Some large deviations in Kingman's coalescent[*]

Andrej Depperschmidt[†]   Peter Pfaffelhuber[‡]   Annika Scheuringer[†]

Abstract

Kingman's coalescent is a random tree that arises from classical population genetic models such as the Moran model. Concerning the structure of the tree-top there are two well-known laws of large numbers: (i) The (shortest) distance, denoted by T_n, from the tree-top to the level when there are n lines in the tree satisfies $nT_n \xrightarrow{n \to \infty} 2$ almost surely; (ii) At time T_n, the population is naturally partitioned in exactly n families where individuals belong to the same family if they have a common ancestor at time T_n in the past. If $F_{i,n}$ denotes the relative size of the ith family, then $n(F_{1,n}^2 + \cdots + F_{n,n}^2) \xrightarrow{n \to \infty} 2$ almost surely. For both laws of large numbers we prove corresponding large deviations results. For (i), the rate of the large deviations is n and we can give the rate function explicitly. For (ii), the rate is n for downwards deviations and $\sqrt{n}$ for upwards deviations. In both cases we give exact rate functions.

1 Introduction

Kingman's coalescent is a random tree introduced in [9] as the genealogy arising in large population genetic models. It has infinitely many leaves and is usually constructed from leaves to the root as follows: given that there are k lines in the tree, after some exponential time with rate $\binom{k}{2}$, two lines are chosen uniformly and merged to one line, leaving the tree with $k - 1$ lines. Due to the quadratic rate $\binom{k}{2}$ the tree immediately comes down from infinitely to finitely many leaves [6]. Since Pitman's seminal paper [10] this random tree has been generalized to other infinite trees arising in population genetics models.

For the Kingman coalescent some laws of large numbers and central limit theorems have been proved. They are nicely summarized in [1], Chapter 4.2; see also Proposition 2.1 below. For $\varepsilon > 0$ let N_ε denote the number of lines time ε in the past. Then, since the Kingman coalescent immediately comes down from infinity, N_ε is finite. Furthermore it is approximately $2/\varepsilon$. Equivalently, the time T_n it takes the coalescent to go from infinitely many lines to n lines is approximately $2/n$ for large n. Going to the fine structure, at time T_n the infinite population is decomposed in n families (whose joint distribution is exchangeable) and every leaf in the tree belongs to exactly one of the n families whose frequencies are denoted by $F_{1,n}, \ldots, F_{n,n}$. It is known that for large n

---

[†]University of Freiburg, Germany. E-mail: depperschmidt@stochastik.uni-freiburg.de
[‡]University of Freiburg, Germany. E-mail: p.p@stochastik.uni-freiburg.de

a randomly chosen $nF_{i,n}$ is approximately exponentially distributed with mean 1. This translates into several laws of large numbers; see e.g. (35) in [1]. In particular the probability of picking (from the initial infinite population) two leaves that belong to the same family, given by $F_{1,n}^2 + \cdots + F_{n,n}^2$, is approximately $2/n$.

The main goal of the present paper is to study the corresponding large deviations results. To the best of our knowledge, except for [2], cf. Remark 2.7 below, results in this direction are not present in the literature. We formulate our results in the next section. Theorem 2.3 gives a full large deviation principle for the distributions of nT_n. The proof, given in Section 3, is an application of the Gärtner-Ellis Theorem. As a byproduct, we derive a large deviation principle for the distributions of $\varepsilon N_\varepsilon$ in Corollary 2.5. Large deviations of $n(F_{1,n}^2 + \cdots + F_{n,n}^2)$ are considered in Theorem 2.8 and exact rate functions for downwards and upwards deviations are given. The proof is given in Section 4.2. For the upward deviations we use a variant of Cramér's theorem for heavy-tailed random variables; see e.g. [8]. For the downward deviations we use a connection to self-normalized large deviations; see [12]. This connection was pointed out to us by Alain Rouault and Nina Gantert. Since the rate function for downward deviations is hard to treat analytically we provide in Theorem 2.10 a simple lower bound. The proof of that bound is given in Section 4.3.

2 Main results

The Kingman coalescent can be seen as a discrete graph, more precisely a discrete tree with infinitely many leaves. Let $S_2, S_3, \ldots$ be independent exponentially distributed variables with mean 1. Then the Kingman coalescent tree can be constructed from the root to the leaves as follows.

1. Start the tree with two lines from the root.
2. For $k \geq 2$ the tree stays with k lines for the amount of time $S_k/\binom{k}{2}$. After that time one of the k lines is randomly chosen. This line splits in two so that the number of lines jumps from k to $k+1$.
3. Stop upon reaching infinitely many lines, which happens after (the almost surely finite) time $T_1 := \sum_{k=2}^{\infty} S_k/\binom{k}{2}$.

The random variable T_1 is the total tree height. Alternatively, T_1 is the time to the most recent common ancestor (MRCA) of the infinite population (of leaves). Counted from the top of the tree at time $\varepsilon > 0$ a random number N_ε of active lines in the Kingman tree is present, i.e.

$$N_\varepsilon := \inf\{n : T_n < \varepsilon\} \qquad \text{for} \qquad T_n := \sum_{k=n+1}^{\infty} \frac{S_k}{\binom{k}{2}}. \tag{2.1}$$

At time T_n every leaf belongs to one of n disjoint families and all members of each such family stem from the same line at time T_n. Let us denote the frequencies of these families (which exist due to exchangeability by de Finetti's Theorem) by $F_{1,n}, \ldots, F_{n,n}$. The following results are well known (see [1] for (2.2) and (2.3) and [7] for (2.4); proofs can also be found in [5].)

Proposition 2.1 (Laws of large numbers).
Let $(T_n)_{n=1,2,\ldots}$, $(N_\varepsilon)_{\varepsilon>0}$ *and* $(F_{1,n}, \ldots, F_{n,n})_{n=1,2,\ldots}$ *be as above. Then*

$$nT_n \xrightarrow{n\to\infty} 2 \quad \text{almost surely,} \tag{2.2}$$

$$\varepsilon N_\varepsilon \xrightarrow{\varepsilon\to 0} 2 \quad \text{almost surely,} \tag{2.3}$$

and

$$n \sum_{k=1}^{n} F_{k,n}^2 \xrightarrow{n \to \infty} 2 \quad \text{almost surely.} \tag{2.4}$$

Remark 2.2 (Interpretation of (2.4)). We note that the left hand side of (2.4) has the interpretation of a *homozygosity by descent* in the following sense: when picking two leaves from the tree at time 0, the probability that both share a common ancestor at time T_n is $\sum_{k=1}^{n} F_{k,n}^2$. Then, the law of large number states that the homozygosity by descent at time T_n is approximately $2/n$ for large n.

