory of inhomogeneous random graphs by Bollobás, Janson and Riordan in [3]. When an edge is passed on by G_k this is almost always because it is contained in $\mathcal{C}_1(G_k)$; it is only rarely because it is contained in some other component, and this case is treatable as a perturbation within the theory. Thus, vertices "appear" in $G_{k+1}(t)$ as governed by $\rho_k(t)$; this is formalized as a "vertex space" in the theory. Once two vertices u and v are both present in $G_{k+1}(t)$, edges between them arrive at rate $1/n$. So, if they arrive at times τ_u and τ_v, the probability they are connected at time t is asymptotically $\frac{1}{n}(t - (\tau_u \vee \tau_v))_+ =: \frac{1}{n}\kappa_t(\tau_u, \tau_v)$; κ_t is the "kernel" in the inhomogeneous random graph framework. The framework then shows that $C_1(G_{k+1}(t))/n$ converges in probability to a certain $\rho(\kappa_t)$, the survival probability of a related inhomogeneous branching process, and this $\rho(\kappa_t)$ is precisely the desired next function $\rho_{k+1}(t)$.

3 Bounds on the expected cost

3.1 Total cost of the first k trees

The following theorem gives lower and upper bounds on the total cost of the first k spanning trees.

▶ **Theorem 12.** *Letting $W_k = \sum_{i=1}^{k} w(T_i)$ be the total cost of the first k spanning trees, for every $k \geqslant 1$,*

$$k^2 \frac{n-1}{n} \leqslant \mathbb{E}\, W_k \leqslant k(k+1)\frac{n-1}{n} < k^2 + k. \tag{3.1}$$

Comparing with Frieze and Johansson [6, Section 3], our upper bound is smaller than their $k^2 + 3k^{5/3}$ despite the fact that they considered a more relaxed minimization problem (see Section 4); as such ours is a strict improvement. In both cases the lower bound is simply the expected total cost of the cheapest $k(n-1)$ edges in G, with (3.2) matching [6, (3.1)].

Proof. The minimum possible cost of the k spanning trees is the cost of the cheapest $k(n-1)$ edges. Since each edge's costs (plural, in our model) are given by a Poisson process of rate 1, the set of all edge costs is given by a Poisson process of rate $\binom{n}{2}$. Recall that in a Poisson process of rate λ, the interarrival times are independent exponential random variables with mean $1/\lambda$, so that the ith arrival, at time Z_i, has $\mathbb{E}\, Z_i = i/\lambda$. It follows in this case that $W_k \geqslant \sum_{i=1}^{k(n-1)} Z_i$ and

$$\mathbb{E}\, W_k \geqslant \sum_{i=1}^{k(n-1)} \frac{i}{\binom{n}{2}} = \frac{(k(n-1))(k(n-1)+1)}{n(n-1)} \geqslant k^2 \frac{n-1}{n}. \tag{3.2}$$

We now prove the upper bound. An arriving edge is rejected from F_i iff both endpoints lie within its "forbidden" set B_i of edges, namely those edges with both endpoints in one component. The nesting property of the components means that $B_1 \supseteq B_2 \supseteq \cdots$. An arriving edge e joins F_k if it is rejected from all previous forests, i.e., $e \in B_{k-1}$ (in which case by the nesting property, e also belongs to all earlier Bs) but can be accepted into F_k, i.e., $e \notin B_k$. The idea of the proof is to show that the first k forests fill reasonably quickly with $n-1$ edges each, and we will do this by coupling the forest-creation process (Kruskal's algorithm) to a simpler, easily analyzable random process.

Let $\mathbf{s}(\tau) = \{s_k(\tau)\}_{k=0}^{\infty}$ denote the vector of the sizes (number of edges) of each forest after arrival of the τ'th edge; we may drop the argument τ when convenient. Let $p_k = |B_k|/\binom{n}{2}$, the rejection probability for F_k. For any τ, by the nesting property of the components and in turn of the B_k,

$$s_1 \geqslant s_2 \geqslant \cdots \qquad \text{and} \qquad p_1 \geqslant p_2 \geqslant \cdots. \tag{3.3}$$

