



**HAL**  
open science

# Random walk in the complete graph: hitting and cover times

François Castella, Bruno Sericola

► **To cite this version:**

François Castella, Bruno Sericola. Random walk in the complete graph: hitting and cover times. Centre Inria. 2024, pp.27. hal-04365696v2

**HAL Id: hal-04365696**

**<https://inria.hal.science/hal-04365696v2>**

Submitted on 17 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Random walk in the complete graph : hitting and cover times

François Castella<sup>1</sup> and Bruno Sericola<sup>2\*</sup>

<sup>1</sup>University of Rennes, Campus de Beaulieu, Rennes, 35042, France.

<sup>2</sup>Inria, Campus de Beaulieu, Rennes, 35042, France.

\*Corresponding author(s). E-mail(s): [bruno.sericola@inria.fr](mailto:bruno.sericola@inria.fr);

Contributing authors: [francois.castella@univ-rennes.fr](mailto:francois.castella@univ-rennes.fr);

## Abstract

We consider the moments and the distribution of the hitting and cover times of a random walk in the complete graph. We study both the time needed to reach any subset of states and the time needed to visit all the states of a subset at least once. We obtain recurrence relations for the moments of all orders and we use these relations to analyze the asymptotic behavior of the hitting and cover times distributions when the number of states tends to infinity.

**Keywords:** Markov chain, Random walk, Hitting times, Cover times, Complete graph, Moments, Asymptotic distribution

## 1 Introduction

In a random walk, the hitting time of a given subset of states is the first time at which the random walk reaches one of its states, while the cover time of a given subset of states is the first time at which the random walk reaches all its states. These random variables are quite important in many areas of distributed systems and networks. Many applications of random walks on graphs are proposed in [1] as the analysis of electrical networks which are detailed in [2]. Conversely, electric networks are of great help to obtain expected values about random walks on graphs that would be more difficult to attain without the electric intuition. They have been extensively used to analyze particular random graphs, see for instance [3], [4] and the references therein. Visiting all the nodes of such systems allows the user to collect information about the performance or the reliability of the system. These visits can be made using random

walks which are a particular case of Markov chains, since the transition probabilities from a given node are uniformly distributed among the number of its neighbours, but they can also be made using general transition probability matrices. Obviously the visiting time of all the nodes is a critical performance measure. For instance, in the context of graph theory for social networks, a notion of centrality has been used to assess the relative importance of nodes in a given network topology and a novel form of centrality based on the second order moments of return times in random walks was proposed in [5].

Several papers have already been written on both the hitting and cover times and most of these papers only consider expected values of these random variables. Very interesting surveys on random walks on graphs are proposed in [6] and [7], while [8] and [9] additionally contains many applications in both computer science and networks. It is shown in [10] and [11] that the expected cover time in a random walk with  $n$  states is at least  $(1 - o(1))n \ln(n)$  and at most  $(4/27 + o(1))n^3$ . In the case of a regular graph, it is at most  $2n^2$ , see [10]. In [12], the authors first prove that for any transition probability matrix, both the expected hitting time and the expected cover time of the path graph is  $\Omega(n^2)$ . They also prove that for some particular Markov chains called degree-biased random walks, the expected hitting time is  $O(n^2)$  and the expected cover time is  $O(n^2 \ln(n))$ . The method proposed in this paper is based on a matrix analysis of the Markov chain evolution equations. The advantage of our method is that it allows to deal with higher moments and with the distribution of the hitting and cover times. This method has been used in [13] to analyze the hitting times on the lollipop graph which consists of a complete graph (a clique) and a path graph. One end of the path graph is connected to one vertex of the clique, and its other end is the target vertex for the hitting times. We have proposed a general recursive way to obtain all the moments of the hitting times on the lollipop graph and we have used this recursion to analyze the asymptotic behavior of both the moments and the distribution of the hitting times when the number of nodes tends to infinity.

In this paper, we consider the moments, the distribution and the asymptotic behavior of the hitting and cover times of a random walk in the complete graph, with and without self-loops. The complete graph or full mesh topology is expensive to design but it clearly provides a high degree of reliability since the failures of several edges will not impact communications. It is quite important as for example in the military domain. Our analysis is based on the general results obtained in [14] for hitting and cover times in general Markov chains.

The remainder of the paper is organized as follows. In section 2, we recall the general results of [14] on the analysis of the hitting time and the cover time of a subset of the state space both in terms of distribution and moments. In section 3, we analyze these results when the graph is a complete graph with self-loops. We obtain simple recurrence relations to compute all the moments of both the hitting and cover times. We also give the asymptotic behavior of these moments when the number of states tends to infinity and we deduce the corresponding asymptotic distributions. We consider in section 4 the case of the complete graph without self-loops. All these results are illustrated using numerical values. Section 5 concludes the paper.

## 2 Basic results

Consider a homogeneous discrete-time Markov chain  $X = \{X_n, n \in \mathbb{N}\}$  over a finite state space  $S$ . We denote by  $P$  the transition probability matrix of  $X$ . We suppose that  $X$  is irreducible and, for every non empty subset  $A$  of  $S$ , we define the hitting time  $T_A$  of subset  $A$  as the first instant at which one state of  $A$  is reached by the Markov chain  $X$ , *i.e.*

$$T_A = \inf\{n \geq 0 \mid X_n \in A\}.$$

The irreducibility of  $X$  implies that, for every  $A \subset S$ ,  $T_A$  is finite almost surely. We also define, for every non empty subset  $A$  of  $S$ , the cover time  $C_A$  of subset  $A$  as the first instant at which all the states of  $A$  have been reached by the Markov chain  $X$ , *i.e.*

$$C_A = \inf\{n \geq 0 \mid A \subseteq \{X_0, \dots, X_n\}\}.$$

The cover time of the Markov chain  $X$ , which is the time needed to reach all the states of  $S$ , is simply denoted by  $C$  and is given by  $C = C_S$ . The complementary subset of  $A$  in  $S$ ,  $S \setminus A$ , is denoted simply by  $A^c$ . The partition  $\{A, A^c\}$  of  $S$  induces a decomposition of  $P$  into four submatrices as

$$P = \begin{pmatrix} P_A & P_{A,A^c} \\ P_{A^c,A} & P_{A^c} \end{pmatrix},$$

where matrix  $P_A$  (resp.  $P_{A^c}$ ) contains the transition probabilities between states of  $A$  (resp.  $A^c$ ) and matrix  $P_{A,A^c}$  (resp.  $P_{A^c,A}$ ) contains the transition probabilities from states of  $A$  (resp.  $A^c$ ) to states of  $A^c$  (resp.  $A$ ). We also denote by  $\mathbb{1}$  the column vector with all its entries equal to 1, its dimension being specified by the context. With these notations, we have, for all  $i \in A^c$ ,  $j \in A$  and  $\ell \geq 1$ , see for instance [15],

$$\mathbb{P}\{T_A = \ell, X_{T_A} = j \mid X_0 = i\} = (P_{A^c}^{\ell-1} P_{A^c,A})_{i,j}. \quad (1)$$

Summing (1) over all  $j \in A$ , leads to

$$\mathbb{P}\{T_A = \ell \mid X_0 = i\} = (P_{A^c}^{\ell-1} P_{A^c,A} \mathbb{1})_i. \quad (2)$$

In the same way, summing (1) over all  $\ell \geq 1$ , yields

$$\mathbb{P}\{X_{T_A} = j \mid X_0 = i\} = \sum_{\ell=1}^{\infty} (P_{A^c}^{\ell-1} P_{A^c,A})_{i,j} = ((I - P_{A^c})^{-1} P_{A^c,A})_{i,j}. \quad (3)$$

From (1), we also obtain

$$\mathbb{E}(T_A \mathbb{1}_{\{X_{T_A}=j\}} \mid X_0 = i) = \sum_{\ell=1}^{\infty} \ell (P_{A^c}^{\ell-1} P_{A^c,A})_{i,j} = ((I - P_{A^c})^{-2} P_{A^c,A})_{i,j}. \quad (4)$$

Using Theorem 1 of [13], the moments of  $T_A$  and  $T_A 1_{\{X_{T_A}=j\}}$  are given, for every  $r \geq 1$ ,  $i \in A^c$  and  $j \in A$ , by

$$\mathbb{E}(T_A^r 1_{\{X_{T_A}=j\}} | X_0 = i) = \sum_{\ell=1}^r c_{r,\ell} \left( (I - P_{A^c})^{-(\ell+1)} P_{A^c, A} \right)_{i,j} \quad (5)$$

and

$$\mathbb{E}(T_A^r | X_0 = i) = \sum_{\ell=1}^r c_{r,\ell} \left( (I - P_{A^c})^{-\ell} \mathbb{1} \right)_i \quad (6)$$

where, for every  $\ell = 1, \dots, r$ ,

$$c_{r,\ell} = (-1)^r \sum_{j=1}^{\ell} (-1)^j \binom{\ell}{j} j^r. \quad (7)$$

This gives for instance

$$\begin{aligned} \mathbb{E}(T_A | X_0 = i) &= \left( (I - P_{A^c})^{-1} \mathbb{1} \right)_i, \\ \mathbb{E}(T_A^2 | X_0 = i) &= - \left( (I - P_{A^c})^{-1} \mathbb{1} \right)_i + 2 \left( (I - P_{A^c})^{-2} \mathbb{1} \right)_i, \\ \mathbb{E}(T_A^3 | X_0 = i) &= \left( (I - P_{A^c})^{-1} \mathbb{1} \right)_i - 6 \left( (I - P_{A^c})^{-2} \mathbb{1} \right)_i + 6 \left( (I - P_{A^c})^{-3} \mathbb{1} \right)_i. \end{aligned}$$

