Deviation results for sparse tables in hashing with linear probing

Thierry Klein, A Lagnoux, P Petit

Deviation results for sparse tables in hashing with linear probing

[1], Thierry Klein[1], Agnès Lagnoux[2], and Pierre Petit[3]

[1]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; ENAC - Ecole Nationale de l'Aviation Civile , Université de Toulouse, France
[2]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT2J, F-31058 Toulouse, France.
[3]Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT3, F-31062 Toulouse, France.

January 14, 2021

Abstract

We consider the model of hashing with linear probing and we establish the moderate and large deviations for the total displacement in sparse tables. In this context, Weibull-like-tailed random variables appear. Deviations for sums of such heavy-tailed random variables are studied in [20, 21]. Here we adapt the proofs therein to deal with conditioned sums of such variables and solve the open question in [8]. By the way, we establish the deviations of the total displacement in full tables, which can be derived from the deviations of empirical processes of i.i.d. random variables established in [28].

Keywords: large deviations, hashing with linear probing, parking problem, Brownian motion, Airy distribution, Łukasiewicz random walk, empirical processes, conditioned sums of i.i.d. random variables, triangular arrays, Weibull-like distribution.
AMS subject classification: 60F10; 60C05; 60G50; 68W40.

1 Introduction

Hashing with linear probing is a classical model in theoretical computer science that appeared in the 50's. It has been studied from a mathematical point of view firstly by Knuth in [14]. Here is a simple description given in [6].

> A table of length m, $T[1..m]$ is set up, as well as a hash function h that maps keys from some domain to the interval $[1..m]$ of table addresses. A collection of n elements with $n \leq m$ are entered sequentially into the table according to the following rule: Each element x is placed at the first unoccupied location starting from $h(x)$ in cyclic order, namely the first of $h(x)$, $h(x)+1$, ..., m, 1, 2, ..., $h(x)-1$.

For more details on the model, we refer to [6, 9, 18, 3, 2, 11]. The length of the move of each element x is called the *displacement* of x and the sum of all displacements, denoted by $d_{m,n}$, is called the *total displacement*. In its seminal papers [14, 15], Knuth assumes that

all the sequence of hash addresses $h(x)$ are independent and uniformly distributed on $[\![1, m]\!]$, computes exact expressions of $\mathbb{E}[d_{m,n}]$ and $\mathrm{Var}(d_{m,n})$, and provides their asymptotic behaviors. The limit distribution of $d_{m,n}$ remains unknown until 1998: in [6], Flajolet, Poblete, and Viola give the limit distribution of $d_{m,n}$ for full tables ($n = m$) and for sparse tables ($n/m = \mu \in (0, 1)$) using combinatorial arguments. In [4], Chassaing and Marckert recover the previous results in the full case via a probabilistic approach. They prove that $d_{m,n}$ is the area under the Łukasiewicz random walk (also called Breadth First Search random walk) associated to a Galton-Watson tree with Poisson progeny. Consequently, the limit distribution of the total displacement $d_{m,m}$ is that of the area under the Brownian excursion, which involves the Airy distribution.

In [9], reformulating the problem in terms of conditioned sums of random variables, Janson establishes the limit distribution of $d_{m,n}$ in all cases with probabilistic tools. In [10], Janson extends the central limit theorem in the sparse case to a general model of conditioned sums of random variables. The corresponding Berry-Esseen bounds are proved by Klein, Lagnoux, and Petit in [12]. Concerning the deviations of such conditioned models, Gamboa, Klein, and Prieur give an answer in the case of light-tailed random variables (see [8]). Unfortunately, their results cannot be applied to the model of hashing with linear probing since this model involves heavy-tailed random variables.

In this paper, we establish the moderate and large deviations for the total displacement $d_{m,n}$ in sparse tables. Deviations for heavy-tailed random variables are studied by several authors (e.g., [24, 17, 20, 21, 22]) and a good survey can be found in Mikosch [19]. In the context of hashing with linear probing, Weibull-like-tailed random variables appear. Deviations for sums of such variables are studied by Nagaev in [20, 21]. Here we adapt its proofs to deal with conditioned sums of such variables. We also need to establish the deviations of $d_{m,n}$ for full tables, which can be derived from the deviations of empirical processes of i.i.d. random variables established by Wu in [28].

The paper is organized as follows. In Section 2, we state the main results for full and sparse tables. The proofs for full tables are given in Section 3 and those for sparse tables can be found in Section 5. In Section 4, we expose Janson's reformulation of the model and provide several useful estimates required in Section 5.

2 Setting and main results

2.1 Model

An equivalent formulation of the problem of hashing can be made in terms of the discrete version of the classical parking problem described, for instance, by Knuth [16]:

> A certain one-way street has m parking spaces in a row numbered 1 to m. A man and his dozing wife drive by, and suddenly, she wakes up and orders him to park immediately. He dutifully parks at the first available space [...].

More precisely, the model describes the following experiment. Let $n \leqslant m$. n cars enter sequentially into a circular parking uniformly at random. The parking spaces are numbered clockwise. A car that intends to park at an occupied space moves to the next empty space, always moving clockwise. The length of the move is called the displacement of the car and we are interested in the sum of all displacements which is a random variable denoted by $d_{m,n}$. When all cars are parked, there are $N = m - n$ empty spaces. These divide the occupied spaces into blocks of consecutive spaces. We consider that the empty space following a block belongs to this block.

For example, assume that $n = 8$, $m = 10$, and $(6, 9, 1, 9, 9, 6, 2, 5)$ are the addresses where the cars land. This sequence of (hash) addresses is called a *hash sequence* of length m and size n.

Let d_i be the displacement of car i. Then $d_1 = d_2 = d_3 = 0$. The car number 4 should park on the 9thspace which is occupied by the 2nd car; thus it moves one space ahead and parks on the 10thspace so that $d_4 = 1$. The car number 5 should park on the 9thspace. Since the 9th, the 10th, and the 1st spaces are occupied, $d_5 = 3$. And so on: $d_6 = 1,\ d_7 = 1,\ d_8 = 0$. Here, the total displacement is equal to $d_{10,8} = 1 + 3 + 1 + 1 = 6$. In our example, there are two blocks: the first one containing spaces 9, 10, 1, 2, 3 (occupied), and space 4 (empty), and the second one containing spaces 5, 6, 7 (occupied), and space 8 (empty).
In this paper, we are interested in the deviations of the total displacement $d_{m,n}$ in sparse tables. To do so, we need the large deviation behavior of $d_{m,n}$ in full tables. By the way, we also established the moderate deviation behavior of $d_{m,n}$ in full tables.

2.2 Deviations in full tables

In this section, we first recall some already existing results for the total displacement $d_{m,m}$ in full tables ($n = m$). As mentioned in the introduction, Knuth in [15] and Flajolet et al. in [6, Theorem 2]) derive the asymptotic behavior of the expectation and the variance of $d_{m,m}$:

$$\mathbb{E}[d_{m,m}] \underset{m\to\infty}{\sim} \frac{\sqrt{2\pi}}{4} m^{3/2} \quad \text{and} \quad \mathrm{Var}(d_{m,m}) \underset{m\to\infty}{\sim} \frac{10 - 3\pi}{24} m^3. \tag{1}$$

The following result was first established in [6, Theorem 3].

Theorem 1 (Standard deviations)**.** *For full tables, the distribution of the total displacement $d_{m,m}/m^{3/2}$ is asymptotically distributed as the area A under the standard Brownian excursion, in the sense that, for all $\delta \geqslant 0$,*

$$\mathbb{P}(d_{m,m} \geqslant m^{3/2}\delta) \xrightarrow[m\to\infty]{} \mathbb{P}(A \geqslant \delta).$$

In this paper, we establish the probabilities of deviation for the total displacement $d_{m,m}$.

Theorem 2 (Moderate deviations)**.** *For all $\alpha \in (3/2, 2)$ and for all $\delta \geqslant 0$,*

$$\frac{1}{m^{2\alpha-3}} \log \mathbb{P}(d_{m,m} \geqslant m^{\alpha}\delta) \xrightarrow[m\to\infty]{} -6\delta^2.$$

Theorem 3 (Large deviations)**.** *For all $\delta \geqslant 0$,*

$$-\frac{1}{m} \log \mathbb{P}(d_{m,m} \geqslant m^2\delta) \xrightarrow[m\to\infty]{} J(\delta) := \begin{cases} (\frac{1}{2} - \delta) \cdot \lambda(\delta) + \log(1 - (\frac{1}{2} + \delta) \cdot \lambda(\delta)) & \text{if } \delta < 1/2 \\ \infty & \text{if } \delta \geqslant 1/2, \end{cases}$$

where $\lambda(\delta)$ is the smallest solution of the equation in λ

$$\left(\lambda \cdot \left(\delta + \frac{1}{2}\right) - 1\right)(1 - e^{\lambda}) = \lambda. \tag{2}$$

Observe that lower deviations are trivial: for all $\alpha \in (3/2, 2]$ and all m large enough, $\mathbb{P}(d_{m,m} - \mathbb{E}[d_{m,m}] \leqslant -m^{\alpha}\delta) = 0$ because of the positiveness of $d_{m,m}$ and using (1). When dealing with the very large deviations, the same trivial behavior occurs both for upper and lower deviations: for all $\alpha > 2$ and all m large enough, $\mathbb{P}(d_{m,m} - \mathbb{E}[d_{m,m}] \geqslant m^{\alpha}\delta) = 0$ and $\mathbb{P}(d_{m,m} - \mathbb{E}[d_{m,m}] \leqslant -m^{\alpha}\delta) = 0$, since $d_{m,m} \leqslant m(m-1)/2$.

Remark 4. If $\delta = 0$, $\lambda(0) = 0$ is the unique solution of Equation (2). If $\delta \in (0, 1/2)$, Equation (2) has two solutions: $\lambda(\delta) < 0$ and 0. Moreover, $J(\delta) \sim 6\delta^2$ as $\delta \to 0$ (we recover the rate function of the moderate deviations in Theorem 2) and $J(\delta) \to +\infty$ as $\delta \to 1/2$.

Remark 5. The conclusions of Theorems 1, 2, and 3 are still valid replacing δ by $\delta + o(1)$, since the rate functions are continuous. In particular, one may replace $d_{m,m}$ by $d_{m,n}$ as soon as $m - n \ll m^{\alpha-1}$ (for instance, in the almost full case where $n = m - 1$). Indeed, naturally coupling $d_{m,n}$ and $d_{m,m}$ by adding $m - n$ balls, one has

$$|d_{m,n} - d_{m,m}| \leqslant (m-1) + (m-2) + \cdots + n \sim m(m-n) \ll m^{\alpha},$$

whence

$$\mathbb{P}(d_{m,n} \geqslant m^{\alpha}\delta) = \mathbb{P}(d_{m,m} \geqslant m^{\alpha}(\delta + o(1))).$$

Moreover, using a new probabilistic approach developed in [4], Theorem 1 was extended in [9, Theorems 1.1 and 2.2] to the case $(m-n)/\sqrt{m} \to a \in [0, \infty)$: for all $\delta \geqslant 0$,

$$\mathbb{P}(d_{m,n} \geqslant m^{3/2}\delta) \xrightarrow[m\to\infty]{} \mathbb{P}(W_a \geqslant \delta),$$

with

$$W_a = \int_0^1 \max_{s \leqslant t}(b(t) - b(s) - a(t-s))dt$$

where b is a Brownian bridge b on $[0, 1]$ periodically extended to $\mathbb{R}$.

2.3 Deviations in sparse tables

In this section, we consider asymptotics in (m, n) with $m \to \infty$ and $n/m \to \mu \in (0, 1)$ (sparse case). This definition of the sparse case is a slight extension of that of [6] ($n/m = \mu \in (0, 1)$). Set $N = m - n$. By [6, Theorem 5],

$$\mathbb{E}[d_{m,n}] \underset{m\to\infty}{\sim} \frac{\mu^2}{2(1-\mu)^2}N \quad \text{and} \quad \mathrm{Var}(d_{m,n}) \underset{m\to\infty}{\sim} \sigma^2(\mu)N, \tag{3}$$

where (cf. [6, Theorem 5])

$$\sigma^2(\mu) := \frac{6\mu^2 - 6\mu^3 + 4\mu^4 - \mu^5}{12(1-\mu)^5}. \tag{4}$$

The following result was first proved in [6, Theorem 6] while another probabilistic proof was given in [9].

Theorem 6 (Standard deviations)**.** *The distribution of the total displacement $d_{m,n}$ is asymptotically Gaussian distributed, in the sense that, for all y,*

$$\mathbb{P}\left(d_{m,n} - \mathbb{E}[d_{m,n}] \leqslant N^{1/2}y\right) \xrightarrow[m\to\infty]{} \mathbb{P}(Z \leqslant y)$$

where $Z \sim \mathcal{N}(0, \sigma^2(\mu))$.

In this paper, we establish the probabilities of deviation of the total displacement $d_{m,n}$ in the sparse case.

Theorem 7 (Lower moderate deviations)**.** *For all $\alpha \in (1/2, 1)$ and for all $y \geqslant 0$,*

$$\frac{1}{N^{2\alpha-1}} \log \mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \leqslant -N^{\alpha}y) \xrightarrow[m\to\infty]{} -\frac{y^2}{2\sigma^2(\mu)}. \tag{5}$$

Theorem 8 (Lower large deviations)**.** *For all $y \geqslant 0$,*

$$\frac{1}{N} \log \mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \leqslant -Ny) \xrightarrow[m\to\infty]{} -\Lambda^*\left(\frac{1}{1-\mu}, \frac{\mu^2}{2(1-\mu)^2} - y\right), \tag{6}$$

where Λ^* is the Fenchel-Legendre transform of the function $\Lambda\colon \mathbb{R}^2 \to (-\infty, \infty]$ defined by

$$\Lambda(s,t) = \log \sum_{l=1}^{\infty} \frac{e^{(s-\mu)l}(\mu l)^{l-1}}{l!} \mathbb{E}[e^{td_{l,l-1}}].$$

For all $\alpha > 1$, we have, asymptotically, $\mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \leqslant -N^\alpha y) = 0$, since $d_{m,n} \geqslant 0$ and $\mathbb{E}[d_{m,n}]$ is asymptotically linear in N.

