Information Transmission under Random Emission Constraints

F R A N C I S   C O M E T S[1],   F R A N Ç O I S   D E L A R U E[2] and R E N É   S C H O T T[3]

[1]Laboratoire de Probabilités et Modèles Aléatoires, Université Paris Diderot–Paris 7, Case 7012,
75205 Paris CEDEX 13, France
(e-mail: `comets@math.univ-paris-diderot.fr`,
`http://www.proba.jussieu.fr/~comets`)

[2]Laboratoire J.-A. Dieudonné, Université de Nice Sophia Antipolis, Parc Valrose,
06108 Nice CEDEX 02, France
(e-mail: `delarue@unice.fr`,
`http://math.unice.fr/~delarue`)

[3]IECL and LORIA, Université de Lorraine, 54506 Vandoeuvre-lès-Nancy, France
(e-mail: `schott@loria.fr`,
`http://www.loria.fr/~schott`)

We model the transmission of a message on the complete graph with *n* vertices and limited resources. The vertices of the graph represent servers that may broadcast the message at random. Each server has a random emission capital that decreases at each emission. Quantities of interest are the number of servers that receive the information before the capital of all the informed servers is exhausted and the exhaustion time. We establish limit theorems (law of large numbers, central limit theorem and large deviation principle), as $n \to \infty$, for the proportion of informed vertices before exhaustion and for the total duration. The analysis relies on a construction of the transmission procedure as a dynamical selection of successful nodes in a Galton–Watson tree with respect to the success epochs of the coupon collector problem.

2010 *Mathematics subject classification*: Primary 90B30
Secondary 05C81, 60F05, 60F10, 60J20, 92D30

1. Introduction

The transmission of information and dissemination of viruses in computer networks has given rise to many practical and theoretical investigations over the past two decades (see [4, 12, 19, 21, 22, 26]).

In this paper we model the transmission of a message on the complete graph with *n* vertices and limited resources. Every vertex represents a server, which can be in one of three states: inactive (it has not yet received the message), active (it has already received

the message and is able to transmit it), exhausted (it has already received the message but can no longer transmit it because it has exhausted its own capital of emissions). Each server S_i has a random emission capital K_i. The message is initially received from outside by one server, which is then turned from the inactive state to the active one (if it has a positive emission capital) or exhausted (if its emission capital is 0), although the $n-1$ other servers are inactive. At each integer time, one of the active servers (say S_i) casts the message, it loses one unit of its own emission capital K_i, and it selects the target at random among the n servers. If the target is inactive, it discovers the information and itself becomes active or exhausted according to its own emission capital. If not, this broadcast is unsuccessful and nothing else happens. When an active server exhausts its emission capital, it enters the exhausted state. The transmission ends at a finite time τ_n, which is at most equal to 1 plus the sum of all initial capitals.

From a practical perspective, the graph may be thought of as a wireless network, the vertices of which are battery-powered sensors with a limited energy capacity. We refer the reader to [4, 12, 19, 21] for applications of graph theory to the performance evaluation of information transmission in wireless networks. We mention that the transmission process can also be interpreted as the busy period of a queue, when the probability that a new customer enters the queue decays linearly as the number of past arrivals increases and finally vanishes after the nth arrival.

Here we describe the asymptotic behaviour of the proportion of informed vertices at the end of the process when n tends to the infinity. The mathematical analysis relies on a twofold structure: a subtree of the Galton–Watson tree, which models the vertices reached by the emission procedure, and the success epochs of the coupon collector problem, which model the successful transmissions. Phrased in a probabilistic way, we propose a coupling of the transmission model as a marginal tree of the Galton–Watson tree, obtained by pruning some of the nodes according to the coupon collector problem. Such a coupling provides a direct interpretation of the scenarios when the network ceases broadcasting at the very beginning of the process: basically, these scenarios correspond to the extinction event in the Galton–Watson tree. For the survival event, we manage to specify the first-order behaviour (in n) of the exhaustion time τ_n and of the proportion of informed nodes. Under suitable integrability conditions on the distribution of the capital of a given vertex, the fluctuations of both the exhaustion time and the proportion of informed nodes are also investigated: a central limit theorem is proved under a square-integrability condition and a somewhat involved large deviation principle is established under an exponential-integrability condition. In particular, when the distribution of the capital of a given vertex is of finite expectation, the probability that all the servers are reached before exhaustion (also referred to as the probability of full transmission) converges to 0 as n tends to the infinity; as a consequence of the large deviation principle we prove here, it decays exponentially fast when the capital has a finite exponential moment. In some cases when the distribution of the capital has a heavy tail, we prove that the limit of the probability of full transmission is different from 0 and, in particular, may coincide with the entire probability of survival of the Galton–Watson tree. We refer to [5, 7, 13, 17, 20] and Chapter 8 of [31] for specific results concerning the coupon collector problem.

The papers [27] and [26] study closely related transmission models. Machado, Machurian and Matzinger [27] consider the case where $K_i = 2$ and prove partial transmission results. Obviously, our approach extends this result, as constant capitals are a particular case of random ones. A specific interest of random capitals consists in allowing K_i to be 0 with a non-trivial probability: as we shall see below, a quick stop of the transmission process then occurs with positive probability, as the extinction event of the Galton–Watson tree. As in [27], Kurtz, Lebensztayn, Leichsenring and Machado [26] investigate the case where the K_i are constant, but possibly larger than 2, for continuous time. In their model, there is one particle at each vertex of the graph at time 0; one of them is active, the others are inactive. The active particle begins to move as a continuous-time, rate 1, random walk on the graph; as soon as any active particle visits an inactive one, the latter becomes active and starts an independent random walk. Each active particle dies at the instant it reaches a total of L jumps (consecutive or not) without activating any particle. Each active particle starts with L lives and loses one life unit whenever it jumps on a vertex which has already been visited by the process. For another similar model with simultaneous jumps in discrete time, the number of informed servers has fluctuations of order $n^{3/4}$ [32].

We also emphasize that the dynamics of the present model is very similar, except for the asynchronization, to the frog model on the complete graph with finite lifetimes: the earliest reference is [1], addressing the question of final coverage. In the frog model on $\mathbb{Z}^d$, there is a phase transition between almost-sure extinction and survival with positive probability according to the underlying death rate [2], and similarly for a continuous-time model [23]. Shape theorems are proved in [3], and also in [30] and [22] in continuous time. Fluctuations are Gaussian in one dimension [8, 9], but unknown when $d \geqslant 2$.

The paper is organized as follows. The basic model is presented in Section 2 together with the main results. In Section 3 we provide an alternative construction based on a pruning procedure of the Galton–Watson tree. The law of large numbers and related fluctuation limit theorems are investigated in Section 4, including the case of heavy tails. The large deviation principle is established in Section 5.

2. The model and main results

2.1. A Markovian definition of the dynamics

From a modelling point of view, we assume that the servers emit between consecutive integers. The global state of the whole system before and after emissions is thus described at integer times. At any time $t \in \mathbb{N}$, $N_n(t)$ denotes the number of servers which have already received the message and the so-called 'total emission capital' $S_n(t)$ denotes the number of available attempts that can be used to deliver the message to a server which has not yet received it. At time 0, only one server detains the information. Its own capital, that is, $S_n(0)$ with our notation, is a random variable, the distribution of which is denoted by μ.

The dynamics of the pair process $(N_n(t), S_n(t))_{t \in \mathbb{N}}$ are then assumed to be Markovian. Conditionally on the states up to time t, the values of $N_n(t + 1)$ and $S_n(t + 1)$ are then

given by

$$S_n(t+1) = S_n(t) + \begin{cases} -1 \\ K(t+1) - 1 \end{cases} \quad \text{with probability} \begin{cases} N_n(t)/n, \\ 1 - N_n(t)/n, \end{cases} \quad (2.1)$$

where $K(t+1)$ is a random variable, with μ as distribution and independent of the past up to time t; respectively, in the above cases,

$$N_n(t+1) = \begin{cases} N_n(t) \\ N_n(t) + 1 \end{cases} \quad \text{accordingly.} \quad (2.2)$$

The Markov chain is absorbed at $S_n = 0$. From a practical point of view, the interpretation is as follows. During the emission that occurs between times t and $t+1$, one server is chosen at random among the n ones; it is referred to as the 'target'. If the target is a server that has already received the information, then the number of informed servers remains the same and the total emission capital decreases by one. Such a scenario happens with the conditional probability $N_n(t)/n$. If the target is a non-informed server, then the number of informed servers increases by one and the total emission capital increases by the own emission capital of the target which has just been activated. This happens with conditional probability $1 - N_n(t)/n$ and $K(t+1)$ then denotes the initial emission capital of the target activated between time t and time $t+1$. It is worth noting that the process $(N_n(t))_{t\in\mathbb{N}}$ is a Markov process itself, known as the 'coupon collector process' in the standard probability literature. It describes the collection, with replacement, of n equally likely coupons.

With such a modelling, the variables $K(t), t \in \mathbb{N}$, are i.i.d., with μ as common distribution, $K(0)$ being equated with $S_n(0)$ and the sequences $(K(t))_{t\in\mathbb{N}}$ and $(N_n(t))_{t\in\mathbb{N}}$ being independent (we emphasize that when the emission between t and $t+1$ is a failure, the variable $K(t+1)$ has no role in the description of the dynamics of the pair process $(N_n(t), S_n(t))_{t\in\mathbb{N}}$). By independence of the two sequences, the sum of all the capitals revealed up to time t has the same law as $R(N_n(t))$, where

$$R(t) = \sum_{s=1}^{t} K(s), \quad t \in \mathbb{N} \setminus \{0\}. \quad (2.3)$$

The transmission process lasts for a duration τ_n, which is the first time t when the emission capital is equal to 0:

$$\tau_n = \min\{t \in \mathbb{N} : S_n(t) = 0\}. \quad (2.4)$$

A natural question consists in determining whether the information will reach all servers or only a proportion of them. We then define the event of full transmission,

$$\text{Trans}_n = \{N_n(\tau_n) = n\},$$

which occurs when all the servers finally receive the information. Then, three regimes of interest can be distinguished, according to

$$\mathbb{P}(\text{Trans}_n) = 0, \text{ or } 1, \text{ or } \in (0,1),$$

which naturally correspond to different tail behaviour of K. In all these cases, we are interested in the large-n asymptotics of τ_n and $N_n(\tau_n)$.

Then, with the notation $s \wedge t = \min\{s, t\}$, the sequence $(S_n(t \wedge \tau_n), N_n(t \wedge \tau_n))_{t \in \mathbb{N}}$ is a Markov chain on $\mathbb{N} \times \{0, 1, \ldots, n\}$ with a non-decreasing second component, and absorption on the vertical axis. The harmonic equations for absorption probabilities are rather intricate for a general K, a natural route being to approximate the process by a differential equation. Here, our analysis will rely on a specific construction of the dynamics obtained by considering the Galton–Watson tree of reproduction law μ.

2.2. Construction as a labelled Galton–Watson tree

In Section 3, we construct the information transmission process with a Galton–Watson tree with degree $K \sim \mu$ and an independent coupon collector process with n coupons. We only give a quick account here. For each n, we prune the tree using the events of the n-coupon collector. Given a realization of the Galton–Watson tree, we visit successively all the nodes starting from the root, we keep (resp. erase) the current node if a new coupon is obtained at that time (resp. if we do not collect a new coupon; then the whole subtree below the current node is deleted). This results in a subtree of the original one, with size $N_n(\tau_n) \leqslant n$.

Here is a precise statement. We denote by Z_k^{GW} the cardinality of the kth generation of the Galton–Watson tree and by $\mathtt{Surv}^{\mathrm{GW}}$ the survival event $\mathtt{Surv}^{\mathrm{GW}} = \{Z_k^{\mathrm{GW}} \geqslant 1, \forall k\}$. It is well known that $\mathbb{P}(\mathtt{Surv}^{\mathrm{GW}}) > 0$ if and only if $\mathbb{E}K > 1$ or $\mathbb{P}(K = 1) = 1$.

Proposition 2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space for which there exist* (i) *a Galton–Watson tree with offspring distribution μ, and* (ii) *for each integer n, a coupon collector process with n coupons, independent of the tree.*

Then, there exist sequences $(S_n(t), t \in \mathbb{N})$ and $(N_n(t), t \in \mathbb{N})$ defined on this probability space with $N_n(0) = 1, S_n(0) \sim \mu$, and such that $((N_n(t), S_n(t)), t \in \mathbb{N})$ is a Markov chain with transitions as in (2.1) *and* (2.2).

Letting

$$Z_{\mathrm{tot}}^{\mathrm{GW}} = \sum_{k \geqslant 0} Z_k^{\mathrm{GW}},$$

the full transmission event is 'asymptotically included' in the survival event in the sense that

$$\mathtt{Trans}_n \subset \{Z_{\mathrm{tot}}^{\mathrm{GW}} \geqslant n\}, \quad \text{where } \{Z_{\mathrm{tot}}^{\mathrm{GW}} \geqslant n\} \searrow \mathtt{Surv}^{\mathrm{GW}} \quad \text{as } n \nearrow \infty,$$

and conversely, the event of termination at time $o(n)$ (where $o(n)$ stands for the Landau notation) in the information process converges to extinction in the Galton–Watson process. Letting Δ denote the symmetric difference,

$$\lim_{\epsilon \to 0^+} \lim_{n \to \infty} \mathbb{P}\big(\{\tau_n \geqslant n\epsilon\} \Delta \mathtt{Surv}^{\mathrm{GW}}\big) = 0. \tag{2.5}$$

The construction is simple and natural, but it seems to be new and it turns out to be a powerful tool for analysing the information process. In all our results below, we consider this particular coupling of the information process with the Galton–Watson tree and the coupon collector.

2.3. Limit for the information coverage and duration

For $\mathbb{E}K > 1$, define $\theta \in (0, \infty)$ as the unique root of the equation

$$\frac{1 - e^{-\theta}}{\theta} = \frac{1}{\mathbb{E}K}. \tag{2.6}$$

Extend this definition by setting $\theta = 0$ if $\mathbb{E}K \leqslant 1$ and $\theta = \infty$ if $\mathbb{E}K = \infty$. The function $\mathbb{E}K \mapsto \theta$ is an increasing bijection from $[0, \infty]$ to $[0, \infty]$. Also, let

$$p = 1 - e^{-\theta} \in [0, 1], \tag{2.7}$$

and note from (2.6) that when $\mathbb{E}K \in (1, \infty)$, then $p = 1 - e^{-\theta} \in (0, 1)$ is the unique solution of

$$p\,\mathbb{E}K = -\ln(1 - p), \tag{2.8}$$

whereas $p = 1$ when $\theta = \infty$. Let $\mathbf{1}_A$ denote the indicator function of the event A.

Theorem 2.2. *Let $\mathbb{E}K \in (0, \infty]$.*

(i) *As $n \to \infty$,*

$$\tau_n/n \longrightarrow \theta\, \mathbf{1}_{\mathrm{Surv}^{\mathrm{GW}}}$$

in probability, with θ defined by (2.6), though

$$N_n(\tau_n)/n \longrightarrow p\, \mathbf{1}_{\mathrm{Surv}^{\mathrm{GW}}}$$

in probability, with $p = 1 - e^{-\theta}$.

(ii) *If $\mathbb{E}K \leqslant 1$ ($\mathbb{P}(K = 1) < 1$ by assumption), then*

$$\lim_{n \to \infty} [\tau_n + 1] = \lim_{n \to \infty} N_n(\tau_n) = Z_{\mathrm{tot}}^{\mathrm{GW}}$$

in probability.

As a straightforward consequence, we observe that $\mathbb{P}(\mathtt{Trans}_n) \to 0$ as $n \to \infty$ whenever $\mathbb{E}K$ is finite, as p is strictly less than 1 in this case. In Theorem 2.8 we shall prove that the rate of decay is exponential in the case when K has a finite exponential moment.

2.4. Gaussian fluctuations in the case of a light tail

Theorem 2.3. *Assume $\mathbb{E}K > 1$ and $\mathbb{E}K^2 < \infty$. Let σ_K^2 denote the variance of K. As $n \to \infty$, conditionally on $\mathrm{Surv}^{\mathrm{GW}}$, we have the following convergence in law:*

$$n^{-1/2}(\tau_n - n\theta) \xrightarrow{\mathrm{law}} \mathcal{N}(0, \sigma_\tau^2),$$

with

$$\sigma_\tau^2 = [(1 - p)\mathbb{E}K - 1]^{-2}[p\sigma_K^2 + (\mathbb{E}K)^2\sigma_N(\theta)^2], \quad \sigma_N(s)^2 = e^{-s}(1 - e^{-s}) - s\,e^{-2s}.$$

Similarly, conditionally on $\mathrm{Surv}^{\mathrm{GW}}$, we have the following convergence in law:

$$n^{-1/2}(N_n(\tau_n) - np) \xrightarrow{\mathrm{law}} \mathcal{N}(0, \sigma_p^2),$$

with

$$\sigma_p^2 = [(1 - p)\mathbb{E}K - 1]^{-2}[p\sigma_K^2 e^{-2\theta} + \sigma_N(\theta)^2].$$

Theorem 2.3 extends results in [26, 27]. We omit further study of random fluctuations, such as corrections to the law of large numbers, but prefer to explore the resulting regimes of full transmission. From Theorem 2.2, when $\mathbb{E}K = \infty$, $N_n(\tau_n)/n \to 1$ in probability, conditional on survival. This leaves open the asymptotics of the probability $\mathbb{P}(\text{Trans}_n)$. Some cases are investigated in the next subsection.