In the present paper we are interested in large deviations results corresponding to the statements of Proposition 2.1. We start with large deviations connected with (2.2). First we introduce some notation. For $n = 1, 2, \dots$ let μ_n denote the distribution of nT_n, i.e. $\mu_n(\cdot) = \mathbf{P}(nT_n \in \cdot)$. Furthermore we denote by $\mathcal{B}(\mathbb{R})$ the Borel σ-algebra on $\mathbb{R}$ and for $\Gamma \in \mathcal{B}(\mathbb{R})$ we denote by Γ° the *interior* and by $\bar{\Gamma}$ the *closure* of Γ. For $x > 0$, let $t_x < 1$ be the unique solution of the equation $x = f(t)$, where the continuous and increasing function $f : (-\infty, 1) \to (0, \infty)$ is defined by (see Figure 1 for a plot)

$$f(t) := \begin{cases} \dfrac{1}{\sqrt{t}} \log \dfrac{1 + \sqrt{t}}{1 - \sqrt{t}} & : \ 0 < t < 1, \\[2mm] 2 & : \ t = 0, \\[2mm] \dfrac{2}{\sqrt{|t|}} \arctan \sqrt{|t|} & : \ t < 0. \end{cases} \tag{2.5}$$

The proof of the following theorem is given in Section 3.1.

Theorem 2.3 (LDP for $(\mu_n)_{n=1,2,\dots}$). *The sequence $(\mu_n)_{n=1,2,\dots}$ satisfies a large deviation principle with scale n and good rate function I given by*

$$I(x) := \begin{cases} \dfrac{t_x}{2} x + \displaystyle\int_1^\infty \log\left(1 - \dfrac{t_x}{y^2}\right) dy & : \ x > 0, \\[4mm] \infty & : \ x \leq 0. \end{cases} \tag{2.6}$$

In other words, for any $\Gamma \in \mathcal{B}(\mathbb{R})$ we have

$$-\inf_{x \in \Gamma^\circ} I(x) \leq \liminf_{n \to \infty} \frac{1}{n} \log \mu_n(\Gamma) \leq \limsup_{n \to \infty} \frac{1}{n} \log \mu_n(\Gamma) \leq -\inf_{x \in \bar{\Gamma}} I(x).$$

Remark 2.4 (Interpretation). Both, the function f from (2.5) and I from (2.6) are plotted in Figure 1. The minimum of the rate function is attained at $x = 2$. This fact is clear from the law of large numbers, (2.2). In addition, $I(x) = \infty$ for $x \leq 0$ because $nT_n > 0$ almost surely.

Let us now have a closer look at the behaviour of $I(x)$ for x near 0 and for large x. Since $2\arctan(t) \xrightarrow{t \to \infty} \pi$, we have that $\sqrt{|t_x|}x \xrightarrow{x \downarrow 0} \pi$, and hence, $t_x \overset{x \downarrow 0}{\approx} -\frac{\pi^2}{x^2}$. In this case,

$$\begin{aligned} xI(x) &\overset{x \downarrow 0}{\approx} -\frac{\pi^2}{2} + x \int_1^\infty \log\left(1 + \frac{\pi^2}{x^2 y^2}\right) dy \\ &\overset{x \downarrow 0}{\approx} -\frac{\pi^2}{2} + \pi \int_0^\infty \log\left(1 + \frac{1}{z^2}\right) dz = \frac{\pi^2}{2}, \end{aligned} \tag{2.7}$$

where the last equality follows from $\int_0^\infty \log(1 + 1/z^2)dz = \pi$. To understand the behaviour for large x, note that since

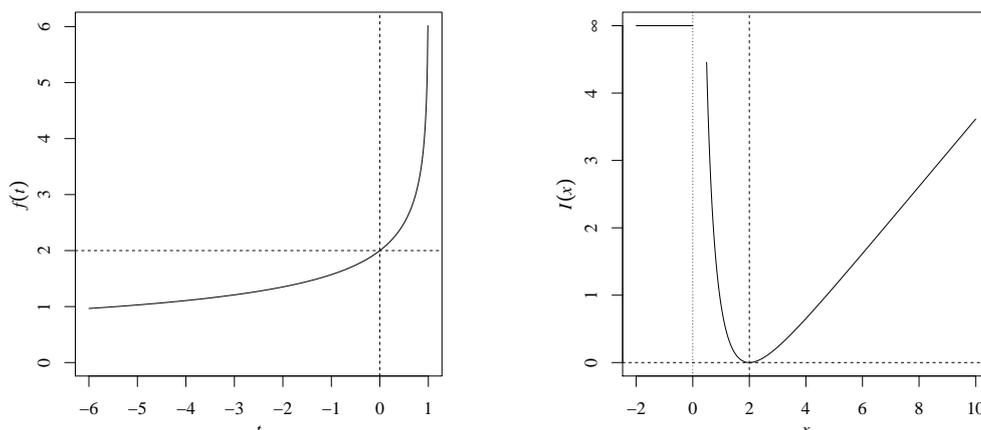$$f(t) \overset{t \uparrow 1}{\approx} \log 2 - \log(1 - \sqrt{t}),$$

Figure 1: This figure displays the functions f and I from (2.5) and (2.6), respectively.

for $x \to \infty$ we have $2/(1 - \sqrt{t_x}) \approx e^x$ and in particular $t_x \approx (1 - 2e^{-x})^2 \approx 1$. It follows

$$\frac{I(x)}{x} \overset{x \to \infty}{\approx} \frac{1}{2} + \frac{1}{x} \int_1^\infty \log(1 - 1/y^2)\, dy \overset{x \to \infty}{\approx} \frac{1}{2}.$$

Note that (2.2) and (2.3) are equivalent. Indeed, $\{T_n \geq \varepsilon\} = \{N_\varepsilon \geq n\}$ (this also holds with $\geq$ replaced by $\leq$) by construction, and $T_n \downarrow 0$ as $n \to \infty$ and $N_\varepsilon \uparrow \infty$ as $\varepsilon \to 0$. Hence, Theorem 2.3 translates into a large deviation principle for $\varepsilon N_\varepsilon$. In the following we denote by ν_ε the distribution of $\varepsilon N_\varepsilon$, i.e. $\nu_\varepsilon(\,\cdot\,) = \mathbf{P}(\varepsilon N_\varepsilon \in \,\cdot\,)$. The proof of the next result is given in Section 3.2; see Figure 2 for a plot of the rate function $\widehat{I}$.