Theorem 9 (Upper deviations)**.**
(i) For all $\alpha \in (1/2, 2/3)$ and for all $y \geqslant 0$,

$$\frac{1}{N^{2\alpha-1}} \log \mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \geqslant N^\alpha y) \xrightarrow[m\to\infty]{} -\frac{y^2}{2\sigma^2(\mu)}. \tag{7}$$

(ii) For all $y \geqslant 0$,

$$\frac{1}{N^{1/3}} \log \mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \geqslant N^{2/3} y) \xrightarrow[m\to\infty]{} -I(y) \tag{8}$$

with

$$I(y) := \begin{cases} \frac{y^2}{2\sigma^2(\mu)} & \text{if } y \leqslant y(\mu), \\ q(\mu)(1-t(y))^{1/2} y^{1/2} + \frac{t(y)^2 y^2}{2\sigma^2(\mu)} & \text{if } y > y(\mu), \end{cases}$$

where

$$q(\mu) := \inf_{0<\delta<1/2} \frac{1}{\sqrt{\delta}} (\kappa(\mu) + J(\delta)),$$

$\kappa(\mu) := \mu - \log(\mu) - 1 \in (0, \infty)$, J has been defined in Theorem 3, $y(\mu) := 3\left(q(\mu)\sigma^2(\mu)\right)^{2/3}/2$, and $t(y)$ is defined for $y > y(\mu)$ as the smallest root of the cubic equation in $t \in [0,1]$

$$t^3 - t^2 + \frac{q^2(\mu)\sigma^4(\mu)}{4y^3} = 0.$$

(iii) For all $\alpha \in (2/3, 2)$ and for all $y \geqslant 0$,

$$\frac{1}{N^{\alpha/2}} \log \mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \geqslant N^\alpha y) \xrightarrow[m\to\infty]{} -q(\mu) y^{1/2}. \tag{9}$$

(iv) For all $y \geqslant 0$,

$$\frac{1}{N} \log \mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \geqslant N^2 y) \xrightarrow[m\to\infty]{} \begin{cases} -\inf_{\delta>0} \left[\sqrt{\frac{y}{\delta}} (\kappa(\mu) + J(\delta)) + \Lambda_0^* \left(\frac{1}{1-\mu} - \sqrt{\frac{y}{\delta}} \right) \right] & \text{if } y < \mu^2/(2(1-\mu)^2) \\ -\infty & \text{if } y \geqslant \mu^2/(2(1-\mu)^2), \end{cases} \tag{10}$$

where Λ_0^* is the Fenchel-Legendre transform of the function $\Lambda_0 \colon \mathbb{R} \to (-\infty, \infty]$ defined by $\Lambda_0(s) := \Lambda(s, 0)$ and Λ has been defined in Theorem 8.

For all $\alpha > 2$, we have, asymptotically, $\mathbb{P}(d_{m,n} - \mathbb{E}[d_{m,n}] \geqslant N^\alpha y) = 0$, since $d_{m,n} \leqslant n(n-1)/2$ and N is asymptotically linear in n.

Remark 10. Observe that, for $\alpha = 2/3$ and for all $y > 0$,

$$I(y) = \inf_{t\in[0,1]} f(t),$$

where

$$f(t) = \left(q(\mu)(1-t)^{1/2} y^{1/2} + \frac{t^2 y^2}{2\sigma^2(\mu)} \right).$$

If $y \leqslant y_1(\mu) = 3\left(q(\mu)\sigma^2(\mu)\right)^{2/3}/2^{4/3}$, then f is decreasing and its minimum $y^2/(2\sigma^2(\mu))$ is attained at $t = 1$. If $y > y_1(\mu)$, f has two local minima, at 1 and at $t(y)$, corresponding to the smallest of the two roots in $[0,1]$ of $f'(t) = 0$ which is equivalent to $t^3 - t^2 + q^2(\mu)\sigma^4(\mu)/(4y^3) =$

0. If $y_1(\mu) < y \leqslant y(\mu)$, the minimum is attained at 1, and at $t(y)$ otherwise. Let $c = q^2(\mu)\sigma^4(\mu)/(4y^3)$. One can prove that $t(y) = 2\mathfrak{Re}(z) + 1/3$ where z is the only complex cube root of

$$\frac{1}{27} - \frac{c}{2} + i\sqrt{\frac{c}{27} - \frac{c^2}{4}}$$

having argument in $(\pi/3, 2\pi/3)$.

Remark 11 (Typical events of deviation in Theorem 9).
(i) All the displacements within the blocks are small but their sum has a Gaussian contribution.
(iii) One block has a large size $\delta^{-1/2}N^{\alpha/2}y^{1/2}$ and the displacement within this block is $N^{\alpha}y$, δ being chosen by the optimization in $q(\mu)$ (two competing terms).
(ii) One block has a large size and the displacement within this block is large $(N_n^{\alpha}(1-t(y))y)$ and the sum of the other displacements has a Gaussian contribution (two extra competing terms).
(iv) One block has a large size $\delta^{-1/2}Ny^{1/2}$ and the displacement within this block is N^2y which forces the sum of the length of the other blocks to be abnormally small, that is $m - \delta^{-1/2}Ny^{1/2} \sim N((1-\mu)^{-1} - \delta^{-1/2}y^{1/2})$ (three competing terms).

3 Proofs: full tables

Here, we take $n = m$. All limits are considered as $m \to \infty$ unless stated otherwise.

3.1 Upper moderate deviations - Theorem 2

For all $m \geqslant 1$, let $(U_{m,i})_{1\leqslant i\leqslant m}$ be a sequence of i.i.d. random variables with uniform distribution $\mathcal{U}$ on $[0,1]$ and let, for all $i \in [\![1,m]\!]$, $V_{m,i} = \lceil mU_{m,i}\rceil$. Note that $(V_{m,i})_{1\leqslant i\leqslant m}$ is a sequence of i.i.d. random variables uniformly distributed on $[\![1,m]\!]$ and, for all $i = 1,\dots,m$, $V_{m,i}$ corresponds to the hash address of item i. Let us define

$$S_m(k) := \sum_{i=1}^{m} \mathbb{1}_{V_{m,i}\leqslant k} = L_m([0,k/m]),$$

where L_m is the empirical measure associated to the random sequence $(U_{m,i})_{1\leqslant i\leqslant m}$.
As in [9, Lemma 2.1], we extend the definition of $S_m(k)$ for $k \in \mathbb{Z}$ so that the sequence $(S_m(k)-k)_{k\in\mathbb{Z}}$ be m-periodic, whence

$$\min_{l<k}\{S_m(l) - l\} = \min_{1\leqslant l\leqslant m}\{S_m(l) - l\}.$$

Thus, by [9, Equation (2.1) and Lemma 2.1], the total displacement $d_{m,m}$ is given by

$$\begin{aligned} d_{m,m} &= \sum_{k=1}^{m}\Big(S_m(k) - k - \min_{1\leqslant l\leqslant m}\{S_m(l)-l\}\Big) \\ &= m\sum_{k=1}^{m}(L_m - \mathcal{U})([0,k/m]) - m^2\min_{1\leqslant l\leqslant m}\{(L_m-\mathcal{U})([0,l/m])\} \\ &= m^2\Big(\int_0^1 (L_m-\mathcal{U})([0,y])dy - \inf_{\substack{0\leqslant y\leqslant 1,\\ y\in\mathbb{Q}}}\{(L_m-\mathcal{U})([0,y])\} + \frac{A_m}{m}\Big) \end{aligned}$$

where $A_m \in [-3/2, -1/2]$. Let $\alpha \in (3/2, 2)$. Then,

$$\begin{aligned} \mathbb{P}\left(d_{m,m} \geqslant m^{\alpha}\delta\right) &= \mathbb{P}\Big(\frac{\sqrt{m}}{m^{\alpha-3/2}}\Big(\int_0^1 (L_m-\mathcal{U})([0,y])dy - \inf_{\substack{0\leqslant y\leqslant 1,\\ y\in\mathbb{Q}}}\{(L_m-\mathcal{U})([0,y])\} + \frac{A_m}{m}\Big) \geqslant \delta\Big) \\ &= \mathbb{P}\Big(\varphi\Big(\frac{\sqrt{m}}{m^{\alpha-3/2}}(L_m-\mathcal{U})\Big) \geqslant \delta_m\Big), \end{aligned}$$

where $\delta_m = \delta - A_m/m^{-(\alpha-1)} \to \delta$ and, for any measure $\nu \in \mathcal{M}([0,1])$ (the space of signed measures on $[0,1]$),

$$\varphi(\nu) := \int_0^1 \nu([0,y])dy - \inf_{\substack{0\leqslant y\leqslant 1,\\ y\in\mathbb{Q}}} \nu([0,y]). \tag{11}$$

Since

$$\varphi(\nu) = \int_0^1 (1-x)d\nu(x) - \inf_{\substack{0\leqslant y\leqslant 1,\\ y\in\mathbb{Q}}} \int_0^1 \mathbb{1}_{[0,y]}(x)d\nu(x)$$

by Fubini's theorem, φ is a measurable function when $\mathcal{M}([0,1])$ is equipped with the σ-algebra generated by the applications $\nu \mapsto \nu(f)$ for all bounded measurable function f on $[0,1]$. By [28, Equation (1.3)], we have

$$\begin{aligned}
&\frac{1}{m^{2\alpha-3}} \log \mathbb{P}\Big(\varphi\Big(\frac{\sqrt{m}}{m^{\alpha-3/2}}(L_m - \mathcal{U})\Big) \geqslant \delta\Big) \\
&\to -\inf\left\{\frac{1}{2}\int_0^1 \Big(\frac{d\nu}{dy}(y)\Big)^2 dy \ ; \ \nu \in \mathcal{M}([0,1]), \nu \ll \mathcal{U}, \varphi(\nu) \geqslant \delta, \nu([0,1]) = 0\right\} \\
&= -\inf\left\{\frac{1}{2}\int_0^1 G'(y)^2 dy \ ; \ G \in \mathrm{AC}([0,1]), \int_0^1 G(y)dy - \min G \geqslant \delta, G(0) = G(1) = 0\right\},
\end{aligned} \tag{12}$$

where $\mathrm{AC}([0,1])$ is the space of absolutely continuous functions on $[0,1]$. We have used the fact that $\inf\{G(y) \ ; \ 0 \leqslant y \leqslant 1, y \in \mathbb{Q}\} = \min G$ by the continuity of G on $[0,1]$. If G is a minimizer in (12) and if $\tilde{y}$ is a minimizer of G, then $\tilde{G}(y) = G(\tilde{y}+y) - G(\tilde{y})$ is also a a minimizer in (12). The addition $\tilde{y}+y$ has to be understood as the addition on the torus $\mathbb{T}$. Moreover, $\min \tilde{G} = 0$ and $\int_0^1 \tilde{G}'(y)^2 dy = \int_0^1 G'(y)^2 dy$. So

$$\begin{aligned}
&\inf\left\{\frac{1}{2}\int_0^1 G'(y)^2 dy \ ; \ G \in \mathrm{AC}([0,1]), \int_0^1 G(y)dy - \min G \geqslant \delta, G(0) = G(1) = 0\right\} \\
&= \inf\left\{\frac{1}{2}\int_0^1 G'(y)^2 dy \ ; \ G \in \mathrm{AC}([0,1]), \int_0^1 G(y)dy \geqslant \delta, G(0) = G(1) = 0\right\}.
\end{aligned} \tag{13}$$

Using the method of Lagrange multipliers, if G is a minimizer of

$$\left\{\frac{1}{2}\int_0^1 G'(y)^2 dy \ ; \ G \in \mathrm{AC}([0,1]), \int_0^1 G(y)dy = \delta\right\}, \tag{14}$$

then there exists $\lambda \in \mathbb{R}$ such that

$$\frac{1}{2}\int_0^1 G'(y)^2 dy + \lambda \int_0^1 G(y)dy = 0.$$

Differentiating with respect to G, we get

$$\forall h \in \mathcal{C}_c^\infty((0,1)) \quad \int_0^1 \Big(G'(y)h'(y) + \lambda h(y)\Big)dy = 0.$$

Integrating by parts, one has

$$\forall h \in \mathcal{C}_c^\infty((0,1)) \quad \int_0^1 \Big(G'(y) - \lambda y\Big)h'(y)dy = 0.$$

By Du Bois-Reymond's lemma in [5, p.184], the function $y \mapsto G'(y) - \lambda y$ is constant, so G is a quadratic polynomial. Getting back to (14), the minimizer is given by

$$G(y) = 6\delta y(1-y)$$

and the infimum is $6\delta^2$. Moreover, we check that $G(0) = G(1) = 0$ as required in (13). Consequently,

$$\frac{1}{m^{2\alpha-3}} \log \mathbb{P}\left(\varphi\left(\frac{\sqrt{m}}{m^{\alpha-3/2}}\left(L_m - \mathcal{U}\right)\right) \geqslant \delta\right) \to -6\delta^2.$$

Using the fact that, for $\Phi = \varphi((\sqrt{m}/m^{\alpha-3/2})\,(L_m - \mathcal{U}))$,

$$\mathbb{P}(\Phi \geqslant \delta + \varepsilon) \leqslant \mathbb{P}(\Phi \geqslant \delta_m) \leqslant \mathbb{P}(\Phi \geqslant \delta - \varepsilon) \tag{15}$$

for all $\varepsilon > 0$ and all m large enough, we conclude to the result in Theorem 2 by continuity of the function $\delta \mapsto -6\delta^2$. □