2.5. Probability of full transmission in the case of a heavy tail

When K has a fat tail, fluctuations will serve the transmission process. The next result shows that the full transmission event converges to the survival event if the tail is very heavy.

Proposition 2.4. *If there exist $c > 0$ and $\alpha \in (0, 1)$ such that*

$$\liminf_{\ell \to \infty} [\ell^\alpha \mathbb{P}(K \geqslant \ell)] \geqslant c,$$

then

$$\mathbb{P}(\text{Trans}_n) \longrightarrow \mathbb{P}(\text{Surv}^{\text{GW}}) \quad \text{as } n \to \infty.$$

Hence, fluctuations of sums of K's variables play a crucial role for the occurrence of full transmission. When μ has a heavy tail, one of the servers activated during the transmission process has a large enough capital K allowing it to contact all the other servers. We now turn to the critical case, when K belongs to the domain of attraction of a stable law of index 1.

Theorem 2.5. *Assume there exists $c > 0$ such that*

$$\mathbb{P}(K \geqslant \ell) \sim \frac{c}{\ell}, \quad \ell \to \infty.$$

Then, as $n \to \infty$,

$$\mathbb{P}(\text{Trans}_n) \longrightarrow \mathbb{P}(\text{Surv}^{\text{GW}}) \times \begin{cases} \mathbb{E}(\exp(-e^{-\mathcal{S}})) & c = 1, \\ 1 & c > 1, \\ 0 & c < 1, \end{cases}$$

where, in the case $c = 1$, $\mathcal{S} = \lim_{n \to \infty} n^{-1}(R(n) - n \ln n)$ is a totally asymmetric Cauchy variable, the centering parameter of which depends on the distribution of K (see (2.3) for the definition of R).

The particular form of the limit relates to the celebrated result of Erdős and Rényi [16] for the coupon collector, that is the time to collect all coupons has a Gumbel limit law.

2.6. Large deviations

Following Chapter 1 in [11], we recall that a random sequence $(Z_n, n \geqslant 1)$, with values in some Polish space $\mathcal{Z}$, obeys a large deviation principle (LDP) with rate function $\mathcal{I}$ and speed n if

- $\mathcal{I} : \mathcal{Z} \to [0, \infty]$ is lower semi-continuous;
- for all closed subset $F \subset \mathcal{Z}$, $\limsup_{n \to \infty} n^{-1} \ln \mathbb{P}(Z_n \in F) \leqslant - \inf\{\mathcal{I}(z); z \in F\}$;
- for all open subset $O \subset \mathcal{Z}$, $\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}(Z_n \in O) \geqslant - \inf\{\mathcal{I}(z); z \in O\}$.

Assume throughout this section that K has exponentially small tails: $\mathbb{E} \exp(a_0 K) < \infty$ for some $a_0 > 0$. Then, by Cramér's theorem (*e.g.*, Section 2.2.1 in [11]), $R(n)/n$ obeys an LDP with rate I,

$$I(u) = \sup\{au - \ln \mathbb{E} \exp(aK); a \in \mathbb{R}\}, \quad u \in [0, \infty), \tag{2.9}$$

and $I(u) = \infty$ for $u < 0$. The function I is convex, lower semi-continuous and has compact level sets $\{u \in \mathbb{R} : I(u) \leqslant c\}$, for $c \in [0, \infty)$. The domain of the rate function I is defined as the set $\mathrm{Dom}(I)$ of reals u with finite $I(u)$. Here, we have $\mathrm{Dom}(I) = [k_*, k^*] \bigcap \mathbb{R}$, with

$$k_* = \min\{k : \mathbb{P}(K = k) > 0\}, \quad k^* = \sup\{k : \mathbb{P}(K = k) > 0\}.$$

We also recall the large deviation principle for the coupon collector process with n coupons from Boucheron, Gamboa and Léonard [7] and Dupuis, Nuzman and Whiting [13]. For all $t > 0$, we have

$$N_n(nt)/n \text{ obeys an LDP with speed } n \text{ and rate function } J_t. \tag{2.10}$$

From (2.7) and Section 4.1 in [13], the rate function J_t is convex; it is finite if and only if $r \in (0, t \wedge 1]$, with $\lim_{r \searrow 0} J_t(r) = \infty$, and it has a finite limit as $r \to t \wedge 1$. For $r \in (0, t \wedge 1)$, it is given by

$$J_t(r) = (1 - r) \ln(1 - r) + (t - r) \ln \rho(r, t) + t e^{-t\rho(r,t)}, \tag{2.11}$$

where $\rho(r, t)$ denotes the unique solution in $(0, \infty)$ of

$$\frac{1 - e^{-t\rho}}{\rho} = r. \tag{2.12}$$

With these ingredients we define the function $\mathcal{F} : \mathbb{R}_+^2 \to [0, \infty]$ by

$$\mathcal{F}(r, t) = \begin{cases} rI(t/r) + J_t(r) & \text{if } r > 0, \\ \infty & \text{if } r = 0, \ t > 0, \\ 0 & \text{if } r = t = 0. \end{cases} \tag{2.13}$$

Lemma 2.6.

(i) *The function $\mathcal{F}$ is lower semi-continuous on $\mathbb{R}_+^2$ with compact level sets $\{(r, t) : \mathcal{F}(r, t) \leqslant c\}$ for non-negative c. Its domain is equal to*

$$\mathrm{Dom}(\mathcal{F}) = \{(r, t) \in \mathbb{R}_+^2 : 0 < r \leqslant t \wedge 1, (k_* \vee 1)r \leqslant t \leqslant k^* r\} \cup \{(0, 0)\}.$$

It is continuous on $\mathrm{Dom}(\mathcal{F}) \setminus \{(0, 0)\}$. It is continuous at the origin if and only if K is bounded.
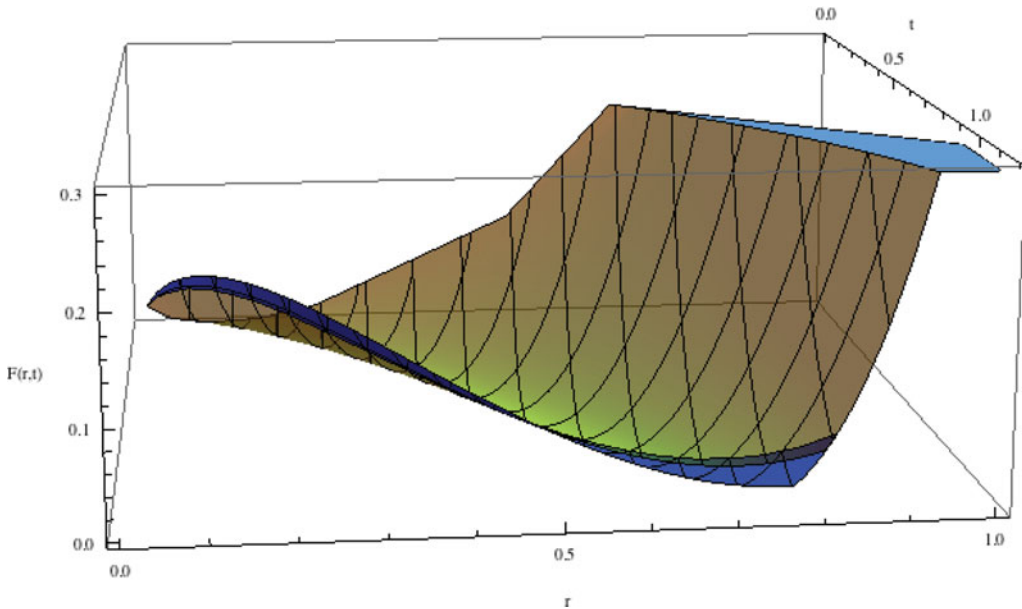
Figure 1. (Colour online) Rate function $\mathcal{F}$ for K Poisson distributed with mean 1.4998. It vanishes at the origin and at $(p, \theta) = (0.5827, 0.8740)$ it is unbounded in neighbourhoods of $(0, 0)$ in its domain. For convenience, large values of $\mathcal{F}$ are truncated, and the graph over the domain $r \leqslant 0.65t$ is not shown. The dark strip corresponds to $0.65t \leqslant r \leqslant 0.67t$, and the remainder of the graph to $r \geqslant 0.67t$.

(ii) *Moreover,*

$$\mathcal{F}\left(1 - e^{-t}, t\right) = (1 - e^{-t})I\left(\frac{t}{1 - e^{-t}}\right),$$

and when $k_* \leqslant 1 < k^*$, we have $\mathcal{F}(1 - e^{-t}, t) \sim tI(1)$ as $t \to 0$.

(iii) *When $\mathbb{E}K > 1$, the function $\mathcal{F}$ is not convex, as it takes the value 0 at points (p, θ) and $(0, 0)$, and is positive elsewhere. When $\mathbb{E}K \leqslant 1$, $\mathcal{F}$ is positive everywhere except at 0.*

The domain of $\mathcal{F}$ might be rather degenerated. For example, in the Bernoulli case $K \in \{0, 1\}$, it reduces to the segment $\{(r, r), 0 \leqslant r \leqslant 1\}$. This function is an intricate combination of the rate functions of the coupon collector process and of the Galton–Watson process. This makes it an interesting rate function in its own right. The shape of the graph of $\mathcal{F}$ is shown in Figure 1 in the case $\mathbb{E}K > 1$.

Theorem 2.7. *The sequence $\left(n^{-1}(N_n(\tau_n), \tau_n); n \geqslant 1\right)$ obeys an LDP with rate function $\mathcal{F}$ and speed n.*

As a corollary, we obtain a variational formula for the probability of full transmission.

Theorem 2.8. *The decay of the probability for all servers to be reached before exhaustion is exponential and given by*

$$\lim_{n \to +\infty} n^{-1} \ln \mathbb{P}(\mathtt{Trans}_n) = -\inf_{s \geqslant 0}\{I\left(\lambda(s)\right) + (\lambda(s) - 1)\ln\left(1 - e^{-s}\right) + \lambda(s)e^{-s}\},$$

with $\lambda(s) = s/(1 - e^{-s})$ for $s > 0$, and $\lambda(0) = 1$. The above right-hand side is negative.

3. Construction from a labelled Galton–Watson tree

Let $\mathcal{W} = \cup_{k \geqslant 0}(\mathbb{N}^*)^k$ be the set of all finite words on the alphabet $\{1, 2, \ldots\}$. Its elements are of the form $w_1 w_2 \ldots w_k, w_i \in \mathbb{N}^*$ when $k \geqslant 1$, and, for $k = 0$, $(\mathbb{N}^*)^k$ reduces to the empty word $\varnothing$, which we call the root. We then denote by $|w| = k$ the length of $w = w_1 w_2 \ldots w_k \in \mathcal{W}$ (with $|\varnothing| = 0$). For $w, w' \in \mathcal{W}$, we write $w < w'$ if $|w| < |w'|$, or $|w| = |w'|$ and $w \leqslant_{\mathrm{lex}} w'$ in the lexicographic order. We let $\preccurlyeq$ denote the usual predecessor relation in $\mathcal{W}$, that is, $w \preccurlyeq w'$ if w is a prefix of w'.

Let $(K(w), w \in \mathcal{W})$ be a family of i.i.d. random variables on $\mathbb{N} = \{0, 1, 2, \ldots\}$ with common law μ. (The notation $K(t), t \in \mathbb{N}$ is also used in Section 2.1. Here we use the same letter K but a different index, as no confusion is possible hereafter.) Assume $\mu(0) < 1$ and $\mu(1) < 1$ for a non-trivial set-up. The associated Galton–Watson tree $\mathcal{T}^{\mathrm{GW}}$ is the set of $w \in \mathcal{W}$ such that $w = \varnothing$ or, for all $i = 1, \ldots, |w|, w_i \leqslant K(v)$ with v the predecessor of w of length $i - 1$. (In other words, given a parent w' at the $(i-1)$th generation, that is, w' is a word of length $i - 1$, the children of w' are the words $w'1, \ldots, w'K(w')$, of length i, obtained by concatenation.) Let Z_k^{GW} denote the size of the kth generation of this tree, $Z_k^{\mathrm{GW}} = \mathrm{card}\,\{w \in \mathcal{T}^{\mathrm{GW}} : |w| = k\}$, given by

$$Z_{k+1}^{\mathrm{GW}} = \sum_{v \in \mathcal{T}^{\mathrm{GW}}, |v| = k} K(v), \quad Z_0^{\mathrm{GW}} = 1.$$

Recalling that $\mu(1) < 1$, it is well known that the survival event

$$\mathtt{Surv}^{\mathrm{GW}} = \bigcap_k \{Z_k^{\mathrm{GW}} \geqslant 1\} = \{\mathrm{card}\,\mathcal{T}^{\mathrm{GW}} = \infty\}$$

has complement probability

$$\sigma^{\mathrm{GW}} = 1 - \mathbb{P}(\mathtt{Surv}^{\mathrm{GW}}) = \begin{cases} = 1 & \text{if } \mathbb{E}K \leqslant 1, \\ < 1 & \text{if } \mathbb{E}K > 1. \end{cases}$$

On the same probability space, we consider an independent coupon collector process with n images ($n \geqslant 1$). Let $\Delta_{i,n}, i = 1, \ldots n - 1$, be independent, geometrically distributed random variables on $\mathbb{N}^*$ with parameter $1 - i/n$ (success probability) respectively. The success epochs are

$$T_{1,n} = 0, \quad T_{i,n} = \sum_{j=1}^{i-1} \Delta_{j,n}, \quad i = 2, \ldots n,$$

and the counting function is

$$N_n(t) = \sum_{i=1}^{n} \mathbf{1}_{\{T_{i,n} \leqslant t\}}, \quad t = 0, 1, \dots.$$

In fact, $N_n(t)$ represents the number of servers having received the information by time t (note that $1 \leqslant N_n(t) \leqslant (t+1) \wedge n$).

For any fixed integer n, with these two ingredients, we can define the transmission process together with the transmission time length τ_n. Let us start with an informal description. We browse a part of the Galton–Watson tree following the order $<$, and we paint the nodes in $\circ$ or in $\triangle$ according to the coupon collector process (success or failure). We only browse nodes which are on stand-by. As soon as a node is painted in $\circ$, its number of children nodes in $\mathcal{T}^{\mathrm{GW}}$ is revealed, and these children are put on stand-by. We then move to the next node on stand-by (next for $<$). The procedure runs until there are no longer any nodes on stand-by.

Here is a precise definition. Recursively for $t = 0, 1, \dots$, we construct $X(t) \in \mathcal{T}^{\mathrm{GW}}$, and disjoint $\mathcal{T}^{\circ}(t), \mathcal{T}^{\boxtimes}(t), \mathcal{T}^{\triangle}(t) \subset \mathcal{T}^{\mathrm{GW}}$ as follows, where $X(t)$ encodes the vertex where the tth tentative emission takes place, $\mathcal{T}^{\circ}(t)$ denotes the set of servers already informed by time t, $\mathcal{T}^{\boxtimes}(t)$ is the set of tentative emissions scheduled but not yet performed at time t, and $\mathcal{T}^{\triangle}(t)$ is set of failed emissions, *i.e.*, those performed before time t for which the target was already informed. We start with

$$X(0) = \varnothing,$$
$$\mathcal{T}^{\circ}(0) = \{\varnothing\},$$
$$\mathcal{T}^{\boxtimes}(0) = \{w \in \mathcal{T}^{\mathrm{GW}} : |w| = 1, 1 \leqslant w_1 \leqslant K(\varnothing)\},$$
$$\mathcal{T}^{\triangle}(0) = \emptyset.$$

Here and below, $\emptyset$ denotes the empty set and will not be confused with the root $\varnothing$ of the tree. With the process $(X(t), \mathcal{T}^{\circ}(t), \mathcal{T}^{\boxtimes}(t), \mathcal{T}^{\triangle}(t))$ at time t, its value at the next step $t+1$ is defined as follows.

- If $\mathcal{T}^{\boxtimes}(t)$ is non-empty, we let $X(t+1)$ be its first element,

$$X(t+1) = \inf\{w \in \mathcal{T}^{\boxtimes}(t)\}, \text{ denoted by } v,$$

  to ease the notations, and we perform a test.