Let us now consider the cover time  $C_A$  of subset  $A$ . First of all, for any event  $u$  and for any random variable  $U$ , we introduce the notation  $\mathbb{P}_i\{u\} = \mathbb{P}\{u | X_0 = i\}$  and  $\mathbb{E}_i(U) = \mathbb{E}(U | X_0 = i)$ . Note that, for every  $j \in S$ , we have

$$C_{\{j\}} = T_{\{j\}}.$$

Moreover, for every  $\ell \geq 0$  and for every  $i \in A$ , with  $|A| \geq 2$ , we have, by definition of  $C_A$ ,

$$\mathbb{P}_i\{C_A = \ell\} = \mathbb{P}_i\{C_{A \setminus \{i\}} = \ell\}.$$

We thus only need to consider the case where the initial state belongs to the subset  $A^c$ . If  $\alpha$  denotes the initial distribution of  $X$ , we get

$$\mathbb{P}\{C_A = \ell\} = \sum_{i \in S} \alpha_i \mathbb{P}_i\{C_A = \ell\} = \sum_{i \in A} \alpha_i \mathbb{P}_i\{C_{A \setminus \{i\}} = \ell\} + \sum_{i \in A^c} \alpha_i \mathbb{P}_i\{C_A = \ell\}.$$

The distribution of the cover time  $C$  of the Markov chain is then given by

$$\mathbb{P}\{C = \ell\} = \sum_{i \in S} \alpha_i \mathbb{P}_i\{C = \ell\} = \sum_{i \in S} \alpha_i \mathbb{P}_i\{C_{S \setminus \{i\}} = \ell\}.$$

The following recursive expression of the distribution of the cover time  $C_A$  of subset  $A$  has been proved in [14]. For every  $A \subset S$  such that  $|A| \geq 2$ ,  $i \in A^c$  and  $\ell \geq 0$ , we have

$$\mathbb{P}_i\{C_A = \ell\} = \begin{cases} 0 & \text{if } 0 \leq \ell < |A| \\ \sum_{h=1}^{\ell-|A|+1} \sum_{j \in A} (P_{A^c}^{h-1} P_{A^c, A})_{i,j} \mathbb{P}_j\{C_{A \setminus \{j\}} = \ell - h\} & \text{if } \ell \geq |A|. \end{cases} \quad (8)$$

We also obtained in [14] the following recursive expression of the moments of the cover time of subset  $A$  starting from state  $i \in A^c$ . For every  $A \subset S$  such that  $|A| \geq 2$ ,  $r \geq 1$  and  $i \in A^c$ , we have

$$\mathbb{E}_i(C_A^r) = \mathbb{E}_i(T_A^r) + \sum_{m=1}^r \binom{r}{m} \sum_{j \in A} \mathbb{E}_i(T_A^{r-m} 1_{\{X_{T_A}=j\}}) \mathbb{E}_j(C_{A \setminus \{j\}}^m). \quad (9)$$

In the following sections, we apply these results to the complete graph to obtain the asymptotic behavior of the moments and distribution of both the hitting and cover times.

### 3 Analysis of the complete graph with self-loops

A random walk on the complete graph with self-loops is a discrete-time Markov chain on the state space  $S = \{0, 1, \dots, n-1\}$ ,  $n \geq 2$ , with transition probability matrix  $P$  given, for every  $u, v \in S$ , by  $P_{u,v} = 1/n$ .

#### 3.1 Hitting times on the complete graph with self-loops

Let  $A$  be a proper subset of  $S$  with  $k$  states, that is  $|A| = k \in \{1, \dots, n-1\}$ . The symmetry of the complete graph implies that the hitting time distribution of subset  $A$  is the same for any subset of  $S$  containing  $k$  states. We thus simply use the notation  $T_k$  for  $T_A$ . Moreover, we suppose without any loss of generality that the initial state always belongs to the complementary subset  $A^c$  of  $A$ . To simplify the notation, using the symmetry again, we will sometimes drop the initial state because the results remain the same for every initial state  $i \in A^c$ .

Using relation (2), we have for every  $\ell \geq 0$  and  $i \in A^c$ ,

$$\mathbb{P}_i\{T_k > \ell\} = \sum_{j=\ell+1}^{\infty} (P_{A^c}^{j-1} P_{A^c, A} \mathbb{1})_i = (P_{A^c}^{\ell} \mathbb{1})_i.$$

Since  $|A^c| = n - k$  we have  $P_{A^c} \mathbb{1} = (1 - k/n) \mathbb{1}$  which recursively leads to

$$P_{A^c}^{\ell} \mathbb{1} = \left(1 - \frac{k}{n}\right)^{\ell} \mathbb{1} \text{ and thus } \mathbb{P}\{T_k > \ell\} = \left(1 - \frac{k}{n}\right)^{\ell}, \quad (10)$$

which means that  $T_k$  is geometrically distributed with parameter  $k/n$ . Using relation (6), we have, for every  $r \geq 1$ ,

$$\mathbb{E}(T_k^r) = \sum_{\ell=1}^r c_{r,\ell} \left(\frac{n}{k}\right)^\ell, \quad (11)$$

where the coefficients  $c_{r,\ell}$  have been defined in (7). The first moments of  $T_k$  are then given, using (7), by

$$\mathbb{E}(T_k) = \frac{n}{k}, \mathbb{E}(T_k^2) = \frac{n(2n-k)}{k^2}, \mathbb{E}(T_k^3) = \frac{n(6n^2 - 6kn + k^2)}{k^3}.$$

Since  $c_{r,r} = r!$  we obtain, for every fixed value of  $k$ ,

$$\mathbb{E}(T_k^r) \underset{n \rightarrow \infty}{\sim} r! \left(\frac{n}{k}\right)^r$$

and, from relation (10), we get, for every fixed value of  $k$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{kT_k/n > t\} = e^{-t},$$

which means that the sequence of random variables  $kT_k/n$  converges in distribution, when  $n$  tends to infinity, to a random variable that is exponentially distributed with rate 1.

### 3.2 Cover times on the complete graph with self-loops

In the same way, if  $A$  is a proper subset of  $S$  with  $k$  states, that is  $|A| = k \in \{1, \dots, n-1\}$  then the symmetry of the complete graph implies that the cover time distribution of subset  $A$  is the same for any subset of  $S$  containing  $k$  states. We thus simply use the notation  $C_k$  for  $C_A$ . Moreover, we suppose without any loss of generality, as we did for the hitting times, that the initial state always belongs to the complementary subset  $A^c$  of  $A$ . In the same way, to simplify the notation, using the symmetry again, we will sometimes drop the initial state because the results remain the same for every initial state  $i \in A^c$ .

We then have for any initial distribution,

$$\mathbb{E}(C) = \mathbb{E}_i(C_{S \setminus \{i\}}) = \mathbb{E}_i(C_{n-1}) = \mathbb{E}(C_{n-1}).$$

For every  $n \geq 1$ , we denote by  $h_n$  the harmonic sum, *i.e.*

$$h_n = \sum_{k=1}^n \frac{1}{k}$$

and we set  $h_0 = 0$ . The symmetry of the complete graph implies that  $\mathbb{E}_i(C_A^r)$  has the same value for every subset  $A$  with  $k$  states and for every  $i \in A^c$ . We then write

$\mathbb{E}_i(C_A^r) = \mathbb{E}(C_k^r)$ . In the same way,  $\mathbb{E}_j(C_{A \setminus \{j\}}^m)$  is related to a set with  $k - 1$  states and is the same for every  $j \in A$ , so we write  $\mathbb{E}_j(C_{A \setminus \{j\}}^m) = \mathbb{E}(C_{k-1}^m)$ . Relation (9) can then be written, for  $k = 1, \dots, n - 1$  and  $r \geq 1$ , as

$$\begin{aligned} \mathbb{E}(C_k^r) &= \mathbb{E}(T_k^r) + \sum_{m=1}^r \binom{r}{m} \mathbb{E}(C_{k-1}^m) \sum_{j \in A} \mathbb{E}_i(T_A^{r-m} 1_{\{X_{T_A}=j\}}) \\ &= \mathbb{E}(T_k^r) + \sum_{m=1}^r \binom{r}{m} \mathbb{E}(C_{k-1}^m) \mathbb{E}(T_k^{r-m}). \end{aligned} \quad (12)$$

The following result gives the mean cover time of a subset of states with  $k$  states in the complete graph.

**Theorem 1.** *For every  $k = 1, \dots, n - 1$ , we have*

$$\mathbb{E}(C_k) = nh_k.$$

*In particular, when  $k = n - 1$ , we have*

$$\mathbb{E}(C) \underset{n \rightarrow \infty}{\sim} n \ln(n).$$

*Proof.* Let  $A$  be a subset of  $S$  with  $k$  states. The expected cover time of subset  $A$  is given by relation (12) with  $r = 1$ . This gives, using the previous results on the hitting times

$$\mathbb{E}(C_k) = \mathbb{E}(T_k) + \mathbb{E}(C_{k-1}) = \frac{n}{k} + \mathbb{E}(C_{k-1}).$$

with  $\mathbb{E}(C_1) = \mathbb{E}(T_1) = n$ . This simple recursion leads to

$$\mathbb{E}(C_k) = nh_k.$$

By taking  $k = n - 1$ , we get  $\mathbb{E}(C) = \mathbb{E}(C_{n-1}) = nh_{n-1}$  and thus

$$\mathbb{E}(C) \underset{n \rightarrow \infty}{\sim} n \ln(n),$$

which completes the proof.  $\square$

Observe that this well-known result can also be obtained using geometric random variables. Indeed, it is easy to check that

$$\mathbb{E}(C) = \mathbb{E}\left(\text{geo}\left(\frac{n-1}{n}\right)\right) + \mathbb{E}\left(\text{geo}\left(\frac{n-2}{n}\right)\right) + \dots + \mathbb{E}\left(\text{geo}\left(\frac{1}{n}\right)\right) = nh_{n-1}.$$

Nevertheless the relations obtained in the proof of Theorem 1 are important because they are needed to get the next results. In the following theorem we consider the second order moment of the cover time of a subset of states with  $k$  states.