3.2 Upper large deviations - Theorem 3

The result for $\delta = 0$ is trivial. Suppose that $\delta > 0$. We have

$$\begin{aligned}\mathbb{P}\left(d_{m,m} \geqslant m^2\delta\right) &= \mathbb{P}\Big(\int_0^1 (L_m - \mathcal{U})([0,y])dy - \inf_{\substack{0\leqslant y\leqslant 1,\\ y\in\mathbb{Q}}} \{(L_m - \mathcal{U})([0,y])\} + \frac{A_m}{m} \geqslant \delta\Big) \\ &= \mathbb{P}\Big(\varphi\left(L_m - \mathcal{U}\right) \geqslant \delta - \frac{A_m}{m}\Big),\end{aligned}$$

where φ is the measurable function defined in (11). By Sanov's theorem (see, e.g., [28, Equation (1.1)]), we have

$$\begin{aligned}&\frac{1}{m} \log \mathbb{P}\left(\varphi\left(L_m - \mathcal{U}\right) \geqslant \delta\right) \\ &\to -\inf\left\{\int_0^1 \Big(\frac{d\nu}{dy}(y)\Big) \log\Big(\frac{d\nu}{dy}(y)\Big)dy \;;\; \nu \in \mathcal{M}_1^+([0,1]), \nu \ll \mathcal{U}, \varphi(\nu - \mathcal{U}) \geqslant \delta\right\} \\ &= -\inf\Big\{\int_0^1 F'(y)\log F'(y)dy \;;\; F \in \mathrm{AC}([0,1]), F' \geqslant 0, \int_0^1 (F - \mathrm{id})(y)dy - \min(F - \mathrm{id}) \geqslant \delta, \\ &\qquad\qquad F(0) = 0, F(1) = 1\Big\},\end{aligned} \tag{16}$$

where $\mathcal{M}_1^+([0,1])$ is the space of probability measures on $[0,1]$. If F is a minimizer in (16) and if $\tilde{y}$ is a minimizer of $F - \mathrm{id}$, then $\tilde{F}(y) = F(\tilde{y} + y) - F(\tilde{y})$ is also a a minimizer in (16). Moreover, $\min(\tilde{F} - \mathrm{id}) = 0$ and $\int_0^1 \tilde{F}'(y)\log\tilde{F}'(y)dy = \int_0^1 F'(y)\log F'(y)dy$. So

$$\begin{aligned}&-\inf\Big\{\int_0^1 F'(y)\log F'(y)dy \;;\; F \in \mathrm{AC}([0,1]), F' \geqslant 0, \int_0^1 (F - \mathrm{id})(y)dy - \min(F - \mathrm{id}) \geqslant \delta, \\ &\qquad\qquad F(0) = 0, F(1) = 1\Big\} \\ &= -\inf\Big\{\int_0^1 F'(y)\log F'(y)dy \;;\; F \in \mathrm{AC}([0,1]), F' \geqslant 0, \int_0^1 F(y)dy \geqslant \delta + \frac{1}{2}, \\ &\qquad\qquad F(0) = 0, F(1) = 1\Big\} \\ &= -\inf K,\end{aligned}$$

where $K\colon \mathrm{AC}([0,1]) \to [0,\infty]$ is the convex function defined by

$$K(F) = \begin{cases} \int_0^1 F'(y)\log(F'(y))dy & \text{if } F' \geqslant 0,\ \int_0^1 F(y)dy \geqslant \delta + \frac{1}{2},\ F(0) = 0,\ F(1) = 1 \\ \infty & \text{otherwise.}\end{cases}$$

It is a standard convex optimization problem, a minimizer of which is

$$\bar{F}(y) = a(1 - e^{\lambda y}), \quad \text{where} \quad \begin{cases} a(1 - e^{\lambda}) = 1 \\ a = \delta + \frac{1}{2} - \frac{1}{\lambda}. \end{cases}$$

(One can see that $a \geqslant 1$ and $\lambda \leqslant 0$.) Indeed, by the definition of a convex function and the subdifferential, it suffices to check that 0 belongs to the subdifferential of K at $\bar{F}$. For all $h \in \mathrm{AC}([0,1])$ and for all $t > 0$,

$$K(\bar{F} + th) = \begin{cases} \int_0^1 (\bar{F}' + th') \log(\bar{F}' + th') & \text{if } F' + th' \geqslant 0,\ \int_0^1 h \geqslant 0,\ h(0) = h(1) = 0 \\ \infty & \text{otherwise.} \end{cases}$$

Differentiating under the integral sign with respect to t and integrating by parts gives

$$K'(\bar{F}; h) = \int_0^1 h'(y)(\log \bar{F}'(y) + 1)dy = -\lambda \int_0^1 h(y)dy \geqslant 0$$

since $h(0) = h(1) = 0$, $\int_0^1 h(y)dy \geqslant 0$, and $\lambda < 0$. It remains to compute the value of K at $\bar{F}$ and to apply the same argument of continuity as in (15) to get the required result. □

4 Interlude

4.1 Janson's reformulation

Here, we consider (m, n) with $m \to \infty$ and $n/m \to \mu \in (0, 1)$. As a consequence, $N = m - n \to \infty$. To make notation clearer, we make quantities depend on N and all limits are considered as $N \to \infty$ unless stated otherwise. In the next section, we are interested in the deviations of d_{m_N, n_N} in that regime, which is called the sparse case (see [6, 9] for this denomination with slight variants). In this section, we introduce a reformulation of the model of hashing with linear probing due to Janson in [9] and prove some preliminary results.

For all $N \geqslant 1$, we consider a vector of random variables (X_N, Y_N) defined as follows. We assume that X_N is distributed according to the Borel distribution with parameter $\mu_N := n_N/m_N \in (0, 1)$, i.e.

$$\forall l \in [\![1, \infty[\![\quad \mathbb{P}(X_N = l) = e^{-\mu_N l} \frac{(\mu_N l)^{l-1}}{l!} \tag{17}$$

(see, e.g., [6] or [9] for more details). In some places, for the ease of computation, we may also use the parametrization $\lambda_N = e^{-\mu_N}\mu_N$ to get an equivalent definition of the Borel distribution:

$$\mathbb{P}(X_N = l) = \frac{1}{T(\lambda_N)} \frac{l^{l-1}\lambda_N^l}{l!}, \tag{18}$$

where T is the tree function (see, e.g., [7, p. 127]). Furthermore, we assume that Y_N given $\{X_N = l\}$ is distributed as $d_{l,l-1}$.
Let $(X_{N,i}, Y_{N,i})_{1 \leqslant i \leqslant N}$ be an i.i.d. sample distributed as (X_N, Y_N) and define, for all $k \in [\![1, N]\!]$,

$$S_{N,k} := \sum_{i=1}^{k} X_{N,i} \quad \text{and} \quad T_{N,k} := \sum_{i=1}^{k} Y_{N,i}.$$

To lighten notation, let $S_N := S_{N,N}$ and $T_N := T_{N,N}$. Notice that, for all $N \geqslant 1$, $\mathbb{E}[X_N] = (1 - \mu_N)^{-1} = m_N/N$, so $\mathbb{E}[S_N] = m_N$. Moreover, $\mathbb{P}(S_N = m_N) > 0$ and we have the following identity (see [9, Lemma 4.1]):

$$\mathcal{L}(d_{m_N, n_N}) = \mathcal{L}(T_N \,|\, S_N = m_N).$$

4.2 Tail estimates

For all $\xi \in (0,1)$, recall that $\kappa(\xi) = \xi - \log(\xi) - 1 \in (0,\infty)$.

Proposition 12 (Tail of X_N)**.** *If $l \geqslant 1/\mu_N$, then*

$$\log \mathbb{P}(X_N = l) \leqslant -\kappa(\mu_N) l. \tag{19}$$

And if $l_N \to \infty$, then

$$\log \mathbb{P}(X_N \geqslant l_N) \sim \log \mathbb{P}(X_N = l_N) \sim -\kappa(\mu) l_N. \tag{20}$$

Proof of Proposition 12. As soon as $\mu_N l \geqslant 1$, and since $\log(l!) \geqslant l(\log(l) - 1)$,

$$\log \mathbb{P}(X_N = l) = -\mu_N l + l \log(\mu_N l) - \log(\mu_N l) - \log(l!) \leqslant -l(\mu_N - \log(\mu_N) - 1) = -\kappa(\mu_N) l.$$

Therefore,

$$\log \mathbb{P}(X_N \geqslant l_N) = \log \sum_{l=l_N}^{\infty} \mathbb{P}(X_N = l_N) \leqslant \log \sum_{l=l_N}^{\infty} e^{-\kappa(\mu_N) l} \sim -\kappa(\mu) l_N.$$

Finally, using Stirling formula, one has

$$\log \mathbb{P}(X_N \geqslant l_N) \geqslant \log \mathbb{P}(X_N = l_N) = -\mu_N l_N + (l_N - 1) \log(\mu_N l_N) - \log(l_N!) \sim -\kappa(\mu) l_N. \qquad \square$$

From the previous proposition and Theorem 3, we deduce the asymptotic behavior of the tail of the pair (X_N, Y_N).

Proposition 13 (Tail of (X_N, Y_N))**.** *Let $l_N \to \infty$ and let p_N be such that $p_N / l_N^2 \to \delta$. Then*

$$\frac{1}{l_N} \log \mathbb{P}(X_N = l_N,\ Y_N \geqslant p_N) \to -(\kappa(\mu) + J(\delta)). \tag{21}$$

Proof of Proposition 13. It suffices to write

$$\begin{aligned} \frac{1}{l_N} \log \mathbb{P}(X_N = l_N,\ Y_N \geqslant p_N) &= \frac{1}{l_N} \log \mathbb{P}(X_N = l_N) + \frac{1}{l_N} \log \mathbb{P}(d_{l_N, l_N - 1} \geqslant p_N) \\ &= \frac{1}{l_N} \log \mathbb{P}(X_N = l_N) + \frac{1}{l_N} \log \mathbb{P}(d_{l_N, l_N} \geqslant l_N^2(\delta + o(1))) \\ &\to -(\kappa(\mu) + J(\delta)) \end{aligned}$$

by Remark 5, Proposition 12, and Theorem 3. $\square$

Lemma 14. *Let $\tilde{J} \leqslant J$ be any nondecreasing function, continuous on $[0, 1/2]$. For all $\varepsilon > 0$, there exists $N_0 \geqslant 1$ and $l_0 \geqslant 1$ such that, for all $N \geqslant N_0$, for all $l \geqslant l_0$, and for all $\delta \geqslant 0$,*

$$\log \mathbb{P}(X_N = l,\ Y_N \geqslant \delta l^2) \leqslant -(\kappa(\mu) + \tilde{J}(\delta) - \varepsilon) l.$$

Proof of Lemma 14. The result is trivial for $\delta \in (1/2, \infty)$. Remember that

$$\mathbb{P}(X_N = l,\ Y_N \geqslant \delta l^2) = \mathbb{P}(X_N = l) \mathbb{P}(d_{l, l-1} \geqslant \delta l^2).$$

On the one hand, let $\varepsilon > 0$. By Proposition 12, if N and l are large enough,

$$\log \mathbb{P}(X_N = l) \leqslant -\kappa(\mu_N) l \leqslant -(\kappa(\mu) - \varepsilon/2) l.$$

On the other hand, the nondecreasing functions $\phi_l \colon \delta \in [0, 1/2] \mapsto \min(-l^{-1} \log \mathbb{P}(d_{l,l-1} \geqslant \delta l^2), \tilde{J}(\delta))$ converge pointwise to the continuous function $\tilde{J}|_{[0,1/2]}$ as $l \to \infty$, by Theorem 3 ; thus the convergence is uniform and the result follows. $\square$

In the sequel, we will also need the asymptotic behavior of the tail of Y_N alone.

Proposition 15 (Tail of Y_N)**.** *If $p_N \to \infty$, then*

$$\frac{1}{\sqrt{p_N}} \log \mathbb{P}(Y_N \geqslant p_N) \to -q(\mu) = -\inf_{0<\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta)\Big]. \tag{22}$$

Proof of Proposition 15. For $\delta > 0$, let $l_N = \Big\lceil (p_N/\delta)^{1/2} \Big\rceil$. Then,

$$\frac{1}{\sqrt{p_N}} \log \mathbb{P}(Y_N \geqslant p_N) \geqslant \frac{l_N}{\sqrt{p_N}} \cdot \frac{1}{l_N} \log \mathbb{P}(X_N = l_N,\ Y_N \geqslant p_N) \to -\frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta)\Big],$$

by Proposition 13. Taking the supremum in $\delta > 0$, one gets

$$\liminf_{N\to\infty} \frac{1}{\sqrt{p_N}} \log \mathbb{P}(Y_N \geqslant p_N) \geqslant -\inf_{0<\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta)\Big]. \tag{23}$$

Now we turn to the upper bound. Let us fix $\beta > 0$ such that $\beta\kappa(\mu) > q(\mu)$. Let $l_N = \Big\lfloor \beta p_N^{1/2} \Big\rfloor$ and write

$$\mathbb{P}(Y_N \geqslant p_N) = \sum_{l=1}^{l_N} \mathbb{P}(X_N = l,\ Y_N \geqslant p_N) + \sum_{l=l_N+1}^{\infty} \mathbb{P}(X_N = l,\ Y_N \geqslant p_N) =: P_N + R_N.$$

First of all, using Proposition 12,

$$\frac{1}{\sqrt{p_N}} \log(R_N) \leqslant \frac{1}{\sqrt{p_N}} \log \mathbb{P}(X_N > l_N) \to -\beta\kappa(\mu) < -q(\mu).$$

Let $\varepsilon > 0$. Taking into account the already proved lower bound, and using Lemma 14 with

$$\tilde{J}(\delta) = J_\varepsilon(\delta) = \begin{cases} J(\delta) \wedge \varepsilon^{-1} & \text{if } \delta \leqslant 1/2 \\ \infty & \text{if } \delta > 1/2, \end{cases}$$

we deduce that

$$\begin{aligned} \limsup_{N\to\infty} \frac{1}{\sqrt{p_N}} \log \mathbb{P}(Y_N \geqslant p_N) &= \limsup_{N\to\infty} \frac{1}{\sqrt{p_N}} \log(P_N) \\ &\leqslant \max_{1\leqslant l\leqslant l_N} -\frac{l}{\sqrt{p_N}} \left[\kappa(\mu) + J_\varepsilon\Big(\frac{p_N}{l^2}\Big) - \varepsilon\right] \\ &\leqslant -\inf_{1/\beta\leqslant\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J_\varepsilon(\delta) - \varepsilon\Big] \\ &=: M_\varepsilon. \end{aligned}$$

Let $\delta_\varepsilon \in (0, 1/2)$ be such that $J(\delta_\varepsilon) = 1/\varepsilon$. We have

$$\inf_{\delta_\varepsilon<\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J_\varepsilon(\delta) - \varepsilon\Big] \geqslant \sqrt{2}\Big(\kappa(\mu) + \varepsilon^{-1} - \varepsilon\Big) \xrightarrow[\varepsilon\to 0]{} \infty.$$

A fortiori, since $J_\varepsilon \leqslant J$,

$$\inf_{\delta_\varepsilon<\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta) - \varepsilon\Big] \xrightarrow[\varepsilon\to 0]{} \infty.$$

So, if ε is small enough,

$$\begin{aligned} M_\varepsilon &= -\inf_{1/\beta\leqslant\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta) - \varepsilon\Big] \\ &\xrightarrow[\varepsilon\to 0]{} -\inf_{1/\beta\leqslant\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta)\Big] \leqslant -\inf_{0<\delta<1/2} \frac{1}{\sqrt{\delta}} \Big[\kappa(\mu) + J(\delta)\Big] \end{aligned}$$

and the result follows. □

4.3 Useful limit theorems

The following lemma is a direct consequence of [9, Lemma 4.3] and Proposition 17.

