Limit laws for the optimal directed tree with random costs

JENNIE C. HANSEN
Actuarial Mathematics and Statistics Department,
Heriot-Watt University,
Edinburgh, Scotland

July, 1996

Abstract

Suppose that $\mathcal{C} = \{c_{ij} : i, j \geq 1\}$ is a collection of i.i.d. nonnegative continuous random variables and suppose T is a rooted, directed tree on vertices labelled 1,2,...,n. Then the 'cost' of T is defined to be $c(T) = \sum_{(i,j)\in T} c_{ij}$, where (i, j) is denotes the directed edge from i to j in the tree T. Let T_n denote the 'optimal' tree, i.e. $c(T_n) = \min\{c(T) : T$ is a directed, rooted tree in with n vertices$\}$. We establish general conditions on the asymptotic behaviour of the moments of the order statistics of the variables $c_{11}, c_{12}, \ldots, c_{in}$ which guarantee the existence of sequences $\{a_n\}, \{b_n\}$, and $\{d_n\}$ such that $b_n^{-1}(c(T_n) - a_n) \to N(0, 1)$ in distribution, $d_n^{-1}c(T_n) \to 1$ in probability, and $d_n^{-1}E(c(T_n)) \to 1$ as $n \to \infty$, and we explicitly determine these sequences. The proofs of the main results rely upon the properties of general random mappings of the set $\{1, 2, \ldots, n\}$ into itself. Our results complement and extend those obtained by McDiarmid [9] for optimal branchings in a complete directed graph.

1 Introduction

In this paper we consider the following optimization problem. Suppose $\mathcal{C} = \{c_{ij} : i, j = 1, 2, \ldots\}$ is a collection of i.i.d. nonnegative continuous random variables with distribution function F. For each $n \geq 1$, let D_n denote the complete directed graph on n vertices, labelled $1, 2, \ldots, n$. A directed edge from i to j in D_n is denoted by the ordered pair (i, j) and the random 'cost' of each edge (i, j) is c_{ij}. A directed, rooted tree T in D_n is a set of directed edges such that if we ignore edge orientations, the corresponding graph is a tree, and such that the out-degree of each vertex in T is at most one. Note that all directed paths in T terminate at the same vertex v_T, which is called the root of the tree T. For any rooted tree T in D_n, define the cost, $c(T)$, of T by $c(T) = \sum_{(i,j)\in T} c_{ij}$. The problem is to find a directed rooted tree T_n on the n vertices of D_n such that $c(T_n)$=min$\{c(T) : T$ is a directed, rooted tree in D_n with n vertices$\}$. We call the tree T_n the optimal tree and we note that it is almost surely unique since the distribution of the edge costs is continuous.

This problem is equivalent to the problem of finding an optimal branching in D_n. McDiarmid [9] considered the optimal branching problem and established that $c(T_n) \to 1$ in mean square as $n \to \infty$ provided the variables $\{c_{ij}\}$ are either exponentially distributed with mean 1 or uniformly distributed on (0,1). This result follows from a more general result by McDiarmid concerning the cost of the random greedy base of a heriditary system on a set E. McDiarmid's general result is proved under the assumption that the costs for elements of E are i.i.d. and exponentially distributed, and the special properties of the exponential distribution are exploited in his proof.

Our approach is different to McDiarmid's. Rather than considering the problem as an example of a general random optimization problem, we investigate the asymptotic distribution of $c(T_n)$ by explicitly constructing a directed rooted tree $\widehat{T}_n$ such that $c(\widehat{T}_n)$ is 'close' to $c(T_n)$. To construct the tree $\widehat{T}_n$, we start with a random mapping on the vertices in D_n and modify the mapping until we obtain the tree $\widehat{T}_n$. The idea behind the construction is similar to the'patching algorithm' described by Karp and Steele [8] which constructs a nearly optimal assignment by patching together the cylcles of a random permutation. In our case, we patch together the components of a random mapping to construct a nearly optimal tree $\widehat{T}_n$. Another feature of our construction is that we do not need to place any restrictions on the distribution of the costs $\{c_{ij}\}$, other than the assumption that the cost distribution is continuous. The analysis of the asymptotic distribution of $c(\widehat{T}_n)$ and $c(T_n)$ depends on various properties of random mappings and on the asymptotic behaviour of the moments of the order statistics of the variables $c_{11}, c_{12}, \ldots, c_{1n}$. In particular, we show how to determine constants a_n, b_n and d_n such that $b_n^{-1}(c(T_n) - a_n) \to N(0,1)$ in distribution, $d_n^{-1}c(T_n) \to 1$ in probability, and $d_n^{-1}E(c(T_n)) \to 1$ as $n \to \infty$.

The paper is organised as follows. In Section 2 we describe the algorithm which constructs the heuristic tree $\widehat{T}_n$ and we establish some useful probability bounds which are needed for the proofs of the main results. In Section 3 we prove the main distributional results for $c(T_n)$ and in Section 4 we provide some examples and further discussion.

2 The Algorithm and Probability Bounds

In this section we describe the algorithm which constructs a nearly optimal directed, rooted tree $\widehat{T}_n$ on the n labelled vertices in D_n. Before proceeding to this description, we introduce some notation. For each $i \geq 1$ and $1 \leq r \leq n$, let $c_i(r:n)$ denote the r^{th} smallest value of $c_{i1}, c_{i2}, \ldots, c_{in}$, i.e. the variables $c_i(1:n), c_1(2:n), ..., c_i(n:n)$ are the order statistics for the variables $c_{i1}, c_{i2}, ..., c_{in}$.

The Algorithm.

Let n, the number of vertices in D_n be fixed.

Step 1: The first step is to define a random mapping $\phi_{n,1} : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. The mapping is defined by setting $\phi_{n,1}(i) = j$ if $c_{ij} = c_i(1:n)$.

Since the variables $\{c_{ij} : i, j = 1, 2, \ldots, n\}$ are independent and identically distributed, $P(\phi_{n,1}(i) = j) = 1/n$ for each $i, j \in \{1, 2, \ldots, n\}$ and, in particular, the variables $\phi_{n,1}(1), \phi_{n,1}(2), \ldots, \phi_{n,1}(n)$ are i.i.d. and uniformly distributed on the set $\{1, 2, \ldots, n\}$. It follows that $P(\phi_{n,1} = f) = 1/n^n$, where $f \in \Pi_n$, the set of all mappings of $\{1, 2, \ldots, n\}$ into itself. The cost, $c(f)$, of a function $f \in \Pi_n$ is defined by $c(f) = \sum_{i=1}^{n} c_{i,f(i)}$, so it follows by construction that $c(\phi_{n,1}) = \min\{c(f) : f \in \Pi_n\}$. Let $\hat{\imath}$ denote the vertex such that $c_{\hat{\imath},\phi_{n,1}(\hat{\imath})} = \max\{c_{i,\phi_1(i)} : i = 1, 2, \ldots, n\}$. Then clearly $c(\phi_{n,1})$-$c_{\hat{\imath},\phi_{n,1}(\hat{\imath})} \leq c(T_n)$.

The mapping $\phi_{n,1}$ has a graphical representation as a directed graph, $G_{\phi_{n,1}}$, on the vertices $1, 2, \ldots, n$. There is a directed edge in $G_{\phi_{n,1}}$ from i to j if $\phi_{n,1}(i) = j$. Each connected component of $G_{\phi_{n,1}}$ consists of a directed cycle with directed trees attached such that all paths in an attached tree terminate at a vertex in the component's cycle. If $G_{\phi_{n,1}}$ has exactly one connected component, then select the vertex i_1 from the unique cycle in $G_{\phi_{n,1}}$ such that $c_{i_1,\phi_{n,1}(i_1)} = \max\{c_{ij} : (i, j)$ is a cyclic edge in $G_{\phi_{n,1}}\}$. Delete the edge $(i_1, \phi_{n,1}(i_1))$ from $G_{\phi_{n,1}}$ to obtain the directed rooted tree $\widehat{T}_n$. If $G_{\phi_{n,1}}$ is not connected proceed to Step 2.

Step 2: Let $m_1(n)$ denote the number of components in $G_{\phi_{n,1}}$ and suppose that $m_1(n) > 1$. Let $C_1, C_2, \ldots, C_{m_1(n)}$ denote these components, and suppose that the components are labelled so that $|C_1| \geq |C_2| \geq \ldots \geq |C_{m_1(n)}|$. Select a cyclic vertex i_1 from the unique cycle in C_1 such that $c_{i_1,\phi_{n,1}(i_1)} = \max\{c_{ij} : (i, j)$ is a cyclic edge in the component $C_1\}$. For the other components C_k, $2 \leq k \leq m_1(n)$, select a cyclic vertex i_k from the unique cycle in C_k such that $i_k = \min\{i : i$ is the label of a cyclic vertex in the component $C_k\}$. The vertex i_1 will be the root of the directed tree $\widehat{T}_n$ which is finally constructed by the algorithm. Let $V_1 = \{i_2, i_3, \ldots, i_{m_1(n)}\}$. We define the new mapping $\phi_{n,2}$ as follows:

(a) For each vertex $i \notin V_1$, set $\phi_{n,2}(i) = \phi_{n,1}(i)$.