**Theorem 2.** For every  $k = 1, \dots, n - 1$ , we have

$$\mathbb{E}(C_k^2) = n^2 h_k^2 + n^2 \sum_{\ell=1}^k \frac{1}{\ell^2} - n h_k.$$

In particular, when  $k = n - 1$ , we have

$$\mathbb{E}(C^2) \underset{n \rightarrow \infty}{\sim} n^2 \ln^2(n).$$

*Proof.* Let  $A$  be a subset of  $S$  with  $k$  states. The second order moment of the cover time of subset  $A$  is given by relation (12) with  $r = 2$ , that is

$$\mathbb{E}(C_k^2) = \mathbb{E}(T_k^2) + 2\mathbb{E}(C_{k-1})\mathbb{E}(T_k) + \mathbb{E}(C_{k-1}^2).$$

The previous results on the hitting times and Theorem 1 which give  $\mathbb{E}(T_k) = n/k$  and  $\mathbb{E}(C_{k-1}) = n h_{k-1}$ , yield

$$\mathbb{E}(C_k^2) = \frac{n(2n - k)}{k^2} + \frac{2n^2 h_{k-1}}{k} + \mathbb{E}(C_{k-1}^2),$$

with  $\mathbb{E}(C_1^2) = \mathbb{E}(T_1^2) = n(2n - 1)$ . This leads to

$$\mathbb{E}(C_k^2) = n \sum_{\ell=1}^k \frac{2n - \ell}{\ell^2} + 2n^2 \sum_{\ell=2}^k \frac{h_{\ell-1}}{\ell} = 2n^2 \left( \sum_{\ell=2}^k \frac{h_{\ell-1}}{\ell} + \sum_{\ell=1}^k \frac{1}{\ell^2} \right) - n h_k. \quad (13)$$

We consider separately the term between parenthesis. Introducing the notation

$$\psi_k = \sum_{\ell=2}^k \frac{h_{\ell-1}}{\ell} \quad \text{and} \quad \varphi_k = \sum_{\ell=1}^k \frac{h_\ell}{\ell},$$

we have

$$\begin{aligned} \psi_k + \sum_{\ell=1}^k \frac{1}{\ell^2} &= \varphi_k = \sum_{\ell=1}^k \frac{1}{\ell} \sum_{r=1}^{\ell} \frac{1}{r} = \sum_{r=1}^k \frac{1}{r} \sum_{\ell=r}^k \frac{1}{\ell} = \sum_{r=1}^k \frac{1}{r} (h_k - h_{r-1}) \\ &= h_k^2 - \sum_{r=2}^k \frac{h_{r-1}}{r} = h_k^2 - \psi_k. \end{aligned}$$

It follows that

$$\psi_k = \frac{h_k^2}{2} - \frac{1}{2} \sum_{\ell=1}^k \frac{1}{\ell^2}.$$

Inserting this relation inside relation (13) leads to

$$\mathbb{E}(C_k^2) = n^2 h_k^2 + n^2 \sum_{\ell=1}^k \frac{1}{\ell^2} - n h_k.$$

By taking  $k = n - 1$ , we get

$$\mathbb{E}(C^2) = \mathbb{E}(C_{n-1}^2) = n^2 h_{n-1}^2 + n^2 \sum_{\ell=1}^{n-1} \frac{1}{\ell^2} - n h_{n-1} \quad \text{and} \quad \mathbb{E}(C^2) \underset{n \rightarrow \infty}{\sim} n^2 \ln^2(n),$$

which completes the proof.  $\square$

The variance of the cover time is then given by

$$\mathbb{V}(C) = \mathbb{E}(C^2) - (\mathbb{E}(C))^2 = n^2 \sum_{\ell=1}^{n-1} \frac{1}{\ell^2} - n h_{n-1}$$

and we have

$$\mathbb{V}(C) \underset{n \rightarrow \infty}{\sim} \frac{n^2 \pi^2}{6}.$$

We illustrate this result through Table 1, which gives the variance of the normalized cover time  $\mathbb{V}(C/n) = \sum_{\ell=1}^{n-1} 1/\ell^2 - h_{n-1}/n$  and its limit  $\pi^2/6$  when  $n$  tends to infinity, for different values of  $n$ .

| $\mathbb{V}(C/10^2)$ | $\mathbb{V}(C/10^3)$ | $\mathbb{V}(C/10^4)$ | $\mathbb{V}(C/10^5)$ | $\mathbb{V}(C/10^6)$ | $\mathbb{V}(C/10^7)$ |
|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 1.5831101250         | 1.6364490958         | 1.6438553112         | 1.6448031654         | 1.6449186741         | 1.6449322973         |

**Table 1** Values of  $\mathbb{V}(C/n)$  for different values of  $n$  and its limiting value  $\pi^2/6 \approx 1.6449340668$  when  $n$  tends to infinity.

Observe that the variance of the normalized cover time  $\mathbb{V}(C/n)$  is very close to its limiting value  $\pi^2/6$  from  $n = 1000$ .

We now consider the third order moment of the cover time of a subset of states with  $k$  states.

**Theorem 3.** *For every  $k = 1, \dots, n - 1$ , we have*

$$\mathbb{E}(C_k^3) = 3n^3 \sum_{\ell=1}^k \frac{h_\ell}{\ell^2} + 3n^3 \sum_{\ell=2}^k \frac{h_{\ell-1}^2}{\ell} + 3n^3 h_k \sum_{\ell=1}^k \frac{1}{\ell^2} - 6n^2 \sum_{\ell=1}^k \frac{h_\ell}{\ell} + n h_k.$$

*In particular, when  $k = n - 1$ , we have*

$$\mathbb{E}(C^3) \underset{n \rightarrow \infty}{\sim} n^3 \ln^3(n).$$

*Proof.* Let  $A$  be a subset of  $S$  with  $k$  states. The third order moment of the cover time of subset  $A$  is given by relation (12) with  $r = 3$ , that is

$$\mathbb{E}(C_k^3) = \mathbb{E}(T_k^3) + 3\mathbb{E}(C_{k-1})\mathbb{E}(T_k^2) + 3\mathbb{E}(C_{k-1}^2)\mathbb{E}(T_k) + \mathbb{E}(C_{k-1}^3),$$

with  $\mathbb{E}(C_1^3) = \mathbb{E}(T_1^3)$ . This gives

$$\mathbb{E}(C_k^3) = \sum_{\ell=1}^k \mathbb{E}(T_\ell^3) + 3 \sum_{\ell=2}^k \mathbb{E}(C_{\ell-1})\mathbb{E}(T_\ell^2) + 3 \sum_{\ell=2}^k \mathbb{E}(C_{\ell-1}^2)\mathbb{E}(T_\ell).$$

All the terms on the right hand side have already been calculated in the previous section on hitting times and in Theorems 1 and 2. We thus have

$$\begin{aligned} \mathbb{E}(C_k^3) &= n \sum_{\ell=1}^k \frac{6n^2 - 6\ell n + \ell^2}{\ell^3} + 3n^2 \sum_{\ell=2}^k \frac{(2n - \ell)h_{\ell-1}}{\ell^2} + 3n^3 \sum_{\ell=2}^k \frac{h_{\ell-1}^2}{\ell} \\ &\quad + 3n^3 \sum_{\ell=2}^k \frac{1}{\ell} \sum_{u=1}^{\ell-1} \frac{1}{u^2} - 3n^2 \sum_{\ell=2}^k \frac{h_{\ell-1}}{\ell} \\ &= 6n^3 \left( \sum_{\ell=2}^k \frac{h_{\ell-1}}{\ell^2} + \sum_{\ell=1}^k \frac{1}{\ell^3} \right) + 3n^3 \sum_{\ell=2}^k \frac{h_{\ell-1}^2}{\ell} + 3n^3 \sum_{\ell=2}^k \frac{1}{\ell} \sum_{u=1}^{\ell-1} \frac{1}{u^2} \\ &\quad - 6n^2 \left( \sum_{\ell=2}^k \frac{h_{\ell-1}}{\ell} + \sum_{\ell=1}^k \frac{1}{\ell^2} \right) + nh_k \\ &= 6n^3 \sum_{\ell=1}^k \frac{h_\ell}{\ell^2} + 3n^3 \sum_{\ell=2}^k \frac{h_{\ell-1}^2}{\ell} + 3n^3 \left( h_k \sum_{\ell=1}^k \frac{1}{\ell^2} - \sum_{\ell=1}^k \frac{h_\ell}{\ell^2} \right) - 6n^2 \sum_{\ell=1}^k \frac{h_\ell}{\ell} + nh_k \\ &= 3n^3 \sum_{\ell=1}^k \frac{h_\ell}{\ell^2} + 3n^3 \sum_{\ell=2}^k \frac{h_{\ell-1}^2}{\ell} + 3n^3 h_k \sum_{\ell=1}^k \frac{1}{\ell^2} - 6n^2 \sum_{\ell=1}^k \frac{h_\ell}{\ell} + nh_k. \end{aligned}$$