Proposition 16. *One has*

$$\mathbb{E}[T_N \,|\, S_N = m_N] = \mathbb{E}[T_N] + o(N^{1/2}).$$

Let (X, Y) be a pair of random variables such that X is distributed according to the Borel distribution with parameter $\mu = \lim \mu_N$ and Y given $\{X = l\}$ is distributed as $d_{l,l-1}$. Let $\lambda = e^{-\mu}\mu$ be the other standard parameter of the Borel distribution as in (18).

Proposition 17 (Moments convergence)**.** $(X_N, Y_N)_{N\geqslant 1}$ *converges to* (X, Y) *in distribution and with all mixed moments of the type* $\mathbb{E}[X_N^p Y_N^q e^{(s+it)X_N}]$*, where* $p \geqslant 0$*,* $q \geqslant 0$*,* $s < -\log(\lambda e)$*, and* $t \in \mathbb{R}$*.*

Proof of Proposition 17. Let $f\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be a bounded measurable function. Using (18), one has

$$\begin{aligned}\mathbb{E}[f(X_N, Y_N)] &= \sum_{l\in\mathbb{N}} \mathbb{E}[f(X_N, Y_N) \,|\, X_N = l]\mathbb{P}(X_N = l)\\ &= \sum_{l\in\mathbb{N}} \mathbb{E}[f(l, d_{l,l-1})]\frac{1}{T(\lambda_N)}\frac{l^{l-1}}{l!}\lambda_N^l.\end{aligned}$$

Since λ_N converges to λ, T is continuous, and

$$\frac{1}{T(\lambda_N)}\frac{l^{l-1}}{l!}\lambda_N^l \leqslant \frac{T(\lambda+\varepsilon)}{T(\lambda-\varepsilon)}\left(\frac{1}{T(\lambda+\varepsilon)}\frac{l^{l-1}}{l!}(\lambda+\varepsilon)^l\right), \tag{24}$$

as soon as $|\lambda_N - \lambda| \leqslant \varepsilon$, we conclude by Lebesgue's dominated convergence theorem that $(X_N, Y_N)_{N\geqslant 1}$ converges in distribution to (X, Y), where X is Borel distributed with parameter λ and $\mathcal{L}(Y \,|\, X = l) = \mathcal{L}(d_{l,l-1})$.
Let $c > 1$ such that $sc < -\log(\lambda e)$ and $(a, b) \in (\mathbb{R}_+^*)^2$ such that $a^{-1} + b^{-1} + c^{-1} = 1$. Hölder's inequality yields

$$|\mathbb{E}[X_N^p Y_N^q e^{(s+it)X_N}]| \leqslant \mathbb{E}[X_N^{ap}]^{1/a}\mathbb{E}[Y_N^{bq}]^{1/b}\mathbb{E}[e^{scX_N}]^{1/c}.$$

Hence the convergence of all mixed moments stems from the convergence of all separate moments of X_N, e^{X_N}, and Y_N. By uniform integrability, it suffices to show that, for each $r > 0$ and $s' < -\log(\lambda e)$, $(\mathbb{E}[X_N^r])_{N\geqslant 1}$, $(\mathbb{E}[e^{s'X_N}])_{N\geqslant 1}$, and $(\mathbb{E}[Y_N^r])_{N\geqslant 1}$ are bounded (see, e.g., [27, Example 2.21]). By (24), $\limsup \mathbb{E}[X_N^r]$ and $\limsup \mathbb{E}[e^{s'X_N}]$ $(s' < -\log(\lambda e))$ are finite. Moreover, since $d_{l,l-1} \leqslant l^2$,

$$\mathbb{E}[Y_N^r] = \sum_{l\in\mathbb{N}} \mathbb{E}[Y_N^r \,|\, X_N = l]\mathbb{P}(X_N = l) \leqslant \sum_{l\in\mathbb{N}} l^{2r}\mathbb{P}(X_N = l) = \mathbb{E}[X_N^{2r}],$$

so $\limsup \mathbb{E}[Y_N^r]$ is finite too. □

Proposition 18 (Local large deviations for S_N)**.** *For any sequence of integers* $(k_N)_{N\geqslant 1}$ *such that* $\lim k_N/N \in (1, \infty)$*, we have*

$$\frac{1}{N}\log \mathbb{P}(S_N = k_N) \sim -\Lambda_{X_N}^*(k_N/N).$$

Proof of Proposition 18. We just check that we can apply [8, Lemma 3.3] to the sequence $(X_N)_{N\geqslant 1}$. First, $\mathrm{Im}(\Lambda'_{X_N}) = (1, \infty) = \mathrm{Im}(\Lambda'_X)$ so, for all N large enough, $k_N/N \in \mathrm{Im}(\Lambda'_{X_N})$ and $\lim k_N/N \in \mathrm{Im}(\Lambda'_X)$. Secondly, $\mathrm{int}(\mathrm{dom}(\Lambda_{X_N})) = (-\infty, -\log(\lambda_N e))$ and $\mathrm{int}(\mathrm{dom}(\Lambda_X)) = (-\infty, -\log(\lambda e))$, so that assumption 1. of [8, Lemma 3.3] holds for all N large enough (since $\lambda_N \to \lambda$). Thirdly, assumption 2. of [8, Lemma 3.3] stems from Proposition 17. □

The following proposition is a non conditioned version of Theorem 9. It stems immediately from [13] (with $\epsilon = 1/2$ and $q = q(\mu)$) and Propositions 15 and 17.

Proposition 19 (Large deviations for T_N)**.**

(i) If $\alpha < 2/3$, then

$$\lim_{N\to\infty} \frac{1}{N^{2\alpha-1}} \log \mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y) = -\frac{y^2}{2\sigma^2(\mu)}. \tag{25}$$

(ii) If $\alpha = 2/3$, then

$$\lim_{N\to\infty} \frac{1}{N^{1/3}} \log \mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^{2/3} y) = -I(y) \tag{26}$$

where

$$I(y) := \begin{cases} \frac{y^2}{2\sigma^2(\mu)} & \text{if } y \leqslant y(\mu), \\ q(\mu)(1-t(y))^{1/2} y^{1/2} + \frac{t(y)^2 y^2}{2\sigma^2(\mu)} & \text{if } y > y(\mu), \end{cases}$$

with $y(\mu) = 3\left(q(\mu)\sigma^2(\mu)\right)^{2/3}/2$ and $t(y)$ being defined for $y > y(\mu)$ as the smallest root of the cubic equation in $t \in [0,1]$:

$$t^3 - t^2 + \frac{q^2(\mu)\sigma^4(\mu)}{4y^3} = 0.$$

(iii) If $\alpha > 2/3$, then

$$\lim_{N\to\infty} \frac{1}{N^{\alpha/2}} \log \mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y) = -q(\mu) y^{1/2}. \tag{27}$$

5 Proofs: sparse tables

5.1 Lower moderate deviations - Theorem 7

One has

$$\begin{aligned} \mathbb{P}(d_{m_N,n_N} - \mathbb{E}[d_{m_N,n_N}] \leqslant -N^\alpha y) &= \mathbb{P}(T_N - \mathbb{E}[T_N \,|\, S_N = m_N] \leqslant -N^\alpha y \,|\, S_N = m_N) \\ &= \mathbb{P}(T_N - \mathbb{E}[T_N] \leqslant -N^\alpha y_N \,|\, S_N = m_N) \end{aligned} \tag{28}$$

where

$$y_N := y - \frac{1}{N^\alpha}(\mathbb{E}[T_N \,|\, S_N = m_N] - \mathbb{E}[T_N]) \to y$$

by Proposition 16. Since the variables are nonnegative, their Laplace transforms are defined on $(-\infty, 0)$ at least. Adapting the proof of [8, Theorem 2.2] to the unilateral case and using [25] (unilateral version of Gärtner-Ellis theorem), we get (5). □

5.2 Lower large deviations - Theorem 8

For any $\mathbb{R}^d$-valued random variable Z, we denote by Λ_Z the log-Laplace transform of Z, i.e. the function defined, for $\lambda \in \mathbb{R}^d$, by

$$\Lambda_Z(\lambda) = \log \mathbb{E}[\exp(\lambda \cdot Z)],$$

and by Λ_Z^* the Fenchel-Legendre transform of the function Λ_Z, i.e. the function defined, for $z \in \mathbb{R}^d$, by

$$\Lambda_Z^*(z) = \sup\left\{\lambda \cdot z - \Lambda_Z(\lambda) \;;\; \lambda \in \mathbb{R}^d\right\}.$$

Proceeding as in the proof of Theorem 7, we get

$$\frac{1}{N}\log\mathbb{P}(d_{m_N,n_N}-\mathbb{E}[d_{m_N,n_N}]\leqslant -Ny)\to -\Lambda^*_{(X,Y)}\Big(\frac{1}{1-\mu},\frac{\mu^2}{2(1-\mu)^2}-y\Big), \tag{29}$$

since $m_N/N\to 1/(1-\mu)$ and admitting that $\Lambda^*_{(X,Y)}((1-\mu)^{-1},\cdot)$ is strictly convex on $(0,\mu^2/(2(1-\mu)^2))$.
Let $x_0=(1-\mu)^{-1}$. Let us prove that $\Lambda^*_{(X,Y)}(x_0,\cdot)$ is strictly convex. Let $y\in(0,\mu^2/(2(1-\mu)^2))$. First, we prove that $\Lambda^*_{(X,Y)}$ is differentiable at (x_0,y). We have $(x_0,y)\in\mathrm{int}(\mathrm{dom}(\Lambda^*_{(X,Y)}))$, therefore the subdifferential $\partial\Lambda^*_{(X,Y)}(x_0,y)$ is nonempty, i.e. there exists $(s,t)\in\partial\Lambda^*_{(X,Y)}(x_0,y)$ (see [26, Theorem 23.4]). It remains to prove that such a point (s,t) is unique. Choosing $\varepsilon>0$ such that $\Lambda^*_Y(y+\varepsilon)>0$,

$$\begin{aligned}-\Lambda^*_{(X,Y)}(x_0,y)&\leqslant\liminf\frac{1}{N}\log\mathbb{P}(S_N\leqslant N(x_0+\varepsilon),T_N\leqslant N(y+\varepsilon))\\&\leqslant\liminf\frac{1}{N}\log\mathbb{P}(T_N\leqslant N(y+\varepsilon))=-\Lambda^*_Y(y+\varepsilon)<0.\end{aligned}$$

Since $\Lambda^*_{(X,Y)}(x_0,\mu^2/(2(1-\mu)^2))=0$ and $\Lambda^*(x_0,\cdot)$ is convex, one has $t<0$. Therefore $(s,t)\in\mathrm{int}(\mathrm{dom}(\Lambda_{(X,Y)}))$. To obtain a local version of [26, Theorem 23.5], we notice that

$$(s,t)\in\partial\Lambda^*_{(X,Y)}(x_0,y)\iff(x_0,y)\in\partial\Lambda^{**}_{(X,Y)}(s,t)=\partial\Lambda_{(X,Y)}(s,t)=\{\nabla\Lambda_{(X,Y)}(s,t)\},$$

since $\Lambda_{(X,Y)}$ is differentiable on $\mathrm{int}(\mathrm{dom}(\Lambda_{(X,Y)}))$. Now,

$$\det(\mathrm{Hess}(\Lambda_{(X,Y)})(\lambda,\rho))=\mathrm{Var}(\tilde{X})\,\mathrm{Var}(\tilde{Y})-\mathrm{Cov}(\tilde{X},\tilde{Y})^2>0$$

where $(\tilde{X},\tilde{Y})$ has a mass function proportional to $e^{\lambda x+\rho y}f_{(X,Y)}(x,y)$ which is not supported by a line, so $\Lambda_{(X,Y)}$ is strictly convex. Thus (s,t) is the unique solution of $(x_0,y)=\nabla\Lambda_{(X,Y)}(s,t)$. Finally, let $y'\neq y$ and $(s',t')=\nabla\Lambda^*_{(X,Y)}(x_0,y')$. Remark that $(x_0,y)=\nabla\Lambda_{(X,Y)}(s,t)$ and $(x_0,y')=\nabla\Lambda_{(X,Y)}(s',t')$ lead to $(s',t')\neq(s,t)$. Therefore, by the strict convexity of $\Lambda_{(X,Y)}$,

$$\begin{aligned}&\langle(x_0,y')-(x_0,y),\nabla\Lambda^*_{(X,Y)}(x_0,y')-\nabla\Lambda^*_{(X,Y)}(x_0,y)\rangle\\&\qquad=\langle\nabla\Lambda_{(X,Y)}(s',t')-\nabla\Lambda_{(X,Y)}(s,t),(s',t')-(s,t)\rangle>0.\end{aligned}$$

