Gaussian tail for empirical distributions of MST on random graphs

Sungchul Lee\textsuperscript{a,\ast,1}, Zhonggen Su\textsuperscript{a,b,2}

\textsuperscript{a}Department of Mathematics, Yonsei University, Seoul 120-749, South Korea
\textsuperscript{b}Department of Mathematics, Zhejiang University, Hangzhou 310028, P.R. China

Received November 2001; received in revised form March 2002

Abstract

Consider the complete graph $K_n$ on $n$ vertices and the $n$-cube graph $Q_n$ on $2^n$ vertices. Suppose independent uniform random edge weights are assigned to each edge in $K_n$ and $Q_n$ and let $\mathcal{T}(K_n)$ and $\mathcal{T}(Q_n)$ denote the unique minimal spanning trees on $K_n$ and $Q_n$, respectively. In this paper we obtain the Gaussian tail for the number of edges of $\mathcal{T}(K_n)$ and $\mathcal{T}(Q_n)$ with weight at most $t/n$.

MSC: primary 60D05; 60F05; secondary 60K35; 05C05; 90C27

Keywords: Empirical distribution; Gaussian tail; Minimal spanning tree

1. Introduction and statement of main results

If $G$ is a connected graph with a weight $w(e)$ assigned to each edge $e$, a spanning tree $\mathcal{F}$ of $G$ has associated total weight $L(\mathcal{F}) = \sum_{e \in \mathcal{F}} w(e)$. A minimal spanning tree (MST) of $G$ is a spanning tree $\mathcal{T}$ such that $L(\mathcal{T}) \leq L(\mathcal{F})$ for any spanning tree $\mathcal{F}$. Minimal spanning trees of certain graphs have been studied by a number of authors. Typically, one is interested in the total weight $L(\mathcal{T})$ of the MST $\mathcal{T}$, the vertex degrees of $\mathcal{T}$, or the empirical distribution of $L(\mathcal{T})$. Among the most studied graphs are the complete graph on $n$ vertices and the $n$-cube graph.

\textsuperscript{\ast}Corresponding author. Tel.: 0082-2365-7966; fax: 0082-2365-2579.
\textit{E-mail addresses:} sungchul@yonsei.ac.kr (S. Lee), zgsu@mail.hz.zj.cn (Z. Su).

\textsuperscript{1}This work was supported by the BK21 project of the Department of Mathematics, Yonsei University, the interdisciplinary research program of KOSEF 1999-2-103-001-5, and Com > 2MaC in POSTECH.

\textsuperscript{2}This work was supported by the BK21 project of the Department of Mathematics, Yonsei University, and by NSFC 10071072.

0167-7152/02/$-see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0167-7152(02)00144-X
This paper is mainly devoted to the study on the asymptotic behaviour of empirical distributions of the edge weights of the MST of the complete graph and the n-cube graph.

Let \(G_n\) be a connected graph with vertex set \(V_n = \{1, 2, \ldots, n\}\) and edge set \(E_n = \{e\}\). We assume that the degree of each vertex \(i \in V_n\) is \(d_n\) so that the number of edges in \(E_n\) is \(nd_n/2\). More specifically, we consider the following two connected graphs. Let \(K_n\) be the complete graph with vertex set \(V_n = \{1, 2, \ldots, n\}\) and edge set \(E_n = \{(i, j): 1 \leq i < j \leq n\}\), and let \(Q_n\) be the n-cube graph with vertex set \(\{0, 1\}^n\) and nearest neighbor edges, i.e., two vertices are taken to be adjacent if they differ in exactly one coordinate. The edge weights \(U(e)\) are defined for each edge \(e\) of \(G_n\), \(K_n\), and \(Q_n\), and taken to be independent uniform random variables on the unit interval \([0, 1]\).

Denote the MST of \(G_n\), \(K_n\), and \(Q_n\) by \(T(G_n)\), \(T(K_n)\), and \(T(Q_n)\), respectively. Since we assign different weights for different edges with probability 1, by Lemma 5 of Lee (1997) we see that \(T(G_n)\), \(T(K_n)\), and \(T(Q_n)\) are uniquely determined.

From now on, we assume that \(T(G_n)\), \(T(K_n)\), and \(T(Q_n)\) are unique.

The one remarkable result concerning the total weight of \(T(K_n)\) is due to Frieze (1985), who gave the exact asymptotic expected value of \(L(T(K_n))\), i.e., as \(n \to \infty\)

\[ EL(T(K_n)) \to \zeta(3). \]

In the case of the n-cube, Penrose (1998) found that the same limit is also valid for \((n/2^n)L(T(Q_n))\).