(b) For each vertex $i_k \in V_1$, do the following:

  i. with probability 1/n, set $\phi_{n,2}(i_k) = \phi_{n,1}(i_k)$, otherwise

  ii. determine j_k such that $c_{i_k,j_k} = c_{i_k}(2 : n)$. If $j_k \in C_k$ (i.e. the vertex j_k is in the same component as i_k), set $\phi_{n,2}(i_k) = \phi_{n,1}(i_k)$. Otherwise, set $\phi_{n,2}(i_k) = j_k$.

If the graph $G_{\phi_{n,2}}$ associated with the mapping $\phi_{n,2}$ is connected, then remove the edge $(i_1, \phi_{n,2}(i_1))$ to obtain the directed rooted tree $\widehat{T}_n$ with root i_1. Otherwise, proceed to Step 3.

Step 3: Let $m_r(n)$ denote the number of components in the graph $G_{\phi_{n,r}}$ and suppose that $m_r(n) > 1$ $(r \geq 2)$. By construction of $\phi_{n,r}$, we must have $m_r(n) \leq m_{r-1}(n)$, where $m_{r-1}(n)$ equals the number of components in $\phi_{n,r-1}$. Let $C_1^r, C_2^r, ..., C_{m_r}^r$ denote the components of $\phi_{n,r}$, labelled so that C_1^r is the component of $G_{\phi_{n,r}}$ which contains the 'special' cyclic vertex i_1 defined in Step 2. (Note that by construction, the vertex i_1 is a cyclic vertex of $\phi_{n,r}$ for each

$r \geq 1$). It follows from the construction of $\phi_{n,r}$ that for each component C_k^r of $G_{\phi_{n,r}}$, there is a vertex in the set V_{r-1} which is *also* a cyclic vertex in C_k^r. Noting this, we can construct the set $V_r \subseteq V_{r-1}$ as follows: For each $2 \leq k \leq m_r(n)$, consider the vertices V_{r-1} which are cyclic in C_k^r and select the vertex which has smallest index in V_{r-1}. Include the selected vertex in the set V_r. Once the set V_r is constructed, reindex the vertices so that $V_r = \{i_2, i_3, ..., i_{m_r(n)}\}$ and $i_k \in C_k^r$ for $2 \leq k \leq m_r(n)$.

Define the new mapping $\phi_{n,r+1}$ as follows:

(a) For each vertex $i \notin V_r$, set $\phi_{n,r+1}(i) = \phi_{n,r}(i)$.

(b) For vertex $i_k \in V_r$, the value of $\phi_{n,r+1}(i_k)$ is determined as follows:

i. Initially, select a vertex j_k at random such that for each $1 \leq m \leq r$, $c_{i_k,j_k} = c_{i_k}(m:n)$ with probability $1/n$ and such that $c_{i_k,j_k} = c_{i_k}(r+1:n)$ with probability $(n-r)/n$.

ii. If $j_k \notin C_k^r$, set $\phi_{n,r+1}(i_k) = j_k$. If $j_k \in C_k^r$ and $c_{i_k,j_k} < \phi_{n,r}(i_k)$, set $\phi_{n,r+1}(i_k) = j_k$; otherwise set $\phi_{n,r+1}(i_k) = \phi_{n,r}(i_k)$.

If the graph $G_{\phi_{n,r+1}}$ associated with the mapping $\phi_{n,r+1}$ is connected, remove the edge $(i_1, \phi_{n,r+1}(i_1))$ to obtain the directed rooted tree $\widehat{T}_n$ with root i_1. If $G_{\phi_{n,r+1}}$ is not connected, then repeat Step 3.

Remark. Here is the idea behind the algorithm. To create the mapping $\phi_{n,r+1}$, we break the cycles of the components of $G_{\phi_{n,r}}$ at each of the vertices in $V_r = \{i_2, i_3, \ldots, i_{m_r(n)}\}$ (the cycle containing the vertex i_1 is never broken). We then map the vertices $i_2, i_3, \ldots, i_{m_r(n)}$ to 'new' vertices and the other edges in the mapping remain unchanged. The new graph $G_{\phi_{n,r+1}}$ typically has fewer connected components than $G_{\phi_{n,r}}$ and if $G_{\phi_{n,r+1}}$ is not connected then at least some of the vertices in the set $V_r = \{i_2, i_3, \ldots, i_{m_r(n)}\}$ will also be cyclic vertices in $G_{\phi_{n,r+1}}$. We construct a new set of vertices V_{r+1} by selecting a subset of V_r which consists of cyclic vertices of $G_{\phi_{n,r+1}}$. One iteration of the algorithm is illustrated by Figures 1 and 2.

In the next section we investigate the asymptotic distribution of the variable $c(T_n)$. Our analysis is based on the observation that

$$c(\phi_{n,1}) - c_{\hat{i},\phi_{n,1}(\hat{i})} \leq c(T_n) \leq c(\widehat{T}_n).$$

The magnitude of the difference $c(\widehat{T}_n) - (c(\phi_{n,1}) - c_{\hat{i},\phi_{n,1}(\hat{i})})$ depends, in part, on the number of iterations, N_n, of the algorithm which are required to construct the heuristic tree $\widehat{T}_n$ and on the number of components, $m_1(n)$, in the first mapping which is constructed by the algorithm. We conclude this section by establishing suitable bounds for the tail probabilities $P(N_n > M)$ and $P(m_1(n) > a)$. These bounds are needed for the calculations in Section 3.

We begin by noting that $N_n > M$ if and only if the graphs $G_{\phi_{n,1}}, G_{\phi_{n,2}}, \ldots, G_{\phi_{n,M}}$ are not connected. Thus, a bound for $P(G_{\phi_{n,1}}, G_{\phi_{n,2}}, \ldots, G_{\phi_{n,M}}$ are not connected) yields a bound for $P(N_n > M)$. To obtain this bound we appeal to the following theorem.

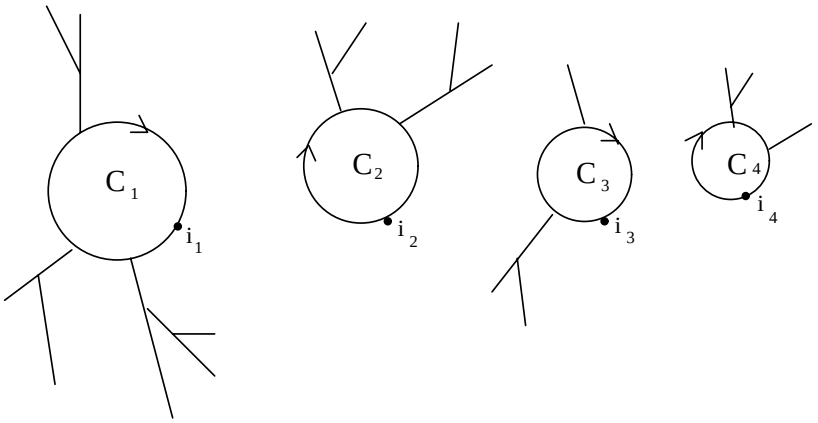

Figure 1: $G_{\phi_{n,1}}$ with $V_1 = \{i_2, i_3, i_4\}$

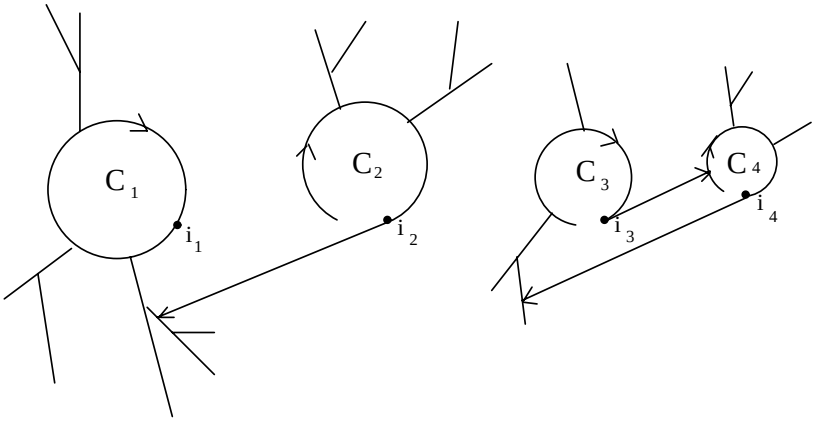

Figure 2: $G_{\phi_{n,2}}$ with $V_2 = \{i_3\}$

Theorem (Ross [10]). *Suppose that $X_1, X_2, ..., X_n$ are i.i.d. random variables such that*

$$P(X_1 = j) = \lambda_j \qquad j = 0, 1, \ldots, n \quad \text{and} \quad \sum_{j=0}^{n} \lambda_j = 1.$$

Let G be the random digraph on the vertices $0, 1, \ldots, n$ obtained by constructing a directed edge from i to j if $X_i = j$. Then $P(G$ is connected$) = \lambda_0$.

Remark. Note that the vertex 0 in Ross's random digraph G always has out-degree zero.