Thus $\Lambda^*_{(X,Y)}(x_0,\cdot)$ is strictly convex. □

5.3 Upper moderate deviations - Theorem 9 (i)

Analogously to (28), one has

$$P_N:=\mathbb{P}(d_{m_N,n_N}-\mathbb{E}[d_{m_N,n_N}]\geqslant N^\alpha y)=\mathbb{P}(T_N-\mathbb{E}[T_N]\geqslant N^\alpha y_N\,|\,S_N=m_N) \tag{30}$$

$$\leqslant\frac{\mathbb{P}(T_N-\mathbb{E}[T_N]\geqslant N^\alpha y_N)}{\mathbb{P}(S_N=m_N)} \tag{31}$$

with $y_N:=y+N^{-\alpha}(\mathbb{E}[T_N\,|\,S_N=m_N]-\mathbb{E}[T_N])\to y$ by Proposition 16. The upper bound then follows from Propositions 19 and 18. As for the lower bound, using (30), one has

$$\begin{aligned}P_N&\geqslant\mathbb{P}(T_N-\mathbb{E}[T_N]\geqslant N^\alpha y_N,\ \forall i,Y_{N,i}<N^{\alpha/2}\,|\,S_N=m_N)\\&=\mathbb{P}(T_N-\mathbb{E}[T_N]\geqslant N^\alpha y_N\,|\,\forall i,Y_{N,i}<N^{\alpha/2},\ S_N=m_N)\mathbb{P}(\forall i,Y_{N,i}<N^{\alpha/2}\,|\,S_N=m_N).\end{aligned}$$

On the one hand, one has

$$\mathbb{P}(\forall i,Y_{N,i}<N^{\alpha/2}\,|\,S_N=m_N)=1-\mathbb{P}(\exists i,Y_{N,i}\geqslant N^{\alpha/2}\,|\,S_N=m_N)\geqslant 1-\frac{N\mathbb{P}(Y_{N,i}\geqslant N^{\alpha/2})}{\mathbb{P}(S_N=m_N)}. \tag{32}$$

Using Propositions 15 and 18, we derive that $\mathbb{P}(\forall i, Y_{N,i} < N^{\alpha/2} \,|\, S_N = m_N) \to 1$. On the other hand, let us turn to the minoration of $\mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y_N \,|\, \forall i, Y_{N,i} < N^{\alpha/2},\ S_N = m_N)$. In order to apply Gärtner-Ellis theorem, we follow the proof of [8, Theorem 2.2] and we introduce

$$g_N(u) = \frac{1}{N^{2\alpha-1}} \log \mathbb{E}\Big[e^{u(T_N - \mathbb{E}[T_N])/N^{1-\alpha}} \,\Big|\, \forall i,\ Y_{N,i} < N^{\alpha/2},\ S_N = m_N\Big].$$

Write

$$\begin{aligned}
&\mathbb{E}\Big[e^{u(T_N - \mathbb{E}[T_N])/N^{1-\alpha}} \,\Big|\, \forall i,\ Y_{N,i} < N^{\alpha/2},\ S_N = m_N\Big] \\
&\quad = \frac{\mathbb{E}\Big[e^{u(T_N - \mathbb{E}[T_N])/N^{1-\alpha}} \mathbb{1}_{S_N = m_N} \,\Big|\, \forall i,\ Y_{N,i} < N^{\alpha/2}\Big]}{\mathbb{P}(S_N = m_N \,|\, \forall i, Y_{N,i} < N^{\alpha/2})} \\
&\quad = \frac{\mathbb{E}\Big[e^{u(T_N^< - \mathbb{E}[T_N])/N^{1-\alpha}} \mathbb{1}_{S_N^< = m_N}\Big]}{\mathbb{P}(S_N^< = m_N)},
\end{aligned}$$

where $S_N^< = \sum_{i=1}^N X_{N,i}^<$, $T_N^< = \sum_{i=1}^N Y_{N,i}^<$, and the random vectors $(X_{N,i}^<, Y_{N,i}^<)$ are independent, each distributed as $\mathcal{L}((X_N, Y_N) \,|\, Y_N < N^{\alpha/2})$. Then,

$$\begin{aligned}
\mathbb{E}\Big[e^{u(T_N^< - \mathbb{E}[T_N])/N^{1-\alpha}} \mathbb{1}_{S_N^< = m_N}\Big] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ism_N} \mathbb{E}\Big[e^{u(T_N^< - \mathbb{E}[T_N])/N^{1-\alpha} + isS_N^<}\Big] ds \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ism_N} \mathbb{E}\Big[e^{u(Y_N^< - \mathbb{E}[Y_N])/N^{1-\alpha} + isX_N^<}\Big]^N ds \\
&= e^{N\Lambda_{Y_N^< - \mathbb{E}[Y_N]}(u/N^{1-\alpha})} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ism_N} \mathbb{E}\Big[e^{is\hat{S}_N^u}\Big]^N ds \\
&= e^{N\Lambda_{Y_N^< - \mathbb{E}[Y_N]}(u/N^{1-\alpha})} \mathbb{P}(\hat{S}_N^u = m_N)
\end{aligned}$$

where $\hat{S}_N^u$ stands for $\sum_{i=1}^N \hat{X}_{N,i}^u$ and the random variables $\hat{X}_{N,i}^u$ are independent copies of $\hat{X}_N^u$, the distribution of which is given by

$$\mathbb{P}(\hat{X}_N^u = x) := e^{-\Lambda_{Y_N^< - \mathbb{E}[Y_N]}(u/N^{1-\alpha})} \mathbb{E}\Big[e^{u(Y_N^< - \mathbb{E}[Y_N])/N^{1-\alpha}} \mathbb{1}_{X_N^< = x}\Big].$$

Consequently,

$$g_N(u) = N^{2-2\alpha} \Lambda_{Y_N^< - \mathbb{E}[Y_N]}(u/N^{1-\alpha}) + \frac{1}{N^{2\alpha-1}} \log \mathbb{P}(\hat{S}_N^u = m_N) - \frac{1}{N^{2\alpha-1}} \log \mathbb{P}(S_N^< = m_N).$$

So, using Lemma 20 below, we get

$$\begin{aligned}
g_N(u) &= -N^{2-2\alpha}\Big(\Lambda^*_{\hat{X}_N^u}(m_N/N) - \Lambda^*_{X_N^<}(m_N/N) - \Lambda_{Y_N^< - \mathbb{E}[Y_N]}(u/N^{1-\alpha})\Big)(1 + o(1)) \\
&= -N^{2-2\alpha}\Big(H_N(u/N^{1-\alpha}) - H_N(0)\Big)(1 + o(1)),
\end{aligned}$$

where

$$H_N(t) = \sup\left\{ s\frac{m_N}{N} - \Lambda_{(X_N^<, Y_N^< - \mathbb{E}[Y_N])}(s,t) \ ;\ s \in \mathbb{R} \right\}.$$

Applying the global version of the inverse function theorem to the function $(s,t) \in \mathbb{R} \times \mathbb{R} \mapsto (t, sm_N/N - \Lambda_{(X_N^<, Y_N^< - \mathbb{E}[Y_N])}(s,t))$ and noting that $\partial_{s,s}\Lambda_{(X_N^<, Y_N^< - \mathbb{E}[Y_N])}(s,t)$ is nonzero since it is the variance of a non constant random variable, there is a unique maximizer $s_N(t)$ in the definition of $H_N(t)$. Moreover, the same algebraic computations as in [8] yield

$$H_N'(0) = \mathbb{E}[Y_N] - \mathbb{E}[\tilde{Y}_N] \quad \text{and} \quad H_N''(0) = \frac{\mathrm{Cov}(\tilde{Y}_N, \tilde{X}_N)^2}{\mathrm{Var}(\tilde{X}_N)} - \mathrm{Var}(\tilde{Y}_N),$$

where the distribution of $(\tilde{X}_N, \tilde{Y}_N)$ is given by

$$\mathbb{P}(\tilde{X}_N = x, \tilde{Y}_N = y) = e^{\tau_N x - \Lambda_{X_N^<}(\tau_N)} \mathbb{P}(X_N^< = x, Y_N^< = y) \tag{33}$$

and τ_N is the unique solution of $\Lambda'_{X_N^<}(\tau_N) = m_N/N$. Then,

$$g_N(u) = -N^{1-\alpha} u H'_N(0) - \frac{u^2}{2} H''_N(0) - \frac{u^3}{6N^{1-\alpha}} H'''_N(z_N)$$

with $z_N \in [0, u/N^{1-\alpha}]$. Using Remark 23 below, one has $N^{1-\alpha} H'_N(0) \to 0$. By Lemma 22 below, by [9, Equation (4.31)], and by (4), we get

$$H''_N(0) \to \frac{\mathrm{Cov}(X,Y)^2}{\mathrm{Var}(X)} - \mathrm{Var}(Y) = -\sigma^2(\mu).$$

As in [8], shortening $\Lambda_{(X_N^<, Y_N^< - \mathbb{E}[Y_N])}$ into Λ and using obvious notation for partial derivatives, one has

$$H'''_N(z_N) = \left(\left(\frac{\Lambda''_{s,t}}{\Lambda''_{s,s}} \right)^3 \Lambda'''_{s,s,s} - 3 \left(\frac{\Lambda''_{s,t}}{\Lambda''_{s,s}} \right)^2 \Lambda'''_{s,s,t} + 3 \frac{\Lambda''_{s,t}}{\Lambda''_{s,s}} \Lambda'''_{s,t,t} - \Lambda'''_{t,t,t} \right) (s_N(z_N), z_N).$$

Let us prove that $s_N(z_N) \to 0$. The sequence of concave functions

$$f_N(s) = s \frac{m_N}{N} - \Lambda_{(X_N^<, Y_N^< - \mathbb{E}[Y_N])}(s, z_N)$$

converges pointwise to the strictly concave function $f(s) := s(1-\mu)^{-1} - \Lambda_X(s)$. This fact follows from the uniform integrability of $\exp(sX_N^< + z_N(Y_N^< - \mathbb{E}[Y_N]))$, which is a consequence of Proposition 17 and the fact that $z_N Y_N^< \leqslant u/N^{1-3\alpha/2}$ is bounded (remember that $\alpha < 2/3$). Now the maximum of f is attained at 0. Let $\varepsilon \in (0, -\log(\lambda e))$. By the strict concavity of f, there exists $\eta > 0$ such that $f(0) - \eta > \max(f(-\varepsilon), f(\varepsilon)) + \eta$. By [26, Theorem 10.8], for all N large enough, $\|f_N - f\|_\infty < \eta$, where $\|\cdot\|_\infty$ is the supremum norm over the compact set $[-\varepsilon, \varepsilon]$. For those N, $s_N(z_N) \in [-\varepsilon, \varepsilon]$. Since ε is arbitrary, we have proved that $s_N(z_N) \to 0$. The uniform integrability of $(X_N^<)^p \left| Y_N^< - \mathbb{E}[Y_N] \right|^q \exp(s_N(z_N) X_N^< + z_N(Y_N^< - \mathbb{E}[Y_N]))$ follows from the same arguments as before and the fact that $s_N(z_N) \to 0$, and Proposition 17 entails

$$\mathbb{E}\Big[(X_N^<)^p \left| Y_N^< - \mathbb{E}[Y_N] \right|^q e^{s_N(z_N) X_N^< + z_N(Y_N^< - \mathbb{E}[Y_N])} \Big] \to \mathbb{E}\Big[X^p \left| Y - \mathbb{E}[Y] \right|^q \Big].$$

Therefore, $H'''_N(z_N)$ is bounded, whence $g_N(u) \to u^2 \sigma^2(\mu)/2$ and (7) follows. □

Lemma 20. *Let $\check{X}_N$ be $X_N^<$ or $\hat{X}_N^u$. Denoting by $\check{S}_N := \check{X}_{N,1} + \ldots \check{X}_{N,N}$, we have, for any sequence of integers $(k_N)_{N \geqslant 1}$ such that $\lim k_N/N \in (1, \infty)$,*

$$\frac{1}{N} \log \mathbb{P}(\check{S}_N = k_N) \sim -\Lambda^*_{\check{X}_N}(k_N/N).$$

Proof of Lemma 20. We just check that we can apply [8, Lemma 3.3] to the sequences $(X_N^<)_{N \geqslant 1}$ and $(\hat{X}_N^u)_{N \geqslant 1}$.

