

Concentration inequalities for sums of random variables with power bounded tails

Oleksii Omelchenko and Andrei A. Bulatov

Abstract

In this paper we consider concentration bounds for sums of random variables, possibly with heavy tails. The paper is mostly based on the results of Borovkov and Borovkov where such issues have been thoroughly studied. While we are not claiming any significant new results, our goal is to make these results easier to use for combinatorial applications such as the Random Satisfiability problem. By slightly restricting the generality of the framework we considerably simplify and ‘discretize’ the proofs. At the same time we improve the bounds in the inequalities.

1 Introduction

Random combinatorial structures and related computational problems have been studied for decades. The majority of research have been focused on structures constructed or selected according to some sort of uniform distribution, as this is usually the most natural approach, and also most of the techniques work best in this case. However, many applied and real world problems are best modelled with non-uniform distributions [1, 2, 5, 13, 16, 7, 3]. So, in recent years there have been an increasing amount of work on structures sampled from less standard distributions, including heavy-tailed ones, such as the power law distribution [11, 8], or somewhat arbitrary distributions [9, 14, 15].

One of the main difficulties in studying distributions similar to the power law is that many standard tools are not available for such distributions. Concentration bounds like Chernoff Bound or Azuma inequality may not apply because heavy-tailed distribution lacks higher moments, or even the second or the first moment. More sophisticated approaches like Fourier analysis may be lost as well, as the Friedgut’s satisfiability threshold [10] demonstrates, that it is notoriously difficult to generalize beyond near-uniform distributions.

Heavy tail distributions have been studied in probability theory for decades [6]. In particular, some (though not very strong) concentration bounds can be found in [4]. It is therefore somewhat surprizing that such bounds are (to our best knowledge) not used in the computer science literature, instead substituted by ad hoc methods or some results working in special cases [12]. Apart from relative obscurity of these results from probability theory, a reason for that may be that the existing bounds tend to be proved in a very general setting, which, although being very powerful, often applies to continuous random variables, or does not give the kind of bounds needed in combinatorics.

In this paper we consider concentration bounds for sums of random variables, possibly with heavy tails. The paper is mostly based on the results of [4] where such issues have been thoroughly studied. While we are not claiming any significant new results, our goal is to make these results easier to use for combinatorial applications such as the Random Satisfiability problem. By slightly restricting the generality of the framework we considerably simplify and ‘discretize’ the proofs. At the same time we improve the bounds in the inequalities.

More precisely, we consider distributions X with integer values, whose tail functions F_{X+} (the right tail) and F_{X-} (the left tail) are majorized by power functions from $Vx^{-\alpha_r}$ and $Wx^{-\alpha_\ell}$, respectively. Note that such a distribution may have no first moment if $\min(\alpha_r, \alpha_\ell) \leq 1$, and it may have no second moment if $\min(\alpha_r, \alpha_\ell) \leq 2$.

We show that if $\min(\alpha_r, \alpha_\ell) \leq 1$, then the sum of such variables w.h.p. does not deviates much from the value one may expect (the mean value of such sum does not exists). Note that we do not assume that these variables are identically distributed.

Theorem 1. Let $S_n = \sum_{i=1}^n X_i$, where for each X_i it holds $F_{X_i+}(x) \leq Vx^{-\alpha_r}$, for some $0 < \alpha_r \leq 1$. Then for any $\epsilon > 0$,

$$\Pr \left[S_n \geq n^{\frac{1}{\alpha_r} + \epsilon} \right] \leq (V + e^{2V})n^{-\alpha_r \epsilon},$$

when $n \rightarrow \infty$.

In the case $\min(\alpha_r, \alpha_\ell) > 1$, the mean of the sum exists and prove a bound on the probability the deviates from it by a certain amount.