  – If $N_n(t+1) = N_n(t) + 1$, we define

$$\mathcal{T}^{\circ}(t+1) = \mathcal{T}^{\circ}(t) \cup \{v\},$$
$$\mathcal{T}^{\boxtimes}(t+1) = \left(\mathcal{T}^{\boxtimes}(t) \setminus \{v\}\right) \cup \{v1, v2, \dots, vK(v)\}, \qquad (3.1)$$
$$\mathcal{T}^{\triangle}(t+1) = \mathcal{T}^{\triangle}(t),$$

  where the notation vk denotes the word of length $|v| + 1$ obtained by concatenation.
  – If $N_n(t+1) = N_n(t)$, we define

$$\mathcal{T}^{\circ}(t+1) = \mathcal{T}^{\circ}(t),$$
$$\mathcal{T}^{\boxtimes}(t+1) = \mathcal{T}^{\boxtimes}(t) \setminus \{v\}, \qquad (3.2)$$
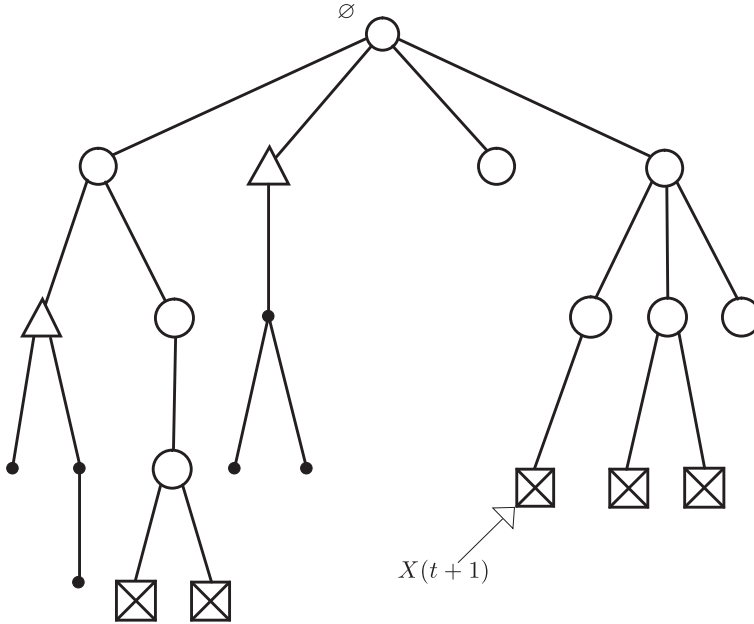$$\mathcal{T}^{\triangle}(t+1) = \mathcal{T}^{\triangle}(t) \cup \{v\}.$$

Figure 2. The Galton–Watson tree is represented up to the fourth generation. $\mathcal{T}^\circ(t)$, $\mathcal{T}^\boxtimes(t)$ and $\mathcal{T}^\triangle(t)$ and $X(t+1)$ are represented at time $t = 10$.

- If $\mathcal{T}^\boxtimes(t)$ is empty, we set $\tau_n = t$, and the construction is stopped (as well as the transmission). The set $\mathcal{T}^\circ(t) = \mathcal{T}^\circ(\tau_n) = \mathcal{T}^\circ(\infty)$ is the set of servers finally informed. Note that $\tau_n \leqslant T_{n,n}$ is a.s. finite.

We observe that for all t, $\mathcal{T}^\circ(t)$ is a tree, as well as $\mathcal{T}^\circ(t) \cup \mathcal{T}^\boxtimes(t)$. Moreover, $\mathcal{T}^\boxtimes(t) \cup \mathcal{T}^\triangle(t)$ is a cutset of $\mathcal{T}^{\mathrm{GW}}$ for its own graph structure.

Figure 2 provides an example of construction of the sets $\mathcal{T}^\circ(t)$, $\mathcal{T}^\boxtimes(t)$ and $\mathcal{T}^\triangle(t)$, for $t \in \{0, 1, \ldots, 10\}$, according to the rules prescribed in (3.1) and (3.2). At time 0, the capital of the initial server is $K(\varnothing) = 4$, so that the root of the tree has four children. Between times 0 and 1, the first emission is a success as the server which is revealed in the tree is painted in $\circ$. This server reads as the first child (starting from the left) at the first generation of the tree. It has two children, that is, $K(\text{`1'}) = 2$, where '1' is here understood as a one-letter word. At time 1, the total emission capital is thus $S(1) = 4 + 2 - 1 = 5$, and $\mathcal{T}^\boxtimes(2)$ contains the three last children at the first generation and the two first children at the second generation. Since the second child at the first generation is painted in $\triangle$, the emission between times 1 and 2 fails, which means that the children of this node are not considered for the rest of the construction. At time 2, $S(2) = 4 + 2 - 2 = 4$. Then, the emission between times 2 and 3 is a success but the server which is informed has no children (third child in the first generation) and the emission between times 3 and 4 is a success as well, with $K(\text{`4'}) = 3$. At time 4, $\mathcal{T}^\boxtimes(4)$ contains five nodes, all of them at the second generation of the tree. The node $X(5)$ is then the first child at the second generation: as it is painted in $\triangle$, the emission between times 4 and 5 fails. And so on up

to time 10. Then, $\mathcal{T}^{\boxtimes}(10)$ contains five nodes, two of them at the fourth generation and three of them at the third generation. With the lexicographic order, $X(11) = \text{`411'}$.

We now relate the above construction to the dynamical model for transmission. Let $\operatorname{card} A$ denote the cardinality of a set A. Consider a new, independent, i.i.d. sequence $(\bar{K}_i)_{i \geqslant 1}$ with law μ, and define, for $i = 1, \ldots n$,

$$K_i = \begin{cases} K(X(T_{i,n})) & \text{if } i \leqslant \operatorname{card} \mathcal{T}^{\circ}(\infty), \\ \bar{K}_i & \text{if } i > \operatorname{card} \mathcal{T}^{\circ}(\infty), \end{cases} \tag{3.3}$$

and also

$$S_n(t) := \sum_{i=1}^{N_n(t)} K_i - t, \quad t \in \mathbb{N} \setminus \{0\}. \tag{3.4}$$

Below, we also write

$$S_n(t) = R(N_n(t)) - t, \quad t \in \mathbb{N}, \quad \text{with} \quad R(m) := \sum_{i=1}^{m} K_i, \quad m \in \mathbb{N} \setminus \{0\}. \tag{3.5}$$

Note that the letter R is also used in (2.3), but as proved below, the two R have the same distribution. Henceforth, we will always refer to (3.5) for the precise definition of R.

By construction, we have

$$N_n(t) = \operatorname{card} \mathcal{T}^{\circ}(t), \quad S_n(t) = \operatorname{card} \mathcal{T}^{\boxtimes}(t). \tag{3.6}$$

Proposition 3.1. *The variables* $(K_i)_{1 \leqslant i \leqslant n}$ *are independent, identically distributed with law* μ, *and* $(K_i)_{1 \leqslant i \leqslant n}$ *is independent of* $(T_{i,n})_{1 \leqslant i \leqslant n}$. *Moreover,*

$$\tau_n = \min\{t \in \mathbb{N} : S_n(t) = 0\}. \tag{3.7}$$

Proof. The formula for τ_n directly follows from (3.6), and the fact that τ_n is finite. We now investigate the distribution of the sequence $(K_i, i \leqslant n)$. Below, we let $\mathcal{F}_w = \sigma(K(w'), w' \leqslant w)$ for $w \in \mathcal{W}$. For the event

$$A = \{(T_{1,n}, \ldots, T_{i,n}) = (k_1, \ldots, k_i), X(k_i - 1) = w, \tau_n \geqslant k_i\},$$
$$0 = k_1 < \cdots < k_i \quad \text{and} \quad w \in \mathcal{W}, \quad |w| \leqslant k_i - 1,$$

$X(T_{i,n})$ coincides with an $\mathcal{F}_w$-measurable random variable, denoted by χ, which satisfies $w < \chi$ almost surely (this follows from the monotonicity of the browsing procedure). Similarly, all the variables $K(X(T_{j,n}))$, $1 \leqslant j \leqslant i - 1$ coincide with $\mathcal{F}_w$-measurable random variables on A. Clearly, $K(\chi)$ is independent of $\mathcal{F}_w \vee \sigma(T_{1,n}, \ldots, T_{n,n})$ and has μ as distribution, since the random variables $(K(w'), w' \in \mathcal{W})$ are i.i.d. and are independent of the success epochs $(T_{1,n}, \ldots, T_{n,n})$. Obviously, the event A belongs to $\mathcal{F}_w \vee \sigma(T_{1,n}, \ldots, T_{n,n})$. This proves that, for any bounded and measurable Borel function ϕ,

$$\mathbb{E}\big[\phi(K_i)\mathbf{1}_{\{T_{i,n} \leqslant \tau_n\}} | (T_{j,n}, 1 \leqslant j \leqslant n), (K_j, 1 \leqslant j \leqslant i - 1)\big] = \mathbf{1}_{\{T_{i,n} \leqslant \tau_n\}} \int_{\mathbb{N}} \phi \, d\mu.$$

For the event $\{(T_{1,n}, \ldots, T_{i,n}) = (k_1, \ldots, k_i), \tau_n < k_i\}$, K_i coincides with $\bar{K}_i$, which is obviously independent of $\sigma((T_{j,n}, 1 \leqslant j \leqslant n), (K_j, 1 \leqslant j \leqslant i-1))$, so that the above equation also holds with $T_{i,n} \leqslant \tau_n$ replaced by $T_{i,n} > \tau_n$. $\qquad\square$

Then the process we have constructed here corresponds to the description of the information transmission process given in Section 2.1.

Proof of Proposition 2.1. Equation (2.5) follows from Lemma 4.2 below and $\{\tau_n \geqslant n\epsilon\} \subset \{Z_{\text{tot}}^{\text{GW}} \geqslant n\epsilon\}$. The other claims directly follow from the construction and Proposition 3.1. $\qquad\square$

4. Proofs of law of large numbers and fluctuations

In all the proofs, we use the following convention. For any discrete process $(V_t, t \in \mathbb{N})$, $(V_t, t \geqslant 0)$ stands for $(V_{\lfloor t \rfloor}, t \geqslant 0)$, where $\lfloor \cdot \rfloor$ is the floor function. We will also use the ceiling function, denoted by $\lceil \cdot \rceil$. For an interval I in $\mathbb{R}$, define the Skorokhod space $D(I)$ as the space of *càdlàg* (right continuous left limited) functions from I to $\mathbb{R}$.

4.1. Proofs of the law of large numbers

Lemma 4.1. *We have the following convergence in law of sequences of processes on the Skorohod space (endowed with the standard J1 topology, keeping in mind that convergence for the J1 topology implies uniform convergence on compacts when the limit function is continuous).*

(i) *For $D([0, 1))$,*

$$n^{-1/2}\left(T_{\lfloor nq \rfloor, n} - n \ln \frac{1}{1-q}\right)_{0 \leqslant q < 1} \xrightarrow{\text{law}} \left[B\left(\sigma_T(q)^2\right)\right]_{0 \leqslant q < 1} \quad \text{as } n \to \infty,$$

where B is a standard Brownian motion, and

$$\sigma_T(q)^2 = \frac{q}{1-q} + \ln(1-q) > 0.$$

(ii) *For $D(\mathbb{R}^+)$,*

$$\left(n^{-1/2}\left(N_n(ns) - n(1 - e^{-s})\right)\right)_{s \geqslant 0} \xrightarrow{\text{law}} \left[B\left(\sigma_N(s)^2\right)\right]_{s \geqslant 0} \quad \text{as } n \to \infty,$$

where B is a standard Brownian motion, and

$$\sigma_N(s)^2 = e^{-s}(1 - e^{-s}) - s\, e^{-2s} > 0.$$

Both limits are independent-increment Gaussian processes with zero mean, and they are martingales.

Proof. Assertion (i) is a direct application of the invariance principle for triangular array of independent, but not identically distributed, square-integrable random variables; see Dacunha-Castelle and Duflo [10, Théorème 7.4.28] or Jacod and Shiryaev [18,

Chapter VII, Theorem 5.4]. The variance is computed as the limit of a Riemann sum:

$$\sigma_T(q)^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{qn} \text{Var}(\Delta_{i,n}) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{qn} \frac{i/n}{(1-i/n)^2} = \int_0^q \frac{y}{(1-y)^2} \, dy.$$

Assertion (ii) follows from (i), using that $N_n(\cdot)$ and $T_{\lfloor n \cdot \rfloor, n}$ are reciprocal in a generalized sense. With $f(q) = -\ln(1-q)$, $f^{-1}(s) = 1 - e^{-s}$, we have (see Billingsley [6, Theorem 17.3])

$$\sigma_N(s)^2 = \sigma_T(f^{-1}(s))^2 \times [f' \circ f^{-1}(s)]^{-2}.$$

$\square$

The next lemma is one of the key arguments of the whole analysis. It shows that when the Galton–Watson tree is infinite, transmission takes place on a macroscopic time level.

Lemma 4.2. *There exists $\varepsilon_0 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_0)$,*

$$\lim_{n \to \infty} \mathbb{P}(\tau_n \geqslant n\varepsilon, \text{Surv}^{\text{GW}}) = \mathbb{P}(\text{Surv}^{\text{GW}}) = 1 - \sigma^{\text{GW}}.$$

Proof. Since the claim is trivial for $\sigma^{\text{GW}} = 1$, we just need to consider the case when $\mathbb{E}K > 1$. Letting $k = \lfloor \ln^2 n \rfloor$, we estimate

$$\mathbb{P}(\tau_n \leqslant n\varepsilon, \text{Surv}^{\text{GW}}) \leqslant \mathbb{P}(N_n(k) \leqslant k) + \mathbb{P}(k \leqslant \tau_n \leqslant n\varepsilon, \text{Surv}^{\text{GW}}), \tag{4.1}$$

using that

$$\{N_n(k) = k + 1, \text{Surv}^{\text{GW}}\} \subset \{\tau_n \geqslant k, \text{Surv}^{\text{GW}}\},$$

which implies that

$$\mathbb{P}(N_n(k) \leqslant k, \text{Surv}^{\text{GW}}) \geqslant \mathbb{P}(\tau_n < k, \text{Surv}^{\text{GW}}).$$

We start with

$$\begin{aligned}
\mathbb{P}(N_n(k) \leqslant k) &= 1 - (1 - 1/n) \times \ldots (1 - k/n) \\
&\leqslant 1 - (1 - k/n)^k \\
&\sim k^2/n \quad \text{as } k^2/n \to 0. \tag{4.2}
\end{aligned}$$

Let $\varepsilon_0 = (1/2)(1 - 1/\mathbb{E}K) > 0$. Fix $\varepsilon \in (0, \varepsilon_0)$, and note that $(1 - 2\varepsilon)\mathbb{E}K > 1$. It remains to prove the convergence

$$\mathbb{P}(k \leqslant \tau_n \leqslant n\varepsilon, \text{Surv}^{\text{GW}}) = \sum_{i=k}^{n\varepsilon} \mathbb{P}(\tau_n = i, \text{Surv}^{\text{GW}}) \leqslant \sum_{i=k}^{n\varepsilon} \mathbb{P}(S_n(i) \leqslant 0) \to 0, \tag{4.3}$$

where the inequality holds by definition of τ_n (2.4). We start by showing that there exists a constant $C_\varepsilon > 0$, independent of n, such that

$$\mathbb{P}(N_n(i) < \lceil (1 - 2\varepsilon)i \rceil) \leqslant \exp(-C_\varepsilon i), \quad \text{for all } i \leqslant n\varepsilon.$$

Indeed, the above probability is equal to

$$\mathbb{P}(N_n(i) < \lceil (1 - 2\varepsilon)i \rceil) = \mathbb{P}(T_{\lceil (1-2\varepsilon)i \rceil, n} > i) \leqslant \mathbb{P}(\bar{T}^\varepsilon_{(1-2\varepsilon)i} > i)$$

with $\bar{T}^{\varepsilon}_{(1-2\varepsilon)i}$ a sum of $\lceil(1-2\varepsilon)i\rceil$ i.i.d. geometric random variables with parameter $1-\varepsilon$. Now the desired estimate follows from Chernov's bound. Next, we note that, for $z \in (0,1)$, $i \leqslant n\varepsilon$ and $G(z) = \mathbb{E}z^K$,

$$
\begin{aligned}
\mathbb{P}\big(S_n(i) \leqslant 0, N_n(i) \geqslant \lceil(1-2\varepsilon)i\rceil\big) &\leqslant \mathbb{E}\big[z^{S_n(i)}; N_n(i) \geqslant \lceil(1-2\varepsilon)i\rceil\big] \\
&\leqslant z^{-i}\mathbb{E}\big[z^{R(\lceil(1-2\varepsilon)i\rceil)}; N_n(i) \geqslant \lceil(1-2\varepsilon)i\rceil\big] \\
&\leqslant z^{-i}G(z)^{\lceil(1-2\varepsilon)i\rceil} \leqslant z^{-i}G(z)^{(1-2\varepsilon)i},
\end{aligned}
$$

where we have used

$$
S_n(i) = R(N_n(i)) - i \geqslant R(\lceil(1-2\varepsilon)i\rceil) - i
$$

for the event $\{N_n(i) \geqslant \lceil(1-2\varepsilon)i\rceil\}$ to pass from the first to the second line, R being given by (3.5). Since $(1-2\varepsilon)\mathbb{E}K > 1$, we have $r := z^{-1}G(z)^{(1-2\varepsilon)} < 1$ by picking $z < 1$ close enough to 1 and then by expanding $G(z)$ as $G(z) = 1 + \mathbb{E}K(z-1) + o(z-1)$. Thus, the left-hand side of (4.3) is bounded by

$$
\mathbb{P}(k \leqslant \tau_n \leqslant n\varepsilon, \mathrm{Surv}^{\mathrm{GW}}) \leqslant \sum_{i=k}^{n\varepsilon}[r^i + \exp(-C_\varepsilon i)] \leqslant 2(1-r_1)^{-1}r_1^{k+1},
$$

with $r_1 = \max\{r, \exp(-C_\varepsilon)\} < 1$. Collecting these estimates in (4.1) and taking $k = \lfloor\ln^2 n\rfloor$, we conclude that $\mathbb{P}(\tau_n \leqslant n\varepsilon, \mathrm{Surv}^{\mathrm{GW}}) = O(n^{-a})$ for all $a \in (0,1)$ ($O(\cdot)$ standing for the Landau notation). $\qquad\square$