We use Ross's theorem to obtain a bound for the *conditional* probability

$$P(G_{\phi_{n,2}}, G_{\phi_{n,3}}, \ldots, G_{\phi_{n,M}} \text{ are not connected} \mid G_{\phi_{n,1}} \text{ is not connected, } |C_1|),$$

where $|C_1|$ is the size of the largest component in $G_{\phi_{n,1}}$. We begin by considering the probability that the graph $G_{\phi_{n,2}}$ is not connected given that the graph $G_{\phi_{n,1}}$ is not connected and given $|C_1|$. To apply the theorem, we treat each component C_k in $G_{\phi_{n,1}}$ as a 'vertex' and we construct a directed graph on these 'vertices' such that there is a directed edge from 'vertex' C_k to 'vertex' C_j if $\phi_{n,2}(i_k) \in C_j$. It follows from the construction of the mapping $\phi_{n,2}$ that

$$P(\text{vertex } C_k \text{ is mapped to vertex } C_j) = P(\phi_{n,2}(i_k) \in C_j) = \frac{|C_j|}{n},$$

and this holds for every k and j. Note that the 'graph' on the 'vertices' $C_1, C_2, .., C_{m_1}$ is connected if and only if the graph $G_{\phi_{n,2}}$ is connected. In this case, the 'vertex' C_1 plays the special role of the vertex 0 in Ross's theorem and, applying the theorem, we have

$$P(G_{\phi_{n,2}} \text{ is not connected} \mid G_{\phi_{n,1}} \text{ is not connected, } |C_1|) = 1 - \frac{|C_1|}{n}.$$

Similarly, for $r > 1$, if $G_{\phi_{n,r}}$ is not connected, we can treat the components $C_1^r, C_2^r, \ldots, C_{m_r}^r$ of $G_{\phi_{n,r}}$ as 'vertices' of a directed graph. In this case, the component C_1^r which contains i_1 is the 'special' vertex, and Ross's theorem yields

$$P(G_{\phi_{n,r+1}} \text{ is not connected} \mid G_{\phi_{n,r}} \text{ is not connected, } |C_1^r|) = 1 - \frac{|C_1^r|}{n}.$$

Note that $|C_1^r| \geq |C_1|$ for $r > 1$, so it follows that

$$P(G_{\phi_{n,2}}, G_{\phi_{n,3}}, \ldots, G_{\phi_{n,M}} \text{ are not connected} \mid G_{\phi_{n,1}} \text{ is not connected, } |C_1|)$$

$$\leq \left(1 - \frac{|C_1|}{n}\right)^{M-1}. \tag{1}$$

It is clear from inequality (1) that we need to determine the distribution of $|C_1| = |C_1(n)|$, the size of the largest component, $C_1(n)$, in $G_{\phi_{n,1}}$. (We

write $C_1(n)$ in order to emphasize that the distribution of $|C_1(n)|$ depends on the number of vertices, n, in the underlying graph $G_{\phi_{n,1}}$.) In particular, we need to approximate $P(|C_1(n)| < m)$ for suitable values of m. As it is difficult to determine $P(|C_1(n)| < m)$ directly, we exploit the following simple observation. Let $\widehat{C}_1(n)$ denote the component in $G_{\phi_{n,1}}$ which contains vertex 1, then $|\widehat{C}_1(n)| \leq |C_1(n)|$. Hence, for all $1 \leq m \leq n$, $P(|\widehat{C}_1(n)| < m) \geq P(|C_1(n)| < m)$. It is relatively easy to obtain an upper bound for $P(|\widehat{C}_1(n)| < m)$, and thus an upper bound for $P(|C_1(n)| < m)$.

To bound $P(|\widehat{C}_1(n)| < m)$, we first determine $P(|\widehat{C}_1(n)| = k)$. The number of ways to construct a connected random mapping on k labelled vertices is given by $(k-1)! \sum_{n=0}^{k-1} k^n/n!$ (see Bollobás [3]). Using this and straightforward counting, we obtain

$$\begin{aligned} P(|\widehat{C}_1(n)| = k) &= \binom{n-1}{k-1}(k-1)! \sum_{n=0}^{k-1} \frac{k^n}{n!} \frac{(n-k)^{n-k}}{n^n} \\ &= \frac{n!(n-k)^{n-k}}{(n-k)!n^{n+1}} \sum_{n=0}^{k-1} \frac{k^n}{n!}. \end{aligned} \tag{2}$$

It follows from the Berry-Esseen theorem (see Feller [5]) that for all $k \geq 1$,

$$\left| \sum_{n=0}^{k-1} \frac{k^n e^{-k}}{n!} - \frac{1}{2} \right| \leq \frac{8}{k^{3/2}},$$

and it follows from this and Stirling's formula that for $1 \leq k \leq n-1$,

$$\left| P(|\widehat{C}_1(n)| = k) - \frac{1}{2n\sqrt{1-k/n}} \right| \leq \frac{C}{\sqrt{n}} \left(\frac{1}{k^{3/2}} + \frac{1}{(n-k)^{3/2}} \right)$$

where C is a constant which does not depend on n or k. So, for $1 < M < \sqrt{n}$, we obtain

$$\begin{aligned} P(|C_1(n)| < \frac{n}{M}) &\leq P(|\widehat{C}_1(n)| < \frac{n}{M}) \\ &\leq \sum_{k<n/M} \frac{1}{2n\sqrt{1-k/n}} + O(1/\sqrt{n}) \\ &\leq \int_0^{1/M} \frac{1}{2\sqrt{1-x}} dx + O(1/\sqrt{n}) \\ &= 1 - \sqrt{1-1/M} + O(\frac{1}{\sqrt{n}}) \\ &\leq \frac{\widetilde{C}}{M} \end{aligned} \tag{3}$$

where the constant $\widetilde{C}$ does not depend on n or M.

The bound obtained above is not tight enough for our purposes, but we can use (3) to obtain a tighter bound. We first introduce some further notation. Let $\widehat{C}_2(n)$ denote the component in $G_{\phi_{n,1}}$ which contains the smallest vertex *not* in $\widehat{C}_1(n)$ (let $\widehat{C}_2(n) = \emptyset$ if $\widehat{C}_1(n) = G_{\phi_{n,1}}$). In general, for $k > 1$, let

$\widehat{C}_k(n)$ denote the component in $G_{\phi_{n,1}}$ which contains the smallest vertex not in $\widehat{C}_1(n) \cup ... \cup \widehat{C}_{k-1}(n)$ provided $\widehat{C}_1(n) \cup ... \cup \widehat{C}_{k-1}(n) \neq G_{\phi_{n,1}}$. Otherwise, let $\widehat{C}_k(n) = \emptyset$. We note that if $|C_1(n)| < \frac{n}{M}$, then $G_{\phi_{n,1}}$ must have at least M components and we must have

$$P(|C_1(n)| < \frac{n}{M}) \leq P(|\widehat{C}_1(n)| < \frac{n}{M}, |\widehat{C}_2(n)| < \frac{n}{M}, \ldots, |\widehat{C}_k(n)| < \frac{n}{M}) \quad (4)$$

for any $1 < k < M < n$. This observation and the following lemma establish a sufficiently tight bound for $P(|C_1(n)| < n/M)$ for suitable values of M.

Lemma 2.1. *For any $k \geq 1$ and all M and n such that $k < M < \sqrt{n}$*

$$P\left(|\widehat{C}_1(n)| < \frac{n}{M}, |\widehat{C}_2(n)| < \frac{n}{M}, \ldots, |\widehat{C}_k(n)| < \frac{n}{M}\right) \leq \left(\frac{\widetilde{C}}{M}\right)^k \prod_{j=1}^{k-1} \left(1 - \frac{j}{M}\right)^{-1}$$

where $\widetilde{C}$ is a constant which does not depend on k, M, or n.

Proof. The proof is by induction on k. For $k = 1$ and $1 < M < \sqrt{n}$, the result follows from (3).