By taking  $k = n - 1$ , we get

$$\mathbb{E}(C^3) = \mathbb{E}(C_{n-1}^3) = 3n^3 \sum_{\ell=1}^{n-1} \frac{h_\ell}{\ell^2} + 3n^3 \sum_{\ell=2}^{n-1} \frac{h_{\ell-1}^2}{\ell} + 3n^3 h_{n-1} \sum_{\ell=1}^{n-1} \frac{1}{\ell^2} - 6n^2 \sum_{\ell=1}^{n-1} \frac{h_\ell}{\ell} + nh_{n-1}.$$

We denote respectively by  $A_{n,1}$ ,  $A_{n,2}$ ,  $A_{n,3}$  and  $A_{n,4}$  the four first sums of the right hand side of this expression. We then have

$$A_{n,1} = \sum_{\ell=1}^{n-1} \frac{h_\ell}{\ell^2}.$$

Since  $h_\ell/\ell^2 \underset{\ell \rightarrow \infty}{\sim} \ln(\ell)/\ell^2$  and the Bertrand series  $\sum_\ell \ln(\ell)/\ell^2$  converges, we have

$$\lim_{n \rightarrow \infty} A_{n,1} = A \approx 2.404.$$

Concerning the second sum

$$A_{n,2} = \sum_{\ell=2}^{n-1} \frac{h_{\ell-1}^2}{\ell},$$

we observe that  $h_\ell^2/\ell \underset{\ell \rightarrow \infty}{\sim} \ln^2(\ell)/\ell$  and the Bertrand series  $\sum_\ell \ln^2(\ell)/\ell$  diverges. Using the fact that the function  $\ln^2(x)/x$  is decreasing for  $x > e^2$ , we obtain

$$\sum_{\ell=2}^{n-1} \frac{\ln^2(\ell)}{\ell} \underset{n \rightarrow \infty}{\sim} \int_1^n \frac{\ln^2(x)}{x} dx = \frac{\ln^3(n)}{3}.$$

Hence

$$A_{n,2} \underset{n \rightarrow \infty}{\sim} \frac{\ln^3(n)}{3}.$$

The third sum  $A_{n,3}$  verifies

$$A_{n,3} = h_{n-1} \sum_{\ell=1}^{n-1} \frac{1}{\ell^2} \underset{n \rightarrow \infty}{\sim} \frac{\pi^2 \ln(n)}{6}.$$

The last sum  $A_{n,4}$  is the term  $\varphi_{n-1}$  already considered in the proof of Theorem 2. It follows that we have

$$A_{n,4} = \sum_{\ell=1}^{n-1} \frac{h_\ell}{\ell} = \frac{h_{n-1}^2}{2} + \frac{1}{2} \sum_{\ell=1}^{n-1} \frac{1}{\ell^2} \underset{n \rightarrow \infty}{\sim} \frac{\ln^2(n)}{2}.$$

Gathering these four results leads to

$$\mathbb{E}(C^3) \underset{n \rightarrow \infty}{\sim} n^3 \ln^3(n),$$

which completes the proof.  $\square$

More generally, the subsequent Theorem 3 allows for a recursive computation of all the moments of the cover time of a subset of states with  $k$  states. It also gives the asymptotic behavior of all the moments. It is based on relation (12). To prove this theorem, we first need the following lemma which gives a recursive expression of the Stirling numbers of the second kind. The Stirling numbers of the second kind can be defined, for  $r \geq 1$  and  $\ell = 1, \dots, r$ , by

$$S(r, \ell) = \frac{1}{\ell!} \sum_{j=1}^{\ell} (-1)^{\ell-j} \binom{\ell}{j} j^r. \quad (14)$$

These numbers are well-known in combinatorial analysis and when  $r \leq \ell$  we have, see for instance relation (12.17) in section II.12 of [16],

$$S(r, \ell) = \begin{cases} 0 & \text{if } r < \ell \\ 1 & \text{if } r = \ell. \end{cases} \quad (15)$$

The following original lemma on the Stirling numbers turns out to be useful in the subsequent analysis.

**Lemma 4.** *For every  $r \geq 2$  and  $\ell = 2, \dots, r$ , we have*

$$S(r, \ell) = \frac{1}{\ell} \sum_{j=\ell-1}^{r-1} \binom{r}{j} S(j, \ell-1). \quad (16)$$

*Proof.* Let  $r \geq 2$  and  $\ell = 2, \dots, r$ . From relation (14) we have

$$\begin{aligned} \frac{1}{\ell} \sum_{j=\ell-1}^{r-1} \binom{r}{j} S(j, \ell-1) &= \frac{1}{\ell!} \sum_{j=\ell-1}^{r-1} \binom{r}{j} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} k^j \\ &= \frac{1}{\ell!} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} \sum_{j=\ell-1}^{r-1} \binom{r}{j} k^j \end{aligned}$$

Observe that this last sum can be written as

$$\sum_{j=\ell-1}^{r-1} \binom{r}{j} k^j = (k+1)^r - k^r - \sum_{j=0}^{\ell-2} \binom{r}{j} k^j.$$

We thus get

$$\begin{aligned} \frac{1}{\ell} \sum_{j=\ell-1}^{r-1} \binom{r}{j} S(j, \ell-1) &= \frac{1}{\ell!} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} [(k+1)^r - k^r] \\ &\quad - \frac{1}{\ell!} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} \sum_{j=0}^{\ell-2} \binom{r}{j} k^j. \end{aligned} \quad (17)$$

The term containing the two sums can be written as

$$\begin{aligned} \frac{1}{\ell!} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} \sum_{j=0}^{\ell-2} \binom{r}{j} k^j &= \frac{1}{\ell!} \sum_{j=0}^{\ell-2} \binom{r}{j} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} k^j \\ &= \frac{1}{\ell!} \left[ \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} + \sum_{j=1}^{\ell-2} \binom{r}{j} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} k^j \right]. \end{aligned}$$

Since  $1 \leq j \leq \ell - 2$ , we have from (15),

$$\sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} k^j = 0.$$

It follows that

$$\frac{1}{\ell!} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} \sum_{j=0}^{\ell-2} \binom{r}{j} k^j = \frac{1}{\ell!} \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} = \frac{(-1)^\ell}{\ell!}.$$

Relation (17) thus becomes

$$\begin{aligned} \frac{1}{\ell} \sum_{j=\ell-1}^{r-1} \binom{r}{j} S(j, \ell-1) &= \frac{1}{\ell!} \left[ (-1)^{\ell-1} + \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} (k+1)^r \right. \\ &\quad \left. - \sum_{k=1}^{\ell-1} (-1)^{\ell-1-k} \binom{\ell-1}{k} k^r \right] \\ &= \frac{1}{\ell!} \left[ (-1)^{\ell-1} + \sum_{k=2}^{\ell} (-1)^{\ell-k} \binom{\ell-1}{k-1} k^r + \sum_{k=1}^{\ell-1} (-1)^{\ell-k} \binom{\ell-1}{k} k^r \right] \\ &= \frac{1}{\ell!} \left[ (-1)^{\ell-1} + \sum_{k=2}^{\ell-1} (-1)^{\ell-k} \binom{\ell}{k} k^r + \ell^r + (-1)^{\ell-1} (\ell-1) \right] \\ &= \frac{1}{\ell!} \left[ (-1)^{\ell-1} + \sum_{k=1}^{\ell} (-1)^{\ell-k} \binom{\ell}{k} k^r - (-1)^{\ell-1} \ell + (-1)^{\ell-1} (\ell-1) \right] \\ &= \frac{1}{\ell!} \sum_{k=1}^{\ell} (-1)^{\ell-k} \binom{\ell}{k} k^r \\ &= S(r, \ell), \end{aligned}$$

which completes the proof.  $\square$

We are now able to prove the following theorem.

**Theorem 5.** For every  $k = 1, \dots, n-1$  and for every  $r \geq 1$ ,  $\mathbb{E}(C_k^r)$  is a  $r$  degree polynomial in  $n$ , i.e.

$$\mathbb{E}(C_k^r) = \sum_{\ell=1}^r a_{r,\ell}^{(k)} n^\ell.$$

Moreover, for every  $\ell = 1, \dots, r$ , we have

$$a_{r,\ell}^{(k)} \underset{k \rightarrow \infty}{\sim} (-1)^{r+\ell} S(r, \ell) \ln^\ell(k).$$

In particular, when  $k = n - 1$ , we have

$$\mathbb{E}(C^r) \underset{n \rightarrow \infty}{\sim} n^r \ln^r(n).$$

*Proof.* Let  $A$  be a subset of  $S$  with  $k$  states. The  $r$ -th order moment of the cover time of subset  $A$  is given by relation (12), that is

$$\begin{aligned} \mathbb{E}(C_k^r) &= \mathbb{E}(T_k^r) + \sum_{m=1}^{r-1} \binom{r}{m} \mathbb{E}(C_{k-1}^m) \mathbb{E}(T_k^{r-m}) \\ &= \mathbb{E}(T_k^r) + \sum_{m=1}^{r-1} \binom{r}{m} \mathbb{E}(C_{k-1}^m) \mathbb{E}(T_k^{r-m}) + \mathbb{E}(C_{k-1}^r), \end{aligned}$$

with  $\mathbb{E}(C_1^r) = \mathbb{E}(T_1^r)$ . This yields

$$\mathbb{E}(C_k^r) = \sum_{j=1}^k \mathbb{E}(T_j^r) + \sum_{m=1}^{r-1} \binom{r}{m} \sum_{j=2}^k \mathbb{E}(C_{j-1}^m) \mathbb{E}(T_j^{r-m}). \quad (18)$$