The MST process can be simulated by using a sequence of i.i.d. random variables $\alpha(\tau) \sim U(0,1)$, incrementing $s_k(\tau)$ if both $\alpha(\tau) \leqslant p_{k-1}(\tau)$ (so that e is rejected from F_{k-1} and thus from all previous forests too) and $\alpha(\tau) > p_k(\tau)$ (so that e is accepted into F_k). We take the convention that $p_0(\tau) = 1$ for all τ. For intuition, note that when $s_k = 0$ an edge is never rejected from F_k ($p_k = 0$, so $\alpha \sim U(0,1)$ is never smaller); when $s_k = 1$ it is rejected with probability $p_k = 1/\binom{n}{2}$; and when $s_k = n-1$ it is always rejected ($|B_k|$ must be $\binom{n}{2}$, so $p_k = 1$).

Given the size $s_k = \sum_{i=1}^{\infty}(C_i(F_k)-1)$ of the kth forest, $|B_k| = \sum_{i=1}^{\infty}\binom{C_i(F_k)}{2}$ is maximized (thus so is p_k) when all the edges are in one component, i.e.,

$$p_k \leqslant \binom{s_k+1}{2} \Big/ \binom{n}{2} \tag{3.4}$$

$$\leqslant \frac{s_k}{n-1} =: \bar{p}_k. \tag{3.5}$$

The size vector $\mathbf{s}(\tau)$ thus determines the values $\bar{p}_k(\tau)$ for all k.

Let $\mathbf{r}(\tau)$ denote a vector analogous to $\mathbf{s}(\tau)$, but with $r_k(\tau)$ incremented if $\hat{p}_k(\tau) < \alpha(\tau) \leqslant \hat{p}_{k-1}(\tau)$, with

$$\hat{p}_k := \frac{r_k}{n-1}. \tag{3.6}$$

By construction,

$$r_1 \geqslant r_2 \geqslant \cdots \qquad \text{and} \qquad \hat{p}_1 \geqslant \hat{p}_2 \geqslant \cdots. \tag{3.7}$$

For intuition, here note that when $r_k = 0$ an arrival is never rejected from r_k ($\bar{p}_k = 0$); when $s_k = 1$ it is rejected with probability $\bar{p}_k = 1/(n-1) > p_k = 1/\binom{n}{2}$; and when $s_k = n-1$ it is always rejected ($\bar{p}_k = 1$).

Taking each $F_i(0)$ to be an empty forest (n isolated vertices, no edges) and accordingly $\mathbf{s}(0)$ to be an infinite-dimensional 0 vector, and taking $\mathbf{r}(0)$ to be the same 0 vector, we claim that for all τ, $\mathbf{s}(\tau)$ majorizes $\mathbf{r}(\tau)$, which we will write as $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$. That is, the prefix sums of $\mathbf{s}$ dominate those of $\mathbf{r}$: for all τ and k, $\sum_{i=1}^{k} s_i(\tau) \geqslant \sum_{i=1}^{k} r_i(\tau)$.

We first prove this; then use it to argue that edge arrivals to the first k forests, i.e., to $\mathbf{s}$, can only precede arrivals to the first k elements of $\mathbf{r}$; and finally analyze the arrival times of all $k(n-1)$ elements to the latter to arrive at an upper bound on the total cost of the first k trees.

We prove $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$ by induction on τ, the base case with $\tau = 0$ being trivial. Figure 1 may be helpful in illustrating the structure of this inductive proof. Suppose the claim holds for τ. The probabilities $p_k(\tau)$ are used to determine the forests $F_k(\tau+1)$ and in turn the size vector $\mathbf{s}(\tau+1)$. Consider an intermediate object $\mathbf{s}'(\tau+1)$, the size vector that would be given by incrementing $\mathbf{s}(\tau)$ using the upper-bound values $\bar{p}_k(\tau)$ taken from $\mathbf{s}(\tau)$ by (3.5). Then, $s_i(\tau+1)$ receives the increment if $p_{i-1} \geqslant \alpha > p_i$, and $s'_j(\tau+1)$ receives the increment if $\bar{p}_{j-1} \geqslant \alpha > \bar{p}_j$; hence, from $\bar{p}_{i-1} \geqslant p_{i-1} \geqslant \alpha$ it is immediate that $i \leqslant j$ and thus $\mathbf{s}(\tau+1) \succeq \mathbf{s}'(\tau+1)$.