Proof of Theorem 2.2. We start with the proof of (i). We assume first that $\mathbb{E}K < \infty$. Then we can apply the law of large numbers to the process

$$
R(m) = \sum_{i=1}^{m} K_i \tag{4.4}
$$

in (3.5), to show that $\mathbb{P}$-a.s., $R(nq)/n \to q\mathbb{E}K$ uniformly on $[0,1]$. Recall from (3.5) that

$$
S_n(t) = R(N_n(t)) - \lfloor t\rfloor. \tag{4.5}
$$

In addition to Lemma 4.1, this shows that, in probability,

$$
S_n(ns)/n \longrightarrow (1-e^{-s})\mathbb{E}K - s, \quad \text{uniformly on compacts of } \mathbb{R}_+ \tag{4.6}
$$

as $n \to \infty$. As a consequence, for any $\delta > 0$,

$$
\mathbb{P}\big(\tau_n > n(\theta+\delta)\big) \leqslant \mathbb{P}\big(\inf[S_n(ns), s \in [0, \theta+\delta]] \geqslant 0\big) \to 0 \quad \text{as } n \to \infty, \tag{4.7}
$$

since $(1-e^{-s})\mathbb{E}K - s < 0$ for $s > \theta$. Now, with $A^{\complement}$ the complement of A, we write

$$
\begin{aligned}
\mathbb{P}\big(|\tau_n - n\theta\,\mathbf{1}_{\mathrm{Surv}^{\mathrm{GW}}}| > n\delta\big) &= \mathbb{P}\big(|\tau_n - n\theta| > n\delta, \mathrm{Surv}^{\mathrm{GW}}\big) + \mathbb{P}\big(\tau_n > n\delta, (\mathrm{Surv}^{\mathrm{GW}})^{\complement}\big) \\
&\leqslant \mathbb{P}\big(\tau_n < n\varepsilon, \mathrm{Surv}^{\mathrm{GW}}\big) + \mathbb{P}\big(|\tau_n - n\theta| > n\delta, \mathrm{Surv}^{\mathrm{GW}}, \tau_n \geqslant n\varepsilon\big) \\
&\quad + \mathbb{P}\big(\tau_n > n\delta, (\mathrm{Surv}^{\mathrm{GW}})^{\complement}\big),
\end{aligned}
$$

where the first term of the right-hand side tends to 0 (as $n \to \infty$) from Lemma 4.2. The last term vanishes because τ_n is smaller than the extinction time of the Galton–Watson

process, which is a.s. finite for the extinction event. Since τ_n is the first time such that $S_n(\tau_n) = 0$, the second term also tends to 0 by (4.6). Indeed, (4.7) yields

$$\limsup_{n\to\infty} \mathbb{P}\big(|\tau_n - n\theta| > n\delta, \mathrm{Surv}^{\mathrm{GW}}, \tau_n \geqslant n\varepsilon\big)$$

$$\leqslant \limsup_{n\to\infty} \mathbb{P}\big(n\varepsilon \leqslant \tau_n \leqslant n\theta - n\delta, \mathrm{Surv}^{\mathrm{GW}}, \tau_n \geqslant n\varepsilon\big)$$

$$\leqslant \limsup_{n\to\infty} \mathbb{P}\big(\inf\big[S_n(ns), s \in [\varepsilon, \theta - \delta]\big] = 0\big) = 0,$$

since $(1 - e^{-s})\mathbb{E}K - s > 0$ for $s \in [\varepsilon, \theta - \delta]$. This ends the proof of the first claim in (i) when $\mathbb{E}K$ is finite. The second claim in (i) is then a straightforward consequence of the first one and of (ii) in Lemma 4.1, from which $N_n(ns)/n \to 1 - e^{-s}$ in probability, uniformly on compacts of $\mathbb{R}_+$.

When $\mathbb{E}K = \infty$, we consider $K^{(L)}(w) = \min\{K(w), L\}$ for a truncation level $L > 0$. By the above proof, we obtain $\lim_{n\to\infty} \tau_n^{(L)}/n = \theta^{(L)}\mathbf{1}_{\mathrm{Surv}^{\mathrm{GW}(L)}}$ with obvious notations. Since

$$\tau_n^{(L)} \leqslant \tau_n, \quad \lim_{L\to\infty} \theta^{(L)} = \theta = \infty \quad \text{and} \quad \mathrm{Surv}^{\mathrm{GW}(L)} \nearrow \mathrm{Surv}^{\mathrm{GW}},$$

as $L \nearrow \infty$, we deduce that $\tau_n/n \to \theta$ on $\mathrm{Surv}^{\mathrm{GW}}$ by letting L tend to ∞. For the extinction event $(\mathrm{Surv}^{\mathrm{GW}})^{\complement}$, we obviously have $\tau_n/n \to 0$, since τ_n is less than the total number of nodes in the tree. This proves the first claim in (i). For the second one, we note in the same way that

$$\lim_{n\to\infty} N_n(\tau_n^{(L)})/n = p^{(L)}\mathbf{1}_{\mathrm{Surv}^{\mathrm{GW}(L)}} \quad \text{and} \quad \lim_{L\to\infty} p^{(L)} = 1.$$

Since $N_n(\tau_n^{(L)}) \leqslant N_n(\tau_n) \leqslant n$, we deduce that $\lim_{n\to\infty} N_n(\tau_n)/n = 1$ on $\mathrm{Surv}^{\mathrm{GW}}$. For the extinction event, it obviously holds that $N_n(\tau_n)/n \to 0$.

We now turn to the proof of assertion (ii). By construction, it must hold that

$$N_n(\tau_n) \leqslant \tau_n + 1 \leqslant Z_{\mathrm{tot}}^{\mathrm{GW}},$$

which is a.s. finite when $\mathbb{E}K \leqslant 1$. By (4.2), for any $A \geqslant 0$,

$$\lim_{n\to\infty} \mathbb{P}(N_n(\tau_n) \leqslant \tau_n, Z_{\mathrm{tot}}^{\mathrm{GW}} \leqslant A) = 0,$$

which proves that, asymptotically, all the emissions before exhaustion of the capital are (almost surely) successful for the event $\{Z_{\mathrm{tot}}^{\mathrm{GW}} \leqslant A\}$. Therefore, asymptotically, all the emissions before exhaustion are (almost surely) successful, which is to say that, asymptotically with probability 1, every node of the tree receives the information. $\qquad\square$

4.2. Proof of the Gaussian fluctuations in the case of a light tail
Proof of Theorem 2.3. By the invariance principle,

$$\big(n^{-1/2}\big(R(nq) - nq\mathbb{E}K\big)\big) \xrightarrow{\text{law}} \bar{B}(q\sigma_K^2))_{q\geqslant 0}$$

in $D(\mathbb{R}_+)$ with $\bar{B}$ a Brownian motion. By independence of $(K_i)_i$ and $(\Delta_{i,n})_i$, we have from Lemma 4.1 that

$$n^{-1/2}\begin{pmatrix} (R(nq) - nq\mathbb{E}K)_{q\geqslant 0} \\ (N_n(ns) - n(1 - e^{-s}))_{s\geqslant 0} \end{pmatrix} \xrightarrow{\text{law}} \begin{pmatrix} \big[\bar{B}\big(q\sigma_K^2\big)\big]_{q\geqslant 0} \\ \big[B\big(\sigma_N(s)^2\big)\big]_{s\geqslant 0} \end{pmatrix} \qquad (4.8)$$

as $n \to \infty$, in $D(\mathbb{R}_+)^2$ endowed with the product topology generated by the J1 topology (the convergence holding true as well for the uniform topology), where B and $\bar{B}$ are independent. Then the convergence

$$n^{-1/2}\big(R(N_n(ns)) - n(1 - e^{-s})\mathbb{E}K\big)_{s \geqslant 0} \xrightarrow{\text{law}} \big[\bar{B}\big((1 - e^{-s})\sigma_K^2\big) + B\big((\mathbb{E}K)^2\sigma_N(s)^2\big)\big]_{s \geqslant 0} \quad (4.9)$$

holds in the Skorohod space. In fact we claim that (4.9) also holds conditionally on Surv^{GW}, that is, under $\mathbb{P}(\cdot|\text{Surv}^{\text{GW}})$. The reason is that the process

$$n^{-1/2}\big(R(N_n(ns)) - n(1 - e^{-s})\mathbb{E}K\big)_{s \geqslant 0}$$

and the event Surv^{GW} are asymptotically independent; see Lemma 4.3 below.

Now, by Theorem 2.2,

$$L_n = \tau_n - n\theta$$

is such that $L_n/n \to 0$ as $n \to \infty$ in probability under $\mathbb{P}(\cdot|\text{Surv}^{\text{GW}})$. Using (4.9) with $ns = \tau_n = n\theta + L_n$, we get

$$n^{-1/2}\big(R(N_n(\tau_n)) - n(1 - e^{-\theta - L_n/n})\mathbb{E}K\big) \xrightarrow{\text{law}} \bar{B}\big(p\sigma_K^2\big) + B\big((\mathbb{E}K)^2\sigma_N(\theta)^2\big) \quad (4.10)$$

under $\mathbb{P}(\cdot|\text{Surv}^{\text{GW}})$. Now, we emphasize from (2.6) that

$$\begin{aligned}
\tau_n - n(1 - e^{-\theta - L_n/n})\mathbb{E}K &= n\theta + L_n - n(1 - e^{-\theta})\mathbb{E}K - ne^{-\theta}(1 - e^{-L_n/n})\mathbb{E}K \\
&= L_n - ne^{-\theta}(1 - e^{-L_n/n})\mathbb{E}K. \quad (4.11)
\end{aligned}$$

Since

$$\mathbb{P}(R(N_n(\tau_n)) = \tau_n|\text{Surv}^{\text{GW}}) \to 1 \quad \text{as } n \to \infty,$$

we deduce from (4.10) and (4.11) that, conditionally on survival,

$$n^{-1/2}\big(L_n - ne^{-\theta}(1 - e^{-L_n/n})\mathbb{E}K\big) \xrightarrow{\text{law}} \bar{B}(p\sigma_K^2) + B((\mathbb{E}K)^2\sigma_N(\theta)^2). \quad (4.12)$$

Since

$$n^{-1/2}L_n = \frac{n^{-1/2}(L_n - ne^{-\theta}(1 - e^{-L_n/n})\mathbb{E}K)}{1 - e^{-\theta}\mathbb{E}K(1 + O(L_n/n))},$$

we deduce from (4.12) that

$$n^{-1/2}L_n \xrightarrow{\text{law}} (1 - e^{-\theta}\mathbb{E}K)^{-1}[\bar{B}(p\sigma_K^2) + B((\mathbb{E}K)^2\sigma_N(\theta)^2)] \quad \text{under } \mathbb{P}(\cdot|\text{Surv}^{\text{GW}}). \quad (4.13)$$

Noting that $1 - e^{-\theta}\mathbb{E}K = 1 - (1 - p)\mathbb{E}K$, this proves the first claim and the value of σ_τ^2.

To prove the second claim, we note that the left-hand sides in (4.13) and in the second line of (4.8) jointly converge as a two-dimensional vector under the conditional law $\mathbb{P}(\cdot|\text{Surv}^{\text{GW}})$. Therefore,

$$n^{-1/2}\big(N_n(\tau_n) - n(1 - e^{-\theta}) - e^{-\theta}L_n\big) \xrightarrow{\text{law}} B\big(\sigma_N(\theta)^2\big) \quad \text{under } \mathbb{P}(\cdot|\text{Surv}^{\text{GW}}),$$

and also

$$n^{-1/2}\big(N_n(\tau_n) - n(1 - e^{-\theta})\big) \xrightarrow{\text{law}} (1 - e^{-\theta}\mathbb{E}K)^{-1}\big[e^{-\theta}\bar{B}(p\sigma_K^2) + B(\sigma_N(\theta)^2)\big],$$

conditionally on $\mathtt{Surv}^{\text{GW}}$. This completes the proof. □

We finally prove the following result.

Lemma 4.3. *Under the assumptions of Theorem 2.3, the process*

$$(n^{-1/2}(R(nq) - nq\mathbb{E}K))_{q \geqslant 0}$$

has the same limits in law under $\mathbb{P}(\cdot|(\mathtt{Surv}^{\text{GW}})^{\complement})$ *and* $\mathbb{P}$.

Proof. It is sufficient to prove that, when $\mathbb{P}(\mathtt{Surv}^{\text{GW}}) < 1$, the process

$$(n^{-1/2}(R(nq) - nq\mathbb{E}K))_{q \geqslant 0}$$

has the same limits in law under $\mathbb{P}(\cdot|(\mathtt{Surv}^{\text{GW}})^{\complement})$ and $\mathbb{P}$. From (3.3), we know that, for the event $\{Z_{\text{tot}}^{\text{GW}} \leqslant A\}$, for $A \geqslant 0$, the variable K_i in the definition of R (see (3.5)) coincides with $\bar{K}_i$ for $i > A$, so that the process

$$\sum_{j=A}^{\lfloor nq \rfloor} \bar{K}_j = R(nq) - \sum_{j=1}^{A} K_j$$

is independent of the Galton–Watson tree. Therefore, the event $\{Z_{\text{tot}}^{\text{GW}} \leqslant A\}$ and the process $(n^{-1/2}(R(nq) - nq\mathbb{E}K))_{q \geqslant 0}$ are asymptotically independent, which is to say that $(\mathtt{Surv}^{\text{GW}})^{\complement}$ and $(n^{-1/2}(R(nq) - nq\mathbb{E}K))_{q \geqslant 0}$ are asymptotically independent. Therefore, the process

$$(n^{-1/2}(R(nq) - nq\mathbb{E}K))_{q \geqslant 0}$$

has the same limits in law under $\mathbb{P}(\cdot|(\mathtt{Surv}^{\text{GW}})^{\complement})$ and $\mathbb{P}$. □

4.3. Proof of the fluctuations in the case of a heavy tail

Proof of Proposition 2.4. Clearly, $\mathbb{E}K = \infty$. By Theorem 2.2, for any $\varepsilon > 0$, as $n \to \infty$,

$$\mathbb{P}(N_n(\tau_n)/n \geqslant 1 - \varepsilon, \mathtt{Surv}^{\text{GW}}) \longrightarrow \mathbb{P}(\mathtt{Surv}^{\text{GW}}).$$

Moreover,

$$\{N_n(\tau_n) = n\} \supset \{N_n(\tau_n) \geqslant n/2, R(\lfloor n/2 \rfloor) \geqslant T_{n,n}\}$$

since

$$R(\lfloor n/2 \rfloor) \geqslant T_{n,n} \Rightarrow R(k) \geqslant k \quad \text{for any } k \in \{\lfloor n/2 \rfloor, \ldots, T_{n,n}\}.$$

Thus, for any $\beta \in (1, 1/\alpha)$,

$$\liminf_{n \to \infty} \mathbb{P}\big(N_n(\tau_n) = n, \mathtt{Surv}^{\text{GW}}\big) \geqslant \liminf_{n \to \infty} \mathbb{P}\big(R(n/2) \geqslant n^{\beta}, T_{n,n} \leqslant n^{\beta}, \mathtt{Surv}^{\text{GW}}\big).$$

By Markov's inequality, $\mathbb{P}(T_{n,n} > n^\beta) \to 0$ as $n \to \infty$, since

$$\mathbb{E} T_{n,n} = \sum_{i=1}^{n-1} n/(n-i) \sim_{n\to\infty} n\ln(n).$$

Moreover, $\mathbb{P}(K_1 \geqslant n^\beta) \geqslant (c/2)n^{-\alpha\beta}$ for n large enough, so that, for n large,

$$\mathbb{P}\left(\bigcap_{i=1}^{\lfloor n/2 \rfloor} \{K_i < n^\beta\}\right) \leqslant \left(1 - \frac{c}{2n^{\alpha\beta}}\right)^{\lfloor n/2 \rfloor} \sim \exp\left(-\frac{c}{4}n^{1-\alpha\beta}\right) \longrightarrow 0 \quad \text{as } n \to \infty.$$

Therefore,

$$\mathbb{P}\left(\sum_{i=1}^{\lfloor n/2 \rfloor} K_i \geqslant n^\beta\right) \to 1.$$