Suppose that the result holds for $k - 1 \geq 1$ and suppose that M and n are such that $k < M < \sqrt{n}$, then we have

$$\begin{aligned} & P\left(|\widehat{C}_1(n)| < \frac{n}{M}, \ldots, |\widehat{C}_k(n)| < \frac{n}{M}\right) \\ = & \sum_{j_1, j_2, \ldots, j_{k-1} < n/M} P\left(|\widehat{C}_1(n)| = j_1, \ldots, |\widehat{C}_{k-1}(n)| = j_{k-1}\right) \\ & \cdot \sum_{j_k < n/M} P\left(|\widehat{C}_k(n)| = j_k \;\middle|\; |\widehat{C}_1(n)| = j_1, \ldots, |\widehat{C}_{k-1}(n)| = j_{k-1}\right) \quad (5) \end{aligned}$$

Let $a(j) = (j-1)! \sum_{n=0}^{j-1} j^n/n!$, the number of ways to construct a connected mapping on j labelled vertices. Then for $j_1, j_2, \ldots, j_k < n/M$, counting yields

$$\begin{aligned} & P\left(|\widehat{C}_k(n)| = j_k \;\middle|\; |\widehat{C}_1(n)| = j_1, \ldots, |\widehat{C}_{k-1}(n)| = j_{k-1}\right) \\ = & \frac{\binom{n-1}{j_1-1}\binom{n-j_1-1}{j_2-1} \cdots \binom{n-j_1-\cdots-j_{k-1}-1}{j_k-1} \prod_{m=1}^{k} a(j_m) \left(n - \sum_{m=1}^{k} j_m\right)^{n-j_1-\cdots-j_k}}{\binom{n-1}{j_1-1} \cdots \binom{n-j_1-\cdots-j_{k-2}-1}{j_{k-1}-1} \prod_{m=1}^{k-1} a(j_m) \left(n - \sum_{m=1}^{k-1} j_m\right)^{n-j_1-\cdots-j_{k-1}}} \\ = & \frac{\binom{n-j_1-\cdots-j_{k-1}-1}{j_k-1} a(j_k) \left(n - \sum^{k} j_m\right)^{n-j_1-\cdots-j_k}}{(n - j_1 - \cdots - j_{k-1})^{n-j_1-\cdots-j_{k-1}}} \\ = & P(|\widehat{C}_1(n - j_1 - \cdots - j_{k-1})| = j_k). \end{aligned}$$

The last equality follows from formula (2). Thus, for any $j_1, \ldots, j_{k-1} < n/M$,

$$\sum_{j_k < n/M} P(|\widehat{C}_k(n)| = j_k \mid |\widehat{C}_1(n)| = j_1, \ldots, |\widehat{C}_{k-1}(n)| = j_{k-1})$$

$$
\begin{aligned}
&= \sum_{j_k<n/M} P(|C^1(n-j_1-\cdots-j_{k-1})| = j_k) \\
&= P\left(|\widehat{C}_1(n-j_1-\cdots-j_{k-1})| < \frac{n}{M}\right) \\
&\leq \frac{\widetilde{C}n}{M(n-j_1-\cdots-j_{k-1})} \\
&\leq \frac{\widetilde{C}}{M(1-(k-1)/M)}.
\end{aligned}
$$

The first inequality follows from the bound (3). Substituting this bound into (5), summing over $j_1, j_2, \ldots, j_{k-1} < n/M$, and using the induction hypothesis, we obtain

$$
\begin{aligned}
& P\left(|\widehat{C}_1(n)| < \frac{n}{M}, \ldots, |\widehat{C}_k(n)| \leq \frac{n}{M}\right) \\
\leq\ & P\left(|\widehat{C}_1(n)| < \frac{n}{M}, \ldots, |\widehat{C}_{k-1}(n)| < \frac{n}{M}\right) \frac{\widetilde{C}}{M(1-(k-1)/M)} \\
\leq\ & \left(\frac{\widetilde{C}}{M}\right)^k \prod_{j=1}^{k-1} (1-\frac{j}{M})^{-1}.
\end{aligned}
$$

It follows from the lemma and inequality (4), that for any $M < \sqrt{n}$

$$
P\left(|C_1(n)| < \frac{n}{M}\right) \leq \left(\frac{\widetilde{C}}{M}\right)^{\lfloor\sqrt{M}\rfloor} \prod_{j=1}^{\lfloor\sqrt{M}\rfloor-1} \left(1-\frac{j}{M}\right)^{-1} \leq e\left(\frac{\widetilde{C}}{M}\right)^{\lfloor\sqrt{M}\rfloor}.
$$

Using this bound and (1), we obtain the following lemma.

Lemma 2.2. *Suppose that M is a positive integer such that $e\widetilde{C} < \sqrt{M} < \sqrt{n}$, where $\widetilde{C}$ is the constant in Lemma 2.1, then*

$$
P(N_n > M) \leq Ke^{-\sqrt[4]{M}},
$$

where K is a constant which does not depend on n (but which may depend on $\widetilde{C}$).

Proof.

$$
\begin{aligned}
& P(N_n > M) = P(G_{\phi_{n,1}}, \ldots, G_{\phi_{n,M}} \text{ are not connected}) \\
=\ & \sum_{m=1}^{n} P(G_{\phi_{n,1}}, \ldots, G_{\phi_{n,M}} \text{ are not connected} \mid |C_1(n)| = m)\, P(|C_1(n)| = m) \\
\leq\ & \sum_{m \geq n/\sqrt{M}} P(G_{\phi_{n,1}}, \ldots, G_{\phi_{n,M}} \text{ are not connected} \mid |C_1(n)| = m)\, P(|C_1(n)| = m) \\
& \quad + \sum_{m<n/\sqrt{M}} P(|C_1(n)| = m)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sum_{m \geq n/\sqrt{M}} \left(1 - \frac{m}{n}\right)^{M-1} P(|C_1(n)| = m) + P\left(|C_1(n)| < n/\sqrt{M}\right) \\
&\leq \left(1 - \frac{1}{\sqrt{M}}\right)^{M-1} + P\left(|C_1(n)| < \frac{n}{\sqrt{M}}\right) \\
&\leq \left(1 - \frac{1}{\sqrt{M}}\right)^{M-1} + e\left(\frac{\widetilde{C}}{\sqrt{M}}\right)^{\lfloor \sqrt[4]{M} \rfloor} \\
&\leq K(e^{-\sqrt[4]{M}})
\end{aligned}
$$

where K is a constant which does not depend on n.

Finally, we state the bound for $P(m_1(n) > a)$ as a lemma.

Lemma 2.3. *Suppose that $a > 1$, then*

$$
P(m_1(n) > a) \leq \widetilde{K} \exp\left(\frac{-\sqrt{2}a}{\sqrt{\log n}} + \sqrt{\frac{\log n}{2}}\right)
$$

where $\widetilde{K}$ is a constant which does not depend on n.

Proof. The random variable $m_1(n)$ equals the number of connected components in a (uniform) random mapping on n labelled vertices and has a moment generating function $\Phi_n(t) = E(e^{tm_1(n)})$, which exists for all $t \in \mathbb{R}$. Flajolet and Soria [6] have shown that as $n \to \infty$

$$
e^{-\sqrt{\log n/2}} \Phi_n\left(\frac{\sqrt{2}t}{\sqrt{\log n}}\right) \to e^{t^2/2}.
$$

It follows that the distribution of $m_1(n)$ has exponential tails. In particular,

$$
\begin{aligned}
P(m_1(n) > a) &= P\left(\frac{\sqrt{2}m_1(n)}{\sqrt{\log n}} > \frac{\sqrt{2}a}{\sqrt{\log n}}\right) \\
&= P\left(\exp\left(\frac{\sqrt{2}m_1(n)}{\sqrt{\log n}}\right) > \exp\left(\frac{\sqrt{2}a}{\sqrt{\log n}}\right)\right) \\
&\leq \exp\left(\frac{-\sqrt{2}a}{\sqrt{\log n}}\right) E\left(\exp\left(\frac{\sqrt{2}m_1(n)}{\sqrt{\log n}}\right)\right) \\
&= \exp\left(\frac{-\sqrt{2}a}{\sqrt{\log n}} + \sqrt{\frac{\log n}{2}}\right) \left(\exp(-\sqrt{\log n/2}) \Phi_n\left(\sqrt{2/\log n}\right)\right) \\
&\leq \widetilde{K} \exp\left(\frac{-\sqrt{2}a}{\sqrt{\log n}} + \sqrt{\frac{\log n}{2}}\right).
\end{aligned}
$$

Remark. Note that the bound obtained in the lemma is meaningful only if $a > (\log n)/2$.

3 Asymptotic Results for $c(T_n)$

In this section we establish asymptotic results for $c(T_n)$. The results are based upon the following observations. First, by construction, $c(\phi_{n,1}) = \sum_{i=1}^{n} c_i(1:n)$, where $c_1(1:n), \ldots, c_n(1:n)$ are i.i.d. random variables, so standard results from probability theory can be employed to establish various limit laws for $c(\phi_{n,1})$. Now provided that $c(T_n)$ is suffificiently close to $c(\phi_{n,1})$ (with high probability) as $n \to \infty$, a limit law which holds for $c(\phi_{n,1})$ can be shown also to hold for $c(T_n)$. Thus, to convert a limit law for $c(\phi_{n,1})$ into a result for $c(T_n)$ we first need a bound for $|c(T_n) - c(\phi_{n,1})|$. The bound that we use is based on the observation that

$$c(\phi_{n,1}) - c_{\hat{\imath},\phi_{n,1}(\hat{\imath})} \le c(T_n) \le c(\widehat{T}_n) = c(\phi_{n,N_n}) - c_{i_1,\phi_{n,1}(i_1)}.$$

It follows that

$$|c(T_n) - c(\phi_{n,1})| \le |c(\phi_{n,N_n}) - c(\phi_{n,1})| + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}. \tag{6}$$

By construction of the mapping ϕ_{n,N_n}, we have

$$c(\phi_{n,N_n}) \le \sum_{i \notin V_1} c_i(1:n) + \sum_{i_k \in V_1} c_{i_k}(N_n:n)$$

and so

$$|c(\phi_{n,N_n}) - c(\phi_{n,1})| \le \sum_{i_k \in V_1} (c_{i_k}(N_n:n) - c_{i_k}(1:n)) \le \sum_{i_k \in V_1} c_{i_k}(N_n:n). \tag{7}$$

Combining (6) and (7) yields the bound

$$|c(T_n) - c(\phi_{n,1})| \le \sum_{i_k \in V_1} c_{i_k}(N_n:n) + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}.$$

Using this bound, we now establish conditions which guarantee that a central limit theorem holds for $c(T_n)$ whenever a central limit theorem holds for $c(\phi_{n,1})$.