We now proceed by recurrence. For  $k = 1$ , we have by relation (11),

$$\mathbb{E}(C_1^r) = \mathbb{E}(T_1^r) = \sum_{\ell=1}^r c_{r,\ell} n^\ell,$$

which is a  $r$  degree polynomial in  $n$  as claimed. Suppose that the result in the theorem is true for indices  $i = 1, \dots, r-1$  and  $j = 1, \dots, k-1$ , i.e. suppose that

$$\mathbb{E}(C_j^i) = \sum_{u=1}^i a_{i,u}^{(j)} n^u, \text{ for } i = 1, \dots, r-1 \text{ and } j = 1, \dots, k-1.$$

We then have from relation (18) and the results on hitting times

$$\begin{aligned} \mathbb{E}(C_k^r) &= \sum_{j=1}^k \mathbb{E}(T_j^r) + \sum_{i=1}^{r-1} \binom{r}{i} \sum_{j=2}^k \mathbb{E}(C_{j-1}^i) \mathbb{E}(T_j^{r-i}) \\ &= \sum_{j=1}^k \sum_{\ell=1}^r c_{r,\ell} \binom{n}{j}^\ell + \sum_{i=1}^{r-1} \binom{r}{i} \sum_{j=2}^k \left( \sum_{u=1}^i a_{i,u}^{(j-1)} n^u \right) \left( \sum_{v=1}^{r-i} c_{r-i,v} \binom{n}{j}^v \right) \\ &= \sum_{\ell=1}^r \left( c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} \right) n^\ell + \sum_{i=1}^{r-1} \binom{r}{i} \sum_{j=2}^k \left( \sum_{u=1}^i a_{i,u}^{(j-1)} n^u \right) \left( \sum_{v=1}^{r-i} \frac{c_{r-i,v}}{j^v} n^v \right), \end{aligned}$$

where the  $c_{r,\ell}$  are given by (7). This formula already establishes that  $\mathbb{E}(C_k^r)$  is a  $r$  degree polynomial in  $n$ . To be able to discuss the value of the  $a_{r,\ell}^{(k)}$ , we need to push the computation a bit further. Observe that the product of the last two parenthesis is

the product of two polynomials with different degrees. This product can be expressed as follows

$$\left( \sum_{u=1}^i a_{i,u}^{(j-1)} n^u \right) \left( \sum_{v=1}^{r-i} \frac{c_{r-i,v}}{j^v} n^v \right) = \sum_{\ell=2}^r \left( \sum_{t=1 \vee (\ell-r+i)}^{(\ell-1) \wedge i} a_{i,t}^{(j-1)} \frac{c_{r-i,\ell-t}}{j^{\ell-t}} \right) n^\ell,$$

where  $x \vee y = \max(x, y)$  and  $x \wedge y = \min(x, y)$ . We thus obtain

$$\begin{aligned} \mathbb{E}(C_k^r) &= \sum_{\ell=1}^r \left( c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} \right) n^\ell + \sum_{i=1}^{r-1} \binom{r}{i} \sum_{j=2}^k \sum_{\ell=2}^r \left( \sum_{t=1 \vee (\ell-r+i)}^{(\ell-1) \wedge i} a_{i,t}^{(j-1)} \frac{c_{r-i,\ell-t}}{j^{\ell-t}} \right) n^\ell \\ &= \sum_{\ell=1}^r \left( c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} \right) n^\ell + \sum_{\ell=2}^r \sum_{i=1}^{r-1} \binom{r}{i} \sum_{j=2}^k \left( \sum_{t=1 \vee (\ell-r+i)}^{(\ell-1) \wedge i} a_{i,t}^{(j-1)} \frac{c_{r-i,\ell-t}}{j^{\ell-t}} \right) n^\ell. \end{aligned} \quad (19)$$

We are now in position to analyze the coefficients  $a_{r,\ell}^{(k)}$  of polynomial  $\mathbb{E}(C_k^r)$  which are given by (19).

For  $\ell = 1$  the coefficient  $a_{r,1}^{(k)}$  of  $n$  is

$$a_{r,1}^{(k)} = c_{r,1} \sum_{j=1}^k \frac{1}{j} = (-1)^{r+1} h_k,$$

and the coefficient  $a_{r,\ell}^{(k)}$  of  $n^\ell$  for  $\ell = 2, \dots, r$  is

$$a_{r,\ell}^{(k)} = c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} + \sum_{i=1}^{r-1} \binom{r}{i} \sum_{j=2}^k \left( \sum_{t=1 \vee (\ell-r+i)}^{(\ell-1) \wedge i} a_{i,t}^{(j-1)} \frac{c_{r-i,\ell-t}}{j^{\ell-t}} \right). \quad (20)$$

We prove by recurrence that for every  $r \geq 1$  and  $\ell = 1, \dots, r-1$ , we have

$$\lim_{k \rightarrow \infty} \frac{a_{r,\ell}^{(k)}}{\ln^\ell(k)} = (-1)^{r+\ell} S(r, \ell),$$

where  $S(r, \ell)$  are the Stirling numbers of the second kind defined by (14). It is well-known that the Stirling numbers are positive.

The result is true for  $\ell = 1$  since  $a_{r,1}^{(k)} = (-1)^{r+1} h_k$ , and  $S(r, 1) = 1$ .

We consider now the case  $\ell = 2, \dots, r$ . Suppose that

$$\lim_{j \rightarrow \infty} \frac{a_{i,t}^{(j)}}{\ln^t(j)} = (-1)^{i+t} S(i, t), \text{ for } i = 1, \dots, r-1 \text{ and } t = 1, \dots, i.$$



Exchanging the sums over  $j$  and  $t$  in relation (20) leads, for  $\ell = 2, \dots, r$ , to

$$a_{r,\ell}^{(k)} = c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} + \sum_{i=1}^{r-1} \binom{r}{i} \sum_{t=1 \vee (\ell-r+i)}^{(\ell-1) \wedge i} c_{r-i,\ell-t} \sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}}.$$

We decompose the sum over index  $i$  in two distinct sums  $i \leq \ell - 2$  and  $i \geq \ell - 1$ . We thus get

$$\begin{aligned} a_{r,\ell}^{(k)} &= c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} + \sum_{i=1}^{\ell-2} \binom{r}{i} \sum_{t=1 \vee (\ell-r+i)}^i c_{r-i,\ell-t} \sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}} \\ &\quad + \sum_{i=\ell-1}^{r-1} \binom{r}{i} \sum_{t=1 \vee (\ell-r+i)}^{\ell-1} c_{r-i,\ell-t} \sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}}. \end{aligned}$$

In the second sum over  $t$ , we extract the term corresponding to  $t = \ell - 1$ . We obtain since  $c_{r-i,1} = (-1)^{r-i+1}$

$$\begin{aligned} a_{r,\ell}^{(k)} &= c_{r,\ell} \sum_{j=1}^k \frac{1}{j^\ell} + \sum_{i=1}^{\ell-2} \binom{r}{i} \sum_{t=1 \vee (\ell-r+i)}^i c_{r-i,\ell-t} \sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}} \\ &\quad + \sum_{i=\ell-1}^{r-1} \binom{r}{i} \sum_{t=1 \vee (\ell-r+i)}^{\ell-2} c_{r-i,\ell-t} \sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}} + (-1)^r \sum_{i=\ell-1}^{r-1} (-1)^{i-1} \binom{r}{i} \sum_{j=2}^k \frac{a_{i,\ell-1}^{(j-1)}}{j}. \end{aligned} \tag{21}$$

We consider separately the four sums over  $j$ . Since  $\ell \geq 2$ , the first sum over  $j$  converges when  $k$  tends to infinity to the Riemann series  $\zeta(\ell)$ , *i.e.*

$$\sum_{j=1}^{\infty} \frac{1}{j^\ell} = \zeta(\ell) < \infty.$$

The second and third sums over  $j$  can be written as

$$\sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}} = \sum_{j=2}^k \frac{a_{i,t}^{(j-1)}}{\ln^t(j)} \frac{\ln^t(j)}{j^{\ell-t}}.$$

The recurrence hypothesis tells us that the ratio  $a_{i,t}^{(j-1)}/\ln^t(j)$  converges to  $(-1)^{i+t}S(i,t)$  when  $j$  tends to infinity. This implies that this ratio is bounded by constant uniformly in  $j$ , *i.e.* there is a constant  $\beta_{i,t}$  such that for all  $j$  we have

$$\left| \frac{a_{i,t}^{(j-1)}}{\ln^t(j)} \right| \leq \beta_{i,t}.$$

We thus have for all  $k \geq 2$ ,

$$\sum_{j=2}^k \left| \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}} \right| \leq \beta_{i,t} \sum_{j=2}^k \frac{\ln^t(j)}{j^{\ell-t}}.$$

The series on the right hand side is a converging Bertrand series whenever  $\ell - t \geq 2$ , a condition that is met in the second and third sums since  $t \leq \ell - 2$ . It follows that the second and third sums converge, *i.e.*

$$\sum_{j=2}^{\infty} \frac{a_{i,t}^{(j-1)}}{j^{\ell-t}} < \infty.$$

We now consider the fourth sum over  $j$ . Since the sequence  $a_{i,\ell-1}^{(j-1)} / \ln^{\ell-1}(j)$  converges to the non zero value  $(-1)^{i+\ell-1} S(i, \ell - 1)$  when  $j$  tends to infinity, it always keeps the same sign as  $(-1)^{i+\ell-1} S(i, \ell - 1)$  after a certain rank.