It suffices then to show that $\mathbf{s}'(\tau+1) \succeq \mathbf{r}(\tau+1)$. These two vectors are obtained respectively from $\mathbf{s}(\tau)$ and $\mathbf{r}(\tau)$, with $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$ by the inductive hypothesis, using probability thresholds $\bar{p}_k(\tau) = f(s_k(\tau))$ and $\hat{p}_k(\tau) = f(r_k(\tau))$ respectively, applied to the common random variable α, where $f(s) = s/(n-1)$ (but all that is important is that f is a monotone function of s). Suppose that

$$f(s_{i-1}) \geqslant \alpha > f(s_i) \qquad \text{and} \qquad f(r_{j-1}) \geqslant \alpha > f(r_j), \tag{3.8}$$

so that elements i in $\mathbf{s}$ and j in $\mathbf{r}$ are incremented. If $i \leqslant j$, we are done. (Prefix sums of $\mathbf{s}(\tau)$ dominated those of $\mathbf{r}(\tau)$, and an earlier element is incremented in $\mathbf{s}'(\tau+1)$ than $\mathbf{r}(\tau+1)$, thus prefix sums of $\mathbf{s}'(\tau+1)$ dominate those of $\mathbf{r}(\tau+1)$.) Consider then the case that $i > j$.

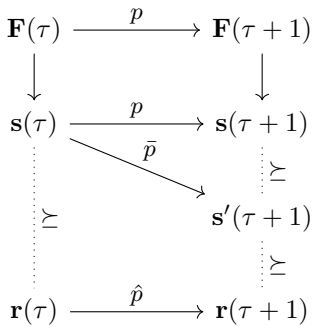

Figure 1 Coupling of the forests' sizes $\mathbf{s}(\tau)$ to a simply analyzable random process $\mathbf{r}(\tau)$, showing the structure of the inductive proof (on τ) that $\mathbf{s}(\tau)$ majorizes $\mathbf{r}(\tau)$.

In both processes the increment falls between indices j and i, so the k-prefix sum inequality continues to hold for $k < j$ and $k \geqslant i$. Thus, for $j \leqslant k < i$,

$$\begin{aligned} \sum_{\ell=1}^{k} s'_\ell(\tau+1) &= \sum_{\ell=1}^{j-1} s_\ell(\tau) + \sum_{\ell=j}^{k} s_\ell(\tau) \\ \sum_{\ell=1}^{k} r_\ell(\tau+1) &= \sum_{\ell=1}^{j-1} r_\ell(\tau) + 1 + \sum_{\ell=j}^{k} r_\ell(\tau). \end{aligned} \tag{3.9}$$

From $j < i$, (3.8), and (3.3) and (3.7) we have that when $j \leqslant \ell \leqslant i-1$,

$$s_\ell \geqslant s_{i-1} \geqslant f^{-1}(\alpha) > r_j \geqslant r_\ell,$$

implying

$$s_\ell \geqslant r_\ell + 1. \tag{3.10}$$

In (3.9), we have $\sum_{\ell=1}^{i-1} s_\ell(\tau) \geqslant \sum_{\ell=1}^{i-1} r_\ell(\tau)$ from the inductive hypothesis that $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$, while using (3.10) gives

$$\sum_{\ell=j}^{k} s_\ell(\tau) \geqslant \sum_{\ell=j}^{k} (1 + r_\ell(\tau)) \geqslant 1 + \sum_{\ell=j}^{k} r_\ell(\tau),$$

from which it follows that $\mathbf{s}'(\tau+1) \succeq \mathbf{r}(\tau+1)$, completing the inductive proof that $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$.

Having shown that the vector $\mathbf{s}(\tau)$ of component sizes majorizes $\mathbf{r}(\tau)$, it suffices to analyze the latter. Until this point we could have used (3.4) rather than (3.5) to define $\bar{p}_k$, $\hat{p}_k$, and the function f, but now we take advantage of the particularly simple nature of the process governing $\mathbf{r}(\tau)$. Recall that a new edge increments r_i for the first i for which the $U(0,1)$ "coin toss" $\alpha(\tau)$ has $\alpha(\tau) > \hat{p}_i \stackrel{\text{def}}{=} r_i/(n-1)$. Equivalently, consider an array of cells $n-1$ rows high and infinitely many columns wide, generate an "arrival" at a random row or "height" $X(\tau)$ uniform on $1, \ldots, n-1$, and let this arrival occupy the first unoccupied cell i at this height, thus incrementing the occupancy r_i of column i. This is equivalent because if r_i of the $n-1$ cells in column i are occupied, the chance that i is rejected – that $X(\tau)$ falls into this set and thus the arrival moves along to test the next column $i+1$ – is $r_i/(n-1)$, matching (3.6).