• First, $\mathrm{Im}(\Lambda'_{X_N^<}) = (1, \infty) = \mathrm{Im}(\Lambda'_X)$ so, for all N large enough, $k_N/N \in \mathrm{Im}(\Lambda'_{X_N^<})$ and $\lim k_N/N \in \mathrm{Im}(\Lambda'_X)$. Secondly, $\mathrm{int}(\mathrm{dom}(\Lambda_{X_N^<})) = \mathbb{R}$ and $\mathrm{int}(\mathrm{dom}(\Lambda_X)) = (-\infty, -\log(\lambda e))$, so that assumption 1. holds. Finally, it remains to check that assumption 2. is satisfied. For $s < -\log(\lambda e)$ and $t \in \mathbb{R}$, we have

$$\begin{aligned} &\left| \mathbb{E}\Big[e^{(s+it)X_N^<} \Big] - \mathbb{E}\Big[e^{(s+it)X} \Big] \right| \\ &\leqslant \frac{\left| \mathbb{E}\Big[e^{(s+it)X_N} \Big] - \mathbb{E}\Big[e^{(s+it)X} \Big] \right| + \mathbb{P}(Y_N > N^{1/2}) \mathbb{E}\Big[e^{sX} \Big] + \mathbb{E}\Big[e^{sX_N} \mathbb{1}_{Y_N > N^{1/2}} \Big]}{\mathbb{P}(Y_N \leqslant N^{1/2})}. \end{aligned}$$

The first term of the numerator converges to 0 uniformly in $t \in \mathbb{R}$ by Proposition 17. So does also the second one by Proposition 15. Finally, by the same arguments and Hölder's inequality, with $1/p + 1/q = 1$ and $ps < -\log(\lambda e)$,

$$\mathbb{E}\Big[e^{sX_N}\mathbb{1}_{Y_N > N^{1/2}}\Big] \leqslant \mathbb{E}\Big[e^{psX_N}\Big]^{1/p}\mathbb{P}(Y_N > N^{1/2})^{1/q} \to 0,$$

leading to the required result.
• First, $\mathrm{Im}(\Lambda'_{\hat{X}_N^u}) = (1,\infty) = \mathrm{Im}(\Lambda'_X)$ so, for all N large enough, $k_N/N \in \mathrm{Im}(\Lambda'_{\hat{X}_N^u})$ and $\lim k_N/N \in \mathrm{Im}(\Lambda'_X)$. Secondly, $\mathrm{int}(\mathrm{dom}(\Lambda_{\hat{X}_N^u})) = \mathbb{R}$ and $\mathrm{int}(\mathrm{dom}(\Lambda_X)) = (-\infty, -\log(\lambda e))$, so that assumption 1. holds. Finally, it remains to check that assumption 2. is satisfied. Using the definition of $\hat{X}_N^u$, one gets

$$\mathbb{E}[e^{(s+it)\hat{X}_N^u}] = e^{-\Lambda_{Y_N^<-\mathbb{E}[Y_N]}(u/N^{1-\alpha})}\mathbb{E}\left[e^{(s+it)X_N^< + \frac{u}{N^{1-\alpha}}(Y_N^<-\mathbb{E}[Y_N])}\right]$$

that converges to $\mathbb{E}[e^{(s+it)X}]$ uniformly in $t \in \mathbb{R}$ by similar arguments. □

Lemma 21. *Remind that τ_N is the unique solution of $\Lambda'_{X_N^<}(\tau_N) = m_N/N$. There exists $c > 0$ such that, for all N large enough, $|\tau_N| \leqslant e^{-cN^{\alpha/4}}$.*

Proof of Lemma 21. First, for all $s < -\log(\lambda e)$,

$$\Big|\Lambda'_{X_N^<}(s) - \Lambda'_{X_N}(s)\Big| = \left|\frac{\mathbb{E}\Big[X_N e^{sX_N}\Big]\mathbb{E}\Big[e^{sX_N}\mathbb{1}_{Y_N \geqslant N^{\alpha/2}}\Big]}{\mathbb{E}\Big[e^{sX_N}\Big]\mathbb{E}\Big[e^{sX_N}\mathbb{1}_{Y_N < N^{\alpha/2}}\Big]} - \frac{\mathbb{E}\Big[X_N e^{sX_N}\mathbb{1}_{Y_N \geqslant N^{\alpha/2}}\Big]}{\mathbb{E}\Big[e^{sX_N}\mathbb{1}_{Y_N < N^{\alpha/2}}\Big]}\right| \leqslant e^{-c_1 N^{\alpha/4}} \tag{34}$$

for some constant $c_1 > 0$ (independent of s and N), using Hölder's inequality and Propositions 15 and 17. Now, write

$$\Lambda'_{X_N}(s) = \mathbb{E}[X_N] + s\,\mathrm{Var}(X_N) + \frac{s^2}{2}\Lambda'''_{X_N}(t)$$

with t between 0 and s. Using Proposition 17, there exists $s_0 > 0$ such that, for all $s \in [-s_0, s_0]$ and for all N large enough,

$$\Big|\Lambda'_{X_N}(s) - \mathbb{E}[X_N] - s\,\mathrm{Var}(X_N)\Big| \leqslant \frac{|s|\,\mathrm{Var}(X_N)}{2}. \tag{35}$$

Since $\Lambda'_{X_N^<}(\tau_N) = \mathbb{E}[X_N]$, (34) and (35) yield $|\tau_N| \leqslant 2e^{-c_1 N^{\alpha/4}}/\,\mathrm{Var}(X_N)$, hence the desired result since $\mathrm{Var}(X_N) \to \mathrm{Var}(X)$. □

Lemma 22. *$(X_N^<, Y_N^<)_{N\geqslant 1}$ and $(\tilde{X}_N, \tilde{Y}_N)_{N\geqslant 1}$ converge to (X, Y) in distribution and with all mixed moments of the type $\mathbb{E}[\check{X}_N^p \check{Y}_N^q e^{s\check{X}_N}]$, where $p \geqslant 0$, $q \geqslant 0$, $s < -\log(\lambda e)$, and $\check{X}$ (resp. $\check{Y}$) stands for $X^<$ or $\tilde{X}$ (resp. $Y^<$ or $\tilde{Y}$).*

Proof of Lemma 22. Following the proof of Proposition 17, we prove separately that $(X_N^<)_{N\geqslant 1}$ and $(\tilde{X}_N)_{N\geqslant 1}$ converge to X in distribution and with all moments and $(Y_N^<)_{N\geqslant 1}$ and $(\tilde{Y}_N)_{N\geqslant 1}$ converge to Y in distribution and with all moments.
The convergence of $(X_N^<)_{N\geqslant 1}$ is a consequence of the proof of Proposition 18. However we give a straightforward proof on which the sequel relies. Let $f\colon \mathbb{R} \mapsto \mathbb{R}$ be a bounded measurable function. By Proposition 17, it suffices to prove that $\mathbb{E}[f(X_N^<)] - \mathbb{E}[f(X_N)]|$ converges to 0. One has

$$|\mathbb{E}[f(X_N^<)] - \mathbb{E}[f(X_N)]| \leqslant (|\mathbb{E}[f(X_N)]| + \mathbb{E}[f(X_N)^2]^{1/2})\frac{\mathbb{P}(Y_N \geqslant N^{\alpha/2})}{\mathbb{P}(Y_N < N^{\alpha/2})}.$$

The result follows from the fact that f is bounded and from Proposition 15. To prove the convergence of the moments of $(X_N^<)_{N\geqslant 1}$ to those of X, it suffices to show that, for all $r>0$, $(\mathbb{E}[(X_N^<)^r])_{N\geqslant 1}$ is bounded (see, e.g., [27, Example 2.21]) and to conclude by uniform integrability. We have

$$\mathbb{E}[(X_N^<)^r] = \mathbb{E}[X_N^r \,|\, Y_N < N^{1/2}] \leqslant \frac{\mathbb{E}[X_N^r]}{\mathbb{P}(Y_N < N^{1/2})},$$

so $\limsup \mathbb{E}[(X_N^<)^r]$ is finite by Propositions 15 and 17. The same calculation leads to the convergence of $(\mathbb{E}[\exp(sX_N^<)])_{N\geqslant 1}$ and the same lines yield the convergence in distribution of $(Y_N^<)_{N\geqslant 1}$ to Y and with all moments.
Let us consider $(\tilde{X}_N)_{N\geqslant 1}$. Let $f\colon \mathbb{R}\mapsto\mathbb{R}$ be a bounded measurable function. One has

$$|\mathbb{E}[f(\tilde{X}_N)] - \mathbb{E}[f(X_N^<)]| \leqslant \frac{\left|\mathbb{E}\left[f(X_N^<)\left(e^{\tau_N X_N^<} - \mathbb{E}\left[e^{\tau_N X_N^<}\right]\right)\right]\right|}{\mathbb{E}\left[e^{\tau_N X_N^<}\right]} \leqslant \|f\|_\infty \frac{\mathbb{E}\left[\left|e^{\tau_N X_N^<} - \mathbb{E}\left[e^{\tau_N X_N^<}\right]\right|\right]}{\mathbb{E}\left[e^{\tau_N X_N^<}\right]}. \tag{36}$$

Using the convergence of $(\mathbb{E}[\exp(sX_N^<)])_{N\geqslant 1}$ together with Lemma 21, Slutsky's Lemma, and uniform integrability, we get $\mathbb{E}\left[e^{\tau_N X_N^<}\right] \to 1$. Similarly, since

$$\left|e^{\tau_N X_N^<} - \mathbb{E}\left[e^{\tau_N X_N^<}\right]\right|^2 = e^{2\tau_N X_N^<} - 2e^{\tau_N X_N^<}\mathbb{E}\left[e^{\tau_N X_N^<}\right] + \mathbb{E}\left[e^{\tau_N X_N^<}\right]^2,$$

the numerator in the right-hand side of (36) converges to 0. Hence $(\tilde{X}_N)_{N\geqslant 1}$ converges in distribution to X and similar arguments show the convergence of all moments of $(\tilde{X}_N)_{N\geqslant 1}$. The same lines lead to the convergence in distribution of $(\tilde{Y}_N)_{N\geqslant 1}$ to Y and with all moments. □

Remark 23. Following the same lines as in the proof of Lemma 22 for $f = \mathrm{id}_{\mathbb{R}}$ and using Proposition 17 instead of the fact that f is bounded and the fact that τ_N converges exponentially rapidly to 0 (by Lemma 21), we get that $\mathbb{E}[\tilde{Y}_N] - \mathbb{E}[Y_N]$ converges exponentially rapidly to 0.

We notice that the rate function in the upper moderate deviations $y^2/(2\sigma^2(\mu))$ in (7) depends on the conditioning on $\{S_N = m_N\}$. As opposed to this, the conditioning does not influence the expression of the rate function in the upper large deviations, due to the fact that the random variables Y_N are heavy-tailed. As a consequence, our proof of (9) in Section 5.5 does not mimic that of [8, Theorem 2.1] but is rather inspired by that of [20, Theorem 5].

5.4 Upper intermediate deviations - Theorem 9 (ii)

The upper bound comes from (31) and Propositions 18 and 19 (ii). Let us turn to the lower bound. We assume that the infimum in the right-hand side of (22) is attained at δ_0. Let $z>0$ and $l_N := \lceil (N^\alpha z/\delta_0)^{1/2}\rceil$. By (30), we have

$$\begin{aligned}
P_N &\geqslant \mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y_N,\ S_N = m_N)\\
&\geqslant \mathbb{P}\Big(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y_N,\ S_N = m_N,\ Y_{N,N} - \mathbb{E}[Y_{N,N}] \geqslant N^\alpha z\Big)\\
&\geqslant \mathbb{P}(T_{N,N-1} - \mathbb{E}[T_{N,N-1}] \geqslant N^\alpha(y_N - z) | S_{N,N-1} = m_N - l_N)\mathbb{P}(S_{N,N-1} = m_N - l_N)\\
&\qquad \mathbb{P}(Y_{N,N} - \mathbb{E}[Y_{N,N}] \geqslant N^\alpha z,\ X_{N,N} = l_N).
\end{aligned}$$

By Proposition 13,

$$\liminf \frac{1}{N^{\alpha/2}} \log \mathbb{P}(Y_N - \mathbb{E}[Y_N] \geqslant N^\alpha z,\ X_N = l_N) = -q(\mu) z^{1/2}$$

and by Proposition 18,

$$\liminf \frac{1}{N^{\alpha/2}} \log \mathbb{P}(S_{N,N-1} = m_N - l_N) \to 0.$$

As for the minoration of the remaining term, we follow the same lines as in the proof of the lower bound in Theorem 9 (i) which remains valid for $\alpha = 2/3$ and $\mathbb{E}[S_{N,N-1}] = m_N + O(N^{1/2})$. Hence,

$$\liminf \frac{1}{N^{\alpha/2}} \log \mathbb{P}(T_{N,N-1} - \mathbb{E}[T_{N,N-1}] \geqslant N^\alpha (y_N - z) \,|\, S_{N,N-1} = m_N - l_N) = -\frac{(y-z)^2}{2\sigma^2(\mu)}.$$

Optimizing in $z = ty$ with $t \in (0,1)$ leads to (8). □

5.5 Upper large deviations - Theorem 9 (iii)

Using (31), the upper bound

$$\limsup_{N\to\infty} \frac{1}{N^{\alpha/2}} \log(P_N) \leqslant \limsup_{N\to\infty} \frac{1}{N^{\alpha/2}} \log\left(\frac{\mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y_N)}{\mathbb{P}(S_N = m_N)}\right) = -q(\mu) y^{1/2},$$

follows from Propositions 15, 18, and 19 (iii). For the lower bound, assume that the infimum in the right-hand side of (22) is attained at δ_0. Let $\varepsilon > 0$ and $l_N := \lceil (N^\alpha(y_N + \varepsilon)/\delta_0)^{1/2} \rceil$. By (28), we have

$$\begin{aligned}
P_N &\geqslant \mathbb{P}(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y_N,\ S_N = m_N) \\
&\geqslant \mathbb{P}\Big(T_N - \mathbb{E}[T_N] \geqslant N^\alpha y_N,\ S_N = m_N,\ Y_{N,N} - \mathbb{E}[Y_{N,N}] \geqslant N^\alpha (y_N + \varepsilon)\Big) \\
&\geqslant \mathbb{P}\Big(T_{N,N-1} - \mathbb{E}[T_{N,N-1}] \geqslant -N^\alpha \varepsilon,\ S_{N,N-1} = m_N - l_N\Big) \\
&\qquad\qquad \mathbb{P}\Big(Y_N - \mathbb{E}[Y_N] \geqslant N^\alpha (y_N + \varepsilon),\ X_N = l_N\Big) \\
&=: P_{N,1} P_{N,2}.
\end{aligned}$$

Applying Proposition 13, one gets

$$\lim_{N\to\infty} \frac{1}{N^{\alpha/2}} \log(P_{N,2}) = -\sqrt{\frac{y+\varepsilon}{\delta_0}}\,(\kappa(\mu) + J(\delta_0)) \underset{\varepsilon\to 0}{\to} -q(\mu) y^{1/2}. \tag{37}$$

Let us turn to the minoration of $P_{N,1}$. One has

$$P_{N,1} \geqslant \mathbb{P}(S_{N,N-1} = m_N - l_N) - \mathbb{P}(T_{N,N-1} - \mathbb{E}[T_{N,N-1}] < -N^\alpha \varepsilon).$$