We deduce that

$$\liminf_{n\to\infty} \mathbb{P}(N_n(\tau_n) = n, \mathrm{Surv}^{\mathrm{GW}}) = \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}}).$$

Finally, on $(\mathrm{Surv}^{\mathrm{GW}})^{\complement}$, we have $N_n(\tau_n)/n \to 0$ in probability, so that

$$\mathbb{P}(N_n(\tau_n) = n, \quad (\mathrm{Surv}^{\mathrm{GW}})^{\complement}) \to 0.$$

$\square$

Proof of Theorem 2.5. By a celebrated result of [16] (see [15, pp. 130–132] for a brief account),

$$G^{(n)} = n^{-1}(T_{n,n} - n\ln n) \xrightarrow{\text{law}} G, \tag{4.14}$$

where the variable G has a Gumbel distribution,

$$\mathbb{P}(G \leqslant x) = e^{-e^{-x}}, \quad x \in \mathbb{R}. \tag{4.15}$$

On the other hand, from the tail assumption for K,

$$\mathcal{S}_c^{(n)} = n^{-1}(R(n-1) - cn\ln n) \xrightarrow{\text{law}} \mathcal{S}_c, \tag{4.16}$$

where $\mathcal{S}_c$ is a totally asymmetric, stable law with index 1 (Cauchy's law), depending upon the parameter c. We define $\mathcal{S} = \mathcal{S}_1$. Precisely, the law of $\mathcal{S}$ is given by

$$\mathbb{E}e^{iu\mathcal{S}} = \exp\left\{\int_0^1 (e^{iux} - 1 - iux)x^{-2}dx + \int_1^\infty (e^{iux} - 1)x^{-2}dx + iuc_0\right\}, \quad u \in \mathbb{R},$$

where $i^2 = -1$ and $c_0 \in \mathbb{R}$ is defined by $c_0 = \lim_{n\to\infty}(\mathbb{E}[K; K \leqslant n] - c\ln n)$. Recall from Theorem 2.2 that

$$\mathbb{P}(\mathrm{Trans}_n \cap \mathrm{Surv}^{\mathrm{GW}}) - \mathbb{P}(\mathrm{Trans}_n) \to 0.$$

We have

$$
\begin{aligned}
\mathrm{Trans}_n \cap \mathrm{Surv}^{\mathrm{GW}} &= \{\tau_n \geqslant T_{n,n}\} \cap \mathrm{Surv}^{\mathrm{GW}} \\
&\subset \{R(n-1) \geqslant T_{n,n}\} \cap \mathrm{Surv}^{\mathrm{GW}} \\
&\subset \{G^{(n)} - \mathcal{S}_c^{(n)} \leqslant (c-1)\ln n\} \cap \mathrm{Surv}^{\mathrm{GW}},
\end{aligned}
$$

where we have used

$$
\tau_n \geqslant T_{n,n} \Rightarrow R(n-1) = R(N_n(T_{n,n}-1)) > T_{n,n}-1
$$

to pass from the first to the second line. The random vector $(G^{(n)}, \mathcal{S}_c^{(n)})$ converges in law to a couple $(G, \mathcal{S}_c)$ with independent components (independence follows from the independence of the processes N_n and R). If $c < 1$, then $\mathbb{P}(G^{(n)} - \mathcal{S}_c^{(n)} \leqslant (c-1)\ln n) \to 0$ so that $\mathbb{P}(\mathrm{Trans}_n) \to 0$. If $c > 1$, it obviously holds that $\limsup_{n\to\infty}\mathbb{P}(\mathrm{Trans}_n) \leqslant \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}})$. To tackle the case when $c = 1$, we observe from Lemma 4.4 below that $\mathrm{Surv}^{\mathrm{GW}}$ and $\mathcal{S}_1^{(n)}$ are asymptotically independent as $n \to \infty$, so that the random vector $(G^{(n)}, \mathcal{S}_1^{(n)}, \mathbf{1}_{\mathrm{Surv}^{\mathrm{GW}}})$ converges in law to a triple $(G, \mathcal{S}, \mathbf{1}_A)$ with independent components, where $\mathbb{P}(A) = \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}})$. As the random variable $G - \mathcal{S}$ has a continuous cumulative distribution function, we get

$$
\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}(\mathrm{Trans}_n) &\leqslant \limsup_{n\to\infty} \mathbb{P}\big(\{G^{(n)} - \mathcal{S}_1^{(n)} \leqslant 0\} \cap \mathrm{Surv}^{\mathrm{GW}}\big) \\
&= \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}})\mathbb{P}(G \leqslant \mathcal{S}) = \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}})\mathbb{E}(\exp(-e^{-\mathcal{S}})),
\end{aligned}
$$

the last equation following from (4.15).

We now turn to the reverse bound in the case when $c \geqslant 1$ (when $c < 1$, the proof is completed). Consider $\varepsilon \in (0,1)$ and a positive sequence $(\delta_n)_{n\geqslant 1}$ such that

$$
\delta_n \ln n \to 0, \quad n\delta_n \to \infty,
$$

as $n \to \infty$. We write

$$
\begin{aligned}
\mathrm{Trans}_n \cap \mathrm{Surv}^{\mathrm{GW}} \supset &\{R(n(1-\delta_n)) > T_{n,n}, \mathrm{Surv}^{\mathrm{GW}}\} \cap \{N_n(\tau_n) \geqslant n(1-\varepsilon), \mathrm{Surv}^{\mathrm{GW}}\} \\
&\cap \{R(n(1-\varepsilon)) > T_{n(1-\delta_n),n}\},
\end{aligned}
$$

the right-hand side being denoted $A \cap B \cap C$. Indeed, on $B \cap C$,

$$
\tau_n < T_{n,n} \Rightarrow \tau_n = R(N_n(\tau_n)) > T_{n(1-\delta_n),n},
$$

so $\tau_n > T_{n(1-\delta_n),n}$. Then, by the same argument, it must hold that $\tau_n \geqslant T_{n,n}$ on $A \cap B \cap C$, as otherwise τ_n would be equal to

$$
R(N_n(\tau_n)) \geqslant R(N_n(T_{n(1-\delta_n),n})) = R(n(1-\delta_n)) > T_{n,n},
$$

yielding a contradiction. Below we will use the estimate

$$
\mathbb{P}(A \cap B \cap C) \geqslant \mathbb{P}(A) - \mathbb{P}(B^{\complement} \cap A) - \mathbb{P}(C^{\complement}).
$$

As above, we write $\{R(n(1 - \delta_n)) > T_{n,n}, \mathrm{Surv}^{\mathrm{GW}}\}$ as

$$\{n^{-1}(R(n(1 - \delta_n)) - cn \ln n - G^{(n)} > -(c - 1) \ln n, \mathrm{Surv}^{\mathrm{GW}}\}.$$

Lemma 4.4 says that $n^{-1}(R(n(1 - \delta_n)) - cn \ln n - G^{(n)}$ converges in law towards $\mathcal{S}_c - G$ as $n \to \infty$. Therefore, when $c > 1$, we get

$$\lim_{n \to \infty} \mathbb{P}\big(R(n(1 - \delta_n)) > T_{n,n}, \mathrm{Surv}^{\mathrm{GW}}\big) = \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}}).$$

When $c = 1$, we make use of Lemma 4.4 again. By asymptotic independence, we obtain, as in the proof of the upper bound,

$$\lim_{n \to \infty} \mathbb{P}\big(R(n(1 - \delta_n)) > T_{n,n}, \mathrm{Surv}^{\mathrm{GW}}\big)$$
$$= \lim_{n \to \infty} \mathbb{P}\left(\left(\frac{R(n(1 - \delta_n))}{n} - \ln n\right) - G^{(n)} > 0, \mathrm{Surv}^{\mathrm{GW}}\right)$$
$$= \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}})\mathbb{P}(G < \mathcal{S}) = \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}})\mathbb{E}(\exp(-e^{-\mathcal{S}})).$$

By Theorem 2.2 with $\mathbb{E}K = \infty$,

$$\mathbb{P}(N_n(\tau_n) \leqslant n(1 - \varepsilon), \mathrm{Surv}^{\mathrm{GW}}) \to 0 \quad \text{as } n \to \infty.$$

It remains to show that

$$\mathbb{P}\big(R(n(1 - \varepsilon)) > T_{n(1 - \delta_n), n}\big) \to 1 \quad \text{as } n \to \infty.$$

We have

$$\mathbb{E}\big(T_{n(1 - \delta_n), n}\big) = \sum_{i=1}^{\lfloor n(1 - \delta_n)\rfloor - 1} \frac{n}{n - i} \sim n \ln 1/\delta_n,$$

$$\mathrm{Var}\big(T_{n(1 - \delta_n), n}\big) \leqslant \sum_{i=1}^{\lfloor n(1 - \delta_n)\rfloor} \frac{n^2}{(n - i)^2} \sim n/\delta_n = o(n^2),$$

as $n\delta_n \to \infty$. The Chebyshev inequality implies that $T_{n(1 - \delta_n), n} = o(n \ln n)$ for $n\delta_n \to \infty$. On the other hand, $R(n(1 - \varepsilon))$ is of order $n \ln n$, which ends the proof. $\qquad \square$

We finally prove the following result.

Lemma 4.4. *Under the assumptions of Theorem 2.5, given a sequence of positive reals* $(\delta_n)_{n \geqslant 1}$ *such that* $\delta_n \ln n \to 0$ *as* $n \to \infty$*, the sequence* $(n^{-1}R(n(1 - \delta_n)) - c\ln(n))_{n \geqslant 1}$ *converges in law towards* $\mathcal{S}_c$*. Moreover, as* $n \to \infty$*, the event* $\mathrm{Surv}^{\mathrm{GW}}$ *and the variable* $n^{-1}R(n(1 - \delta_n)) - c\ln n$ *become independent.*

Proof. In order to prove that the sequence $(n^{-1}R(n(1 - \delta_n)) - c\ln(n))_{n \geqslant 1}$ converges in law towards $\mathcal{S}_c$, it is sufficient to check that $(n^{-1}[R(n - 1) - R(n(1 - \delta_n))])_{n \geqslant 1}$ converges

towards 0 in probability. For any $\varepsilon > 0$ and $A > 0$, we have

$$\mathbb{P}\big(R(n-1) - R(n(1-\delta_n)) \geqslant n\varepsilon\big) \leqslant \mathbb{P}\big(\exists \ell = \lfloor n(1-\delta_n) \rfloor + 1, \ldots, n-1 : K_\ell \geqslant A\big)$$

$$+ \mathbb{P}\left(\sum_{i=\lceil n(1-\delta_n) \rceil}^{n} K_i \mathbf{1}_{\{K_i \leqslant A\}} \geqslant n\varepsilon \right)$$

$$\leqslant \mathbb{P}\big(\exists \ell = \lfloor n(1-\delta_n) \rfloor + 1, \ldots, n-1 : K_\ell \geqslant A\big)$$

$$+ \varepsilon^{-1}\delta_n \mathbb{E}\big[K\mathbf{1}_{\{K \leqslant A\}}\big]$$

$$\leqslant 1 - \big(1 - \mathbb{P}(K \geqslant A)\big)^{n\delta_n} + c'\varepsilon^{-1}\delta_n \ln(A),$$

for some constant c' independent of n and A. For $A = n$, we have $1 - \mathbb{P}(K \geqslant n) \sim 1 - c/n$ by assumption and thus $(1 - \mathbb{P}(K \geqslant n))^{n\delta_n} \sim \exp(-c\delta_n)$, which tends to 1 as $n \to \infty$. As $\delta_n \ln n \to 0$ as $n \to \infty$, we deduce that $(n^{-1}[R(n-1) - R(n(1-\delta_n))])_{n\geqslant 1}$ indeed converges towards 0 in probability.

The asymptotic independence is proved as in the proof of Lemma 4.3. $\qquad\square$

5. Proofs of large deviations

Before starting the proofs we recall a few facts about the LDP for the coupon collector. As r is increased from 0 to $t \wedge 1$, the function $r \mapsto \rho(r,t)$ in (2.12) decreases from $+\infty$ to zero when $t \leqslant 1$ or to a positive value otherwise. The value $\rho = 1$ corresponds to the typical case $r = 1 - e^{-t}$; and $\rho > 1$ (resp. $\rho < 1$) to large deviations with $N_n(nt)$ much smaller (resp. larger) than $\mathbb{E}[N_n(nt)]$ (see [13, Appendix A.2.1]). The LDP holds at the process level for $(N(nt); t \geqslant 0)$, and the overwhelming contribution to the probability of the event $\{N_n(nt) \in n(r-\varepsilon, r+\varepsilon)\}$ occurs in the neighbourhood of the optimal path $N_n(ns) = ng(s)$ with

$$g(s) = \frac{1 - e^{-s\rho}}{\rho}, \quad s \in [0,t], \tag{5.1}$$

and $\rho = \rho(r,t)$ as before (see also Lemma 5.1 below for the uniqueness of the optimal path).

Proof of Lemma 2.6. **(i)** Lower semi-continuity is easily checked. Boundedness of the level sets follows from the lower bound

$$rI\left(\frac{t}{r}\right) \geqslant r\left(a_0 \frac{t}{r} - \ln \mathbb{E}\exp(a_0 K)\right) \geqslant a_0 t - \ln \mathbb{E}\exp(a_0 K)),$$

which implies for a well-chosen value of a_0 that t must be bounded when $\mathcal{F}(r,t)$ is bounded. Boundedness of r easily follows since $r \leqslant t$ when $\mathcal{F}(r,t)$ is finite.

To determine the domain of $\mathcal{F}$, we recall the expression of $\mathrm{Dom}(I)$ in terms of k_*, k^*, and write

$$\mathrm{Dom}(\mathcal{F}) = \{t/r \in \mathrm{Dom}(I), r/t \in]0,1]\} \cup \{(0,0)\}.$$

Continuity on the first set in the above union follows from that of I on its domain. If $k^* = \infty$, then $\mathcal{F}$ is unbounded in any neighbourhood of the origin, so it is not continuous

at this point. If $k^* < \infty$, then I is bounded on its domain and continuity at the origin of $\mathcal{F}$ easily follows.

(ii) Part (ii) trivially holds.

(iii) Finally, roots of $\mathcal{F}$ must satisfy $r = 0$ or $t/r = \mathbb{E}K$, for $0 < r \leqslant t \wedge 1$, as $\mathbb{E}K$ is the only zero of I. When $r = 0$, t must be zero as well. When $0 < r \leqslant t \wedge 1$, the condition $t/r = \mathbb{E}K$ requires $\mathbb{E}K$ to be strictly greater than 1. (When $\mathbb{E}K = 1$, the condition $t = r > 0$ implies $t \leqslant 1$ and $J_t(r) = J_t(t) = t > 0$.) When $\mathbb{E}K > 1$, $t/r = \mathbb{E}K$ implies $t\rho = \theta$ (compare with (2.6)) and $J_t(r) = 0$ implies $r = 1 - e^{-t}$, that is, $\rho = 1$ and thus $t = \theta$ and $r = p$ (see (2.7)). Conversely, it is readily checked that 0 is a root of $\mathcal{F}$ and that (p, θ) is a root as well when $\mathbb{E}K > 1$. $\qquad\square$

Proof of Theorem 2.7.
Upper bound. We start with the case when r, t satisfy $0 < r < t \wedge 1$. Given $\varepsilon, \delta > 0$, we need to prove the local upper bound

$$\limsup_{\varepsilon,\delta \to 0^+} \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\big(N_n(\tau_n)/n \in [r - \varepsilon, r + \varepsilon], \tau_n/n \in [t - \delta, t + \delta]\big) \leqslant -\mathcal{F}(r, t). \quad (5.2)$$

We first tackle the case $r < 1 - e^{-t}, t/r < \mathbb{E}K$, which is relevant only when $\mathbb{E}K > 1$. Without any loss of generality, we can assume that ε and δ satisfy

$$r + \varepsilon < 1 - e^{-(t-\delta)}, \quad (t + \delta)/(r - \varepsilon) < \mathbb{E}K. \quad (5.3)$$

We then define the events

$$A = \{N_n(n(t - \delta)) \leqslant n(r + \varepsilon)\}, \quad B = \{R(n(r - \varepsilon)) \leqslant n(t + \delta)\}. \quad (5.4)$$

Since N_n, R are non-decreasing,

$$\{N_n(\tau_n) \in [n(r - \varepsilon), n(r + \varepsilon)], \tau_n \in [n(t - \delta), n(t + \delta)]\} \subset A \cap B,$$

and by independence of the processes N_n and R, we get

$$\mathbb{P}\big(N_n(\tau_n)/n \in [r - \varepsilon, r + \varepsilon], \tau_n/n \in [t - \delta, t + \delta]\big) \leqslant \mathbb{P}(A)\mathbb{P}(B). \quad (5.5)$$

From the LDPs (2.9) and (2.10), and by (5.3), we have

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\big(N_n(\tau_n)/n \in [r - \varepsilon, r + \varepsilon], \tau_n/n \in [t - \delta, t + \delta]\big)$$
$$\leqslant -J_{t-\delta}(r + \varepsilon) - (r - \varepsilon)I((t + \delta)/(r - \varepsilon)).$$

By lower semi-continuity of I on the whole $\mathbb{R}$ and by continuity of $(r, t) \mapsto J_t(r)$ at the prescribed value of (r, t), we obtain (5.2). With similar arguments, one easily obtains the same result in all other cases of $0 < r < t \wedge 1$.