Recalling that the cost of an edge arriving at time t is t/n in the original graph problem, the combined cost W_k of the first k spanning trees is $1/n$ times the sum of the arrival times of their $k(n-1)$ edges. The majorization $\sum_{i=1}^{k} s_i(\tau) \geqslant \sum_{i=1}^{k} r_i(\tau)$ means that the ℓ'th arrival

to the first k forests comes no later than the ℓ'th arrival to the first k columns of the cell array. Thus, the cost W_k of the first k trees is at most $1/n$ times the sum of the times of the $k(n-1)$ arrivals to the array's first k columns.

The continuous-time edge arrivals are a Poisson process with intensity $1/n$ on each of the $\binom{n}{2}$ edges, thus intensity $(n-1)/2$ in all; it is at the Poisson arrival times that the discrete time τ is incremented and $X(\tau)$ is generated. Subdivide the "X" process into the $n-1$ possible values that X may take on, so that arrivals at each value (row in the cell array) are a Poisson process of intensity $\lambda = \frac{1}{2}$. The sum of the first k arrival times in a row is the sum of the first k arrival times in its Poisson process. The ith such arrival time is the sum of i exponential random variables, and has expectation i/λ. The expected sum of k arrival times of a line is thus $\binom{k+1}{2}/\lambda = k(k+1)$, and (remembering that cost is time divided by n), the expected total cost of all $n-1$ lines is

$$\frac{n-1}{n}k(k+1),$$

yielding the upper bound in (3.1) and completing the proof of the theorem. ◀

▶ **Corollary 13.** *Let* $\Gamma_k := \sum_{i=1}^{k} \gamma_i$. *Then, for every* $k \geqslant 1$,

$$k^2 \leqslant \Gamma_k = \sum_{i=1}^{k} \gamma_i \leqslant k^2 + k. \tag{3.11}$$

Proof. Immediate from Theorems 12 and 3. ◀

▶ **Example 14.** In particular, Corollary 13 gives $1 \leqslant \gamma_1 \leqslant 2$ and $4 \leqslant \gamma_1 + \gamma_2 \leqslant 6$. In fact, we know that $\gamma_1 = \zeta(3) \doteq 1.2021$ [7] and $\gamma_1 + \gamma_2 > 4.1704$ by [6] and Section 4, see Corollary 20. Numerical estimates suggest $\gamma_1 + \gamma_2 \doteq 4.30$; see Section 5, including Table 1, for various estimates of γ_2. ◀

3.2 Corollaries and conjectures for the kth tree

Turning to individual γ_k instead of their sum Γ_k, we obtain Corollary 7, namely that $2k - 2k^{1/2} < \gamma_k < 2k + 2k^{1/2}$.

Proof of Corollary 7. For the upper bound, we note that obviously $\gamma_1 \leqslant \gamma_2 \leqslant \dots$, and thus, for any $\ell \geqslant 1$, using both the upper and lower bound in (3.11),

$$\begin{aligned}\ell\,\gamma_k &\leqslant \sum_{i=k}^{k+\ell-1} \gamma_i = \Gamma_{k+\ell-1} - \Gamma_{k-1} \leqslant (k+\ell-1)(k+\ell) - (k-1)^2 \\ &= \ell^2 + \ell(2k-1) + k - 1 \end{aligned} \tag{3.12}$$

and hence

$$\gamma_k \leqslant 2k - 1 + \ell + \frac{k-1}{\ell}. \tag{3.13}$$

Choosing $\ell = \lceil \sqrt{k} \rceil$ gives the upper bound in (1.1).