Observe now that the function  $f$  defined by

$$f(x) = \frac{\ln^{\ell-1}(x)}{x}$$

is non negative and decreasing on the interval  $[e^{\ell-1}, +\infty)$ . We thus have

$$\sum_{j=2}^k \frac{\ln^{\ell-1}(j)}{j} \underset{k \rightarrow \infty}{\sim} \int_1^k \frac{\ln^{\ell-1}(x)}{x} dx = \frac{\ln^{\ell}(k)}{\ell}.$$

It follows that

$$\frac{a_{i,\ell-1}^{(j-1)}}{j} = \frac{a_{i,\ell-1}^{(j-1)}}{\ln^{\ell-1}(j)} \frac{\ln^{\ell-1}(j)}{j} \underset{j \rightarrow \infty}{\sim} \frac{(-1)^{i+\ell-1} S(i, \ell - 1) \ln^{\ell-1}(j)}{j}.$$

The series  $\sum_j \ln^{\ell-1}(j)/j$  being divergent we have

$$\sum_{j=2}^k \frac{a_{i,\ell-1}^{(j-1)}}{j} \underset{k \rightarrow \infty}{\sim} (-1)^{i+\ell-1} S(i, \ell - 1) \sum_{j=2}^k \frac{\ln^{\ell-1}(j)}{j} \underset{k \rightarrow \infty}{\sim} \frac{(-1)^{i+\ell-1} S(i, \ell - 1) \ln^{\ell}(k)}{\ell}.$$

Using these four results in relation (21) and applying Lemma 4, we obtain

$$\lim_{k \rightarrow \infty} \frac{a_{r,\ell}^{(k)}}{\ln^{\ell}(k)} = \frac{(-1)^{r+\ell}}{\ell} \sum_{i=\ell-1}^{r-1} \binom{r}{i} S(i, \ell - 1) = (-1)^{r+\ell} S(r, \ell).$$

By taking  $k = n - 1$ , we obtain, since  $S(r, r) = 1$ ,

$$\mathbb{E}(C^r) = \mathbb{E}(C_{n-1}^r) = \sum_{\ell=1}^r a_{r,\ell}^{(n-1)} n^\ell \underset{n \rightarrow \infty}{\sim} n^r \ln^r(n),$$

which completes the proof.  $\square$

We can now determine the limiting behavior of the distribution of the cover time  $C = C_{n-1}$ .

**Theorem 6.** *For all  $t \in \mathbb{R} \setminus \{1\}$ , we have,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C}{n \ln(n)} > t \right\} = \begin{cases} 1 & \text{if } t > 1 \\ 0 & \text{if } t < 1. \end{cases}$$

*Proof.* From the second part of Theorem 5, we have for every  $r \geq 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \left( \frac{C}{n \ln(n)} \right)^r \right) = 1.$$

Let  $Z$  be the constant random variable defined by  $Z = 1$ . For every  $r \geq 0$ , we have  $\mathbb{E}(Z^r) = 1$ . Thus the power series  $\sum_{r \geq 0} \mathbb{E}(Z^r) x^r / r!$  has a positive radius (equal to  $+\infty$ ). This implies, from Theorem 30.1 of [7], that the random variable  $Z$  is uniquely determined by its moments. It follows from Theorem 30.2 of [7] that

$$\frac{C}{n \ln(n)} \xrightarrow{\mathcal{L}} Z = 1, \text{ as } n \rightarrow \infty,$$

which means that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C}{n \ln(n)} > t \right\} = \begin{cases} 1 & \text{if } t < 1 \\ 0 & \text{if } t > 1, \end{cases}$$

which completes the proof.  $\square$

Observe that the previous theorem does not allow us to conclude when  $t = 1$  because the convergence in distribution occurs only at the points  $t$  for which the limiting distribution is continuous.

We can solve this problem by considering the result of [14], which tell us that in the complete graph with self-loops we have

$$\mathbb{P}\{C > \ell\} = \begin{cases} 1 & \text{if } 0 \leq \ell < n - 1 \\ \sum_{j=0}^{n-2} (-1)^j \binom{n-1}{j+1} \left( \frac{n-(j+1)}{n} \right)^\ell & \text{if } \ell \geq n - 1. \end{cases} \quad (22)$$

Using this result, we obtain the following limiting distribution which completes the result of Theorem 6 at point  $t = 1$ .

**Theorem 7.** For every  $t \in \mathbb{R}$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C - n \ln(n)}{n} > t \right\} = 1 - \exp(-e^{-t}),$$

which is known as the standard Gumbel distribution. At point  $t = 0$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C}{n \ln(n)} > 1 \right\} = 1 - e^{-1} \approx 0.6321205588.$$

*Proof.* Taking  $\ell = n \ln(n) + nt$  in (22), we obtain

$$\mathbb{P}\{C > n \ln(n) + nt\} = \sum_{j=0}^{n-2} (-1)^j \binom{n-1}{j+1} \left( \frac{n-(j+1)}{n} \right)^{n \ln(n) + nt}.$$

Introducing the notation

$$u_{n,j} = (-1)^j \binom{n-1}{j+1} \left( \frac{n-(j+1)}{n} \right)^{n \ln(n) + nt} \mathbf{1}_{\{j \leq n-2\}},$$

we have

$$u_{n,j} = (-1)^j \frac{(n-1) \cdots (n-(j+1))}{(j+1)!} \exp \left( [n \ln(n) + nt] \ln \left( 1 - \frac{j+1}{n} \right) \right) \mathbf{1}_{\{j \leq n-2\}}.$$

For any fixed  $j$ , we have

$$u_{n,j} \underset{n \rightarrow \infty}{\sim} (-1)^j \frac{n^{j+1}}{(j+1)!} e^{-(j+1)[\ln(n)+t]} \mathbf{1}_{\{j \leq n-2\}} = \frac{(-1)^j e^{-(j+1)t}}{(j+1)!} \mathbf{1}_{\{j \leq n-2\}}.$$

We thus have, for every  $j \geq 0$ ,

$$\lim_{n \rightarrow \infty} u_{n,j} = \frac{(-1)^j e^{-(j+1)t}}{(j+1)!}.$$

Moreover, since  $\ln(1-x) \leq -x$  for all  $x \in [0, 1)$ , we obtain

$$\begin{aligned} |u_{n,j}| &= \frac{(n-1) \cdots (n-(j+1))}{(j+1)!} \exp \left( [n \ln(n) + nt] \ln \left( 1 - \frac{j+1}{n} \right) \right) \mathbf{1}_{\{j \leq n-2\}} \\ &\leq \frac{n^{j+1}}{(j+1)!} e^{-(j+1)[\ln(n)+t]} = \frac{e^{-(j+1)t}}{(j+1)!} \end{aligned}$$

and the series  $\sum_{j \geq 0} e^{-(j+1)t}/(j+1)!$  is convergent with

$$\sum_{j=0}^{\infty} \frac{e^{-(j+1)t}}{(j+1)!} = \exp(e^{-t}) - 1.$$

The Lebesgue dominated convergence theorem for series allows us to conclude that

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{\infty} u_{n,j} = \sum_{j=0}^{\infty} \frac{(-1)^j e^{-(j+1)t}}{(j+1)!} = 1 - \exp(-e^{-t}),$$

which is the complementary standard Gumbel distribution function. Take  $t = 0$  to complete the proof.  $\square$

Observe that, since the limit  $Z = 1$  is constant, we also have the convergence in probability of  $C/(n \ln(n))$  to  $Z = 1$ , that is, for any  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \left| \frac{C}{n \ln(n)} - 1 \right| > \varepsilon \right\} = 0.$$

We plot in Figure 1 the complementary cumulative distribution function  $F_n(t)$  defined by

$$F_n(t) = \mathbb{P} \left\{ \frac{C}{n \ln(n)} > t \right\}$$

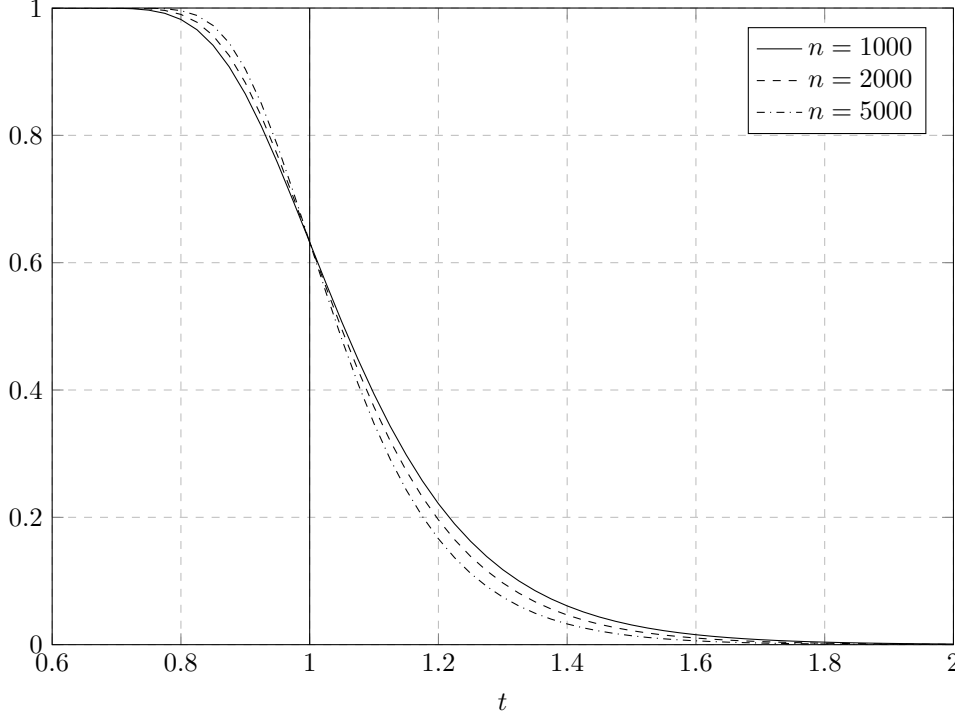
for the values of  $n = 1000, 2000, 5000$ . To compute these functions we used relation 22. The limiting value  $F$  of these functions is given by Theorem 6, for  $t \neq 1$  and by Theorem 7 for  $t = 1$ , that is

$$F(t) = \begin{cases} 1 & \text{if } t < 1 \\ 1 - e^{-1} \approx 0.6321205588 & \text{if } t = 1 \\ 0 & \text{if } t > 1. \end{cases}$$