Applying Proposition 18, we derive that

$$\frac{1}{N^{\alpha/2}} \log \mathbb{P}(S_{N,N-1} = m_N - l_N) \sim -N^{1-\alpha/2} \Lambda^*_{X_N}\Big(\frac{m_N - l_N}{N}\Big) \sim -\frac{N^{1-\alpha/2}}{2\sigma_X^2}\Big(\frac{l_N}{N}\Big)^2 = O\Big(\frac{1}{N^{1-\alpha/2}}\Big)$$

which converges to 0 since $\alpha < 2$. Applying a unilateral version of [8, Theorem 2.2], we get

$$\frac{1}{N^{\alpha/2}} \log \mathbb{P}(T_{N,N-1} - \mathbb{E}[T_{N,N-1}] < -N^\alpha \varepsilon) \begin{cases} \sim -c_\varepsilon N^{3\alpha/2-1} \to -\infty & \text{if } \alpha \in (2/3, 1] \\ = -\infty & \text{if } \alpha > 1 \end{cases}$$

for some $c_\varepsilon > 0$. Eventually, $N^{-\alpha/2} \log(P_{N,1}) \to 0$, which leads, together with (37), to

$$\liminf \frac{1}{N^{\alpha/2}} \log(P_N) \geqslant -q(\mu) y^{1/2}.$$ □

5.6 Upper large deviations for $\alpha = 2$ - Theorem 9 (iv)

By (30) and Proposition 18, making the change of variable $c = (y/\delta)^{1/2}$ in the infimum, and setting $x = (1-\mu)^{-1}$, it suffices to prove that

$$\begin{aligned} &\frac{1}{N}\log \mathbb{P}(S_N = m_N,\ T_N \geqslant N^2 y) \\ &\qquad \to \begin{cases} -\inf\limits_{c>0}\Big[c(\kappa(\mu) + J(y/c^2)) + \Lambda_0^*(x-c)\Big] & \text{if } y < (x-1)^2/2 \\ -\infty & \text{if } y \geqslant (x-1)^2/2 \end{cases} \end{aligned} \tag{38}$$

to establish (10). Assume that (38) holds for $y < (x-1)^2/2$. Observe that the probability in the left-hand side of (38) is decreasing in y. Moreover, if $y < (x-1)^2/2)$, then

$$\begin{aligned} \inf_{c>0}\Big[c(\kappa(\mu) + J(y/c^2)) + \Lambda_0^*(x-c)\Big] &= \inf_{\sqrt{2y}<c\leqslant x-1}\Big[c(\kappa(\mu) + J(y/c^2)) + \Lambda_0^*(x-c)\Big] \\ &\geqslant \inf_{\sqrt{2y}<c\leqslant x-1} cJ(y/c^2) \\ &\geqslant \sqrt{2y}J(y/(x-1)^2) \\ &\to \infty \quad \text{as } y \to (x-1)^2/2. \end{aligned}$$

So (38) holds for $y \geqslant (x-1)^2/2$.

Proof of (38) – *Lower bound for* $y < (x-1)^2/2$. Let $c > 0$ and $l_N = \lfloor cN \rfloor$. Let also $\varepsilon > 0$. We have

$$\begin{aligned} &\frac{1}{N}\log \mathbb{P}(S_N = m_N,\ T_N \geqslant N^2 y) \\ &\qquad \geqslant \frac{1}{N}\log \mathbb{P}(X_N = l_N, Y_N \geqslant N^2 y) + \frac{1}{N}\log \mathbb{P}(S_{N,N-1} = m_N - l_N) \\ &\qquad \to -c(\kappa(\mu) + J(y/c^2)) - \Lambda_0^*(x-c), \end{aligned}$$

by Propositions 13 and 18. Taking the supremum in $c > 0$ yields the desired lower bound. □

Proof of (38) – *Upper bound for* $y < (x-1)^2/2$. Let us write

$$\mathbb{P}(S_N = m_N,\ T_N \geqslant N^2 y) = P_{N,0} + P_{N,1}$$

where

$$P_{N,0} = \mathbb{P}(S_N = m_N,\ T_N \geqslant N^2 y,\ \forall i \in [\![1, N]\!] \quad Y_{N,i} \leqslant N^2 y)$$

and

$$P_{N,1} = \mathbb{P}(S_N = m_N,\ T_N \geqslant N^2 y,\ \exists i \in [\![1, N]\!] \quad Y_{N,i} > N^2 y).$$

Behavior of $P_{N,0}$ Let us apply the exponential version of Chebyshev's inequality. Let $(s,t) \in (-\infty, 0] \times [0, +\infty)$. We have

$$\begin{aligned} P_{N,0} &\leqslant \mathbb{E}[\mathbb{1}_{S_N - m_N \leqslant 0}\ \mathbb{1}_{T_N/N - Ny \geqslant 0}\ \mathbb{1}_{\forall i \in [\![1,N]\!] \quad Y_{N,i} \leqslant N^2 y}] \\ &\leqslant \mathbb{E}[e^{s(S_N - m_N) + t(T_N/N - Ny)}\ \mathbb{1}_{\forall i \in [\![1,N]\!] \quad Y_{N,i} \leqslant N^2 y}] \\ &= e^{-N(sm_N/N + ty)}\mathbb{E}[e^{sX_N + tY_N/N}\ \mathbb{1}_{Y_N \leqslant N^2 y}]^N. \end{aligned} \tag{39}$$

Let us write

$$\begin{aligned} \mathbb{E}[e^{sX_N + tY_N/N}\ \mathbb{1}_{Y_N \leqslant N^2 y}] &= \mathbb{E}[e^{sX_N + tY_N/N}\ \mathbb{1}_{Y_N \leqslant N^{1/2}}] + \mathbb{E}[e^{sX_N + tY_N/N}\ \mathbb{1}_{N^{1/2} < Y_N \leqslant N^2 y}] \\ &=: E_1 + E_2. \end{aligned} \tag{40}$$

First, by uniform integrability (see Proposition 17),

$$E_1 \to \mathbb{E}[e^{sX}]. \tag{41}$$

Secondly, remembering that $\mathbb{P}(X_N = l,\ Y_N = p) = 0$ if $l(l-1)/2 < p$,

$$\begin{aligned} E_2 &= \mathbb{E}[e^{sX_N + tY_N/N}\ \mathbb{1}_{N^{1/2} < Y_N \leqslant N^2 y}] \\ &= \sum_{\substack{l > N^{1/4} \\ N^{1/2} < p \leqslant N^2 y}} e^{sl + tp/N} \mathbb{P}(X_N = l,\ Y_N = p) \\ &= \sum_{l > N^{1/4}} e^{sl} \Bigg(\sum_{N^{1/2} < p \leqslant N^2 y} e^{tp/N}(1 - e^{-t/N}) \mathbb{P}(X_N = l,\ Y_N \geqslant p) \\ &\qquad + \mathbb{P}(X_N = l,\ Y_N > N^{1/2}) e^{t/N^{1/2}} - \mathbb{P}(X_N = l,\ Y_N > N^2 y) e^{tNy} \Bigg), \end{aligned} \tag{42}$$

after a summation by parts. By uniform integrability (see Proposition 17),

$$\sum_{l > N^{1/4}} e^{sl} \mathbb{P}(X_N = l,\ Y_N > N^{1/2}) e^{t/N^{1/2}} = \mathbb{E}[e^{sX_N + t/N^{1/2}} \mathbb{1}_{Y_N > N^{1/2}}] \to 0. \tag{43}$$

Lemma 24. *The function $K\colon \delta \in [0, 1/2) \mapsto \delta^{-1/2} J(\delta)$ is increasing, convex, and $K(\delta) \to \infty$ as $\delta \to 1/2$.*

For all $\varepsilon \in (0, 1/2)$, we introduce the function

$$\begin{aligned} J_\varepsilon\colon\ & [0, \infty) \to \mathbb{R} \\ & \delta \mapsto \begin{cases} J(\delta) & \text{if } \delta \leqslant 1/2 - \varepsilon \\ \delta^{1/2}[K(1/2 - \varepsilon) + (\delta - 1/2 + \varepsilon) K'(1/2 - \varepsilon)] & \text{if } 1/2 - \varepsilon < \delta \leqslant 1/2 \\ \infty & \text{if } \delta > 1/2. \end{cases} \end{aligned}$$

The following lemma is a straightforward consequence of Lemma 24.

Lemma 25. *The function J_ε is non decreasing and less than J. Moreover, the function $\delta \in [0, 1/2] \mapsto \delta^{-1/2} J_\varepsilon(\delta)$ is a bounded convex function.*

Let $\varepsilon \in (0, \kappa(\mu) \wedge 1/2)$. As a consequence of (42), (43), and Lemmas 14 and 25, we get, for any $a > 0$,

$$\begin{aligned} \limsup E_2 &\leqslant \limsup \sum_{l > N^{1/4}} e^{sl} \sum_{N^{1/2} < p \leqslant N^2 y} e^{tp/N}(1 - e^{-t/N}) e^{-l(\kappa(\mu) + J_\varepsilon(p/l^2) - \varepsilon)} \\ &\leqslant \limsup \Bigg(\sum_{\substack{N^{1/4} < l \leqslant Na \\ N^{1/2} < p \leqslant N^2 y}} e^{l(s - \kappa(\mu) - J_\varepsilon(p/l^2) + \varepsilon) + tp/N} + N^2 y \sum_{l > Na} e^{-(\kappa(\mu) - \varepsilon) l + tNb} \Bigg) \\ &\leqslant \limsup N^3 a y \cdot \exp\Bigg(N \cdot \max_{\substack{N^{1/4} < l \leqslant Na \\ N^{1/2} < p \leqslant N^2 y}} \Big[\frac{l}{N} \Big(s - \kappa(\mu) - J_\varepsilon\Big(\frac{p}{l^2}\Big) + \varepsilon \Big) + \frac{tp}{N^2} \Big] \Bigg), \end{aligned} \tag{44}$$

as soon as $(\kappa(\mu) - \varepsilon) a > ty$. Now,

$$\begin{aligned} &\max_{\substack{N^{1/4} < l \leqslant Na \\ N^{1/2} < p \leqslant N^2 y}} \Big[\frac{l}{N} \Big(s - \kappa(\mu) - J_\varepsilon\Big(\frac{p}{l^2}\Big) + \varepsilon \Big) + \frac{tp}{N^2} \Big] \\ &\qquad \leqslant \sup_{\substack{0 < \delta < 1/2 \\ N^{-3/4} < u \leqslant \sqrt{y\delta^{-1}}}} \Big[u(s - \kappa(\mu) - J_\varepsilon(\delta) + \varepsilon) + t\delta u^2 \Big] =: S. \end{aligned} \tag{45}$$

As soon as

$$\sup_{0<\delta<1/2}\Big[\sqrt{b\delta^{-1}}(s-\kappa(\mu)-J_\varepsilon(\delta)+\varepsilon)+ty\Big]<0,$$

i.e. as soon as

$$t<\inf_{0<\delta<1/2}\frac{\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon-s}{\sqrt{\delta y}}, \tag{46}$$

and, for all N large enough,

$$S=\sup_{0<\delta<1/2}\Big[N^{-3/4}(s-\kappa(\mu)-J_\varepsilon(\delta)+\varepsilon)+t\delta N^{-3/2}\Big]\leqslant-\frac{(\kappa(\mu)-\varepsilon)N^{-3/4}}{2}. \tag{47}$$

Therefore, under (46), using (44), (45), and (47),

$$\limsup E_2\leqslant\limsup N^3 ay\cdot\exp\Big(-\frac{(\kappa(\mu)-\varepsilon)N^{1/4}}{2}\Big)=0. \tag{48}$$

Combining (39), (40), (41), and (48), we get

$$\begin{aligned}\limsup\frac{1}{N}\log P_{N,0}&\leqslant\inf\Big\{\Lambda_0(s)-sx-ty\ ;\ s\leqslant 0,\ t<\inf_{0<\delta<1/2}(\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon-s)(\delta y)^{-1/2}\Big\}\\&=\inf_{s\leqslant 0}\sup_{0<\delta<1/2}\Lambda_0(s)-sx-(\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon-s)(y/\delta)^{1/2}\\&=:M'_\varepsilon.\end{aligned}$$

Lemma 26. *The function*

$$\begin{aligned}f\colon\ (0,1/2)\times(-\infty,0]&\to\mathbb{R}\\(\delta,s)&\mapsto\Lambda_0(s)-sx-(\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon-s)(y/\delta)^{1/2}\end{aligned}$$

is concave in δ and convex in s. Moreover, $f(\delta_0,\cdot)$ is bounded from below for some $\delta_0\in(0,1/2)$.

Thanks to Lemma 26, the minimax theorem of [23] applies and yields

$$\begin{aligned}M'_\varepsilon&=\sup_{0<\delta<1/2}\inf_{s\leqslant 0}\Big[\Lambda_0(s)-sx-(\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon-s)(y/\delta)^{1/2}\Big]\\&=-\inf_{0<\delta<1/2}\Big[(y/\delta)^{1/2}(\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon)+\Lambda_0^*\Big(x-(y/\delta)^{1/2}\Big)\Big],\end{aligned}$$

since $x-(y/\delta)^{1/2}\leqslant x-\sqrt{2y}\leqslant(1-\mu)^{-1}$. Notice that

$$\begin{aligned}&\inf_{1/2-\varepsilon<\delta<1/2}\Big[(y/\delta)^{1/2}(\kappa(\mu)+J_\varepsilon(\delta)-\varepsilon)+\Lambda_0^*\Big(x-(y/\delta)^{1/2}\Big)\Big]\\&\qquad\geqslant\sqrt{2y}(\kappa(\mu)-\varepsilon)+\sqrt{y}K(1/2-\varepsilon)+\Lambda_0^*\Big(x-\sqrt{2y}\Big)\xrightarrow[\varepsilon\to 0]{}\infty\end{aligned}$$

by Lemma 24. A fortiori, since $J_\varepsilon\leqslant J$,

$$\inf_{1/2-\varepsilon<\delta<1/2}\Big[(y/\delta)^{1/2}(\kappa(\mu)+J(\delta)-\varepsilon)+\Lambda_0^*\Big(x-(y/\delta)^{1/2}\Big)\Big]\xrightarrow[\varepsilon\to 0]{}\infty.$$

So, if ε is small enough, i.e. $\varepsilon\in(0,\varepsilon_0)$,

$$M'_\varepsilon=-\inf_{0<\delta<1/2}\Big[(y/\delta)^{1/2}(\kappa(\mu)+J(\delta)-\varepsilon)+\Lambda_0^*\Big(x-(y/\delta)^{1/2}\Big)\Big].$$

Finally, again applying the minimax theorem of [23], we get

$$\begin{aligned}\inf_{0<\varepsilon<\varepsilon_0} M'_\varepsilon &= -\sup_{0<\varepsilon<\varepsilon_0}\inf_{0<\delta<1/2}\Big[(y/\delta)^{1/2}(\kappa(\mu)+J(\delta)-\varepsilon)+\Lambda_0^*\Big(x-(y/\delta)^{1/2}\Big)\Big]\\ &= -\inf_{0<\delta<1/2}\Big[(y/\delta)^{1/2}(\kappa(\mu)+J(\delta))+\Lambda_0^*\Big(x-(y/\delta)^{1/2}\Big)\Big]\\ &= -\inf_{\sqrt{2y}<c\leqslant x-1} c(\kappa(\mu)+J(y/c^2))+\Lambda_0^*(x-c),\end{aligned}$$

since $\Lambda_0^*(x-(y/\delta)^{1/2})=\infty$ for $\delta < y/(x-1)^2$.