For the lower bound we similarly have, for $1 \leqslant \ell \leqslant k$,

$$\ell\gamma_k \geqslant \Gamma_k - \Gamma_{k-\ell} \geqslant k^2 - (k-\ell)(k-\ell+1) = -\ell^2 - (2k+1)\ell - k \tag{3.14}$$

and hence

$$\gamma_k \geqslant 2k + 1 - \ell - \frac{k}{\ell}. \tag{3.15}$$

Choosing, again, $\ell = \lceil \sqrt{k} \rceil$ gives the lower bound in (1.1). ◀

▶ Remark 15. For a specific k, we can improve (1.1) somewhat by instead using (3.13) and (3.15) with $\ell = \lfloor\sqrt{k}\rfloor$ or $\ell = \lceil\sqrt{k}\rceil$. For example, for $k = 2$, taking $\ell = 1$ yields $2 \leqslant \gamma_2 \leqslant 5$. For $k = 3$, taking $\ell = 2$ yields $3.5 \leqslant \gamma_3 \leqslant 8$. ◀

Besides these rigorous results, taking increments of the left and right-hand sides of (3.11) also suggests the following conjecture.

▶ **Conjecture 16.** *For $k \geqslant 1$, $2k - 1 \leqslant \gamma_k \leqslant 2k$.*

▶ Remark 17. Moreover, if $\gamma_k = 2k + \delta + o(1)$ holds, as conjectured in Conjecture 5, then $\Gamma_k = k^2 + k(\delta + 1) + o(k)$, and thus necessarily $\delta \in [-1, 0]$ as a consequence of Corollary 13. In fact, the numerical estimates described in Section 5, suggest that $\delta = -1$; see Conjecture 6. ◀

3.3 Improved upper bounds

The upper bounds in Theorem 12 and Corollary 13 were proved using the bound (3.5). A stronger, but less explicit, bound can be proved by using instead the sharper (3.4). That is, we consider the random vectors $\mathbf{r}(\tau)$ defined as above but with (3.6) replaced by

$$\hat{p}_k := \binom{r_k + 1}{2} \Big/ \binom{n}{2}. \tag{3.16}$$

As remarked before (3.4), this approximation comes from imagining all edges in each F_k to be in a single component; this overestimates the probability that an arriving edge is rejected from F_k and, as developed in the previous subsection, gives $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$ just as when $\hat{p}_k$ was defined by (3.5).

Using for consistency our usual time scaling in which edges arrive at rate $(n-1)/2$, by a standard martingale argument one can show that, for each $k \geqslant 1$,

$$\frac{1}{n} r_k(\lfloor \tfrac{1}{2} n t \rfloor) \xrightarrow{\text{p}} g_k(t), \qquad \text{uniformly for } t \geqslant 0, \tag{3.17}$$

for some continuously differentiable functions $g_k(t)$ satisfying the differential equations, with $g_0(t) := 1$,

$$g_k'(t) = \tfrac{1}{2}\left(g_{k-1}(t)^2 - g_k(t)^2\right), \qquad g_k(0) = 0, \qquad k \geqslant 1. \tag{3.18}$$

Moreover, using $\mathbf{s}(\tau) \succeq \mathbf{r}(\tau)$ and taking limits, it can be shown that

$$\Gamma_k := \sum_{i=1}^{k} \gamma_i \leqslant \frac{1}{2} \int_0^\infty t\big(1 - g_k(t)^2\big)\, \mathrm{d}t. \tag{3.19}$$

We omit the details, but roughly, in time dt, $\frac{1}{2} n\, \mathrm{d}t$ edges arrive, all costing about t/n, and a $g_k(t)^2$ fraction of them pass beyond the first k graphs (to the degree that we are now modeling graphs).

For $k = 1$, (3.18) has the solution $g_1(t) = \tanh(t/2)$, and (3.19) yields the bound $\Gamma_1 = \gamma_1 \leqslant 2 \ln 2 \doteq 1.3863$. This is better than the bound 2 given by (3.11), but still far from precise since $\gamma_1 = \zeta(3) \doteq 1.2021$.

For $k \geqslant 2$ we do not know any exact solution to (3.18), but numerical solution of (3.18) and calculation of (3.19) (see Section 5) suggests that $\Gamma_k < k^2 + 1$. We leave the proof of this as an open problem. If proved, this would be a marked improvement on $\Gamma_k \leqslant k^2 + k$, which was the exact expectation of the random process given by (3.5) (that part of the analysis was tight). In particular, it would establish that $2k - 2 \leqslant \gamma_k \leqslant 2k$.