Observe in this figure that the convergence of functions  $F_n$  to function  $F$  seems to be slower outside the point  $t = 1$  and much more faster at point  $t = 1$ . This is due to the discontinuity of function  $F$  at this point. Indeed for  $t = 1$ , some numerical values of  $F_n(1)$  and  $F(1)$  are shown in table 2 for several values of  $n$ .

| $F_{200}(1)$ | $F_{500}(1)$ | $F_{1000}(1)$ | $F_{2000}(1)$ | $F_{5000}(1)$ | $F(1)$       |
|--------------|--------------|---------------|---------------|---------------|--------------|
| 0.6323958940 | 0.6319687752 | 0.6322128195  | 0.6321761524  | 0.6321547507  | 0.6321205588 |

**Table 2** Values of  $F_n(1)$  for different values of  $n$  and its limiting value  $F(1) = 1 - e^{-1}$  when  $n$  tends to infinity.



**Fig. 1** Complementary cumulative distribution functions  $F_n(t)$  of  $C/(n \ln(n))$  for  $n = 1000$ ,  $n = 2000$  and  $n = 5000$  and their limiting value  $F(t)$  when  $n \rightarrow \infty$ .

We plot in Figure 2 the complementary cumulative distribution function  $G_n(t)$  defined by

$$G_n(t) = \mathbb{P} \left\{ \frac{C - n \ln(n)}{n} > t \right\}$$

for the values of  $n = 10, 20$ . To compute these functions we used relation 22. The limiting value  $G$  of these functions is given by Theorem 7, that is

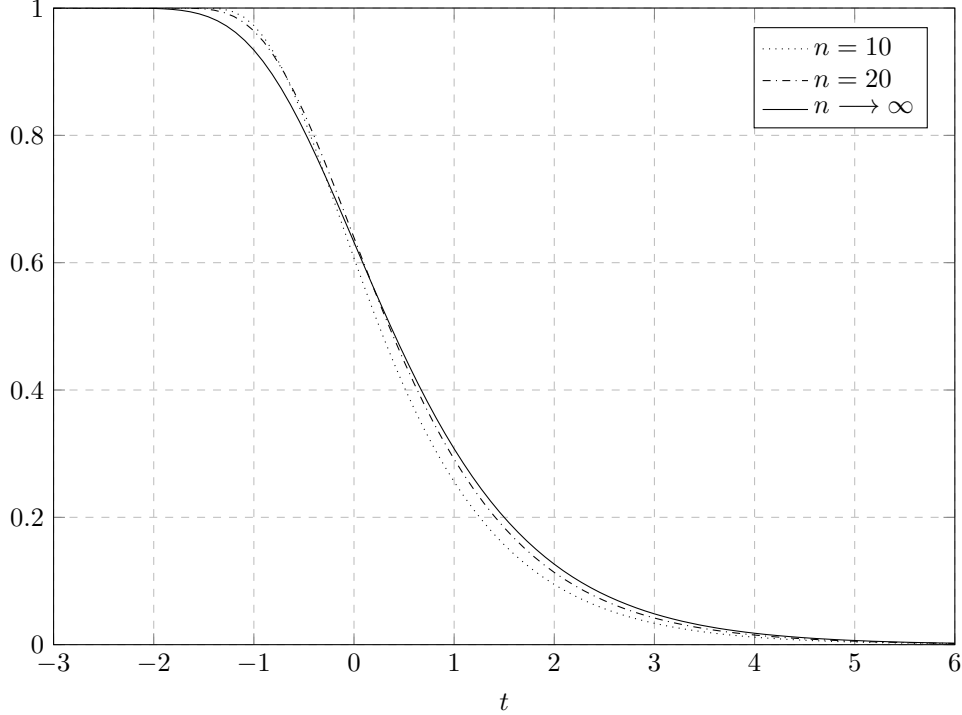
$$G(t) = 1 - \exp(-e^{-t}).$$

Observe in this figure that the convergence of functions  $G_n$  to function  $G$  is quite fast. Indeed, in that case the limiting distribution  $G$  is infinitely differentiable.

Some numerical values of  $G_n(1)$  and  $G(1)$  are shown in table 3 for several values of  $n$ .

| $G_{10}(1)$  | $G_{20}(1)$  | $G_{100}(1)$ | $G_{200}(1)$ | $G_{1000}(1)$ | $G(1)$       |
|--------------|--------------|--------------|--------------|---------------|--------------|
| 0.2559540800 | 0.2912861575 | 0.3024095098 | 0.3050352510 | 0.3071453686  | 0.3077993724 |

**Table 3** Values of  $G_n(1)$  for different values of  $n$  and its limiting value  $G(1) = 1 - \exp(-e^{-1})$  when  $n$  tends to infinity.



**Fig. 2** Complementary cumulative distribution functions  $G_n(t)$  of  $(C - n \ln(n))/n$  for  $n = 10$  and  $n = 20$  and their limiting distribution  $G(t)$  when  $n \rightarrow \infty$ .

## 4 Analysis of the complete graph without self-loops

A random walk on the complete graph without self-loops is a discrete-time Markov chain on the state space  $S = \{0, 1, \dots, n-1\}$ ,  $n \geq 2$ , with transition probability matrix  $P$  given, for every  $u \neq v$ , by  $P_{u,v} = 1/(n-1)$  and  $P_{u,u} = 0$ . All the results of this section are close to those of the complete graph with self-loops.

### 4.1 Hitting times on the complete graph without self-loops

Let  $A$  be a proper subset of  $S$  with  $k$  states, that is  $|A| = k \in \{1, \dots, n-1\}$ . To simplify the notation, using again the symmetry, we will drop sometimes the initial state because the results remain the same for every initial state  $i \in A^c$ .

Using relation (2), we have for every  $\ell \geq 0$  and  $i \in A^c$ ,

$$\mathbb{P}_i\{T_k > \ell\} = \sum_{j=\ell+1}^{\infty} (P_{A^c}^{j-1} P_{A^c, A} \mathbf{1})_i = (P_{A^c}^{\ell} \mathbf{1})_i.$$

Since  $|A^c| = n - k$  we have  $P_{A^c} \mathbb{1} = (1 - k/(n - 1)) \mathbb{1}$  which recursively leads to

$$P_{A^c}^\ell \mathbb{1} = \left(1 - \frac{k}{n-1}\right)^\ell \mathbb{1} \text{ and thus } \mathbb{P}\{T_k > \ell\} = \left(1 - \frac{k}{n-1}\right)^\ell, \quad (23)$$

which means that  $T_k$  is geometrically distributed with parameter  $k/(n - 1)$ . Using relation (6), we have, for every  $r \geq 1$ ,

$$\mathbb{E}(T_k^r) = \sum_{\ell=1}^r c_{r,\ell} \left(\frac{n-1}{k}\right)^\ell, \quad (24)$$

a relation that is to be compared with relation (11). The first moments of  $T_k$  are then given by replacing  $n$  by  $n - 1$  in those obtained for the complete graph with self-loops, *i.e.* by

$$\begin{aligned} \mathbb{E}(T_k) &= \frac{n-1}{k}, \mathbb{E}(T_k^2) = \frac{(n-1)(2(n-1) - k)}{k^2}, \\ \mathbb{E}(T_k^3) &= \frac{(n-1)(6(n-1)^2 - 6k(n-1) + k^2)}{k^3}. \end{aligned}$$

In the same way, we have, for every fixed value of  $k$ ,

$$\mathbb{E}(T_k^r) \underset{n \rightarrow \infty}{\sim} r! \left(\frac{n}{k}\right)^r \text{ and } \lim_{n \rightarrow \infty} \mathbb{P}\{kT_k/n > t\} = e^{-t},$$

which means that the sequence of random variables  $kT_k/n$  converges in distribution, when  $n$  tends to infinity, to a random variable exponentially distributed with rate 1.

The results on the cover times are also very close to those on the complete graph with self-loops.

## 4.2 Cover times on the complete graph without self-loops

Here again, the results on the cover times are also very close to those on the complete graph with self-loops. We consider the first moments of the cover times.