There are also interesting results regarding the vertex degrees in \(T(K_n)\) and \(T(Q_n)\). For these results see Aldous (1990), Penrose (1998), and Kim and Lee (2001).

Now, let's consider the empirical distributions. For \(0 < f(n, t) < 1\), we define

\[ N(G_n, t) = \sum_{e \in T(G_n)} 1(U(e) \leq f(n, t)). \]

For \(K_n\) and \(Q_n\) we consider a more specific form of the empirical distributions. For \(t > 0\), we define

\[ N(K_n, t) = \sum_{e \in T(K_n)} 1(U(e) \leq \frac{t}{n}), \quad N(Q_n, t) = \sum_{e \in T(Q_n)} 1(U(e) \leq \frac{t}{n}) \]

and

\[ F(K_n, t) = \frac{N(K_n, t)}{n - 1}, \quad F(Q_n, t) = \frac{N(Q_n, t)}{2^n - 1}. \]

As a consequence of a deep result on the asymptotic empirical distribution of edge lengths of Aldous (1990), one can easily see that

\[ F(K_n, t) \to \frac{1}{2} \int_0^t (1 - \bar{\psi}(s)) \, ds \quad \text{in probability}, \]

where \(\bar{\psi}(t)\) is the survival probability for a Galton–Watson branching process from a single ancestor with Poisson\((t)\) offspring. In the case of the n-cube, Penrose (1998) showed that the same limit is also valid for \(F(Q_n, t)\).

In this paper, as a first step toward the central limit theorems we obtain the Gaussian tail bound for \(N(K_n, t)\) and \(N(Q_n, t)\). The central limit theorems for \(N(K_n, t)\) and \(N(Q_n, t)\) are largely open. Our main results are as follows.
Theorem 1. For any strictly positive but finite constant $C_1$ such that $e^x \leq 1 + 1.5x$, $0 \leq x \leq C_1$, and for any $0 \leq \lambda \leq (2C_1)^{1/2}nd_nf(n,t)$
\[ P(|N(G_n,t) - EN(G_n,t)| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{4nd_nf(n,t)}\right). \]

Corollary 1. For any strictly positive but finite constant $C_1$ such that $e^x \leq 1 + 1.5x$, $0 \leq x \leq C_1$, and for any $0 \leq \varepsilon \leq (8C_1n)^{1/2}$
\[ P(|N(K_n,t) - EN(K_n,t)| \geq \varepsilon \sqrt{n}) \leq 2 \exp\left(-\frac{1}{4t} \varepsilon^2\right). \]

Corollary 2. For any strictly positive but finite constant $C_1$ such that $e^x \leq 1 + 1.5x$, $0 \leq x \leq C_1$, and for any $0 \leq \varepsilon \leq (C_12^{n+1})^{1/2}$
\[ P(|N(Q_n,t) - EN(Q_n,t)| \geq \varepsilon \sqrt{2^n}) \leq 2 \exp\left(-\frac{1}{4t} \varepsilon^2\right). \]

As a simple corollary to Corollaries 1 and 2, by the Borel–Cantelli lemma we have the following. We leave this as an easy exercise to the reader.

Corollary 3. For $t > 0$, as $n \to \infty$
\[ F(K_n,t) \to \frac{1}{2} \int_0^t (1 - \psi^2(s)) \, ds \quad \text{a.s.} \]
and
\[ F(Q_n,t) \to \frac{1}{2} \int_0^t (1 - \psi^2(s)) \, ds \quad \text{a.s.} \]

2. Gaussian tail

In this section, we introduce a martingale inequality. We also introduce an algorithm for the construction of a minimal spanning tree. Using these two tools we obtain Theorem 1. From Theorem 1 we obtain the Gaussian tail bound for $N(Q_n,t)$ and $N(K_n,t)$. So, here we mainly concentrate on Theorem 1.

Let’s start with a martingale inequality. Here is the idea behind the inequality. Recall that $K_n$ is the complete graph on $n$ vertices with vertex set $V_n$ and edge set $E_n$. Also recall that $\{U(e): e \in E_n\}$ are iid uniform on $[0, 1]$. For an easy presentation, we rename $\{U(e): e \in E_n\}$ as $\{U_k: 1 \leq k \leq m\}$ where $m = (\binom{n}{2})$. Let for $i = 1, \ldots, m,$
\[ N = N(K_n,t) = (n - 1)F(K_n,t) = \sum_{e \in \mathcal{E}(K_n)} 1 \left(U(e) \leq \frac{t}{n}\right) \]
and let

\[ d_i = E(N|U_1, \ldots, U_i) - E(N|U_1, \ldots, U_{i-1}) \]