Theorem 3.1. *For $n \ge 1$, let the variables T_n, $\{c_i(k:n) : 1 \le i \le n, 1 \le k \le n,\ n = 1, 2, \ldots\}$,and $c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}$ be as defined in Section 2. Suppose that $Var(c_1(1:n)) < \infty$ and that*

$$\frac{c(\phi_{n,1}) - E(c(\phi_{n,1}))}{\sqrt{Var(c(\phi_{n,1})}} = \frac{c(\phi_{n,1}) - nE(c_1(1:n))}{\sqrt{n}\sqrt{Var(c_1(1:n))}}$$

converges in distribution to the Normal(0,1) distribution as $n \to \infty$. If

- *(a)*
$$\lim_{n\to\infty} \frac{E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})})}{\sqrt{n}(Var(c_1(1:n))^{1/2}} = 0, \qquad \textit{and}$$
- *(b) if for each integer $M > 0$*
$$\lim_{n\to\infty} \frac{(\log n)E(c_1(M:n))}{\sqrt{n}(Var(c_1(1:n))^{1/2}} = 0,$$

then

$$\frac{c(T_n) - E(c(\phi_{n,1}))}{\sqrt{Var(c(\phi_{n,1}))}}$$

also converges in distribution to the $N(0,1)$ *distribution as* $n \to \infty$.

Remark. To show that

$$\frac{(c(\phi_{n,1}) - E(c(\phi_{n,1}))}{\sqrt{Var(c(\phi_{n,1}))}}$$

converges weakly to the standard normal distribution, one can appeal to Liapounov's theorem (see Chung [4]). In particular, it suffices to check that

$$\lim_{n\to\infty} \frac{E(c_1(1:n))^3}{\sqrt{n}(Var(c_1(1:n)))^{3/2}} = 0.$$

Proof. Since

$$\frac{c(\phi_{n,1}) - E(c(\phi_{n,1}))}{\sqrt{Var(c(\phi_{n,1}))}}$$

converges in distribution to the standard normal distribution, it suffices to show that

$$\frac{|c(T_n) - c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} \to 0$$

in probability. Let $A_n = \{N_n < M, m_1(n) < \log n\}$ where M is an arbitrary positive integer and let the ordered triple (j, m, S) denote the event $\{N_n = j,\ m_1(n) = m,\ V_1 = S\}$ where $|S| = m - 1$. Then for any $\epsilon > 0$

$$\begin{aligned} &P\left(\frac{|c(T_n) - c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon\right) \\ &\leq \sum_{(j,m,S)\in A_n} P\left(\frac{|c(T_n) - c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon \;\middle|\; (j,m,S)\right) P((j,m,S)) + P(A_n^c). \end{aligned} \tag{8}$$

In order to bound the conditional probabilities in the sum on the RHS of (8), we first establish that the variables $\{c_i(k:n) : 1 \leq i \leq n,\ 1 \leq k \leq n\}$ are independent of the event $\{N_n = j,\ m_1(n) = m,\ V_1 = S\}$ for each $j, m > 0$ and $S \subseteq \{1, 2, \ldots, n\}$ such that $|S| = m - 1$.

For each $n \geq 1$, define variables $X_i(k:n)$ for $1 \leq i \leq n,\ 1 \leq k \leq n,$ by setting $X_i(k:n) = j$ if $c_{ij} = c_i(k,n)$. Note that for each $n \geq 1$, the σ-algebras generated by $\{c_i(k:n) : 1 \leq i \leq n,\ 1 \leq k \leq n\}$ and $\{X_i(k:n) : 1 \leq i \leq n,\ 1 \leq k \leq n\}$, respectively, are independent. Also, given the values of $X_i(k:n)$ for $1 \leq i \leq n$ and $1 \leq k \leq n$, then the set V_1 and the values of N_n and $m_1(n)$ are completely determined, i.e. the event $\{V_1 = S, N_n = j, m_1(n) = k\} \in \sigma\{X_i(k:$

$n) : 1 \le i \le n,\ 1 \le k \le n\}$, and hence this event is independent of the variables $\{c_i(k : n) : 1 \le n,\ 1 \le k \le n\}$.

Now suppose that $V_1 = S = \{i_2, i_3, \ldots, i_{m_1(n)}\}, N_n = j \le M$, and $m_1(n) = m$. Then it follows from (6) and (7) that

$$\begin{aligned} |c(T_n) - c(\phi_{n,1}| &\le |c(\phi_{n,j}) - c(\phi_{n,1})| + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})} \\ &\le \sum_{i_k \in V_1} (c_{i_k}(j : n) + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})} \\ &\le \sum_{i_k \in V_1} c_{i_k}(M : n) + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}. \end{aligned}$$

Thus, since the event $\{V_1 = S, N_n = j, m_1(n) = m\}$ and the variables $\{c_i(k : n) : 1 \le i \le n, 1 \le k \le n\}$ are independent, we have

$$\begin{aligned} & P\left(\frac{|c(T_n) - c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon \,\middle|\, (j, m, S)\right) \\ \le\ & P\left(\frac{\sum_{i_k \in S} c_{i_k}(M : n) + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon \,\middle|\, (j, m, S)\right) \\ =\ & P\left(\frac{\sum_{i_k \in S} c_{i_k}(M : n) + c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon\right) \\ \le\ & \frac{mE(c_1(M : n)) + E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})})}{\epsilon\sqrt{nVar(c_1(1 : n))}} \\ \le\ & \frac{(\log n)E(c_1(M : n)) + E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})})}{\epsilon\sqrt{nVar(c_1(1 : n))}}. \end{aligned} \tag{9}$$

The last inequality holds since $m < \log n$. We substitute the bound given by (9) into the terms in the sum on the RHS of inequality (8) to obtain

$$\begin{aligned} & P\left(\frac{|c(T_n) - c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon\right) \\ \le\ & \left(\frac{(\log n)E(c_1(M : n)) + E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})})}{\epsilon\sqrt{nVar(c_1(1 : n))}}\right) P(A_n) + P(A_n^c). \end{aligned}$$

It follows from the above inequality and the hypotheses that

$$\limsup_{n\to\infty} P\left(\frac{|c(T_n) - c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon\right) \le \limsup_{n\to\infty} P(A_n^c).$$

Since A_n^c corresponds to the event '$N_n \ge M$ or $m_1(n) \ge \log n$',

$$P(A_n^c) \le P(N_n \ge M) + P(m_1(n) \ge \log n).$$

It follows from Lemmas 2.2 and 2.3 that $\limsup_{n\to\infty} P(N_n \ge M) \le Ke^{-\sqrt[4]{M}}$ and $\lim_{n\to\infty} P(m_1(n) \ge \log n) = 0$, so

$$\limsup_{n\to\infty} P(A_n^c) \le Ke^{-\sqrt[4]{M}}.$$

Since M was arbitray, we conclude that

$$\limsup_{n\to\infty} P\left(\frac{|c(T_n)-c(\phi_{n,1})|}{\sqrt{Var(c(\phi_{n,1}))}} > \epsilon\right) = 0.$$

Adapting the arguments from the proof of Theorem 3.1, we can also establish the following result.

Theorem 3.2. *For $n \geq 1$, let the variables T_n, $c_{\hat{i},\phi_{n,1}(\hat{i})}$ and $\{c_i(k:n) : 1 \leq i \leq n, 1 \leq k \leq n, n \geq 1\}$ be as in Section 2. Suppose that $E(c_1(1:n)) < \infty$ and that*

$$\frac{c(\phi_{n,1})}{E(c(\phi_{n,1}))} = \frac{c(\phi_{n,1})}{nE(c_1(1:n))} \to 1$$

in probability as $n \to \infty$. If

- *(a)*

$$\lim_{n\to\infty} \frac{E(c_{\hat{i},\phi_{n,1}(\hat{i})})}{nE(c_1(1:n))} = 0, \quad \textit{and}$$

- *(b) if for each integer $M > 0$*

$$\lim_{n\to\infty} \frac{(\log n)E(c_1(M:n))}{nE(c_1(1:n))} = 0,$$

then we also have

$$\frac{c(T_n)}{nE(c_1(1:n))} \to 1$$

in probability as $n \to \infty$.

Proof. Since

$$\frac{c(\phi_{n,1}) - c_{\hat{i},\phi_{n,1}(\hat{i})}}{nE(c_1(1:n))} \leq \frac{c(T_n)}{nE(c_1(1:n))} \leq \frac{c(\phi_{n,1}) + \sum_{i_k\in V_1} c_{i_k}(N_n:n)}{nE(c_1(1:n))}, \tag{10}$$

it suffices to show that, in probability, both $c_{\hat{i},\phi_{n,1}(\hat{i})}\{nE(c_1(1:n))\}^{-1} \to 0$ and $\sum_{i_k\in V_1} c_{i_k}(N_n:n)\{nE(c_1(1:n))\}^{-1} \to 0$ as $n \to \infty$.