For $k = 2$, the numerical calculations in Section 5 give $\gamma_1 + \gamma_2 \leqslant 4.5542\ldots$ and thus $\gamma_2 \leqslant 3.3521\ldots$. The same value was also obtained using Maple's numerical differential equation solver, with Maple giving greater precision but the two methods agreeing in the digits shown here.

4 A related problem by Frieze and Johansson

As said in the introduction, Frieze and Johansson [6] recently considered the problem of finding the minimum total cost of k edge-disjoint spanning trees in K_n, for a fixed integer $k \geqslant 2$. (They used random costs with the uniform model; we may consider all three models described in Section 1.1.) We denote this minimum cost by mst_k, following [6]. Trivially,

$$\mathrm{mst}_k \leqslant \sum_{i=1}^{k} w(T_i), \tag{4.1}$$

and as said in the introduction, it is easy to see that strict inequality may hold when $k \geqslant 2$, i.e., that our greedy procedure of choosing $T_1, T_2, \ldots$ successively does not yield the minimum cost set of k disjoint spanning trees.

We assume in this section that $n \geqslant 2k$; then k edge-disjoint spanning trees exist and thus $\mathrm{mst}_k < \infty$.

▶ Remark 18. As observed by Frieze and Johansson [6], the problem is equivalent to finding the minimum cost of a basis in the matroid $\mathcal{M}_k$, defined as the union matroid of k copies of the cycle matroid of K_n. This means that the elements of $\mathcal{M}_k$ are the edges in K_n, and a set of edges is independent in $\mathcal{M}_k$ if and only if it can be written as the union of k forests, see e.g. [20, Chapter 8.3]. (Hence, the bases, i.e., the maximal independent sets, are precisely the unions of k edge-disjoint spanning trees. For the multigraph version in the Poisson model, of course we use instead the union matroid of k copies of the cycle matroid of K_n^∞; we use the same notation $\mathcal{M}_k$.) We write r_k for rank in this matroid. ◀

For $k = 2$, Frieze and Johansson [6] show that

$$\mathbb{E}\,\mathrm{mst}_2 \to \mu_2 \doteq 4.1704. \tag{4.2}$$

This is strictly smaller than our numerical estimates from Table 1 for the total cost of two edge-disjoint spanning trees chosen successively, $\gamma_1 + \gamma_2 \doteq 1.20 + 3.09 > 4.29$; we show this calculation to only two digits as we are confident of this level of precision. This would show that choosing minimum spanning trees one by one is not optimal, even asymptotically, except that our estimates are not rigorous. The following theorem is less precise but establishes rigorously that the values are indeed different. (We rely only on $\sigma_2 < \mu_2$, coming from the estimate of μ_2 above, and our estimate $\sigma_2 \doteq 2.69521$, obtained as the numerical solution to a differential equation; see the full version for details.)

▶ **Theorem 19.** *There exists $\delta > 0$ such that, for any of the three models, w.h.p. $w(T_1) + w(T_2) \geqslant \mathrm{mst}_2 + \delta$.*

This can be restated in the following equivalent form.

▶ **Corollary 20.** $\gamma_1 + \gamma_2 > \mu_2$.

Proof. The equivalence of the statements in Theorem 19 and Corollary 20 is immediate since $w(T_1) \xrightarrow{\mathrm{p}} \gamma_1$ and $w(T_2) \xrightarrow{\mathrm{p}} \gamma_2$ by Theorem 1 or 2 (depending on the choice of model), and $\mathrm{mst}_2 \xrightarrow{\mathrm{p}} \mu_2$ by [6] and justification that this holds in all three models (see the full version). ◀

▶ Remark 21. Numerically, $\gamma_2 > 2.9683$. This is immediate from Corollary 20, the value of μ_2 given by [6], and (by [7]) $\gamma_1 = \zeta(3)$. ◀

The proof of Theorem 19 is based on the fact that many edges are rejected from T_1 and T_2 after time σ_2, but none is rejected from the union matroid until a later time c_3, namely the threshold for appearance of a 3-core in a random graph.