Behavior of $P_{N,1}$ Let $\varepsilon>0$. We have

$$\begin{aligned}&P_{N,1}=\mathbb{P}(S_N=m_N,\ T_N\geqslant N^2y,\ \exists i\in[\![1,N]\!]\quad Y_{N,i}>N^2y)\\ &\leqslant N\mathbb{P}(S_N=m_N,\ Y_{N,N}>N^2y)\\ &\leqslant N\sum_{l=K}^{L}\mathbb{P}\Big(m_N-Nl\varepsilon\leqslant S_{N,N-1}<m_N-N(l-1)\varepsilon\Big)\mathbb{P}(N(l-1)\varepsilon<X_N\leqslant Nl\varepsilon,\ Y_N>N^2y)\\ &\leqslant N\sum_{l=K}^{L}P_{N,1,l},\end{aligned}$$

where

$$P_{N,1,l}:=\mathbb{P}(S_{N,N-1}<m_N-N(l-1)\varepsilon)\mathbb{P}(N(l-1)\varepsilon<X_N\leqslant Nl\varepsilon,\ Y_N>N^2y),$$

K is the smallest integer such that $(K\varepsilon)^2/2\geqslant y$ (remember that $Y_N\leqslant X_N^2/2$), and L is a fixed integer such that, for all $N\geqslant 1$, $m_N/N-(L-1)\varepsilon\leqslant 0$ (remember that $m_N/N\to x$). Let us evaluate each term. For all $l\in[\![K,L]\!]$,

$$\limsup\frac{1}{N}\log P_{N,1,l}\leqslant -\Lambda_0^*(x-(l-1)\varepsilon)-l\varepsilon\Big(\kappa(\mu)+J\Big(\frac{y}{(l\varepsilon)^2}\Big)\Big).$$

applying the exponential version of Chebyshev's inequality, Proposition 13, and the principle of the largest term. Applying one more time the principle of the largest term, we get:

$$\limsup\frac{1}{N}\log P_{N,1}\leqslant -\inf_{c>\sqrt{2y}}\Big[\Lambda_0^*(x-c+\varepsilon)+c\Big(\kappa(\mu)+J\Big(\frac{y}{c^2}\Big)\Big)\Big],$$

and we get the desired upper bound for $P_{N,1}$ when $\varepsilon\to 0$, since $\Lambda_0^*(x-c+\varepsilon)$ converges uniformly in $c\in(\sqrt{2y},x-1]$ to $\Lambda_0^*(x-c)$ and since

$$\begin{aligned}\inf_{x-1<c\leqslant x-1+\varepsilon}\Big[\Lambda_0^*(x-c+\varepsilon)+c\Big(\kappa(\mu)+J\Big(\frac{y}{c^2}\Big)\Big)\Big]&\geqslant\Lambda_0^*(1+\varepsilon)+\inf_{x-1<c\leqslant x-1+\varepsilon}c\Big(\kappa(\mu)+J\Big(\frac{y}{c^2}\Big)\Big)\\ &\xrightarrow[\varepsilon\to 0]{}\Lambda_0^*(1)+(x-1)\Big(\kappa(\mu)+J\Big(\frac{y}{(x-1)^2}\Big)\Big).\end{aligned}$$

□

It remains to prove Lemmas 24 and 26. Let us define the function

$$\begin{aligned}\delta\colon\ (-\infty,0]&\to[0,1/2)\\ \lambda&\mapsto\begin{cases}0 & \text{if } \lambda=0\\ \frac{1}{\lambda}+\frac{1}{1-e^\lambda}-\frac{1}{2} & \text{if } \lambda\in(-\infty,0).\end{cases}\end{aligned}$$

An easy computation shows that the function δ is a smooth nonnegative and concave decreasing bijection. The function $\lambda\colon [0,1/2) \to (-\infty, 0]$ defined by (2) is the inverse bijection of δ and thus is a smooth decreasing function. Now let us introduce the functions

$$\begin{aligned} H\colon (-\infty,0] &\to [0,+\infty) \\ \lambda &\mapsto \lambda \cdot \Big(\frac{1}{2} - \delta(\lambda)\Big) + \log\Big(1 - \lambda \cdot \Big(\frac{1}{2} + \delta(\lambda)\Big)\Big) \end{aligned}$$

and

$$\begin{aligned} F\colon (-\infty,0] &\to [0,+\infty) \\ \lambda &\mapsto \begin{cases} 0 & \text{if } \lambda = 0 \\ \delta(\lambda)^{-1/2} H(\lambda) & \text{if } \lambda \in (-\infty, 0). \end{cases} \end{aligned}$$

Proof of Lemma 24. The fact that $K(\delta) \to \infty$ as $\delta \to 1/2$ follows from the already mentioned fact that $J(\delta) \to \infty$ as $\delta \to 1/2$. We want to prove that $\delta \mapsto K(\delta) = \delta^{-1/2} J(\delta) = F(\lambda(\delta))$ is an increasing convex function. Since $K(\delta) \sim 6\delta^{3/2}$ as $\delta \to 0$, $K'(0) = 0$, so it only remains to prove that $K'' > 0$ over $(0, 1/2)$. Using the expressions of the first and second derivatives of the inverse function λ of δ, one gets:

$$\begin{aligned} K''(\delta) &= \frac{d^2}{d\delta^2} F(\lambda(\delta)) \\ &= \lambda''(\delta) F'(\lambda(\delta)) + (\lambda'(\delta))^2 F''(\lambda(\delta)) \\ &= \frac{1}{\delta'(\lambda(\delta))} \left(-\frac{\delta''(\lambda(\delta))}{\delta'(\lambda(\delta))^2} F'(\lambda(\delta)) + \frac{1}{\delta'(\lambda(\delta))} F''(\lambda(\delta)) \right) \\ &= \frac{1}{\delta'(\lambda(\delta))} \left(\frac{F'}{\delta'} \right)' (\lambda(\delta)). \end{aligned}$$

Hence, since $\delta' < 0$, our study reduces to show that F'/δ' is a decreasing function over $(-\infty, 0)$. Now,

$$\frac{F'(\lambda)}{\delta'(\lambda)} = -\delta(\lambda)^{1/2} \cdot \frac{\lambda\delta(\lambda) + \frac{1}{2} H(\lambda)}{\delta(\lambda)^2} =: -\delta(\lambda)^{1/2} f(\lambda).$$

The identity $H'(\lambda) = -\lambda\delta'(\lambda)$ entails

$$f'(\lambda) = \frac{\delta(\lambda)^2 - \frac{3}{2}\lambda\delta'(\lambda)\delta(\lambda) - H(\lambda)\delta'(\lambda)}{\delta(\lambda)^3} =: \frac{k(\lambda)}{\delta(\lambda)^3}.$$

Differentiating the function k, we get

$$k'(\lambda) = \frac{1}{2}\delta'(\lambda)\Big(\delta(\lambda) - \lambda\delta'(\lambda)\Big) - \delta''(\lambda)\Big(H(\lambda) + \frac{3}{2}\lambda\delta(\lambda)\Big).$$

On the one hand, $H(\lambda) + \frac{3}{2}\lambda\delta(\lambda) < 0$, because

$$\frac{d}{d\lambda}\Big(H(\lambda) + \frac{3}{2}\lambda\delta(\lambda)\Big) = \frac{e^{2\lambda}(-3\lambda + 4) + e^{\lambda}(2\lambda^2 - 8) + 3\lambda + 4}{4\lambda(1 - e^{\lambda})^2},$$

the sign of which is easy to find (by differentiating several times). On the other hand, $\delta(\lambda) - \lambda\delta'(\lambda) > 0$, because

$$\delta(\lambda) - \lambda\delta'(\lambda) = \frac{e^{2\lambda}(-\lambda + 4) + e^{\lambda}(-2\lambda^2 - 8) + \lambda + 4}{2\lambda(1 - e^{\lambda})^2},$$

the sign of which is easy to find (similarly). Henceforth, $k' < 0$ and $k > 0$ over $(-\infty, 0)$ (k is decreasing on $(-\infty, 0]$ and $k(0) = 0$). Finally, f is an increasing and nonpositive function. Together with the fact that $\lambda \mapsto -\delta(\lambda)^{1/2}$ is also an increasing and nonpositive function, we finally get that F'/δ' is a decreasing function. □

Proof of Lemma 26. The concavity of $f(\cdot, s)$ follows from Lemma 25 and the fact that $\kappa(\mu) - \varepsilon - s \geqslant 0$ and $y \geqslant 0$. The convexity of $f(\delta, \cdot)$ follows from the convexity of Λ_0. Finally, since $y < (x-1)^2/2$, one can choose $\delta_0 \in (0, 1/2)$ such that $x - (y/\delta_0)^{1/2} > 1$. Using the definition of the Fenchel-Legendre transform and Cramér's theorem for the random variable $X \geqslant 1$,

$$\Lambda_0(s) \geqslant s - \Lambda_0^*(1) = s + \log \mathbb{P}(X = 1),$$

so the function $f(\delta_0, \cdot)$ is bounded from below. □

References

[1] M. S. Bartlett. The Characteristic Function of a Conditional Statistic. *J. London Math. Soc.*, S1-13(1):62, 1938.

[2] P. Chassaing and P. Flajolet. Hachage, arbres, chemins & graphes. *Gazette des Mathématiciens*, 95:29–49, 01 2003.

[3] P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms*, 21(1):76–119, 2002.

[4] P. Chassaing and J.-F. Marckert. Parking functions, empirical processes, and the width of rooted labeled trees. *Electron. J. Combin.*, 8(1):Research Paper 14, 19, 2001.

[5] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264 of *Graduate Texts in Mathematics*. Springer, London, 2013.

[6] P. Flajolet, P. Poblete, and A. Viola. On the analysis of linear probing hashing. *Algorithmica*, 22(4):490–515, 1998. Average-case analysis of algorithms.

[7] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, New York, NY, USA, 1 edition, 2009.

[8] F. Gamboa, T. Klein, and C. Prieur. Conditional large and moderate deviations for sums of discrete random variables. Combinatoric applications. *Bernoulli*, 18(4):1341–1360, 2012.

[9] S. Janson. Asymptotic distribution for the cost of linear probing hashing. *Random Structures Algorithms*, 19(3-4):438–471, 2001. Analysis of algorithms (Krynica Morska, 2000).

[10] S. Janson. Moment convergence in conditional limit theorems. *J. Appl. Probab.*, 38(2):421–437, 2001.

[11] S. Janson. Individual displacements for linear probing hashing with different insertion policies. *ACM Trans. Algorithms*, 1(2):177–213, 2005.

[12] T. Klein, A. Lagnoux, and P. Petit. A conditional Berry–Esseen inequality. *Journal of Applied Probability*, 56(1):76–90, 2019.

[13] T. Klein, A. Lagnoux, and P. Petit. Large deviation results for triangular arrays of semiexponential random variables. Preprint, 2020.

[14] D. Knuth. Notes on "open" addressing, 1963.

[15] D. E. Knuth. Computer science and its relation to mathematics. *Amer. Math. Monthly*, 81:323–343, 1974.

[16] D. E. Knuth. *The art of computer programming. Vol. 3.* Addison-Wesley, Reading, MA, 1998. Sorting and searching, Second edition [of MR0445948].

[17] J. V. Linnik. On the probability of large deviations for the sums of independent variables. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II*, pages 289–306. Univ. California Press, Berkeley, Calif., 1961.

[18] J.-F. Marckert. Parking with density. *Random Structures Algorithms*, 18(4):364–380, 2001.

[19] T. Mikosch and A. V. Nagaev. Large deviations of heavy-tailed sums with applications in insurance. *Extremes*, 1(1):81–110, 1998.

[20] A. Nagaev. Integral Limit Theorems Taking Large Deviations into Account when Cramér's Condition Does Not Hold. I. *Theory of Probability and Its Applications*, 14(1):51–64, 1969.

[21] A. Nagaev. Integral Limit Theorems Taking Large Deviations Into Account When Cramér's Condition Does Not Hold. II. *Theory of Probability and Its Applications*, 14(2):193–208, 1969.

[22] S. V. Nagaev. Large deviations of sums of independent random variables. *The Annals of Probability*, pages 745–789, 1979.

[23] V. Perchet and G. Vigeral. A minmax theorem for concave-convex mappings with no regularity assumptions. *Journal of Convex Analysis*, 22(2):537–540, 2015. 4 pages.

[24] V. V. Petrov. Generalization of Cramér's limit theorem. *Uspehi Matem. Nauk (N.S.)*, 9(4(62)):195–202, 1954.

[25] D. Plachky and J. Steinebach. A theorem about probabilities of large deviations with an application to queuing theory. *Period. Math. Hungar.*, 6(4):343–345, 1975.

[26] R. T. Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1970.

[27] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[28] L. M. Wu. Large deviations, moderate deviations and LIL for empirical processes. *Ann. Probab.*, 22(1):17–27, 1994.