It is immediate, by assumption (a), that $c_{\hat{i},\phi_{n,1}(\hat{i})}\{nE(c_1(1:n))\}^{-1} \to 0$ in probability. To show that $\sum_{i_k\in V_1} c_{i_k}(N_n:n)\{nE(c_1(1:n))\}^{-1} \to 0$ in probability, we repeat the conditioning argument given in the proof of Theorem 3.1 to establish that

$$P\left(\frac{\left|\sum_{i_k\in V_1} c_{i_k}(N_n:n)\right|}{nE(c_1(1:n))} > \epsilon\right) \leq \frac{(\log n)E(c_1(M:n))}{\epsilon nE(c_1(1:n))} P(A_n) + P(A_n^c)$$

where $A_n = \{N_n < M,\ m_1(n) < \log n\}$ and M is an arbitrary postive integer. Using assumption (b) and repeating the argument from the proof of Theorem 3.1 we obtain

$$\limsup_{n\to\infty} P\left(\frac{\left|\sum_{i_k\in V_1} c_{i_k}(N_n:n)\right|}{nE(c_1(1:n))} > \epsilon\right) \leq \limsup_{n\to\infty} P(A_n^c) \leq Ke^{-\sqrt[4]{M}}.$$

Since M was arbitrary, the result follows.

Finally, we establish a result for the asymptotic value of $E(c(T_n))$.

Theorem 3.3. *For $n \geq 1$, let the variables T_n, $c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}$ and $\{c_i(k:n) : 1 \leq i \leq n, 1 \leq k \leq n, n \geq 1\}$ be as in Section 2. Suppose that $E(c_1(n:n)) < \infty$ for all $n \geq 1$ and that*

- *(i)*
$$\lim_{n\to\infty} \frac{E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})})}{nE(c_1(1:n))} = 0.$$

In addition, suppose there exist sequences $\{a_n\}$ and $\{M_n\}$ such that $a_n \to \infty$ and $M_n \to \infty$ as $n \to \infty$, and such that

- *(ii)*
$$\lim_{n\to\infty} \frac{a_n E(c_1(M_n:n))}{nE(c_1(1:n))} = 0,$$

- *(iii)*
$$\lim_{n\to\infty} \frac{E(c_1(n:n))}{E(c_1(1:n))} P(m_1(n) > a_n) = 0, \qquad \text{and}$$

- *(iv)*
$$\lim_{n\to\infty} \frac{E(c_1(n:n))}{E(c_1(1:n))} P(N_n > M_n) = 0.$$

Then we have
$$\lim_{n\to\infty} \frac{E(c(T_n))}{E(c(\phi_{n,1}))} = 1.$$

Remark. To apply Theorem 3.3, some care must be used when choosing the sequences $\{a_n\}$ and $\{M_n\}$. The choice of a_n and M_n will depend on the distribution of the cost variables c_{ij}. Both a_n and M_n must grow slowly enough to ensure the that condition (ii) holds whilst growing fast enough to guarantee that (iii) and (iv) also are satisfied.

Proof. As before, it follows from the construction of the algorithm and from (10) that

$$1 - \frac{E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})})}{nE(c_1(1:n))} \leq \frac{E(c(T_n))}{nE(c_1(1:n))} \leq 1 + \frac{E\left(\sum_{i_k \in V_1} c_{i_k}(N_n:n)\right)}{nE\left(c_1(1:n)\right)}.$$

By condition (i), we have $1 \leq \liminf_{n\to\infty}\{E(c(T_n))/E(c(\phi_{n,1}))\}$, so it suffices to show that

$$\lim_{n\to\infty} \frac{E\left(\sum_{i_k \in V_1} c_{i_k}(N_n:n)\right)}{nE(c_1(1:n))} = 0.$$

Let $B_n = \{m_1(n) \leq a_n, N_n \leq M_n\}$. Then by suitably modifying the conditioning argument given in the proof of Theorem 3.1, we obtain

$$\frac{E\left(\sum_{i_k \in V_1} c_{i_k}(N_n:n)\right)}{nE(c_1(1:n))}$$

$$
\begin{aligned}
&\leq \frac{a_n E(c_1(M_n:n))}{nE(c_1(1:n))} + \frac{E\left(\sum_{i_k\in V_k} c_{i_k}(N_n:n) \mid B_n^c\right)}{nE(c_1(1:n))} P(B_n^c) \\
&\leq \frac{a_n E(c_1(M_n:n))}{nE(c_1(1:n))} + \frac{E\left(\sum_{i=1}^n c_i(n:n)\right)}{nE(c_1(1:n))} P(B_n^c) \\
&\leq \frac{a_n E(c_1(M_n:n))}{nE(c_1(1:n))} + \frac{E(c_1(n:n))}{E(c_1(1:n))} (P(m_1(n) > a_n) + P(N_n > M_n)) \quad (11)
\end{aligned}
$$

Hence, it follows from (ii), (iii), and (iv) that the limit of the LHS of (11) is 0 as $n \to \infty$, and this completes the proof of the theorem.

Remark. The theorems proved above are not necessarily the most general results that can be obtained, but the hypotheses of the theorems are usually straightforward to check. In particular, we only need to verify certain moment conditions for the order statistics of the variables $c_{11}, c_{12}, ...c_{1n}$. Nevertheless, verifying these conditions for some cost distributions can still be quite tricky. In the next section we consider two examples, costs with a power distribution and costs with a Weibull distribution, and we indicate some of the calculations that are typically used to verify the hypotheses of the theorems. Calculations for other distributions are left to the reader.

4 Examples and Discussion

4.1 Random costs with a power distribution

Suppose that for each $i, j \geq 1$, the cost variable c_{ij} has a power distribution with parameter $\nu > 0$ and density $f(x) = \nu x^{\nu-1}$ for $0 < x < 1$. To apply Theorem 3.1, we first verify that the Central Limit Theorem holds for the array $\{c_i(1:n) : 1 \leq i \leq n, n \geq 1\}$. By Liapounov's theorem, it suffices to show that $\lim_{n\to\infty} E(c_1(1:n)^3)/\sqrt{n(Var(c_1(1:n)))^3} = 0$. It is known (see [1]) that for $1 \leq k \leq n$ and $m \geq 1$,

$$
E\left((c_1(k:n))^m\right) = \frac{\Gamma(n+1)\Gamma(k+m/\nu)}{\Gamma(n+1+m/\nu)\Gamma(k)}
$$

and Stirling's formula yields

$$
\begin{aligned}
E(c_1(1:n)) &= \frac{\Gamma(n+1)\Gamma(1+1/\nu)}{\Gamma(n+1+1/\nu)} \sim \frac{\Gamma(1+1/\nu)}{n^{1/\nu}} \quad (12) \\
Var(c_1(1:n)) &= \frac{\Gamma(n+1)\Gamma(1+2/\nu)}{\Gamma(n+1+2/\nu)} - \left(\frac{\Gamma(n+1)\Gamma(1+1/\nu)}{\Gamma(n+1+1/\nu)}\right)^2 \\
&\sim \frac{\Gamma(1+2/\nu) - (\Gamma(1+1/\nu))^2}{n^{2/\nu}} \quad (13) \\
E((c_1(1:n))^3) &= \frac{\Gamma(n+1)\Gamma(1+3/\nu)}{\Gamma(n+1+3/\nu)} \sim \frac{\Gamma(1+3/\nu)}{n^{3/\nu}}. \quad (14)
\end{aligned}
$$

Thus, it follows from Liapounov's theorem that the Central Limit Theorem holds for the array.

Next we check that conditions (a) and (b) of Theorem 3.1 are satisfied. To establish condition (a), note that for any $\beta < 0$

$$\begin{aligned} E(c_{\hat{i},\phi_{n,1}(\hat{i})}) &\leq n^\beta + \int_{n^\beta}^1 P(c_{\hat{i},\phi_{n,1}(\hat{i})} > x)dx \\ &\leq n^\beta + n\int_{n^\beta}^1 P(c_1(1:n) > x)dx \\ &= n^\beta + n\int_{n^\beta}^1 (P(c_{11} > x))^n dx \\ &= n^\beta + n\int_{n^\beta}^1 (1-x^\nu)^n dx \\ &\leq n^\beta + n(1-n^{\beta\nu})^n. \end{aligned}$$

It follows from this and from the asymptotics for $Var(c_1(1:n))$ that

$$\frac{E(c_{\hat{i},\phi_{n,1}(\hat{i})})}{\sqrt{nVar(c_1(1:n))}} \leq C(n^{\beta-1/2+1/\nu} + n^{1/2-1/\nu}(1-n^{\beta\nu})^n) \tag{15}$$

where C is a positive constant which may depend on ν, but which does not depend on n. Choose β so that $-\frac{1}{\nu} < \beta < \min(0, 1/2 - 1/\nu)$, then the RHS of (15) goes to 0 as $n \to \infty$, and this establishes condition (a).