Corollary 2.5 (LDP for $(\nu_\varepsilon)_{\varepsilon>0}$)**.** *For $\varepsilon \downarrow 0$ the family $(\nu_\varepsilon)_{\varepsilon>0}$ satisfies a large deviation principle with scale $1/\varepsilon$ and good rate function $\widehat{I}$ given by*

$$\widehat{I}(x) = \begin{cases} xI(x) & :\ x > 0, \\ \dfrac{\pi^2}{2} & :\ x = 0, \\ \infty & :\ x < 0, \end{cases} \tag{2.8}$$

with I from (2.6). In particular, for $\Gamma \in \mathcal{B}(\mathbb{R})$ we have

$$-\inf_{x \in \Gamma^\circ} \widehat{I}(x) \leq \liminf_{\varepsilon \to 0} \varepsilon \log \nu_\varepsilon(\Gamma) \leq \limsup_{\varepsilon \to 0} \varepsilon \log \nu_\varepsilon(\Gamma) \leq -\inf_{x \in \overline{\Gamma}} \widehat{I}(x).$$

Remark 2.6 (The full distribution of N_ε)**.** The distributions $\nu_\varepsilon, \varepsilon > 0$ (as well as $\mu_n, n = 1, 2, \dots$) have been described explicitly in the literature. In Section 6 in [13] it is shown

$$\mathbf{P}(N_\varepsilon = n) = \sum_{k=n}^\infty e^{-\binom{k}{2}\varepsilon} \frac{(-1)^{k-n}(2k-1) \cdot n \cdots (n+k-2)}{n!(k-n)!}.$$

In principle, this formula must also give the large deviations for the measures ν_ε, but this does not seem straight-forward.

Remark 2.7 (The rate function $\widehat{I}$ and comparison with [2])**.** Although the main goal in [2] was the analysis of spatial Λ-coalescents, also some large deviations bounds for Kingman's coalescent are provided. These bounds are mainly based on Markov inequality. Precisely, in Lemma 2.2 in [2] it is shown that for $0 < x < \frac{1}{2}$

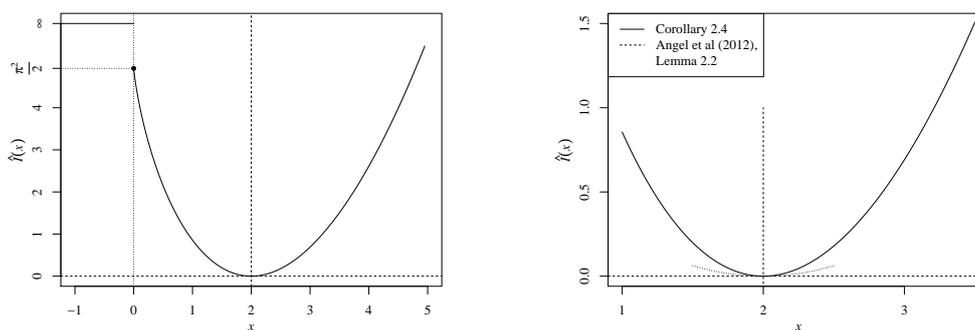$$\mathbf{P}(|\varepsilon N_\varepsilon - 2| > x) < e^{-\frac{x^2}{4\varepsilon}}$$

Figure 2: The figure on the left shows the rate function $\widehat{I}$ from Corollary 2.5. The figure on the right is a comparison of $\widehat{I}$ with the lower bound obtained from [2].

and therefore

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(|\varepsilon N_\varepsilon - 2| > x) \leq -\frac{x^2}{4}.$$

In the neighbourhood of 2 the last inequality translates easily into a bound for the rate function $\widehat{I}$ from (2.8); see Figure 2. Namely, for $x \in (1.5, 2.5)$ we have

$$\widehat{I}(x) \geq \frac{(x-2)^2}{4}.$$

Next, we state some large deviations results connected to (2.4). For

$$W_n := n \sum_{k=1}^{n} F_{k,n}^2$$

we know from (2.4) that $W_n \xrightarrow{n \to \infty} 2$ holds almost surely. The proof of this result is based on the well-known fact (see e.g. Section 5 in [9]) that the distribution of W_n can be derived using uniform order statistics: Let $U_1, \ldots, U_{n-1}$ be independent and uniformly distributed on $[0, 1]$, and $0 < U_{(1)} < \cdots < U_{(n-1)} < 1$ be their order statistics. Additionally, let $R_1, \ldots, R_n$ be independent exponentially distributed random variables with mean 1. Then,

$$\begin{aligned}(F_{1,n}, \ldots, F_{n,n}) &\overset{d}{=} \left(U_{(1)}, U_{(2)} - U_{(1)}, \ldots, U_{(n-1)} - U_{(n-2)}, 1 - U_{(n-1)}\right) \\ &\overset{d}{=} \left(\frac{R_1}{\sum_{j=1}^{n} R_j}, \ldots, \frac{R_n}{\sum_{j=1}^{n} R_j}\right).\end{aligned} \tag{2.9}$$

Here the second equality in distribution is one of the well known representations of uniform spacings; see e.g. Section 4.1 in [11]. It follows

$$W_n \overset{d}{=} n \frac{\sum_{k=1}^{n} R_k^2}{\left(\sum_{j=1}^{n} R_j\right)^2} = \frac{\frac{1}{n} \sum_{k=1}^{n} R_k^2}{\left(\frac{1}{n} \sum_{j=1}^{n} R_j\right)^2}. \tag{2.10}$$

We will use this representation to obtain large deviations results for W_n. In particular we show that upwards large deviations of W_n are on the scale $\sqrt{n}$ while downwards large deviations are on the scale n. The proof is given in Section 4.2.

Theorem 2.8 (Large deviations of W_n). *For each $x \geq 2$, we have*

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \log \mathbf{P}\left(W_n \geq x\right) = -\sqrt{x - 2}. \tag{2.11}$$

Furthermore $\mathbf{P}(W_n < 1) = 0$ *and for each* $1 < x < 2$, *we have*

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbf{P}\left(W_n \leq x\right) = -\widetilde{I}(x). \tag{2.12}$$

The function $\widetilde{I}(x)$ *is positive for* $1 < x < 2$ *and is given by*

$$\widetilde{I}(x) := \sup_{c \geq 0} \inf_{t \geq 0} M(x, c, t). \tag{2.13}$$

Here $M : (1, 2) \times [0, \infty) \times [0, \infty) \to \mathbb{R}$ *is a function of the form* $M := M_1 + M_2 + M_3$ *with*

$$M_1(x, c, t) := \frac{1}{2} \log(2\pi) + \frac{1}{4} \log x - \frac{1}{2} \log t$$

$$M_2(x, c, t) := \frac{(tc - 1)^2 x - t^2 c^2}{2t\sqrt{x}}$$

$$M_3(x, c, t) := \log \Phi\left(\frac{(tc - 1)x^{1/4}}{\sqrt{t}}\right)$$

where Φ *denotes the distribution function of the one dimensional standard Gaussian distribution.*

Though the rate function in (2.12) is exact it is hard to treat analytically. For this reason we provide in Theorem 2.10 a much simpler lower bound for downwards large deviations of W_n. For the proof we use the following lemma which provides another representation of W_n in terms of exponential random variables (see Section 4 for proofs).