5 Computational results

A variety of computations were performed, all of which will be mentioned here but only one presented in any detail; for the rest see [12].

1. We performed naive simulations, generating edge-weighted random graphs and finding the successive trees.
2. We performed a similar simulation, but instead of introducing edges in order of increasing weight, we simply generate random edges. The details are below.
3. We solved the differential equations (3.18) numerically up to $k = 50$, to get upper bounds on Γ_k as in (3.19). The results suggest that $\Gamma_k < k^2 + 1$ (perhaps $\Gamma_k < k^2 + 0.743$). If proved, this would be a marked improvement on $\Gamma_k \leqslant k^2 + k$, which was the exact expectation of the random process given by (3.5) (that part of the analysis was tight). In particular, it would establish that $2k - 2 \leqslant \gamma_k \leqslant 2k$.
4. Finally, the functions $\rho_k(t)$ can be obtained, recursively on k, through the solution to certain functional fixed-point equations. We solved these numerically, getting results consistent with those in the set of simulations listed as (2) above.

We now detail the set of simulations listed as (2) above, done with reference to the Poisson multigraph model introduced in Section 2.2 and used throughout. We begin with k empty graphs of order n. At each step we introduce a random edge e and, in the first graph G_i for which e does not lie within a component, we merge the two components given by its endpoints. (If this does not occur within the k graphs under consideration, we do nothing, just move on to the next edge.) For each graph we simulate only the components (i.e., the sets of vertices comprised by each component); there is no need for any more detailed structure. The edge arrivals should be regarded as occurring as a Poisson process of intensity $(n-1)/2$ but instead we simply treat them as arriving at times $2/n$, $4/n$, etc.

Figure 2 depicts the result of a single such simulation with $n = 1\,000\,000$, showing for each k from 1 to 5 the size of the largest component of G_k (as a fraction of n) against time. Similar experiments with multiple simulations and larger values of n support Conjecture 6 that $\gamma_k = 2k - 1 + o(1)$. The largest experiment's results are shown in part in Table 1; its support for the conjecture continues through $k = 29$, the last value for which it gives good data.

Table 1 Estimates of $\gamma_1, \ldots, \gamma_9$ from 10 simulations each with $n = 10\,000\,000$, through time $t = 40$.

10 simulations each with $n = 10\,000\,000$									
	γ_1	γ_2	γ_4	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9
mean	1.2020	3.0921	5.0482	7.0253	9.0169	11.0091	13.0067	15.0035	17.0039
std err	0.0002	0.0003	0.0005	0.0008	0.0010	0.0012	0.0016	0.0010	0.0015

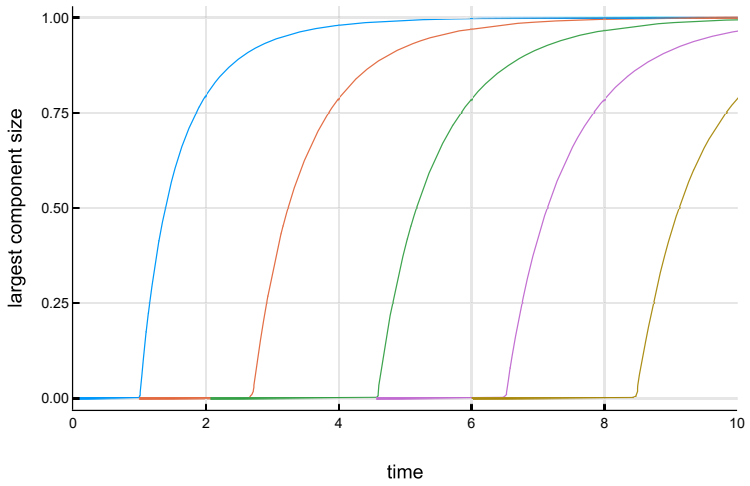

Figure 2 Largest component sizes, as a fraction of n, for graphs $G_1, \dots, G_5$, based on a single simulation with $n = 1\,000\,000$.

6 Open questions

We would be delighted to confirm the various conjectures above, in particular Conjectures 4–6, and to get a better understanding of (and ideally a closed form for) ρ_∞ (provided it exists).