Next we verify that consition (b) is satisfied. Stirling's formula yields

$$\frac{(\log n)E(c_1(M:n))}{\sqrt{nVar(c_1(1:n))}} \sim \frac{\Gamma(M+1/\nu)}{\Gamma(M)\sqrt{\Gamma(1+2/\nu) - (\Gamma(1+1/\nu))^2}} \frac{\log n}{\sqrt{n}}$$

so the condition is satisfied and we have

$$\frac{c(T_n) - nE(c_1(1:n))}{\sqrt{nVar(c_1(1:n))}} \to N(0,1)$$

in distribution as $n \to \infty$. It also follows from the asymptotic formulas for $E(c_1(1:n))$ and $Var(c_1(1:n))$, that

$$\frac{n^{1/\nu-1/2}c(T_n)}{\sqrt{\Gamma(1+2/\nu) - (\Gamma(1+1/\nu))^2}} - \frac{\sqrt{n}\Gamma(1+1/\nu)}{\sqrt{\Gamma(1+2/\nu) - (\Gamma(1+1/\nu))^2}} \to N(0,1)$$

in distribution as $n \to \infty$.

It is easy to check (the calculations are similar to those given above) that the hypotheses of Theorem 3.2 are satisfied and so $c(T_n)/\{nE(c_1(1:n))\}^{-1} \to 1$ in probability as $n \to \infty$. Since $E(c_1(1:n)) \sim \Gamma(1+1/\nu)n^{-1/\nu}$, we also have

$$\frac{c(T_n)}{\Gamma(1+1/\nu)n^{1-1/\nu}} \to 1$$

in probability.

Finally, the conditions of Theorem 3.3 are satisfied if we let $a_n = (1/\nu)(\log n)^{3/2}$ and $M_n = \lceil (\log n)^5 \rceil$. It is easy to check that conditions (i) and (ii) hold. To verify that conditions (iii) and (iv) hold, note that the variables $\{c_{ij}\}$ are bounded, and in particular, $E(c_1(n:n) \leq 1$ for all $n \geq 1$. Thus

$$\frac{E(c_1(n:n))P(m_1(n) > a_n)}{E(c_1(1:n))} \leq \frac{P(m_1(n) > a_n)}{E(c_1(1:n))} \sim \frac{n^{1/\nu}P(m_1(n) > a_n)}{\Gamma(1+1/\nu)}. \tag{16}$$

Applying Lemma 2.3 with $a_n = (1/\nu)(\log n)^{3/2}$, we obtain

$$n^{1/\nu} P(m_1(n) > (1/\nu)(\log n)^{3/2}) \leq \widetilde{K} n^{1/\nu} \exp(-\sqrt{2}(1/\nu)(\log n) + \sqrt{\log n/2})$$

where $\widetilde{K}$ is a positive constant that may depend on ν but which does not depend on n. Thus, when $a_n = (1/\nu)(\log n)^{3/2}$, the RHS of (16) goes to 0 as $n \to \infty$, and condition (iii) is satisfied. Likewise,

$$\frac{E(c_1(n:n))P(N_n > M_n)}{E(c_1(1:n))} \leq \frac{P(N_n > M_n)}{E(c_1(1:n))} \sim \frac{n^{1/\nu} P(N_n > M_n)}{\Gamma(1+1/\nu)}, \tag{17}$$

and it follows from Lemma 2.2, that for $M_n = \lceil (\log n)^5 \rceil$,

$$n^{1/\nu} P(N_n > M_n) \leq K n^{1/\nu} \exp(-(\log n)^{5/4})$$

where K is a positive constant which may depend on ν, but which does not depend on n. Hence the RHS of (17) goes to 0 as $n \to \infty$, condition (iv) is satisfied, and we obtain

$$\lim_{n\to\infty} \frac{E(c(T_n)}{\Gamma(1+1/\nu) n^{1-1/\nu}} = 1$$

4.2 Random costs with a Weibull distribution

Suppose that for each $i, j \geq 1$, c_{ij} has a Weibull distribution with parameter $\delta > 0$ and density $f(x) = \delta x^{\delta-1} \exp(-x^\delta)$ for $x \geq 0$. To apply Theorem 3.1, we must verify that the Central Limit Theorem holds for the array $\{c_i(1:n) : 1 \leq i \leq n, n \geq 1\}$. Recall that it suffices to show that $\lim_{n\to\infty} E(c_1(1:n)^3)/\sqrt{n(Var(c_1(1:n)))^3} = 0$. It is known (see [1]) that for $1 \leq k \leq n$ and $m \geq 1$,

$$E\left((c_1(k:n))^m\right) = \frac{n!}{(k-1)!(n-k)!} \Gamma(1+m/\delta) \sum_{r=1}^{k-1} \binom{k-1}{r} (n-k+r+1)^{-1-m/\delta}.$$

Thus

$$\begin{aligned} E(c_1(1:n)) &= \frac{\Gamma(1+1/\delta)}{n^{1/\delta}}, \\ Var(c_1(1:n)) &= \frac{\Gamma(1+2/\delta) - (\Gamma(1+1/\delta))^2}{n^{2/\delta}}, \\ E((c_1(1:n))^3) &= \frac{\Gamma(1+3/\delta)}{n^{3/\delta}} \end{aligned}$$

and hence the Central Limit Theorem holds for the array.

Next we check that conditions (a) and (b) of the theorem are satisfied. For all the calculations given below, let $\gamma = (\Gamma(1+2/\delta) - (\Gamma(1+1/\delta))^2)^{-1/2}$. To establish condition (a), note that for any $\beta \in R$,

$$E(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})}) \leq n^\beta + \int_{n^\beta}^{\infty} P(c_{\hat{\imath},\phi_{n,1}(\hat{\imath})} > x) dx$$

$$
\begin{aligned}
&\leq n^{\beta} + n\int_{n^{\beta}}^{\infty} P(c_1(1:n) > x)dx \\
&= n^{\beta} + n\int_{n^{\beta}}^{\infty} (P(c_{11} > x))^n dx \\
&= n^{\beta} + n\int_{n^{\beta}}^{\infty} e^{-nx^{\delta}} dx \\
&= n^{\beta} + n^{1+1/\delta}\int_{n^{\beta+1/\delta}}^{\infty} e^{-u^{\delta}} du.
\end{aligned}
$$

Thus

$$
\frac{E(c_{\hat{i},\phi_{n,1}(\hat{i})})}{\sqrt{nVar(c_1(1:n))}} \leq \gamma\left(n^{\beta-1/2+1/\delta} + n^{1/2+2/\delta}\int_{n^{\beta+1/\delta}}^{\infty} e^{-u^{\delta}} du\right), \tag{18}
$$

and, provided we choose β so that $-\frac{1}{\delta} < \beta < \frac{1}{2} - \frac{1}{\delta}$, the RHS of (18) goes to 0 as $n \to \infty$. This establishes condition (a).

To check that condition (b) is satisfied we must obtain a bound for $E(c_1(M : n)$. It is difficult to work directly with the exact expression for $E(c_1(M : n))$, so we employ a different approach. For each $x > 0$ and $1 \leq j \leq n$, define the indicator function $I_j(x)$ by setting $I_j(x) = 1$ if $c_{1j} \leq x$ (and $I_j(x) = 0$ otherwise). Let $S_n(x) = \sum_{j=1}^{n} I_j(x)$, then for any $\beta \in R$

$$
\begin{aligned}
E(c_1(M:n)) &\leq n^{\beta} + \int_{n^{\beta}}^{\infty} P(c_1(M:n) > x)dx \\
&= n^{\beta} + \int_{n^{\beta}}^{\infty} P(S_n(x) < M)dx.
\end{aligned}
$$

There are two cases to consider. First, suppose that $\delta > 2$ and choose β such that $0 < \beta < 1/2 - 1/\delta$, then

$$
\frac{\log n E(c_1(M:n))}{\sqrt{nVar(c_1(1:n))}} \leq \gamma(\log n)n^{\beta-1/2+1/\delta} + \gamma(\log n)n^{1/\delta-1/2}\int_{n^{\beta}}^{\infty} P(S_n(x) < M)dx. \tag{19}
$$

Clearly the first term on the RHS of (19) goes to zero as $n \to \infty$ (by choice of β). Next, note that $E(S_n(x)) = n(1-e^{-x^{\delta}})$ and $Var(S_n(x)) = ne^{-x^{\delta}}(1-e^{-x^{\delta}})$, so

$$
\begin{aligned}
&\gamma(\log n)n^{1/\delta-1/2}\int_{n^{\beta}}^{\infty} P(S_n(x) < M)dx \\
\leq\ &\frac{\gamma(\log n)}{n^{1/2-1/\delta}}\int_{n^{\beta}}^{\infty} P(|S_n(x) - n(1-e^{-x^{\delta}})| > n(1-e^{-x^{\delta}}) - M)dx \\
\leq\ &\frac{\gamma(\log n)}{n^{1/2-1/\delta}}\int_{n^{\beta}}^{\infty} \frac{ne^{-x^{\delta}}(1-e^{-x^{\delta}})}{(n(1-e^{-x^{\delta}}) - M)^2}dx \\
\leq\ &\frac{\gamma(\log n)}{(1-\exp(-n^{\beta\delta}) - M/n)^2 n^{3/2-1/\delta}}\int_{n^{\beta}}^{\infty} e^{-x^{\delta}}dx.
\end{aligned} \tag{20}
$$

Thus the second term on the RHS of (19) goes to 0 as $n \to \infty$ since the RHS of (20) goes to 0 as $n \to \infty$.