Lemma 2.9 (Representation of W_n). *Let* $R_1, \dots, R_n$ *be independent exponentially distributed random variables with mean 1. Then,*

$$W_n \overset{d}{=} \frac{\frac{1}{n}\left(2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}\right)}{\left(\frac{1}{n} \sum_{k=1}^{n} R_k\right)^2}. \tag{2.14}$$

Theorem 2.10 (Lower bound on downwards large deviations of W_n). *For* $1 < x < 2$ *we have*

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbf{P}\left(W_n \leq x\right) \geq 1 - \frac{1}{\sqrt{x - 1}}. \tag{2.15}$$

Remark 2.11 (Rationale and use of the representation in Lemma 2.9). The main point in the proof of Lemma 2.9 is that W_n does not depend on the order of the R_k and hence we can as well order them according to their size.

Let us briefly explain how we will use (2.14) in the proof of in (2.15). Since W_n is minimal if $R_1 = \cdots = R_n$ (whence $W_n = 1$), we have to look for possibilities that all R_k's are of about the same size in order to obtain a large deviations result for W_n. Let $R_{(1)}, \dots, R_{(n)}$ denote the above exponential random variables ordered in increasing order, i.e. $R_{(i)}$ is the ith smallest value. Using "competing exponential clocks" arguments (see also the proof of the lemma) one can see that $R_{(i)} - R_{(i-1)}$ is exponentially distributed with mean $1/(n - i + 1)$. Hence, one way of obtaining similar values for all R_k's arises if $R_{(1)}$ is particularly large, which then leads to a large deviations result for W_n.

Remark 2.12 (Interpretation, uniform spacings and the Poisson-Dirichlet distribution). 1. Let us give some heuristics about the rates arising in Theorem 2.8. For (2.11), we have to ask ourselves about the easiest way W_n becomes too large. From (2.9), we see that this is the case if one of the R_k's is too large, making this kind of deviations a local property in the sense that only a single of the R_k's has to show some untypical
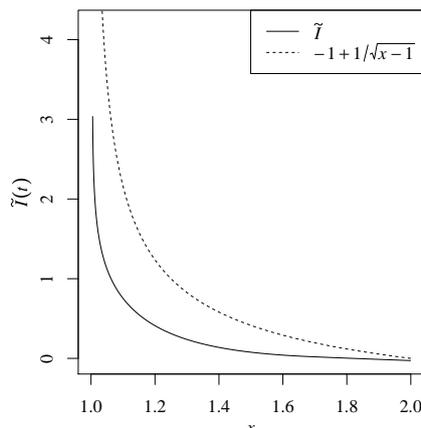
Figure 3: Numerical comparison of the exact rate function for downwards large deviations of W_n from (2.12) in Theorem 2.8 and the lower bound from (2.15) in Theorem 2.10.

behavior. This is different when looking at (2.12), i.e. too small values of W_n. First, observe that W_n is small only if all (or many) families have about equal sizes (extreme case $F_{1,n} = \cdots = F_{n,n} = \frac{1}{n}$ gives the minimal value $W_n = 1$). Hence, such downward deviations require to study a global property of the random variable W_n, which is significantly harder. For the proof of (2.12) we will interpret W_n as a self-normalised sum and use a result on large deviations result for such sums from [12].

2. From (2.9), we see that in fact W_n is a function of uniform order statistics, which, for instance, have been studied in detail (although no large deviations results were given) in [11]. Hence, Theorem 2.8 may as well be interpreted as a large deviations result for uniform order statistics.

3. As stated in Remark 2.2, W_n/n can be interpreted as homozygosity at time T_n. Using a Poisson process along the tree with intensity $\theta/2$, we can ask for the probability of picking two leaves from the tree which are not separated by a Poisson mark, denoted by *homozygosity in state*, abbreviated by H_θ/θ. This quantity is closely related to the Poisson-Dirichlet distribution and some large deviations (in the limit of large θ) were derived in [3]. It is shown there in Theorem 5.1 that $H_\theta/\theta \xrightarrow{\theta \to \infty} 0$ and that

$$\mathbf{P}(H_\theta > \theta x) = e^{-\theta(I(\sqrt{x}) + o(1))}$$

for $I(x) = -\log(1 - x)$. However, a large deviation principle for the quantity H_θ (noting that $H_\theta \xrightarrow{\theta \to \infty} 1$), which corresponds to the results from Theorem 2.8, could not be obtained in [3]. At least, it was shown that its scale cannot be larger than $\sqrt{\theta}$.

3 Proof of Theorem 2.3 and Corollary 2.5

3.1 Proof of Theorem 2.3

The proof of Theorem 2.3 is an application of the Gärtner-Ellis theorem; see for instance Section 2.3 in [4].

Let $\Lambda_n(t) := \log \mathbf{E}[e^{tnT_n}]$ and $\mu_n(\cdot) = \mathbf{P}(nT_n \in \cdot)$. To show that the sequence $\mu_1, \mu_2, \ldots$ satisfies a large deviation principle with scale n and a good rate function we need to check the following three conditions.

GE1 $\Lambda(t) := \lim_{n \to \infty} \frac{1}{n}\Lambda_n(nt)$ exists for all t as a limit in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Furthermore $t \to \Lambda(t)$ is lower-semicontinuous, $0 \in \mathcal{D}_\Lambda^\circ$, where $\mathcal{D}_\Lambda := \{t : \Lambda(t) < \infty\}$.

GE2 Λ is differentiable on $\mathcal{D}_\Lambda^\circ$.

GE3 Λ is *steep*, i.e. $\Lambda'(t_n) \xrightarrow{n \to \infty} \infty$ whenever $t_1, t_2, \cdots \in \mathcal{D}_\Lambda^\circ$ and $t_n \xrightarrow{n \to \infty} t \in \partial \mathcal{D}_\Lambda$.

Then the good rate function is given by

$$x \mapsto I(x) = \sup_{t \in \mathbb{R}} (tx - \Lambda(t)). \tag{3.1}$$

We proceed in three steps. First, we compute $\Lambda(t) := \lim_{n \to \infty} \frac{1}{n} \Lambda_n(nt)$. Second, we check the further assumptions of the Gärtner-Ellis theorem and obtain I as the Fenchel-Legendre transform of Λ. In the third step, for the rate function I from (3.1) we obtain its simplified form given in Theorem 2.3.