It is also of natural interest to ask this kth-minimum question for structures other than spanning trees. Subsequent to this work, the length X_k of the kth shortest s–t path in a complete graph with random edge weights has been studied by Mezei, Gerke and Sorkin [8]. They show that $X_k/(2k/n + \ln n/n) \xrightarrow{\mathrm{p}} 1$ for all k from 1 to $n-1$. In particular, the first few paths all cost nearly identical amounts, quite different from the situation for successive MSTs.

The "random assignment problem" is to determine the cost of a minimum-cost perfect matching in a complete bipartite graph with random edge weights. A great deal is known about it, by a variety of methods; for one relatively recent work, with references to others, see Wästlund [19]. It would be interesting to understand the kth cheapest matching.

It could also be interesting to consider other variants of all these questions. Frieze and Johansson [6] considered the k disjoint structures which together have the smallest possible total cost, where we consider disjoint structures generated successively. In either case, instead of asking for disjoint structures, we could require structures which are merely distinct, or perhaps which differ in some adversarially specified elements.

References

1 Aaron Archer and Éva Tardos. Frugal path mechanisms. *ACM Transactions on Algorithms*, 3(1):3:1–3:22, 2007.

2 Béla Bollobás and Alan M. Frieze. On matchings and Hamiltonian cycles in random graphs. In Michał Karoński and Andrzej Ruciński, editors, *Random Graphs '83 (Poznań, 1983), Ann. Discrete Math.* ***28***, volume 118 of *North-Holland Mathematics Studies*, pages 23–46. North-Holland, 1985.

3 Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Alg.*, 31(1):3–122, 2007.

4 Prasad Chebolu, Alan Frieze, Páll Melsted, and Gregory B Sorkin. Average-case analyses of Vickrey costs. In *Proceedings of APPROX / RANDOM 2009*, volume 5687 of *Lecture Notes in Comput. Sci.*, pages 434–447. Springer, Berlin, Heidelberg, 2009.

5 Colin Cooper, Alan Frieze, Nate Ince, Svante Janson, and Joel Spencer. On the length of a random minimum spanning tree. *Combin. Probab. Comput.*, 25(1):89–107, 2016.

6 Alan Frieze and Tony Johansson. On edge disjoint spanning trees in a randomly weighted complete graph. *Combin. Probab. Comput.*, 27(2):228–244, 2018.

7 Alan M. Frieze. On the value of a random minimum spanning tree problem. *Discrete Applied Mathematics*, 10:47–65, 1985.

8 Stefanie Gerke, Balázs Mezei, and Gregory B. Sorkin. Successive shortest paths, 2019. Manuscript in preparation.

9 Svante Janson. The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. *Random Struct. Alg.*, 7:337–355, 1995.

10 Svante Janson, Donald E. Knuth, Tomasz Łuczak, and Boris Pittel. The birth of the giant component. *Random Struct. Alg.*, 4(3):231–358, 1993.

11 Svante Janson and Gregory B. Sorkin. VCG auction mechanism cost expectations and variances, 2013. `arXiv:1310.1777`.

12 Svante Janson and Gregory B Sorkin. Successive minimum spanning trees, 2019. `arXiv:1310.1777`.

13 Svante Janson and Johan Wästlund. Addendum to "The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph". *Random Struct. Alg.*, 28(4):511–512, 2006.

14 Anna R. Karlin, David Kempe, and Tami Tamir. Beyond VCG: Frugality of truthful mechanisms. In *Proceedings of 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 615–624. IEEE, 2005.

15 Joseph B. Kruskal, Jr. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Amer. Math. Soc.*, 7:48–50, 1956.

16 Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory.* Cambridge Univ. Press, New York, 2007.

17 Kunal Talwar. The price of truth: frugality in truthful mechanisms. In *Proceedings of STACS 2003, Lecture Notes in Comput. Sci.*, volume 2607, pages 608–619. Springer, 2003.

18 Johan Wästlund. Evaluation of Janson's constant for the variance in the random minimum spanning tree problem, 2005. *Linköping Studies in Mathematics* No. 7.

19 Johan Wästlund. An easy proof of the $\zeta(2)$ limit in the random assignment problem. *Elect. Comm. in Probab.*, 14:261—269, 2009.

20 Dominic J. A. Welsh. *Matroid theory.* Academic Press, New York, 1976. Reprinted, Dover, Mineola, New York, 2010.