If $\delta \leq 2$, choose β such that $-1/\delta < \beta < 1/2 - 1/\delta < 0$. Again, by choice of β, the first term on the RHS of (19) goes to 0 as $n \to \infty$. Next, observe that $P(S_n(x) < M)$ is a decreasing function of x, so

$$\begin{aligned} & \gamma(\log n) n^{1/\delta - 1/2} \int_{n^\beta}^{\infty} P(S_n(x) < M) dx \\ \leq \; & \gamma(\log n) n^{\alpha + 1/\delta - 1/2} P(S_n(n^\beta) < M) + \gamma(\log n) n^{1/\delta - 1/2} \int_{n^\alpha}^{\infty} P(S_n(x) < M) dx, \end{aligned} \tag{21}$$

for any $\alpha > 0$. Observe that $1 - e^{-n^{\beta\delta}} \leq n^{\beta\delta}$ since $\beta < 0$. So

$$\begin{aligned} P(S_n(n^\beta) < M) &= \sum_{m=0}^{M-1} \binom{n}{m} (1 - e^{-n^{\beta\delta}})^m (e^{-n^{\beta\delta}})^{n-m} \\ &\leq \sum_{m=0}^{M-1} (n^{1+\beta\delta})^m (e^{-n^{\beta\delta}})^{n-m} \\ &\leq \log n (n^{(1+\beta\delta)\log n}) (e^{-(1/2)n^{1+\beta\delta}}). \end{aligned}$$

Since $1 + \beta\delta > 0$, it follows from this inequality that the first term on the RHS of (21) tends to 0 as $n \to \infty$. To bound the second term on the RHS of (21), we repeat the calculations used to obtain inequality (20) to obtain

$$\frac{\gamma(\log n)}{n^{1/2 - 1/\delta}} \int_{n^\alpha}^{\infty} P(S_n(x) < M) dx \leq \frac{\gamma(\log n)}{(1 - \exp(-n^{\alpha\delta}) - M/n)^2 n^{3/2 - 1/\delta}} \int_{n^\alpha}^{\infty} e^{-x^\delta} dx. \tag{22}$$

Thus the second term on the RHS of (21) goes to 0 as $n \to \infty$, since the RHS of (22) goes to 0 as $n \to \infty$. So for all $\delta > 0$, condition (b) of the theorem is satisfied and we have

$$\frac{c(T_n) - \Gamma(1 + 1/\delta) n^{1 - 1/\delta}}{\sqrt{n Var(c_1(1:n))}} \to N(0, 1)$$

in distribution as $n \to \infty$.

The weak law for $c(T_n)$ follows immediately from the above calculations. In particular, it is trivial to check that $\frac{c(\phi_{n,1})}{\Gamma(1+1/\delta) n^{1-1/\delta}} \to 1$ in probability as $n \to \infty$. Also, both conditions (a) and (b) of Theorem 3.2 are satisfied since the hypotheses of Theorem 3.1 are satisfied and since $nE(c_1(1:n)) > \sqrt{nVar(c_1(1:n))}$ for all sufficiently large n.

To apply Theorem 3.3, let $a_n = (1+1/\delta)(\log n)^{3/2}$ and $M_n = \lceil (\log n)^5 \rceil$, then condition (i) is satisfied and, by suitably modifying the estimates calculated above, it is easy to check that condition (ii) is satisfied. To verify that conditions (iii) and (iv) hold, let $F(x)$ denote the distribution function of c_{11} and note that

$$\begin{aligned} E(c_1(n:n)) &= \int_0^\infty P(c_1(n:n) > x) dx \\ &= \int_0^\infty 1 - (F(x))^n dx \\ &\leq \lim_{x \to \infty} x P(c_1(n:n) > x) + n \int_0^\infty x f(x) dx \\ &= nE(c_{11}). \end{aligned}$$

Therefore, it follows from Lemma 2.3 that

$$\begin{aligned}\frac{E(c_1(n:n))}{E(c_1(1:n))}P(m_1(n) > a_n) &\leq \frac{n^{1+1/\delta}E(c_{11})}{\Gamma(1+1/\delta)}P(m_1(n) > a_n) \\ &\leq \widetilde{K}n^{1+1/\delta}\exp(-\sqrt{2}(1+1/\delta)\log n + \sqrt{\log n}/2)\end{aligned} \tag{23}$$

where $\widetilde{K}$ is a constant which may depend on δ, but which does not depend on n. Thus condition (iii) is satisfied since the RHS of (23) goes to 0 as $n \to \infty$. Similarly, it follows from Lemma 2.2, that

$$\begin{aligned}\frac{E(c_1(n:n))}{E(c_1(1:n))}P(N_n > M_n) &\leq \frac{n^{1+1/\delta}E(c_{11})}{\Gamma(1+1/\delta)}P(N_n > M_n) \\ &\leq Kn^{1+1/\delta}\exp(-(\log n)^{5/3})\end{aligned} \tag{24}$$

where K is a positive constant which may depend on δ, but which does not depend on n. Thus condition (iv) is satisfied since the RHS of (24) goes to 0 as $n \to \infty$ and we have

$$\frac{E(c(T_n))}{\Gamma(1+1/\delta)n^{1-1/\delta}} \to 1$$

as $n \to \infty$.

Remark. Note that the uniform distribution corresponds to a power distribution with $\nu = 1$ and the exponential distribution with mean 1 corresponds to a Weibull distribution with $\delta = 1$. So the results obtained above for random costs with either a power distribution or a Weibull distribution complement and extend those obtained by McDiarmid [9].

There are various ways in which the results in this paper could be extended and generalized. For example, much of the analysis of the algorithm remains the same under the assumption that for each $n \geq 1$ and $1 \leq i \leq n$, the *vectors* $(c_{i1}, c_{i2}, \ldots, c_{in})$ are i.i.d. and with a common nondegenerate, continuous and exchangeable joint distribution. In particular, the distribution of $m_1(n)$ and N_n remain the same and the statements of the theorems are unchanged. However, it becomes more difficult to check the moment conditions for the order statistics of the costs under the assumption of exchangeability.

It would be interesting to determine whether the methods of this paper could be applied to other problems. Basically, results for $c(T_n)$ can be obtained because $c(T_n)$ is bounded below by a sum of random variables whose distribution is known and it is bounded above by a sum of variables which is obtained by changing a relatively small number of values in the sum which bounds $c(T_n)$ from below. The magnitude of the difference between $c(T_n)$ and its lower bound is a function of the number of terms in the sum that are altered (which is just $m_1(n)$) and of the typical magnitude of an alteration of any given term (which depends on the number of iterations, N_n, of the algorithm). This method works because, for random mappings, $m_1(n)$=$O(\log n)$ with high probability and because size of the largest component of a random mapping is typically

$\Omega(n)$, which guarantees, by Ross's theorem, that with high probability the number of iterations of the algorithm will be relatively small compared to n.

Random mappings are one example of a random logarithmic combinatorial structure (other examples include random permutations and 2-regular graphs). A random logarithmic structure of size n decomposes into disjoint 'components' such that, with high probability, the number of components is $O(\log n)$ and the size of the largest component is $\Omega(n)$ (see [2], [6], [7]). Hence, it may be possible to devise other algorithms to solve other optimisation problems so that the algorithm sequentially modifies a random logarithmic structure and it can be analyzed by methods similar to those used above. In fact, the 'patching algorithm' of Karp and Steele [8] which modifies a minimal cost random permutation to obtain a nearly optimal assignment does this but unfortunately precise results for the cost of the minimal random permutation are not yet known.

Acknowledgement. The author would like to thank Colin McDiarmid for suggesting simplifications of the statements of Theorems 3.1 and 3.2.

References

[1] Arnold, B. C., Balakrishan, N. and Nagaraja, H. N. (1992) *A First Course in Order Statistics*. Wiley.

[2] Arratia, R., Stark, D. and Tavare, S. (1995) Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies. *Ann. of Probab.* **23** 1347-1388.

[3] Bollobás, B. (1985) *Random Graphs*. Academic Press.

[4] Chung, K. L. (1974) *A Course in Probability Theorey*, 2nd Ed. Academic Press.

[5] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, Vol. 2. Wiley.

[6] Flajolet, P. and Soria, M. (1993) General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Math.* **114** 159-180.

[7] Hansen, J. C. (1994) Order statstics for decomposable combinatorial structures. *Rand. Structures and Algorithms* **5** 517-533.

[8] Karp, R. M. and Steele, J. M. (1985) Probabalistic analysis of heuristics. In E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan and D. B. Shmoys (editors), *The Travelling Salesman Problem* 181-205. Wiley.

[9] McDiarmid, C. J. M. (1986) On the greedy algorithm with random costs. *Math. Program.* **36** 245-255.

[10] Ross, S. M. (1981) A random graph. *J. Appl. Probab.* **18** 309-315.