\[ = \int \cdots \int (N(U_1, \ldots, U_i, U_{i+1}', \ldots, U_m') - N(U_1, \ldots, U_{i-1}, U_i', \ldots, U_m')) \, dU_i' \cdots dU_m' \]  \hspace{1cm} (2.1)

and

\[ \Delta_i = N(U_1, \ldots, U_i, U_{i+1}', \ldots, U_m') - N(U_1, \ldots, U_{i-1}, U_i', \ldots, U_m'), \]  \hspace{1cm} (2.2)

where \( \{U_k': 1 \leq k \leq m\} \) is an independent copy of \( \{U_k: 1 \leq k \leq m\} \). In our study \( \Delta_i \) has values 1, 0, and -1, and \( P(\Delta_i = 1) = P(\Delta_i = -1) \) is small. In this case with some extra technical conditions the following martingale inequality is quite powerful.

**Lemma 1.** For an arbitrary random variable \( X \in L^1(\Omega, \mathcal{F}, P) \) and for an arbitrary filtration \( \{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_m = \mathcal{F} \). We let \( d_i, 1 \leq i \leq m \), be the martingale difference, that is \( d_i = E(X|\mathcal{F}_i) - E(X|\mathcal{F}_{i-1}) \). Suppose that \( d_i = E(\Delta_i|\mathcal{F}_i) \) for some \( \Delta_i \in L^1(\Omega, \mathcal{F}, P) \). In addition, assume that for a fixed \( c > 0 \), \( \Delta_i \) takes only \( c, 0, \) and \(-c\), and that for a fixed \( p \) and for \( 1 \leq i \leq m \),

\[ P(\Delta_i = c|\mathcal{F}_{i-1}) \leq p \quad \text{a.s.} \]  \hspace{1cm} (2.3)

Then, for any strictly positive but finite constant \( C_1 \) such that \( e^x \leq 1 + 1.5x, \) \( 0 \leq x \leq C_1 \), and for any \( 0 \leq \lambda \leq (8C_1)^{1/2}mc\),

\[ P(|X - EX| \geq \lambda) \leq 2 \exp \left( -\frac{\lambda^2}{8mc^2}p \right). \]  \hspace{1cm} (2.4)

**Remark.** If one uses Azuma’s inequality in the situation considered in the above lemma, one has

\[ P(|X - EX| \geq \lambda) \leq 2 \exp \left( -\frac{\lambda^2}{2mc^2} \right). \]  \hspace{1cm} (2.5)

If \( p \) is small, then (2.4) beats (2.5) and actually this is the case in our study. For Azuma’s inequality and its application see Steele (1997) and Yukich (1998).

**Proof.** We just estimate the moment generating function \( Ee^{\lambda(X - EX)} \) and apply the Markov inequality to get (2.4). One can also get the similar inequality by Lemma 1.6 of Ledoux and Talagrand (1991) with \( b^2 = mpc^2 \). We leave these to the interested reader. \( \Box \)

Now we recall the add and delete algorithm for the construction of an MST which one can find in Kim and Lee (2001). Here is the situation. Let \( G = (V, E, w) \) be a connected weighted graph with \( |V| < \infty, |E| < \infty \) where \( w: E \to [0, \infty) \) is a weight function defined on the edge set \( E \). Let \( \mathcal{F} \) be an MST on \( G \). Now, suppose only one edge \( e \) changes its weight from \( w(e) \) to \( w'(e) \), that is suppose we are facing a new connected weighted graph \( G' = (V, E, w') \) where

\[ w'(e') = \begin{cases} w(e') & \text{for } e' \neq e, \\ w'(e) & \text{for } e' = e. \end{cases} \]
In order to construct an MST $\mathcal{T}'$ on $G'$ we use the following add and delete algorithm:

1. If $e \in \mathcal{T}$ and $w'(e) \leq w(e)$, we just let $\mathcal{T}' = \mathcal{T}$.
2. If $e \in \mathcal{T}$ and $w'(e) > w(e)$, we delete $e$ from $\mathcal{T}$. Then two connected components are generated. Choose $f$ such that $(\mathcal{T} \setminus \{e\}) \cup \{f\}$ is connected and
   \[ w'(f) = \min\{w'(e') : (\mathcal{T} \setminus \{e\}) \cup \{f\} \text{ is connected} \} . \]

Now, we let $\mathcal{T}' = (\mathcal{T} \setminus \{e\}) \cup \{f\}$.