Step 1. The limit of $\frac{1}{n}\Lambda_n(nt)$: We will show that

$$\Lambda(t) = \lim_{n \to \infty} \frac{1}{n} \Lambda_n(nt) = \begin{cases} -\int_1^\infty \log\left(1 - \frac{2t}{x^2}\right) dx & : t \leq \frac{1}{2}, \\ \infty & : t > \frac{1}{2}. \end{cases} \tag{3.2}$$

For this, recall from (2.1) that $T_n = \sum_{k=n+1}^\infty S_k / \binom{k}{2}$ where $S_k/\binom{k}{2}$ is exponentially distributed with rate $\binom{k}{2}$ as well as independent of S_ℓ for all $\ell \neq k$. Furthermore recall that the moment generating function of an exponentially distributed random variable R with rate $\lambda > 0$ is given by

$$\mathbf{E}\left[e^{tR}\right] = \begin{cases} \frac{\lambda}{\lambda - t}, & \text{if } t < \lambda, \\ \infty, & \text{if } t \geq \lambda. \end{cases} \tag{3.3}$$

Hence, for each $n \in \mathbb{N}$ and $t \in \mathbb{R}$ we obtain by the monotone convergence theorem

$$\varphi_n(nt) := \mathbf{E}\left[e^{tn^2 T_n}\right] = \mathbf{E}\left[e^{tn^2 \sum_{k=n+1}^\infty S_k/\binom{k}{2}}\right] = \prod_{k=n+1}^\infty \mathbf{E}\left[e^{tn^2 S_k/\binom{k}{2}}\right]. \tag{3.4}$$

We have to consider two cases $t > \frac{1}{2}$ and $t \leq \frac{1}{2}$ separately. First suppose that $t > \frac{1}{2}$. Then there exists $n_0 \in \mathbb{N}$ so that for all $n \geq n_0$ we have

$$2t \geq \frac{n+1}{n}, \quad \text{i.e.} \quad \frac{tn^2}{\binom{n+1}{2}} = 2t \frac{n}{n+1} > 1.$$

Consequently, using (3.3), we obtain $\mathbf{E}[e^{tn^2 T_n}] = \infty$ for each $n \geq n_0$. Hence, $\varphi_n(nt) = \infty$ and $\Lambda_n(nt) = \log \varphi_n(nt) = \infty$ for n large enough. Thus, we have

$$\Lambda(t) = \lim_{n \to \infty} \frac{1}{n} \Lambda_n(nt) = \infty \quad \text{for all} \ \ t > \frac{1}{2}.$$

Now suppose that $t \leq \frac{1}{2}$. For $n \in \mathbb{N}$ and $k \geq n+1$ we have $\binom{k}{2} \geq tn^2$. Furthermore using (3.4) and (3.3) we can write

$$\varphi_n(nt) = \prod_{k=n+1}^\infty \frac{1}{1 - \frac{2tn^2}{k(k-1)}}.$$

Using this we can rewrite $\frac{1}{n}\Lambda_n(nt)$ for $t \leq \frac{1}{2}$ as

$$\frac{1}{n}\Lambda_n(nt) = \frac{1}{n} \log\left(\prod_{k=n+1}^\infty \frac{1}{1 - \frac{2tn^2}{k(k-1)}}\right) = -\frac{1}{n} \sum_{k=n}^\infty \log\left(1 - \frac{2t}{\frac{k}{n}\frac{k+1}{n}}\right)$$

$$= -\frac{1}{n} \sum_{x \in \left\{1, 1+\frac{1}{n}, 1+\frac{2}{n}, \ldots\right\}} \log\left(1 - \frac{2t}{x\left(x + \frac{1}{n}\right)}\right) = -\int_1^\infty \log\left(1 - \frac{2t}{\frac{\lfloor xn \rfloor}{n}\frac{\lfloor xn \rfloor + 1}{n}}\right) dx$$

and by the dominated convergence theorem we obtain

$$\frac{1}{n}\Lambda_n(nt) \xrightarrow{n\to\infty} -\int_1^\infty \log\left(1 - \frac{2t}{x^2}\right)\,dx.$$

Hence, GE1 is shown with Λ as in (3.2). Moreover, we have $\mathcal{D}_\Lambda = (-\infty, \frac{1}{2}]$, $\Lambda(\frac{1}{2}) = \int_1^\infty \log(1 - \frac{1}{x^2}, dx = \pi$ and Λ is lower-semi-continuous.

Step 2. Further assumptions of the Gärtner-Ellis theorem: We proceed by checking the assumptions GE2 and GE3. For differentiability of Λ for $t < \frac{1}{2}$ consider for $-\infty < r < 0 < s < \frac{1}{2}$ the function

$$f : (1, \infty) \times (r, s) \to \mathbb{R}$$

$$(x, t) \qquad \mapsto -\log\left(1 - \frac{2t}{x^2}\right).$$

We have $\int_1^\infty |f(x, t)|\,dx < \infty$ for $t \in (r, s)$ and the derivative

$$\frac{d}{dt} f(x, t) = \frac{2}{x^2 - 2t}$$

exists for each $x \in (1, \infty)$ and is continuous in t. Hence, we can interchange differentiation and integration and obtain

$$\Lambda'(t) = \int_1^\infty \frac{2}{x^2 - 2t}\,dx.$$

Furthermore, for a sequence $t_1, t_2, \ldots$ with $t_n \uparrow \frac{1}{2}$ we obtain

$$\lim_{n\to\infty} \Lambda'(t_n) = \lim_{n\to\infty} \int_1^\infty \frac{2}{x^2 - 2t_n}\,dx$$

$$= \lim_{n\to\infty} \frac{1}{\sqrt{2t_n}}\left(\log\left(1 + \sqrt{2t_n}\right) - \log\left(1 - \sqrt{2t_n}\right)\right) = \infty,$$

i.e. condition GE3 is also satisfied.