We are now left with proving the local upper bound on the boundary. For $(r, t) = (0, 0)$, there is nothing to prove since the rate function is zero. The case $r = 0 < t$ is simple since it is enough to bound

$$\mathbb{P}\big(N_n(\tau_n)/n \leqslant \varepsilon, \tau_n/n \in [t - \delta, t + \delta]\big) \leqslant \mathbb{P}\big(N_n(n(t - \delta))/n \leqslant \varepsilon\big),$$

to get a rate of decay $J_{t-\delta}(\varepsilon)$ which tends to ∞ as both δ and ε vanish. The cases $r = 1 < t$ and $r = t \leqslant 1$ use arguments similar to those above in the general case.

We have proved the upper bound for compact sets. To extend it to closed sets, it is enough to show that

$$\lim_{t\to\infty} \limsup_{n\to\infty} \frac{1}{n} \ln \mathbb{P}(\tau_n/n \geqslant t) = -\infty. \tag{5.6}$$

For this, we observe that

$$\{\tau_n \geqslant nt\} \subset \{R(N_n(nt)) \geqslant \lfloor nt \rfloor\} \subset \{R(n) \geqslant \lfloor nt \rfloor\}.$$

For $t > \mathbb{E}K$, the probability of the last event can be estimated by Cramér's bound, yielding

$$\mathbb{P}(\tau_n \geqslant nt) \leqslant \exp(-nI(t)).$$

Since $I(t)$ tends to ∞ as $t \to \infty$, we obtain (5.6).

Lower bound. The lower bound is subtle. It is sufficient to show that, for all r, t with $\mathcal{F}(r, t) < \infty$ and all $\eta > 0$,

$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbb{P}\big(N_n(\tau_n)/n \in [r - \eta, r + \eta], \tau_n/n \in [t - \eta, t + \eta]\big) - \mathcal{F}(r, t). \tag{5.7}$$

We start with the general case when r, t satisfy $0 < r < t \wedge 1$ and $I(t/r) < \infty$. (If $I(t/r) = \infty$, the lower bound above is trivial.)

First step. The proof can be divided into several steps. The first one is to bound from below the left-hand side above by the probability of an event depending in a separate way on the dynamics of the coupon collector on the one hand and on the capitals of the servers and the Galton–Watson tree on the other hand. We are thus given $0 < \eta' < \eta < r$ and $0 < \varepsilon' < 2\varepsilon$ such that $r + \eta + \varepsilon < 1 \wedge t$, $\varepsilon < r\eta/(4t) < \eta/4$, $\eta' < (1 - r/t)\eta$ and $\eta'(1 + t/r) < \varepsilon'$. For an integer $m \in [n\varepsilon', n(\varepsilon - \varepsilon')]$, we define the events

$A = \{N_n(nt + n\eta) \leqslant nr + n\varepsilon\},$
$A' = \{N_n(n\varepsilon) = \lfloor n\varepsilon \rfloor + 1\},$
$A'' = \{N_n(\ell) \geqslant \ell r/t + n\varepsilon(1 - r/t) - n\eta'; n\varepsilon \leqslant \ell \leqslant nt - n\eta\} \cap \{N_n(nt - n\eta) \leqslant nr + n\varepsilon\},$
$B = \{R(nr + n\varepsilon) - R(m) \leqslant nt + n\eta/2\},$
$B'' = \{R(\ell) - R(m) \geqslant (\ell - m)t/r - n\eta'; m \leqslant \ell \leqslant nr + n\varepsilon\}.$

Recall that $N_n(s) = N_n(\lfloor s \rfloor)$ and similarly for R. We also define

$$B' = \{R(m) \in [n\varepsilon, 2n\varepsilon], Z_{\mathrm{tot}}^{\mathrm{GW}} \geqslant n\varepsilon\}. \tag{5.8}$$

For A', it holds that $T_{i+1,n} = i(i \leqslant n\varepsilon)$, so the emission process cannot stop before $n\varepsilon$ unless the total size of the Galton–Watson tree is strictly less than $n\varepsilon$. Therefore, $A' \cap B' \subset \{\tau_n \geqslant n\varepsilon\}$. Further, $A' \cap A'' \cap B' \cap B'' \subset \{\tau_n > n(t - \eta)\}$. Indeed, on $A' \cap A'' \cap B' \cap B''$, for

$n\varepsilon \leqslant \ell \leqslant nt - n\eta$,

$$R\big(N_n(\ell)\big) \tag{5.9}$$
$$\geqslant R\big(\ell r/t + n\varepsilon(1 - r/t) - n\eta'\big)$$
$$\geqslant \big(\ell r/t + n\varepsilon(1 - r/t) - n\eta' - m\big)t/r + R(m) - n\eta'$$
$$\geqslant \ell + (n\varepsilon - m)(t/r - 1) + R(m) - m - n\eta'(1 + t/r) > \ell + n\varepsilon' - n\eta'(1 + t/r) > \ell,$$

where, to pass from the second to the third line, we use the fact that, for the prescribed values of ℓ, we have $\ell r/t + n\varepsilon(1 - r/t) - n\eta' \in [m, nr + n\varepsilon]$.

Moreover, on $A \cap B \cap B'$,

$$R\big(N_n(nt + n\eta)\big) \leqslant R\big(nr + n\varepsilon\big) \leqslant nt + n\eta/2 + R(m) \leqslant nt + n\eta/2 + 2n\varepsilon < nt + n\eta,$$

so that $\tau_n \leqslant n(t + \eta)$ on $A \cap B \cap B'$. Therefore,

$$\big(A \cap A' \cap A'' \cap B \cap B' \cap B''\big) \subset \{N_n(\tau_n) \in [n(r - \eta), n(r + \eta)], \tau_n \in [n(t - \eta), n(t + \eta)]\}, \tag{5.10}$$

since, on $A \cap A' \cap A'' \cap B \cap B' \cap B''$,

$$N_n(\tau_n) \geqslant N_n(n(t - \eta)) \geqslant n(t - \eta)r/t + n\varepsilon(1 - r/t) - n\eta'$$
$$\geqslant nr - n\eta r/t - n\eta' \geqslant nr - n\eta, \tag{5.11}$$
$$N_n(\tau_n) \leqslant N_n(n(t + \eta)) \leqslant nr + n\varepsilon \leqslant nr + n\eta.$$

By independence of the processes N_n and R, we get

$$\mathbb{P}(N_n(\tau_n)/n \in [r - \eta, r + \eta], \tau_n/n \in [t - \eta, t + \eta]) \geqslant \mathbb{P}\big(A \cap A' \cap A''\big)\mathbb{P}\big(B \cap B' \cap B''\big). \tag{5.12}$$

Second step. We estimate the first factor on the right-hand side by using Markov's property:

$$\mathbb{P}\big(A \cap A' \cap A''\big) \geqslant \inf_{i \in \hat{I}} \mathbb{P}\big(A | \hat{A}_i\big)\mathbb{P}\big(A'' | A'\big)\mathbb{P}\big(A'\big),$$
$$\text{with } \hat{A}_i = \{N_n(nt - n\eta) = i\}, \tag{5.13}$$
$$\hat{I} = \{\lceil n(r - \eta r/t) + n\varepsilon(1 - r/t) - n\eta'\rceil, \ldots, \lfloor nr + n\varepsilon\rfloor\}.$$

Clearly, $\mathbb{P}(A')$ is bounded from below by the probability for the binomial distribution with parameters $\lfloor n\varepsilon \rfloor$ and ε to be equal to zero. We get

$$\mathbb{P}\big(A'\big) \geqslant e^{-n\varepsilon|\ln(1 - \varepsilon)|}. \tag{5.14}$$

Similarly, $\inf_{i \in \hat{I}} \mathbb{P}(A | \hat{A}_i)$ is bounded from below by the probability for the binomial distribution with parameters $\lfloor 2n\eta \rfloor + 1$ and $1 - r + \eta r/t - \varepsilon(1 - r/t) + \eta' < 1 - r + \eta$ to be equal to 0. Therefore,

$$\inf_{i \in \hat{I}} \mathbb{P}\big(A | \hat{A}_i\big) \geqslant e^{-(2n\eta + 1)|\ln(r - \eta)|}. \tag{5.15}$$

We now turn to $\mathbb{P}(A'' | A')$. We observe that the conditional probability $\mathbb{P}(\cdot | A')$ is (up to a shift in time) the law of the coupon collector with n images when starting from $n\varepsilon$ different images, already collected at the initial time. Since

$$n(r - \eta r/t) + n\varepsilon(1 - r/t) - n\eta' \leqslant nr + n(\varepsilon - \eta r/t) - n\eta' \leqslant nr - n\eta',$$

we deduce from the large deviation lower bound in [13]:

$$\mathbb{P}(\hat{A}|A') \geqslant \exp\{-nJ^{\varepsilon}_{t-(\eta+\varepsilon)}(r) + o(n)\} \tag{5.16}$$

$$\text{with } \hat{A} = \{n(r-\eta r/t) + n\varepsilon(1-r/t) - n\eta' \leqslant N_n(nt-n\eta) \leqslant nr + n\varepsilon\} = \bigcup_{i\in\hat{I}} \hat{A}_i,$$

where the rate function $J^{\varepsilon}_{t-(\varepsilon+\eta)}(r)$ is given by Theorem 2.7 in [13] (with $I = 0$, $\alpha_0 = 1-\varepsilon$ and $\omega_0 = 1-r$ therein). Precisely, (5.16) follows from the LDP for the 'time-shifted' variable $N_n(n(t-(\eta+\varepsilon)))$ with $N_n(0) = \lfloor n\varepsilon \rfloor + 1$ as initial condition. Using the same notation as in [13], $J^{\varepsilon}_t(r)$ may be expressed as the relative entropy:

$$J^{\varepsilon}_t(r) = (1-\varepsilon)\left[\frac{1-r}{1-\varepsilon}\ln\left(\frac{1-r}{(1-\varepsilon)e^{-t}}\right)\right] + (1-\varepsilon)C^{\varepsilon}\sum_{j\geqslant 1}\mathcal{P}_j(\rho^{\varepsilon}t)\ln\left[C^{\varepsilon}\frac{\mathcal{P}_j(\rho^{\varepsilon}t)}{\mathcal{P}_j(t)}\right]$$

$$+ \varepsilon\sum_{j\geqslant 0}\mathcal{P}_j(\rho^{\varepsilon}t)\ln\left[\frac{\mathcal{P}_j(\rho^{\varepsilon}t)}{\mathcal{P}_j(t)}\right],$$

where $\mathcal{P}_j(\alpha)$ is the jth weight of the Poisson distribution of parameter α, ρ^{ε} is the unique root of

$$\frac{1-r}{1-\varepsilon} + \frac{1-\varepsilon\rho^{\varepsilon}}{(1-\varepsilon)\rho^{\varepsilon}}(1-e^{-\rho^{\varepsilon}t}) = 1, \tag{5.17}$$

and C^{ε} reads

$$C^{\varepsilon} = \frac{1-\varepsilon\rho^{\varepsilon}}{(1-\varepsilon)\rho^{\varepsilon}}.$$

It is plain to check that, as ε tends to 0, ρ^{ε} converges towards $\rho = \rho(r,t)$ and C^{ε} towards $1/\rho$. Moreover, standard computations yield

$$\sum_{j\geqslant 0}\mathcal{P}_j(\rho^{\varepsilon}t)\ln\left[\frac{\mathcal{P}_j(\rho^{\varepsilon}t)}{\mathcal{P}_j(t)}\right] = -(\rho^{\varepsilon}-1)t + \rho^{\varepsilon}\ln(\rho^{\varepsilon})t,$$

$$\sum_{j\geqslant 1}\mathcal{P}_j(\rho^{\varepsilon}t)\ln\left[\frac{\mathcal{P}_j(\rho^{\varepsilon}t)}{\mathcal{P}_j(t)}\right] = -(\rho^{\varepsilon}-1)t + \rho^{\varepsilon}\ln(\rho^{\varepsilon})t + (\rho^{\varepsilon}-1)t\,e^{-\rho^{\varepsilon}t},$$

so that, by the definition of $\rho(r,t)$,

$$\lim_{\varepsilon\to 0}J^{\varepsilon}_t(r) = (1-r)\ln(1-r) + (1-r)t - \frac{\ln\rho}{\rho}\left(1-e^{-\rho t}\right) - \frac{\rho-1}{\rho}t + t\ln\rho + \frac{\rho-1}{\rho}t\,e^{-\rho t}$$

$$= J_t(r).$$

Similarly,

$$J^{\varepsilon}_{t-(\eta+\varepsilon)}(r) \longrightarrow J_{t-\eta}(r), \quad \text{as } \varepsilon \to 0.$$

Therefore, (5.16) may be expressed as

$$\mathbb{P}(\hat{A}|A') \geqslant \exp\{-n[J_t(r) + \delta(\eta) + \delta'(\varepsilon;\eta)] + o(n)\}, \tag{5.18}$$

where $\delta(\eta)$ is a generic term that tends to 0 with η and $\delta'(\varepsilon;\eta)$ is a generic term that tends to 0 with ε when $\eta > 0$ is given. We then claim that $\mathbb{P}(A''|A')$ satisfies the same lower

bound, that is,

$$\mathbb{P}(A''|A') \geqslant \exp\{-n[J_t(r) + \delta(\eta) + \delta'(\varepsilon; \eta)] + o(n)\}. \tag{5.19}$$

Basically, it arises from the fact that the optimal path explaining the large deviation of the random variable $N_n(n(t - \eta))$ in the neighbourhood of r given the initial condition in A' lies above the linear constraint in A''. As already explained, the correct optimal path to consider is the one when the collector has already collected $\lfloor n\varepsilon \rfloor + 1$ coupons at (rescaled) time ε. We denote it by $g^\varepsilon : [\varepsilon, t - \eta] \to \mathbb{R}$. By Theorem 2.8 in [13] and Lemma 5.1 below, the optimal limit path on $[0, t]$ (as $n \to \infty$) for the coupon collector running from the proportion ε to the proportion r of collected coupons in time t is unique and reads as follows (we put a tilde over g^ε below to emphasize that the interval is shifted in time and that the terminal time is t and not $t - (\eta + \varepsilon)$):

$$\tilde{g}^\varepsilon : [0, t] \ni s \mapsto \varepsilon + \frac{1 - \varepsilon \rho^\varepsilon}{\rho^\varepsilon}(1 - e^{-\rho^\varepsilon s}), \tag{5.20}$$

which is equal to r at time t. By concavity of $\tilde{g}^\varepsilon$,

$$n\tilde{g}^\varepsilon(s) \geqslant n\frac{s}{t}\tilde{g}^\varepsilon(t) + n\left(1 - \frac{s}{t}\right)\tilde{g}^\varepsilon(0) = n\frac{s}{t}(r - \varepsilon) + n\varepsilon, \quad 0 \leqslant s \leqslant t.$$

Coming back to g^ε, we deduce from the inequality $t - (\varepsilon + \eta) \leqslant t - t\varepsilon/r$ that

$$ng^\varepsilon(s) \geqslant n\frac{s - \varepsilon}{t - (\varepsilon + \eta)}(r - \varepsilon) + n\varepsilon \geqslant n\frac{(s - \varepsilon)r}{t} + n\varepsilon, \quad \varepsilon \leqslant s \leqslant t - \eta, \tag{5.21}$$

which proves that g^ε is strictly above the linear constraint in A''. Optimality of g^ε then says that (compare with (5.16))

$$\limsup_{n \to \infty} n^{-1} \ln \mathbb{P}(\hat{A}, \exists s \in [\varepsilon, t - \eta] : N_n(ns) \leqslant n(g^\varepsilon(s) - \eta')|A') < -J_{t-(\eta+\varepsilon)}^\varepsilon(r). \tag{5.22}$$

Following the proof of (5.18), we deduce from (5.21) and (5.22) that (5.19) holds. From (5.13), (5.14), (5.15) and (5.19), we deduce that

$$\mathbb{P}(A \cap A' \cap A'') \geqslant \exp\{-n[J_t(r) + \delta(\eta) + \delta'(\varepsilon; \eta)] + o(n)\}, \tag{5.23}$$

with $\delta(\eta) \to 0$ as $\eta \to 0$ and $\delta'(\varepsilon; \eta) \to 0$ as $\varepsilon \to 0$ for a given $\eta > 0$.