3. If $e \notin \mathcal{T}$ and $w'(e) > w(e)$, we just let $\mathcal{T}' = \mathcal{T}$.
4. If $e \notin \mathcal{T}$ and $w'(e) < w(e)$, then add $e$ to $\mathcal{T}$, i.e., form $\mathcal{T} \cup \{e\}$. $\mathcal{T} \cup \{e\}$ contains a unique circuit $C$. Choose an edge $f \in C$ such that $w'(f) = \max\{w'(e') : e' \in C\}$. Delete $f$ from $\mathcal{T} \cup \{e\}$, i.e., we let $\mathcal{T}' = (\mathcal{T} \cup \{e\}) \setminus \{f\}$.

**Lemma 2.** The spanning tree $\mathcal{T}'$ is a minimal spanning tree of $G'$.

**Proof.** This is Lemma 2.2 of Kim and Lee (2001). □

We need the following property of an MST.

**Lemma 3.** Let $G = (V, E, w)$ be a connected weighted graph with $|V| < \infty$ and $|E| < \infty$. If there exists a path $\pi = (v_1, \ldots, v_n)$ in $G$ from $v_1 \in V$ to $v_2 \in V$ such that $w(e_i) \leq v$, $1 \leq i \leq n$, then in any MST $\mathcal{T}$ on $G$ there exists a path $\pi' = (v_1, \ldots, v_m)$ in $\mathcal{T}$ from $v_1$ to $v_2$ such that $w(e_j') \leq v$, $1 \leq j \leq m$.

**Proof.** This is Lemma 2 of Kesten and Lee (1996). □

Now we are ready for the main argument of this paper.

**Proof of Theorem 1.** Let’s recall $d_i$ and $\Delta_i$ from (2.1) and (2.2), that is $d_i$ is some kind of integration of $\Delta_i$ where

\[ \Delta_i = N(U_1, \ldots, U_i, U_{i+1}', \ldots, U_m') - N(U_1, \ldots, U_{i-1}, U_i', \ldots, U_m') . \]

Let $\mathcal{T}'$ and $\mathcal{T}$ be MST under the new configuration $(U_1, \ldots, U_i, U_{i+1}', \ldots, U_m')$ and the old configuration $(U_1, \ldots, U_{i-1}, U_i', \ldots, U_m')$, respectively. We also let $e_j$ be the edge where the weight $U_j$ or $U_j'$ is attached to. We claim that $\Delta_i$ has values 1, 0, and $-1$, and that $P(\Delta_i = 1) = P(\Delta_i = -1)$ is small. Furthermore, it satisfies technical conditions in Lemma 1 so that in our study we can use the martingale inequality in Lemma 1.

First, one can easily see that by Lemma 2, $\Delta_i$ takes only 1, 0, and $-1$. Suppose that $U_j' > f(n, t)$ and $U_i > U_i'$. There are two cases we have to consider; the first case is that $e_i \notin \mathcal{T}$ and the second is that $e_i \notin \mathcal{T}$. If $e_i \notin \mathcal{T}$, then by Lemma 2 we first delete $e_i$ from $\mathcal{T}$. Then, we add the smallest edge $f$ which makes $\mathcal{T} \setminus \{e_i\} \cup \{f\}$ connected. Under the new configuration $(U_1, \ldots, U_i, U_{i+1}', \ldots, U_m')$ the weight attached to the edge $f$ is either $U(f')$ or $U'(f)$. For simplicity we call this weight $U$. By Lemma 3 the weight $U$ is greater than $U_i'$. Therefore, under the new configuration $1(U \leq f(n, t)) = 0$. Since under the old configuration, $1(U_i' \leq f(n, t)) = 0$, we have $\Delta_i = 0$. If $e_i \notin \mathcal{T}$, then by Lemma 2,
$\mathcal{T}' = \mathcal{T}$ and of course $A_i = 0$. By the similar arguments we see that $A_i = 1$ only when $U_i' > f(n,t)$ and $U_i \leq f(n,t)$. Therefore with the filtration $\mathcal{F}_{i-1} = \sigma(U_1, \ldots, U_{i-1})$

$$P(A_i = 1 | \mathcal{F}_{i-1}) \leq P(U_i \leq f(n,t)) = f(n,t) \quad \text{a.s.}$$

Now, we apply Lemma 1 for $m = nd_n/2$, $c = 1$, $p = f(n,t)$, to obtain Theorem 1.

**Proof of Corollaries 1 and 2.** We apply Theorem 1 for $\lambda = \varepsilon \sqrt{n}$, $nd_n/2 = \binom{n}{2}$, $f(n,t) = t/n$, to obtain Corollary 1. To obtain Corollary 2, we apply Theorem 1 for $\lambda = \varepsilon \sqrt{2^n}$, $nd_n/2 = 2^{n-1}n$, $f(n,t) = t/n$.

**Acknowledgements**

The authors would like to thank David Aldous for helpful discussions regarding the branching process argument of Aldous (1990) and the referee for his thoughtful comments on an earlier version.

**References**