Step 3. Properties of I: Applying the Gärtner-Ellis theorem reveals that the sequence of distributions of $nT_n, n = 1, 2, \ldots$ satisfies a large deviation principle with good rate function

$$I(x) = \sup_{t \leq \frac{1}{2}} \left[tx + \int_1^\infty \log\left(1 - \frac{2t}{y^2}\right)\,dy \right] = \sup_{t \leq 1} \left[\frac{t}{2}x + \int_1^\infty \log\left(1 - \frac{t}{y^2}\right)\,dy \right].$$

In order to compute that supremum, we write for $t \geq 0$

$$2\frac{\partial}{\partial t}\left[\frac{t}{2}x + \int_1^\infty \log\left(1 - \frac{t}{y^2}\right)\,dy \right] = x - 2\int_1^\infty \frac{1}{y^2 - t}dy$$

$$= x + \frac{1}{\sqrt{t}}\int_1^\infty \frac{1}{y + \sqrt{t}} - \frac{1}{y - \sqrt{t}}dy$$

$$= x - \frac{1}{\sqrt{t}}\log\frac{1 + \sqrt{t}}{1 - \sqrt{t}}$$

while for $t \leq 0$

$$2\frac{\partial}{\partial t}\left[\frac{t}{2}x + \int_1^\infty \log\left(1 - \frac{t}{y^2}\right)\,dy \right] = x - 2\int_1^\infty \frac{1}{y^2 + |t|}dy$$

$$= x - \frac{2}{\sqrt{|t|}}\arctan\sqrt{|t|}.$$

It is easy to see that the second derivative is negative throughout, such that the supremum is attained at t_x given by the solution of $f(t_x) = x$ for f as in (2.5). Finally we note that for $t \in [0, 1)$ the range of $t \mapsto \frac{1}{\sqrt{t}} \log \frac{1+\sqrt{t}}{1-\sqrt{t}}$ is $[2, \infty)$ and for $t \in (-\infty, 0]$ the range of $t \mapsto \frac{2}{\sqrt{|t|}} \arctan \sqrt{|t|}$ is $(0, 2]$. Hence, the scale function I is of the form given in (2.6).

3.2 Proof of Corollary 2.5

The proof is based on the fact that $\{T_n \geq \varepsilon\} = \{N_\varepsilon \geq n\}$. Thus, for $x \geq 2$ we have

$$xI(x) = x \lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}(nT_n \geq x) = \lim_{n\to\infty} \frac{x}{n} \log \mathbf{P}(\frac{x}{n} N_{x/n} \geq x) = \lim_{\varepsilon\to 0} \varepsilon \log \mathbf{P}(\varepsilon N_\varepsilon \geq x)$$

and for $0 < x \leq 2$

$$xI(x) = x \lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}(nT_n \leq x) = \lim_{n\to\infty} \frac{x}{n} \log \mathbf{P}(\frac{x}{n} N_{x/n} \leq x) = \lim_{\varepsilon\to 0} \varepsilon \log \mathbf{P}(\varepsilon N_\varepsilon \leq x).$$

The value $\widehat{I}(0)$ follows from (2.7). Since the rate function I attains its minimum at $x = 2$, is decreasing below and increasing above 2, the result follows.

4 Proof of Lemma 2.9, Theorem 2.8 and Theorem 2.10

4.1 Proof of Lemma 2.9

When looking at (2.10), note that W_n does not depend on the order of the R_k's. Therefore, it is possible to order them according to their size. Precisely, let $0 < R_{(1)} < \cdots < R_{(n)}$ be their order statistics. Then it is well-known that

$$(R_{(n)}, R_{(n-1)}, \ldots, R_{(1)}) \stackrel{d}{=} \left(\sum_{k=1}^n \frac{R_k}{k}, \sum_{k=2}^n \frac{R_k}{k}, \ldots, \frac{R_n}{n} \right), \quad \text{i.e.} \quad R_{(n-k+1)} \stackrel{d}{=} \sum_{i=k}^n \frac{R_i}{i}.$$

Indeed, the smallest of n independent exponentially distributed mean 1 random variables is exponentially distributed with mean $\frac{1}{n}$ (as does $\frac{R_n}{n}$), and the second smallest then has the same distribution as $\frac{R_n}{n} + \frac{R_{n-1}}{n-1}$ etc. Now, we obtain (2.14) as follows

$$
\begin{aligned}
W_n &\stackrel{d}{=} \frac{\frac{1}{n}\sum_{k=1}^n R_{(n-k+1)}^2}{\left(\frac{1}{n}\sum_{j=1}^n R_{(n-j+1)}\right)^2} \stackrel{d}{=} \frac{\frac{1}{n}\sum_{k=1}^n \left(\sum_{i=k}^n \frac{R_i}{i}\right)^2}{\left(\frac{1}{n}\sum_{j=1}^n \sum_{i=j}^n \frac{R_i}{i}\right)^2} \\
&= \frac{\frac{1}{n}\left(2\sum_{k=1}^n \sum_{i=k}^n \sum_{j=i}^n \frac{R_i R_j}{ij} - \sum_{k=1}^n \sum_{i=k}^n \frac{R_i^2}{i^2}\right)}{\left(\frac{1}{n}\sum_{i=1}^n \sum_{j=1}^i \frac{R_i}{i}\right)^2} \\
&= \frac{\frac{1}{n}\left(2\sum_{j=1}^n \sum_{i=1}^j \sum_{k=1}^i \frac{R_i R_j}{ij} - \sum_{i=1}^n \sum_{k=1}^i \frac{R_i^2}{i^2}\right)}{\left(\frac{1}{n}\sum_{i=1}^n R_i\right)^2} \\
&= \frac{\frac{1}{n}\left(2\sum_{j=1}^n \sum_{i=1}^j \frac{R_i R_j}{j} - \sum_{i=1}^n \frac{R_i^2}{i}\right)}{\left(\frac{1}{n}\sum_{i=1}^n R_i\right)^2}.
\end{aligned}
$$

4.2 Proof of Theorem 2.8

We start by proving (2.11). Let $x \geq 2$ and let $R_1, R_2, \ldots$ be independent exponential random variables with mean 1. In what follows we set

$$X_n := \frac{1}{n}\sum_{k=1}^n R_k \quad \text{and} \quad Z_n := \frac{1}{n}\sum_{k=1}^n R_k^2. \tag{4.1}$$

According to (2.10), it suffices to show that

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \log \mathbf{P}\left(\frac{Z_n}{X_n^2} \geq x\right) = -\sqrt{x-2}. \tag{4.2}$$

To this end we will show that for all $0 < \varepsilon < 1$,

$$\limsup_{n \to \infty} \frac{1}{\sqrt{n}} \log \mathbf{P}\left(\frac{Z_n}{X_n^2} \geq \frac{x}{1-\varepsilon}\right) \leq -\sqrt{x-2} \tag{4.3}$$

as well as

$$\liminf_{n \to \infty} \frac{1}{\sqrt{n}} \log \mathbf{P}\left(\frac{Z_n}{X_n^2} \geq x\right) \geq -\sqrt{(1+\varepsilon)x-2} \tag{4.4}$$

and obtain (4.2) by letting $\varepsilon \to 0$. For (4.3) we have

$$\mathbf{P}\left(\frac{Z_n}{X_n^2} \geq \frac{x}{1-\varepsilon}\right) \leq \mathbf{P}\left(Z_n \geq x\right) + \mathbf{P}\left(X_n \leq \sqrt{1-\varepsilon}\right). \tag{4.5}$$