Third step. We now provide a lower bound for $\mathbb{P}(B \cap B' \cap B'')$ in (5.12). To this end, we shall use the stopping time $T_{m,n} = \inf\{t \geqslant 0 : N_n(t) = m\}$ together with the set $C = \{\tau_n \geqslant T_{m,n}\}$. For C, we have $N_n(\tau_n) \geqslant m$, that is, $m \leqslant \text{card } \mathcal{T}^\circ(\infty)$ (see (3.3) for the notation). Therefore, on C, $R(m)$ has the form

$$R(m) = \sum_{i=1}^{m} K(X(T_{i,n})).$$

In particular, on C,

$$R(m) \geqslant n\varepsilon \Rightarrow Z_{\text{tot}}^{\text{GW}} \geqslant n\varepsilon,$$

which implies that $C \cap B' = C \cap \{R(m) \in [n\varepsilon, 2n\varepsilon]\}$. Since C may be also expressed as

$$C = \{R(N_n(\ell)) > \ell; \ell = 0, \ldots, T_{m,n} - 1\} \quad \text{and} \quad \ell < T_{m,n} \Rightarrow N_n(\ell) \leqslant m,$$

we deduce that

$$C \cap B' \in \sigma(T_{1,n}, \ldots, T_{n,n}, K_1, \ldots, K_m). \tag{5.24}$$

By Proposition 3.1, we deduce that $C \cap B'$ is independent of $B \cap B''$, so that

$$\mathbb{P}(B \cap B' \cap B'') \geqslant \mathbb{P}(C \cap B')\mathbb{P}(B \cap B''). \tag{5.25}$$

Now, we emphasize that $\tau_n = N_n(\tau_n)$ on $C^{\complement}$ (which means that the capital is exhausted at some time less than or equal to n). Thus, on $C^{\complement}$,

$$\tau_n = N_n(\tau_n) \leqslant N_n(T_{m,n}) = m \leqslant n\varepsilon.$$

Therefore, by Lemma 4.2, for $\varepsilon < \varepsilon_0$,

$$\lim_{n \to \infty} \mathbb{P}(C^{\complement} \cap \mathrm{Surv}^{\mathrm{GW}}) = 0.$$

Finally, since $\mathbb{P}(E \cap F) \geqslant \mathbb{P}(E) - \mathbb{P}(F^c)$,

$$\begin{aligned}
\mathbb{P}(C \cap B') &\geqslant \mathbb{P}(C \cap \mathrm{Surv}^{\mathrm{GW}} \cap \{R(m) \in [n\varepsilon, 2n\varepsilon]\}) \\
&\geqslant \mathbb{P}(C \cap \mathrm{Surv}^{\mathrm{GW}}) - \mathbb{P}(R(m) \notin [n\varepsilon, 2n\varepsilon]) \\
&= \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}}) - \mathbb{P}(C^{\complement} \cap \mathrm{Surv}^{\mathrm{GW}}) - \mathbb{P}(R(m) \notin [n\varepsilon, 2n\varepsilon]),
\end{aligned} \tag{5.26}$$

the second term in the last line converging to 0 as n tends to ∞.

Fourth step. We now complete the proof when $\mathbb{E}K > 1$, which is a simpler case to handle than the opposite case $\mathbb{E}K \leqslant 1$. We then choose

$$m = \left\lfloor \frac{\zeta}{\mathbb{E}K} n\varepsilon \right\rfloor, \quad \text{with } \zeta = \min\left(\frac{1 + \mathbb{E}K}{2}, 3/2\right).$$

Clearly,

$$m \in [n\varepsilon/\mathbb{E}K, (1 + \mathbb{E}K)/(2\mathbb{E}K)\varepsilon] \subset [n\varepsilon', n(\varepsilon - \varepsilon')]$$

as required, for some well-chosen ε'. Moreover, the typical values of $R(m)$ are in the neighbourhood of $\zeta n\varepsilon \in (n\varepsilon, 2n\varepsilon)$. Therefore, by the law of large numbers, the third term on the right-hand side of (5.26) tends to 0 as $n \to \infty$. We deduce

$$\liminf_{n \to \infty} \mathbb{P}(C \cap B') \geqslant \mathbb{P}(\mathrm{Surv}^{\mathrm{GW}}) > 0, \tag{5.27}$$

since $\mathbb{E}K > 1$. Moreover, by Mogulskii's theorem (see Theorem 5.1.2 in [11]), we have a lower bound for the probability that the process $(n^{-1}R(ns))_{0 \leqslant s \leqslant r+\varepsilon}$ is in the neighbourhood of the path $[0, r + \varepsilon] \ni s \mapsto st/r$. We then observe that

$$B \cap B'' \supset \{(\ell - m)t/r - n\eta' \leqslant R(\ell) - R(m) \leqslant (\ell - m)t/r + n\eta/4; m \leqslant \ell \leqslant nr + n\varepsilon\},$$

since the upper condition in B can be reformulated as

$$R(nr + n\varepsilon) - R(m) \leqslant (nr + n\varepsilon - m)t/r + n(\eta/2 - t\varepsilon/r) + mt/r,$$

and $\eta/2 - t\varepsilon/r \geqslant \eta/4$. By Mogulskii's theorem, we get

$$\mathbb{P}(B \cap B'') \geqslant \exp\{-n(r + \varepsilon)I(t/r) + o(n)\}. \tag{5.28}$$

Collecting (5.12), (5.23), (5.25), (5.27) and (5.28), we deduce that (for a possibly new choice of $\delta'(\varepsilon;\eta)$)

$$\lim_{n\to\infty} n^{-1}\ln\mathbb{P}\big(N_n(\tau_n)/n \in [r-\eta, r+\eta], \tau_n/n \in [t-\eta, t+2\eta]\big)$$
$$\geqslant -\big[J_t(r) + rI(t/r)\big] - \delta(\eta) - \delta'(\varepsilon;\eta),$$

which tends to $-[J_t(r) + rI(t/r)] - \delta(\eta)$ as $\varepsilon \searrow 0$. Since η can be chosen as small as needed, we complete the proof of (5.7) in the case $\mathbb{E}K > 1$.

Fifth step. We now investigate the case $\mathbb{E}K \leqslant 1$. The above argument fails since m cannot be chosen as $\lfloor \zeta n\varepsilon/\mathbb{E}K \rfloor$ on the one hand and since $\mathbb{P}(\text{Surv}^{\text{GW}}) = 0$ on the other hand. We thus need to give a relevant version of the previous step. We first tackle the case $\mathbb{P}(K \geqslant 2) > 0$. The point is to change the probability measure in order to switch back to the case $\mathbb{E}K > 1$. The change of probability relies on the same trick as in the proof of Cramér's theorem. Since $k_* = 0$ and $k^* \geqslant 2$, it is standard that, for any $\xi \in (1, 2)$, there exists $\alpha \in \mathbb{R}$ such that $\varphi'(\alpha) = \xi$, where

$$\varphi(\alpha) = \ln\big[\mathbb{E}\exp(\alpha K)\big].$$

We now choose $\xi = 1 + \varepsilon$ and set $\zeta' = (1 + \xi)/2 \in (1, 3/2)$. For $m = \lfloor \zeta' n\varepsilon/\xi \rfloor$ (where m is in $[n\varepsilon', n(\varepsilon - \varepsilon')]$ for some well-chosen ε' and plays the same role below as in the first step), we then define $\tilde{\mathbb{P}}$ as

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \exp\left(\sum_{i=1}^{m} \{\alpha K_i - \varphi(\alpha)\}\right).$$

It is plain to see that, under $\tilde{\mathbb{P}}$, the variables $(K_i)_{i\geqslant 1}$ are independent, the variables $(K_i)_{i\geqslant m+1}$ having the same distribution as they have under $\mathbb{P}$ and the variables $(K_i)_{1\leqslant i\leqslant m}$ being identically distributed with

$$\tilde{\mathbb{E}}K_i = \xi, \quad i = 1, \dots, m,$$

where $\tilde{\mathbb{E}}$ denotes the expectation under $\tilde{\mathbb{P}}$. Following (5.25), we write

$$\mathbb{P}(C \cap B') = \tilde{\mathbb{E}}\left[\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\mathbf{1}_{C\cap B'}\right]$$
$$= \tilde{\mathbb{E}}\big[\exp\big(-\alpha R(m) + m\varphi(\alpha)\big)\mathbf{1}_{C\cap B'}\big]$$
$$\geqslant \exp\big(-2n\alpha\varepsilon + m\varphi(\alpha)\big)\tilde{\mathbb{P}}(C \cap B'). \tag{5.29}$$

By (5.24), the probability of $C \cap B'$ under $\tilde{\mathbb{P}}$ coincides with the probability of $C \cap B'$ under $\mathbb{P}$ when the expectation of the reproduction law of the Galton–Watson tree is strictly larger than 1. Since $m = \lfloor \zeta' n\varepsilon/\tilde{\mathbb{E}}K \rfloor$, with $\zeta' = \min[(1 + \tilde{\mathbb{E}}K)/2, 3/2]$, we know from the case $\mathbb{E}K > 1$ that

$$\lim_{n\to\infty} \tilde{\mathbb{P}}(C \cap B') = \tilde{\sigma} > 0,$$

where $\tilde{\sigma}$ stands for the probability that the Galton–Watson tree survives when the reproduction is governed by the law of K under $\tilde{\mathbb{P}}$. Finally, we get from (5.29):

$$\liminf_{n\to\infty} n^{-1}\ln\mathbb{P}\big(C \cap B'\big) \geqslant -\varepsilon\xi^{-1}\big(2\alpha\xi - \zeta'\varphi(\alpha)\big) = \varepsilon G(\xi), \tag{5.30}$$

where $G(\xi)$ remains bounded as ξ ranges over any compact subset of $[1, 2]$. In particular, the right-hand side of (5.30) tends to 0 with ε. The completion of the proof is then the same as in the case when $\mathbb{P}(\mathrm{Surv}^{\mathrm{GW}}) > 0$.

In the case $\mathbb{P}(K \leqslant 1) = 1$ we have $I(t/r) = \infty$, and thus $\mathcal{F}(t, r) = \infty$ too, so that the lower bound is obvious.

Final step. Now we prove the lower bound on the boundary. For $r = t = 0$, we introduce $k_0 = \min\{k \geqslant 0 : \mathbb{P}(K = k) > 0\}$, and we write

$$\mathbb{P}\big(N_n(\tau_n)/n \leqslant \eta, \tau_n/n \leqslant \eta\big) \geqslant \mathbb{P}(K_1 = k_0, N_n(k_0) = 1) = \mathbb{P}(K = k_0)n^{-k_0},$$

showing the bound with $\mathcal{F}(0, 0) = 0$.

It remains to tackle the cases $0 < r = t \leqslant 1$ and $0 < r = 1 < t$. Without any loss of generality, we can assume that $I(t/r) < \infty$ as otherwise the bound is obvious. Given an open set $O \ni (r, t)$ ($O \not\ni (0, 0)$), we deduce from Lemma 2.6 that $\mathcal{F}$ is continuous on

$$O \cap \{(r', t') \in \mathbb{R}^2 : 0 < r' \leqslant t' \wedge 1, t'/r' \in \mathrm{Dom}(I)\}.$$

If $0 < r = 1 < t$, then we can find a sequence $(r_n, t_n)_{n \geqslant 1}$, converging towards (r, t), such that $r_n \nearrow r$, $t_n \nearrow t$ and $t_n/r_n = t/r$, with $r_n < 1 < t_n$ for any $n \geqslant 1$. Thus,

$$\mathcal{F}(r, t) \geqslant \inf\{\mathcal{F}(r', t'), (r', t') \in O'\}, \quad O' = O \cap \{(r', t') \in \mathbb{R}^2 : 0 < r' < t' \wedge 1\}. \tag{5.31}$$

Assuming without any loss of generality that

$$O \subset \{(r', t') \in \mathbb{R}^2 : 1 - \eta < r' < 1 + \eta < t - \eta < t' < t + \eta\}, \quad \text{for some } \eta > 0,$$

we deduce that

$$\begin{aligned}
\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big((N_n(\tau_n)/n, \tau_n/n) \in O\big) &\geqslant \liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big((N_n(\tau_n)/n, \tau_n/n) \in O'\big) \\
&\geqslant -\inf\{\mathcal{F}(r', t'), (r', t') \in O'\} \\
&= -\inf\{\mathcal{F}(r', t'), (r', t') \in O\}, \tag{5.32}
\end{aligned}$$

the second line following from the LDP we proved above for (r, t) satisfying $0 < r < t \wedge 1$. This is enough to conclude the proof.

Assume $0 < r = t \leqslant 1$, and $I(1) < \infty$ since otherwise the bound is obvious. If in addition $k^* \geqslant 2$, (5.31) still holds, and we can repeat (5.32).

It thus remains to handle the case when K has a Bernoulli distribution of parameter $p \in (0, 1]$ (if $p = 0$, $K \equiv 0$ and the lower bound is obvious) and O is an open set containing some point (r, t) with $0 < r = t \leqslant 1$. As above, O might intersect the line $(1, t')$ for t' in the neighbourhood of 1; thanks to (5.32), this has no real consequences. Then, for some small $\eta > 0$,

$$\begin{aligned}
\mathbb{P}\big(N_n(\tau_n) &\in [n(t - \eta), n(t + \eta)], \tau_n \in [n(t - \eta, t + \eta)]\big) \\
&\geqslant \mathbb{P}\big(N_n(\lfloor n(t - \eta) \rfloor) = \lfloor n(t - \eta) \rfloor + 1, N_n(\lfloor n(t + \eta) \rfloor) \leqslant \lfloor n(t + \eta) \rfloor\big) \tag{5.33} \\
&\quad \times \mathbb{P}\big(R(\lfloor n(t - \eta) \rfloor + 1) = \lfloor n(t - \eta) \rfloor + 1\big).
\end{aligned}$$

Clearly,

$$\mathbb{P}\big(R(\lfloor n(t - \eta) \rfloor + 1) = \lfloor n(t - \eta) \rfloor + 1\big) = p^{\lfloor n(t - \eta) \rfloor + 1} \geqslant p^{n(t - \eta)} = \exp\big[-n(t - \eta)I(1)\big], \tag{5.34}$$

as $I(x) = x \ln(x/p) + (1-x)\ln[(1-x)/(1-p)]$, for $x \in [0,1]$, in the Bernoulli case. Moreover,

$$\mathbb{P}\big(N_n(\lfloor n(t-\eta)\rfloor) = \lfloor n(t-\eta)\rfloor + 1, N_n(\lfloor n(t+\eta)\rfloor) \leqslant \lfloor n(t+\eta)\rfloor\big) \qquad (5.35)$$

$$\geqslant (t-\eta-1/n) \prod_{i=1}^{\lfloor n(t-\eta)\rfloor} (1-i/n) = (t-\eta-1/n)\exp\bigg(\sum_{i=1}^{\lfloor n(t-\eta)\rfloor} \ln(1-i/n)\bigg).$$

It remains to see that

$$\lim_{n\to\infty} -n^{-1}\sum_{i=1}^{\lfloor n(t-\eta)\rfloor} \ln(1-i/n) = -\int_0^{t-\eta} \ln(1-u)du = (1+\eta-t)\ln(1+\eta-t)+t-\eta$$

$$\longrightarrow (1-t)\ln(1-t)+t = J_t(t) \quad \text{as } \eta \to 0, \qquad (5.36)$$

using $\rho(t, t-\delta) \sim 2t^{-2}\delta$ for $t \leqslant 1$ and $\delta \searrow 0$ in (2.11). By (5.33), (5.34), (5.35) and (5.36), the proof is easily completed. $\qquad\square$

In the second step of the previous proof, we used the following.

Lemma 5.1. *The path g in (5.1) is the only optimal path minimizing the limit cost for getting r as proportion of collected coupons over the rescaled time interval* $[0,t]$.

Proof. The proof is an adaptation of Section A.4 in [13]. It is sufficient to prove that the function $\varepsilon \mapsto G[\varepsilon]$ therein (see also the expression below) is strictly convex in the neighbourhood of 0_+ whenever $\tilde{\gamma}$ is different from γ, where $\gamma = 1 - g$ and $\tilde{\gamma}$ stands for the proportion of non-collected coupons along another path with the same boundary conditions as γ at times 0 and t. We notice that $G[\varepsilon]$ has the following form (see Section A.1.2 in [13]):

$$G[\varepsilon] = \int_0^t \bigg[-\big(\dot{\gamma}_s + \varepsilon\dot{\eta}_s\big)\ln\bigg(-\frac{\dot{\gamma}_s + \varepsilon\dot{\eta}_s}{\gamma_s + \varepsilon\eta_s}\bigg) + \big(1+\dot{\gamma}_s + \varepsilon\dot{\eta}_s\big)\ln\bigg(\frac{1+\dot{\gamma}_s + \varepsilon\dot{\eta}_s}{1-(\gamma_s + \varepsilon\eta_s)}\bigg)\bigg]ds,$$

where $\eta = \tilde{\gamma} - \gamma$. Therefore, $G[\varepsilon]$ can be split into three terms:

$$G[\varepsilon] = \int_0^t \big[-(\dot{\gamma}_s + \varepsilon\dot{\eta}_s)\ln(-\dot{\gamma}_s - \varepsilon\dot{\eta}_s) + (1+\dot{\gamma}_s + \varepsilon\dot{\eta}_s)\ln(1+\dot{\gamma}_s + \varepsilon\dot{\eta}_s)\big]ds$$

$$+ \int_0^t \big[(\dot{\gamma}_s + \varepsilon\dot{\eta}_s)\ln(\gamma_s + \varepsilon\eta_s) - (\dot{\gamma}_s + \varepsilon\dot{\eta}_s)\ln(1-(\gamma_s + \varepsilon\eta_s))\big]ds$$

$$- \int_0^t \ln(1-(\gamma_s + \varepsilon\eta_s))ds$$

$$= G_1[\varepsilon] + G_2[\varepsilon] + G_3[\varepsilon].$$

It is readily seen that $G_1''[\varepsilon]$ is well-defined in the neighbourhood of 0 and is always (strictly) positive unless $\dot{\eta} \equiv 0$. Similarly, $G_3''[\varepsilon]$ is non-negative in the neighbourhood of 0. Finally,

$$G_2[\varepsilon] = \Big[(\gamma_s + \varepsilon\eta_s)\ln(\gamma_s + \varepsilon\eta_s) + \big[1 - (\gamma_s + \varepsilon\eta_s)\big]\ln\big[1 - (\gamma_s + \varepsilon\eta_s)\big]\Big]_0^t$$

is independent of ε as $\eta_t = \eta_0 = 0$. $\qquad\square$

Before proving Theorem 2.8, we deduce the LDP for the sequence $(N_n(\tau_n/n))_{n\geqslant 1}$ as an application of Theorem 2.7.