**Theorem 8.** *For every  $k = 1, \dots, n - 1$ , we have*

$$\mathbb{E}(C_k) = (n - 1)h_k.$$

*In particular, when  $k = n - 1$ , we have*

$$\mathbb{E}(C) \underset{n \rightarrow \infty}{\sim} n \ln(n).$$

*Proof.* Let  $A$  be a subset of  $S$  with  $k$  states. The expected cover time of subset  $A$  is given by relation (12) with  $r = 1$ . This gives, using the previous results on the hitting times

$$\mathbb{E}(C_k) = \mathbb{E}(T_k) + \mathbb{E}(C_{k-1}) = \frac{n-1}{k} + \mathbb{E}(C_{k-1}).$$



with  $\mathbb{E}(C_1) = \mathbb{E}(T_1) = n-1$ . This simple recursion simply leads to  $\mathbb{E}(C_k) = (n-1)h_k$ . By taking  $k = n-1$ , we get  $\mathbb{E}(C) = \mathbb{E}(C_{n-1}) = (n-1)h_{n-1}$  and thus

$$\mathbb{E}(C) \underset{n \rightarrow \infty}{\sim} n \ln(n),$$

which completes the proof.  $\square$

Observe that this well-known result can also be obtained using geometric random variables. Indeed, it is easy to check that

$$\mathbb{E}(C) = \mathbb{E}\left(\text{geo}\left(\frac{n-1}{n-1}\right)\right) + \dots + \mathbb{E}\left(\text{geo}\left(\frac{1}{n-1}\right)\right) = (n-1)h_{n-1}.$$

In the following theorem we consider the second order moment of the cover time of a subset of states with  $k$  states.

**Theorem 9.** *For every  $k = 1, \dots, n-1$ , we have*

$$\mathbb{E}(C_k^2) = (n-1)^2 h_k^2 + (n-1)^2 \sum_{\ell=1}^k \frac{1}{\ell^2} - (n-1)h_k.$$

*In particular, when  $k = n-1$ , we have*

$$\mathbb{E}(C^2) \underset{n \rightarrow \infty}{\sim} n^2 \ln^2(n).$$

*Proof.* It suffices to follow the same steps of the proof of Theorem 2.  $\square$

We consider now the third order moment of the cover time of a subset of states with  $k$  states.

**Theorem 10.** *For every  $k = 1, \dots, n-1$ , we have*

$$\begin{aligned} \mathbb{E}(C_k^3) &= 3(n-1)^3 \sum_{\ell=1}^k \frac{h_\ell}{\ell^2} + 3(n-1)^3 \sum_{\ell=2}^k \frac{h_{\ell-1}^2}{\ell} + 3(n-1)^3 h_k \sum_{\ell=1}^k \frac{1}{\ell^2} \\ &\quad - 6(n-1)^2 \sum_{\ell=1}^k \frac{h_\ell}{\ell} + (n-1)h_k. \end{aligned}$$

*In particular, when  $k = n-1$ , we have*

$$\mathbb{E}(C^3) \underset{n \rightarrow \infty}{\sim} n^3 \ln^3(n).$$

*Proof.* It suffices to follow the same steps of the proof of Theorem 3.  $\square$

More generally, we have the result of the following theorem which allows a recursive computation of all the moments of the cover time of a subset of states with  $k$  states. It also give the asymptotic behavior of all the moments. It is based on relation (12).

**Theorem 11.** For every  $k = 1, \dots, n-1$  and for every  $r \geq 1$ ,  $\mathbb{E}(C_k^r)$  is a  $r$  degree polynomial in  $n$ , i.e.

$$\mathbb{E}(C_k^r) = \sum_{\ell=1}^r a_{r,\ell}^{(k)} (n-1)^\ell.$$

For every  $\ell = 1, \dots, r$ , we have

$$a_{r,\ell}^{(k)} \underset{k \rightarrow \infty}{\sim} (-1)^{r+\ell} S(r, \ell) \ln^\ell(k).$$

Moreover, we have

$$\mathbb{E}(C^r) \underset{n \rightarrow \infty}{\sim} n^r \ln^r(n).$$

*Proof.* It suffices to consider the polynomial in  $n-1$  instead of  $n$  and to follow the same steps of the proof of Theorem 5.  $\square$

We can now determine the limiting behavior of the distribution of the cover time  $C = C_{n-1}$ .

**Theorem 12.** For all  $t \in \mathbb{R} \setminus \{1\}$ , we have,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C}{n \ln(n)} > t \right\} = \begin{cases} 1 & \text{if } t > 1 \\ 0 & \text{if } t < 1. \end{cases}$$

*Proof.* It is the same proof as the one of Theorem 6.  $\square$

Observe that the previous theorem does not allow us to conclude when  $t = 1$  because the convergence in distribution occurs only at the points  $t$  for which the limiting distribution is continuous.

We can solve this problem by considering the result of [14], which tell us that in the complete graph without self-loops we have

$$\mathbb{P}\{C > \ell\} = \begin{cases} 1 & \text{if } 0 \leq \ell < n-1 \\ \sum_{j=0}^{n-2} (-1)^j \binom{n-1}{j+1} \left( \frac{n-1-(j+1)}{n-1} \right)^\ell & \text{if } \ell \geq n-1. \end{cases} \quad (25)$$

Using this result, we obtain the following limiting distribution which completes the result of Theorem 12 at point  $t = 1$ .

**Theorem 13.** For every  $t \in \mathbb{R}$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C - n \ln(n)}{n} > t \right\} = 1 - \exp(-e^{-t}),$$

which is known as the standard Gumbel distribution. At point  $t = 0$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \frac{C}{n \ln(n)} > 1 \right\} = 1 - e^{-1} \approx 0.6321205588.$$

*Proof.* The proof is identical to the one of Theorem 7 by using relation (25).  $\square$

Observe that, since the limit  $Z = 1$  is constant, we also have the convergence in probability of  $C/(n \ln(n))$  to  $Z = 1$ , that is, for any  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left\{ \left| \frac{C}{n \ln(n)} - 1 \right| > \varepsilon \right\} = 0.$$

## 5 Conclusion

We considered in this paper a random walk on the complete graph and we obtained both the limiting distribution and the moments of the hitting and cover times of a subset of states of the complete graph. Further research will be to analyze other particular graphs such as the Barbell graph, the star graph or the Tadpole graph.

## 6 Declarations

**Funding :** There has been no support for this work.

**Conflicts of interest/Competing interests :** there are neither Conflicts of interest nor Competing interests for this work.

**Availability of data and material :** Not applicable.

**Authors' contributions :** The two authors, who are François Castella and Bruno Sericola, contributed equally to this work.

## References

- [1] Aldous, D.: Applications of Random Walks on Finite Graphs. <https://www.google.com/search?client=firefox-b-d&q=Applications+of+random+walks+on+graphs>, preprint, (1989)
- [2] Doyle, P.G., Snell, J.L.: Random Walks and Electrical Networks. The carus mathematical monographs, p. 155. Mathematical Assoc. of America, Washington, DC (1984)
- [3] Palacios, J.L.: Cover times for the stars. Math. Sci. **19**, 103–107 (1994)
- [4] Palacios, J.L.: Expected hitting and cover times of random walks on some special graphs. Random Struct. Algorithms **5**, 173–182 (1994)
- [5] Kermarrec, A.-M., Le Merrer, E., Sericola, B., Trédan, G.: Second order centrality: distributed assessment of nodes importance in complex networks. Computer Communications **34**(5), 619–628 (2011). Special issue on Complex Networks
- [6] Lovász, L.: Random walks on graphs: A survey. In: Miklós, D., Sós, V.T., Szönyi, T. (eds.) Combinatorics, Paul Erdős Is Eighty, Volume 2, Keszthely (Hungary) (1996)

- [7] Aldous, D., Fill, J.A.: Reversible Markov Chains and Random Walks on Graphs. <https://www.stat.berkeley.edu/~aldous/RWG/book.pdf>, unfinished monograph, (2002)
- [8] Motwani, R., Raghavan, P.: Randomized Algorithms, p. 467. Cambridge University Press, New York (1995)
- [9] Xia, F., Liu, J., Nie, H., Fu, Y., Wan, L., Kong, X.: Random walks: A review of algorithms and applications. *IEEE Transactions on Emerging Topics in Computational Intelligence* **4**(2), 95–107 (2020)
- [10] Feige, U.: A tight upper bound on the cover time for random walks on graphs. *Random Structures and Algorithms* **6**(1), 51–54 (1995)
- [11] Feige, U.: A tight lower bound on the cover time for random walks on graphs. *Random Structures and Algorithms* **6**(4), 433–438 (1995)
- [12] Ikeda, S., Kubo, I., Yamashita, M.: The hitting and cover times of random walks on finite graphs using local degree information. *Theoretical Computer Science* **410**(1), 94–100 (2009)
- [13] Sericola, B., Castella, F.: Hitting times on the lollipop graph. Technical report, Centre Inria de l’université de Rennes (June 2023). <https://inria.hal.science/hal-04143403/file/lolli.pdf>. <https://inria.hal.science/hal-04143403>
- [14] Sericola, B.: On cover times of Markov chains. Technical report, Inria Rennes - Bretagne Atlantique (November 2022). <https://hal.inria.fr/hal-03837323/file/cover.pdf>. <https://hal.inria.fr/hal-03837323>
- [15] Sericola, B.: Markov Chains. Theory, Algorithms and Applications. Applied stochastic methods series, p. 416. Wiley-ISTE, London (2013)
- [16] Feller, W.: An Introduction to Probability Theory and Its Applications, 3rd edn. John Wiley and Sons Inc., New York (1968)