We consider the two terms on the right hand side of the last display separately and start with the first one. Observe that $\mathbf{E}[e^{\lambda R_1^2}] = \infty$ for $\lambda > 0$, $\mathbf{E}[R_1^2] = 2$ and $\mathbf{P}\left(R_1^2 \geq t\right) = e^{-\sqrt{t}}$ for $t \geq 0$. We use a variant of Cramér's theorem for heavy-tailed random variables from [8]. In particular, we refer to the statement around equation (1.2) there (the assumption there is fulfilled with X_1 replaced by R_1^2 and $r = \frac{1}{2}$, $m = 2$ and $c = 1$). We obtain

$$\mathbf{P}\left(Z_n \geq x\right) = e^{-\sqrt{n}\left(\sqrt{x-2}+o(1)\right)} \quad \text{as } n \to \infty. \tag{4.6}$$

For the second term on the right hand side of (4.5) by the (classical) Cramér theorem we obtain

$$\mathbf{P}\left(X_n \leq \sqrt{1-\varepsilon}\right) = e^{-n I_{\exp}(\sqrt{1-\varepsilon})(1+o(1))}, \quad \text{as } n \to \infty, \tag{4.7}$$

where

$$I_{\exp}(y) := y - 1 - \log(y) \tag{4.8}$$

is the Fenchel-Legendre transform of the function $t \mapsto \log \mathbf{E}[e^{\lambda R_1}]$. Now, using (4.5), (4.6) and (4.7) we obtain

$$\limsup_{n \to \infty} \frac{1}{\sqrt{n}} \log \mathbf{P}\left(\frac{Z_n}{X_n^2} \geq \frac{x}{1-\varepsilon}\right)$$

$$\leq \limsup_{n \to \infty} \frac{1}{\sqrt{n}} \log \left(e^{-\sqrt{n}(\sqrt{x-2}+o(1))} + e^{-n I_{\exp}(\sqrt{1-\varepsilon})(1+o(1))}\right) = -\sqrt{x-2},$$

which shows (4.3). For the proof of (4.4) we write

$$\begin{aligned} \mathbf{P}\left(\frac{Z_n}{X_n^2} \geq x\right) &\geq \mathbf{P}\left(\frac{Z_n}{X_n^2} \geq x, X_n^2 \leq 1+\varepsilon\right) \\ &\geq \mathbf{P}\left(Z_n \geq x(1+\varepsilon), X_n^2 \leq 1+\varepsilon\right) \\ &\geq \mathbf{P}\left(Z_n \geq x(1+\varepsilon)\right) - \mathbf{P}\left(X_n \geq \sqrt{1+\varepsilon}\right). \end{aligned} \tag{4.9}$$

Again we consider both terms in the last line separately. For the first term, as in (4.6) we obtain

$$\mathbf{P}\left(Z_n \geq x(1+\varepsilon)\right) = e^{-\sqrt{n}\left(\sqrt{x(1+\varepsilon)-2}+o(1)\right)}, \quad \text{as } n \to \infty. \tag{4.10}$$

For the second term, we use the same argument as for (4.7) and get

$$\mathbf{P}\left(X_n \geq \sqrt{1+\varepsilon}\right) \leq e^{-nI_{\exp}(\sqrt{1+\varepsilon})(1+o(1))}, \quad \text{as } n \to \infty. \tag{4.11}$$

Combining (4.10) and (4.11) with (4.9) now gives (4.4) which proves (2.11).

Since the minimum of W_n is 1 (when $F_{k,n} = 1/n$ for all k) the assertion $\mathbf{P}(W_n < 1) = 0$ is clear. It remains to prove (2.12), show that the rate function is of the form (2.13) and justify the positivity of $\widetilde{I}(x)$ for $x \in (1,2)$.

For $x \in (1,2)$ using (2.10) we obtain

$$\mathbf{P}(W_n \leq x) = \mathbf{P}\left(\frac{\sum_{j=1}^n R_j}{\sqrt{n}\sqrt{\sum_{k=1}^n R_k^2}} \geq \frac{1}{\sqrt{x}}\right). \tag{4.12}$$

Furthermore, for $x \in (1,2)$ we have $1/\sqrt{x} > 1/\sqrt{2} = \mathbf{E}[R_1]/\sqrt{\mathbf{E}[R_1^2]}$. Thus, we can use Theorem 1.1 from [12] and obtain

$$\mathbf{P}(W_n \leq x)^{1/n} = \sup_{c \geq 0} \inf_{t \geq 0} \mathbf{E}\left[\exp\left(t(cR_1 - \frac{1}{2\sqrt{x}}(R_1^2 + c^2))\right)\right]. \tag{4.13}$$

Now we have

$$E\left[\exp\left(t(cR_1 - \frac{1}{2\sqrt{x}}(R_1^2 + c^2))\right)\right] = \int_0^\infty \exp\left(-y + t(cy - \frac{1}{2\sqrt{x}}(y^2 + c^2))\right) dy$$

and elementary integration yields

$$= \frac{\sqrt{2\pi}x^{1/4}}{\sqrt{t}} \exp\left(\frac{(tc-1)^2 x - t^2 c^2}{2t\sqrt{x}}\right) \Phi\left(\frac{(tc-1)x^{1/4}}{\sqrt{t}}\right),$$

where Φ denotes the distribution function of the one dimensional standard Gaussian distribution. Taking $\log$ of the last term we obtain (2.13).