Proposition 5.2. *With the same notation as in the statement of Theorem 2.8, the sequence $(N_n(\tau_n)/n)_{n\geqslant 1}$ obeys an LDP with rate function*

$$\mathcal{G}(r) = (1-r)\ln(1-r)$$
$$+ r \inf_{s\geqslant 0}\{I(\lambda(s)) + (\lambda(s) - 1)\ln((1 - \exp(-s))/r) + \lambda(s)\exp(-s)\},$$

if $r \in [0,1]$, and $\mathcal{G}(r) = \infty$ otherwise, and speed n.

Proof. By Varadhan's contraction principle (see [11, Theorem 4.2.1]), we know that

$$\mathcal{G}(r) = \inf\{\mathcal{F}(r,t), t \geqslant 0\}.$$

Therefore, $\mathcal{G}(r) = 0$ if $r = 0$, and $\mathcal{G}(r) = \infty$ if $r > 1$, as stated. In the case $r \in (0,1]$, the infimum above can be restricted to the values of t in $[r,\infty)$. Then, $\mathcal{F}(r,t)$ reads

$$\mathcal{F}(r,t) = (1-r)\ln(1-r) + r[I(\lambda) + (\lambda - 1)\ln(t\rho/\lambda r) + \lambda\exp(-t\rho)], \qquad (5.37)$$

with $\lambda = t/r$, $t\rho$ solving the equation

$$\frac{1 - \exp(-t\rho)}{t\rho} = \lambda^{-1}.$$

If $\lambda = 1$, then $(\lambda - 1)\ln(t\rho/\lambda r)$ and ρ in (5.37) are both considered as 0. Letting $s = t\rho$, we note that s is the unique root of the equation

$$\frac{1 - \exp(-s)}{s} = \lambda^{-1},$$

with $\lambda = 1$ if $s = 0$. As λ ranges over $[1,\infty]$, R ranges over $[0,\infty)$. Expressing λ in terms of s, the proof is easily completed. $\qquad\square$

Proof of Theorem 2.8. The LDP for $N_n(\tau_n)/n$ yields

$$\limsup_{n\to+\infty} n^{-1}\ln\mathbb{P}(N_n(\tau_n) = n) \leqslant -\mathcal{G}(1),$$

with

$$\mathcal{G}(1) = \inf_{s\geqslant 0}\{I(\lambda(s)) + (\lambda(s) - 1)\ln(1 - \exp(-s)) + \lambda(s)\exp(-s)\}.$$

The point is thus to prove the lower bound, which cannot be proved from the LDP directly since the lower bound in the LDP holds for open subsets only. Then, we can focus on the case $\mathbb{P}(K \geqslant 1) > 0$ as otherwise both sides in the statement of Theorem 2.8 are infinite. We then follow the proof of the lower bound in the proof of Theorem 2.7. With the same notation as in the first step of the proof (in particular, given $0 < r < 1 \wedge t$), we already know that

$$\liminf_{n\to\infty} n^{-1}\ln\mathbb{P}(A \cap A' \cap A'' \cap B \cap B' \cap B'') \geqslant -\mathcal{F}(r,t) + \delta(\eta) + \delta'(\varepsilon;\eta),$$

where $\delta(\eta) \to 0$ as η tends to 0 and $\delta'(\varepsilon;\eta) \to 0$ as ε tends to 0 when $\eta > 0$ is given. (Here ε and η play the same role as in the first step of the proof of the lower bound in Theorem 2.7, as do r and t.) Now define the new events

$$A''' = \{N_n(\ell+1) = N_n(\ell) + 1; \lfloor n(t-\eta) \rfloor \leqslant \ell < T_{n,n}\},$$
$$B''' = \{K_\ell \geqslant 1; \lfloor n(r+\varepsilon) \rfloor + 1 \leqslant \ell \leqslant n\}.$$

We claim that, on $A' \cap A'' \cap A''' \cap B' \cap B'' \cap B'''$, it holds that $T_{n,n} \leqslant \tau_n$. Indeed, by definition of A'', $\lfloor n(t-\eta) \rfloor \leqslant T_{\lfloor nr+n\varepsilon \rfloor,n}$, so that, for $\ell \in [\lfloor n(t-\eta) \rfloor + 1, T_{\lfloor nr+n\varepsilon \rfloor+1,n} - 1]$, we have (in the same way as in (5.9))

$$\begin{aligned}
R(N_n(\ell)) &\geqslant R(m) + (N_n(\ell) - m)t/r - n\eta' \\
&\geqslant R(m) + (N_n(\lfloor nt-n\eta \rfloor) - m)t/r - n\eta' + (\ell - \lfloor nt - n\eta \rfloor)t/r \\
&> \lfloor nt - n\eta \rfloor + (\ell - \lfloor nt - n\eta \rfloor)t/r > \ell,
\end{aligned}$$

and thus, for $\ell \in [T_{\lfloor nr+n\varepsilon \rfloor+1,n}, T_{n,n}]$, we also have

$$\begin{aligned}
R(N_n(\ell)) &= R(\lfloor nr + n\varepsilon \rfloor) + R(N_n(\ell)) - R(\lfloor nr + n\varepsilon \rfloor) \\
&= R(N_n(T_{\lfloor nr+n\varepsilon \rfloor,n})) + R(N_n(\ell)) - R(N_n(T_{\lfloor nr+n\varepsilon \rfloor,n})) \\
&\geqslant R(N_n(T_{\lfloor nr+n\varepsilon \rfloor,n})) + N_n(\ell) - N_n(T_{\lfloor nr+n\varepsilon \rfloor,n}) \\
&> T_{\lfloor nr+n\varepsilon \rfloor,n} + \ell - T_{\lfloor nr+n\varepsilon \rfloor,n} = \ell.
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{P}(N_n(\tau_n) = n) &\geqslant \mathbb{P}(A' \cap A'' \cap A''' \cap B' \cap B'' \cap B''') \\
&\geqslant \mathbb{P}(A' \cap A'' \cap A''')\mathbb{P}(B' \cap B'' \cap B''') \\
&\geqslant \mathbb{P}(A' \cap A'' \cap A''')\mathbb{P}(C \cap B' \cap B'' \cap B'''),
\end{aligned} \tag{5.38}$$

with C as in the third step of the proof of the lower bound in Theorem 2.7.

From (5.11), it is plain to see that

$$\mathbb{P}(A'''|A' \cap A'') \geqslant \exp\left[\sum_{i=\lfloor n(r-\eta) \rfloor}^{\infty} \ln(1-i/n)\right],$$

so that

$$\liminf_{n\to\infty} n^{-1} \ln \mathbb{P}(A'''|A' \cap A'') \geqslant \int_{r-\eta}^{1} \ln(1-u)du = \delta''(r,\eta), \tag{5.39}$$

where here and below, $\delta''(r,\eta)$ stands for a generic term such that $\delta''(r,\eta) \to 0$ as $(r,\eta) \to (1,0)$. Similarly, by (5.24),

$$\mathbb{P}(B'''|B' \cap B'' \cap C) \geqslant \mathbb{P}\{K \geqslant 1\}^{n-\lfloor n(r+\varepsilon) \rfloor},$$

so that

$$\liminf_{n\to\infty} n^{-1} \ln \mathbb{P}(B'''|C' \cap B' \cap B'') \geqslant (1-r)\ln \mathbb{P}(K \geqslant 1) = \delta'''(r), \tag{5.40}$$

with $\delta'''(r) \to 0$ as $r \to 1$. Therefore, from (5.38), (5.39) and (5.40),

$$\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big(N_n(\tau_n) = n\big) \geqslant \liminf_{n \to \infty} n^{-1} \big[\ln \mathbb{P}(A' \cap A'') + \ln \mathbb{P}(C \cap B' \cap B'')\big]$$
$$+ \delta''(r, \eta) + \delta'''(r).$$

By the proof of the lower bound in Theorem 2.7, we know that

$$\liminf_{n \to \infty} n^{-1} \big[\ln \mathbb{P}(A' \cap A'') + \ln \mathbb{P}(C \cap B' \cap B'')\big] \geqslant -\mathcal{F}(r, t) + \delta(\eta) + \delta'(\varepsilon; \eta).$$

Finally we deduce that

$$\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big(N_n(\tau_n) = n\big) \geqslant -\mathcal{F}(r, t) + \delta(\eta) + \delta'(\varepsilon; \eta) + \delta''(r, \eta),$$

for $0 < r < t \wedge 1$. Assume then that $t > 1$ and $t \in \mathrm{Dom}^\circ(I)$ (interior of $\mathrm{Dom}(I)$). Then, by continuity of $\mathcal{F}$ (see Lemma 2.6), we can let r tend to 1. Letting η and ε also tend to 0, we get

$$\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big(N_n(\tau_n) = n\big) \geqslant -\mathcal{F}(1, t).$$

If $t \geqslant 1$ and $t \in \mathrm{Dom}(I)$, the above inequality still holds, by continuity as well, provided that $\mathrm{Dom}^\circ(I) \cap (1, \infty) \neq \emptyset$. If $t \notin \mathrm{Dom}(I)$, the result obviously holds, so that, in the case when $\mathrm{Dom}^\circ(I) \cap (1, \infty) \neq \emptyset$,

$$\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big(N_n(\tau_n) = n\big) \geqslant -\inf_{t \geqslant 1} \mathcal{F}(1, t) = -\mathcal{G}(1).$$

It thus remains to tackle the case when $\mathrm{Dom}^\circ(I) \cap (1, \infty) = \emptyset$. In fact this is the case when K is a Bernoulli random variable. Then, we can follow the special case we discussed in the final step of the proof of Theorem 2.7. Indeed, for $0 < r < 1$, we deduce from (5.33), (5.34) and (5.35) (with $t = 1$ and $r = t - \eta$ therein),

$$\mathbb{P}\big(N_n(\tau_n)/n \geqslant r\big) \geqslant (r - 1/n) \exp\bigg(-\sum_{i=1}^{\lfloor nr \rfloor} \ln(1 - i/n) - nrI(1)\bigg).$$

Letting r tend to 1, we deduce that

$$\mathbb{P}\big(N_n(\tau_n) = n\big) \geqslant \exp\bigg(-\sum_{i=1}^{n-1} \ln(1 - i/n) - nI(1)\bigg),$$

so that

$$\liminf_{n \to \infty} n^{-1} \ln \mathbb{P}\big(N_n(\tau_n) = n\big) \geqslant -\int_0^1 \ln(1 - u) du - I(1) = -(1 + I(1)).$$

It then remains to check that it is equal to $-\mathcal{G}(1)$. Clearly, the infimum in the definition of $\mathcal{G}(1)$ reduces to that value of s for which $\lambda(s) = 1$, that is, $s = 0$. We easily deduce that $\mathcal{G}(1) = 1 + I(1)$.

This completes the proof of the variational formula. Since it has compact level sets, the lower semi-continuous function $\mathcal{F}$, when restricted to the set of points (r, t) with $r = 1$, achieves its minimum, and the value of the minimum is non-zero. Hence $\mathcal{G}(1) > 0$. $\quad\square$

Acknowledgements

References

[1] Alves, O., Lebensztayn, E., Machado, F. and Martinez, M. (2006) Random walks systems on complete graphs. *Bull. Braz. Math. Soc.* (NS) **37** 571–580.

[2] Alves, O., Machado, F. and Popov, S. (2002) Phase transition for the frog model. *Electron. J. Probab.* **7** #16.

[3] Alves, O., Machado, F. and Popov, S. (2002) The shape theorem for the frog model. *Ann. Appl. Probab.* **12** 533–546.

[4] Baccelli, F., Blaszczyszyn, B. and Mirsadeghi, M. (2011) Optimal paths on the space–time SINR random graph, *Adv. Appl. Probab.* **43** 131–150.

[5] Baum, L. and Billingsley, P. (1965) Asymptotic distributions for the coupon collector's problem. *Ann. Math. Statist.* **36** 1835–1839.

[6] Billingsley, P. (1968) *Convergence of Probability Measures*, Wiley.

[7] Boucheron, S., Gamboa, F. and Léonard, C. (2002) Bins and balls: Large deviations of the empirical occupancy process. *Ann. Appl. Probab.* **12** 607–636.

[8] Comets, F., Quastel, J. and Ramírez, A. (2007) Fluctuations of the front in a stochastic combustion model. *Ann. Inst. H. Poincaré Probab. Statist.* **43** 147–162.

[9] Comets, F., Quastel, J. and Ramírez, A. (2009) Fluctuations of the front in a one dimensional model of $X + Y \to 2X$. *Trans. Amer. Math. Soc.* **361** 6165–6189.

[10] Dacunha-Castelle, D. and Duflo, M. (1983) *Probabilités et Statistiques 2: Temps Mobile*, Masson.

[11] Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Applications*, second edition, Springer.

[12] Ding, L. and Guan, Z.-H. (2008) Modeling wireless sensor networks using random graph theory. *Physica A* **387** 3008–3016.

[13] Dupuis, P., Nuzman, C. and Whiting, P. (2004) Large deviation asymptotics for occupancy problems. *Ann. Probab.* **32** 2765–2818.

[14] Duquesne, T. and Le Gall, J.-F. (2002) Random trees, Lévy processes and spatial branching processes. *Astérisque* **281**.

[15] Durrett, R. (1995) *Probability: Theory and Examples*, second edition, Duxbury.

[16] Erdős, P. and Rényi, A. (1961) On a classical problem of probability theory. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6** 215–220.

[17] Flajolet, P., Gardy, D. and Thimonier, L. (1992) Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discrete Appl. Math.* **39** 207–229.

[18] Jacod, J. and Shiryaev, A. N. (2002) *Limit Theorems for Stochastic Processes*, second edition, Springer.

[19] Jia, X. (2004) Wireless networks and random geometric graphs. In *Proc. 2004 International Symposium on Parallel Architectures, Algorithms and Networks: ISPAN'04*, pp. 575–580.

[20] Kan, N. (2002) The martingale approach to the coupon collection problem. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* **294** 113–126. Translation in *J. Math. Sci.* **127** (2005) 1737–1744.

[21] Kawahigashi, H., Terashima, Y., Miyauchi, N. and Nakakawaji, T. (2005) Modeling ad hoc sensor networks using random graph theory. In *Proc. Second IEEE Consumer Communications and Networking Conference: CCNC 2005*.

[22] Kesten, H. and Sidoravicius, V. (2005) The spread of a rumor or infection in a moving population. *Ann. Probab.* **33** 2402–2462.

[23] Kesten, H. and Sidoravicius, V. (2006) A phase transition in a model for the spread of an infection. *Illinois J. Math.* **50** 547–634.

[24] Kesten, H. and Stigum, B. P. (1966) A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37** 1211–1223.

[25] Kurkova, I., Popov, S. and Vachkovskaia, M. (2004) On infection spreading and competition between independent random walks. *Electron. J. Probab.* **9** 1–22.

[26] Kurtz, T., Lebensztayn, E., Leichsenring, A. R. and Machado, F. P. (2008) Limit theorems for an epidemic model on the complete graph. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 45–55.

[27] Machado, F., Machurian, H. and Matzinger, H. (2011) CLT for the proportion of infected individuals for an epidemic model on a complete graph. *Markov Proc. Rel. Fields* **17** 209–224.

[28] Neveu, J. (1986) Arbres et processus de Galton–Watson. *Annales de l'IHP* B **22** 199–207.

[29] Pitman, J. (2006) *Combinatorial Stochastic Processes: St. Flour 2002*, Vol. 1875 of *Lecture Notes in Mathematics*, Springer.

[30] Ramírez, A. and Sidoravicius, V. (2004) Asymptotic behavior of a stochastic combustion growth process. *J. Eur. Math. Soc.* **6** 293–334.

[31] Sedgewick, R. and Flajolet, P. (1996) *An Introduction to the Analysis of Algorithms*, Addison-Wesley.

[32] Zhukovskiĭ, M. E. (2012) The law of large numbers for an epidemic model (Russian). *Dokl. Akad. Nauk* **442** 736–739. Translation in *Dokl. Math.* **85** (2012) 113–116.