Now we fix $x \in (1,2)$ and show that $\widetilde{I}(x)$ is positive. In the sequel we write

$$h(r,c) := cr - \frac{1}{2\sqrt{x}}(r^2 + c^2).$$

We have

$$\inf_{t \geq 0} \mathbf{E}\left[\exp\left(th(R_1, c)\right)\right]$$

$$\geq \mathbf{E}\left[\inf_{t \geq 0} \exp\left(th(R_1, c)\right)\right]$$

$$= \mathbf{E}\left[\mathbb{1}_{\{h(R_1,c)<0\}} \inf_{t \geq 0} \exp\left(th(R_1, c)\right)\right] + \mathbf{E}\left[\mathbb{1}_{\{h(R_1,c)\geq 0\}} \inf_{t \geq 0} \exp\left(th(R_1, c)\right)\right]$$

$$= \mathbf{P}\left(h(R_1, c) \geq 0\right).$$

The function $r \mapsto h(r,c)$ is non-negative on the interval $[r_1, r_2]$ where $r_{1/2} = r_{1/2}(c) := c(\sqrt{x} \pm \sqrt{x-1})$ are the zeros of the function. It follows

$$\mathbf{E}\left[\inf_{t \geq 0} \exp\left(th(R_1, c)\right)\right] = \mathbf{P}\left(r_1 \leq R_1 \leq r_2\right) = e^{-c(\sqrt{x}-\sqrt{x-1})} - e^{-c(\sqrt{x}+\sqrt{x-1})}.$$

Finally, by elementary calculation we obtain

$$\sup_{c \geq 0}\left(e^{-c(\sqrt{x}-\sqrt{x-1})} - e^{-c(\sqrt{x}+\sqrt{x-1})}\right) = \frac{2\sqrt{x-1}}{\sqrt{x}+\sqrt{x-1}}\left(\frac{\sqrt{x}-\sqrt{x-1}}{\sqrt{x}+\sqrt{x-1}}\right)^{\frac{\sqrt{x}-\sqrt{x-1}}{2\sqrt{x-1}}}.$$

This expression (and therefore also $\widetilde{I}(x)$) is positive for $x \in (1,2)$. Thus, the proof of Theorem 2.8 is concluded.

4.3 Proof of Theorem 2.10

We prove the inequality (2.15) using Lemma 2.9. Let $1 < x < 2$ and set $y = \frac{1}{\sqrt{x-1}} - 1$. For $\varepsilon > 0$ we have

$$
\begin{aligned}
&\frac{1}{n} \log \mathbf{P}\big(W_n \leq x + \varepsilon\big) \\
&= \frac{1}{n} \log \mathbf{P}\left(n \frac{2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}}{\left(\sum_{k=1}^{n} R_k\right)^2} \leq x + \varepsilon \right) \\
&\geq \frac{1}{n} \log \mathbf{P}\left(n \frac{2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}}{\left(\sum_{k=1}^{n} R_k\right)^2} \leq x + \varepsilon, R_n > ny \right) \\
&= \frac{1}{n} \log \left\{ \mathbf{P}\big(R_n > ny\big) \mathbf{P}\left(n \frac{2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}}{\left(\sum_{k=1}^{n} R_k\right)^2} \leq x + \varepsilon \;\middle|\; R_n > ny \right) \right\}.
\end{aligned}
$$

Now $\frac{1}{n} \log \mathbf{P}\big(R_n > ny\big) = -y$, and conditioning in the second factor in the curly braces can be removed by using the fact that conditioned on $R_n > ny$ the exponential random variable R_n has the same distribution as $ny + R_n$. After some elementary calculations we see that the last line of the above display equals

$$
-y + \frac{1}{n} \log \mathbf{P}\left(\frac{\frac{1}{n}\left(2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}\right) + 2y \frac{1}{n} \sum_{k=1}^{n} R_k + y^2}{\left(\frac{1}{n} \sum_{k=1}^{n} R_k\right)^2 + 2y \frac{1}{n} \sum_{k=1}^{n} R_k + y^2} \leq x + \varepsilon \right).
$$

From the strong law of large numbers and (2.4) with Lemma 2.9 we know that

$$
\frac{1}{n} \sum_{k=1}^{n} R_k \xrightarrow{n \to \infty} 1, \quad \text{and} \quad \frac{1}{n}\left(2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}\right) \xrightarrow{n \to \infty} 2 \quad \text{almost surely.}
$$

It follows that almost surely

$$
\frac{\frac{1}{n}\left(2 \sum_{l=1}^{n} \sum_{k=1}^{l} \frac{R_k R_l}{l} - \sum_{k=1}^{n} \frac{R_k^2}{k}\right) + 2y \frac{1}{n} \sum_{k=1}^{n} R_k + y^2}{\left(\frac{1}{n} \sum_{k=1}^{n} R_k\right)^2 + 2y \frac{1}{n} \sum_{k=1}^{n} R_k + y^2} \xrightarrow{n \to \infty} \frac{2 + 2y + y^2}{1 + 2y + y^2} = x.
$$

Thus,

$$
\liminf_{n \to \infty} \frac{1}{n} \log \mathbf{P}\big(W_n \leq 2 - x + \varepsilon\big) \geq -y = 1 - \frac{1}{\sqrt{x-1}}.
$$

The rest follows by letting $\varepsilon \downarrow 0$.

References

[1] David J. Aldous, *Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists*, Bernoulli **5** (1999), no. 1, 3–48. MR-1673235

[2] Omer Angel, Nathanaël Berestycki, and Vlada Limic, *Global divergence of spatial coalescents*, Probab. Theory Related Fields **152** (2012), no. 3-4, 625–679. MR-2892958

[3] Donald A. Dawson and Shui Feng, *Asymptotic behavior of the Poisson-Dirichlet distribution for large mutation rate*, Ann. Appl. Probab. **16** (2006), no. 2, 562–582. MR-2244425

[4] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010, Corrected reprint of the second (1998) edition. MR-2571413

[5] Andrej Depperschmidt, Andreas Greven, and Peter Pfaffelhuber, *Path-properties of the tree-valued Fleming-Viot process*, Electron. J. Probab. **18** (2013), no. 84, 1–47. MR-3109623

[6] Peter Donnelly, *Weak convergence to a Markov chain with an entrance boundary: ancestral processes in population genetics*, Ann. Probab. **19** (1991), no. 3, 1102–1117. MR-1112408

[7] Steven N. Evans, *Kingman's coalescent as a random metric space*, Stochastic models (Ottawa, ON, 1998), CMS Conf. Proc., vol. 26, Amer. Math. Soc., Providence, RI, 2000, pp. 105–114. MR-1765005

[8] Nina Gantert, Kavita Ramanan, and Franz Rembart, *Large deviations for weighted sums of stretched exponential random variables*, Electron. Commun. Probab. **19** (2014), no. 41, 14. MR-3233203

[9] J. F. C. Kingman, *The coalescent*, Stochastic Process. Appl. **13** (1982), no. 3, 235–248. MR-671034

[10] Jim Pitman, *Coalescents with multiple collisions*, Ann. Probab. **27** (1999), no. 4, 1870–1902. MR-1742892

[11] R. Pyke, *Spacings. (With discussion.)*, J. Roy. Statist. Soc. Ser. B **27** (1965), 395–449. MR-0216622

[12] Qi-Man Shao, *Self-normalized large deviations*, Ann. Probab. **25** (1997), no. 1, 285–328. MR-1428510

[13] Simon Tavaré, *Line-of-descent and genealogical processes, and their applications in population genetics models*, Theoret. Population Biol. **26** (1984), no. 2, 119–164. MR-770050