Theorem 2. Let $S_n = \sum_{i=1}^n X_i$, where for each X_i it holds $F_{X_i+}(x) \leq Vx^{-\alpha_r}$, $F_{X_i-}(x) \leq Wx^{-\alpha_\ell}$, for some $\alpha_r, \alpha_\ell > 1$. Then letting $\alpha = \min(\alpha_r, \alpha_\ell)$, for any $\epsilon > 0$,

$$\Pr \left[S_n - \mathbb{E}S_n \geq n^{\max(1/\alpha, 1/2) + \epsilon} \right] \leq Vn^{1 - \max(1, \alpha/2) - \alpha\epsilon} + e^{2V}n^{-\alpha\epsilon},$$

when $n \rightarrow \infty$.

Bounds for the left tail are similar.

The methods we use are fairly standard and boil down to careful evaluation of the tails of the sum of the X_i 's.

2 Notation and preliminaries

We say that some sequence of events $\{A_n\}$ happens with high probability (**w.h.p.**), when

$$\lim_{n \rightarrow \infty} \Pr[A_n] = 1.$$

Let $S_n = \sum_{i=1}^n X_i$ be the sum of n independent *not necessary* identically distributed integer-valued random variables X_i 's with tails that do not depend on n (probably this constraint could be relaxed to some extent, however, in this work we will deal only with variables, which constitute the sum, that do not depend on n).

First, we need to introduce a couple of useful concepts that we will exploit heavily in the subsequent chapters. Primarily, we will need *right-* and *left-tail functions*:

Definition 1. Let X be some random variable with support on $\mathbb{S} \subseteq \mathbb{R}$. Then the function

$$F_{X+}(x) := \Pr[X \geq x], \text{ where } x > 0$$

is the right-tail function of the r.v. X . Similarly, the left-tail function of the r.v. X is

$$F_{X-}(x) := \Pr[X \leq -x], \text{ where } x > 0.$$

Note, though, that the above definition is slightly broader than we need, since it applies to any numerical variable. However, in this paper, we deal with random variables that take integer values, meaning their support $\mathbb{S} = \mathbb{Z}$.

We do not require variables X_i 's to be identically distributed, but we do need some property that they all share in order to work with different distributions in a simple and unified way. For that purpose we harness the concept of *majorization* (or domination):

Definition 2. We say that a function $g(x)$ majorizes a function $f(x)$, if $g(x) \geq f(x)$ for every x from the domain of f .

We will focus on variables X_i 's with tails that can be majorized by some power functions with negative exponents. As it will be shown later, the values of these exponents play a critical role in the behaviour of the sum S_n .

Definition 3. Random variable X has a probability distribution function that belongs to the set $\mathbb{D}(\alpha_l, \alpha_r)$, if there exist constants $V > 0$ and $W > 0$ (which we call left- and right-tail constants of the respective r.v.), such that $F_{X^+}(x) \leq Vx^{-\alpha_r}$ and $F_{X^-}(x) \leq Wx^{-\alpha_l}$ for $x > 0$, where $\alpha_l, \alpha_r > 0$ are constants (we will call them left and right tail exponents or powers).

The fact that the distribution of a r.v. X is from $\mathbb{D}(\alpha_l, \alpha_r)$ will be denoted as

$$X_i \sim \mathbb{D}(\alpha_l, \alpha_r).$$

Although, it may seem that the variables from $\mathbb{D}(\alpha_l, \alpha_r)$ belong to a rather restrictive class of random variables, however, such majorization can be applied to a very broad range of variables, including gaussian, subexponential, heavy-tailed, and, obviously, power-law random variables. Therefore, all the results of this paper are applicable to these classes of random variables.

We will also say that the variable's distribution is from $\mathbb{D}(\cdot, \alpha_r)$ (or $\mathbb{D}(\alpha_l, \cdot)$), if the right tail is majorized by some power function with exponent $-\alpha_r$, while the left tail is arbitrary (or, respectively, when the left tail can be majorized by some power function with exponent $-\alpha_l$, while the right tail is arbitrary). Moreover, by $\mathbb{D}(> 1, \cdot)$ (or $\mathbb{D}(\cdot, > 1)$) we denote the set of distributions with left-tail (or right-tail) functions that can be bounded by some power function $Cx^{-\alpha}$, where $\alpha > 1$.

Additionally, when finite sequence $\{X_i\}_{i=1}^n$ of random variables consists of r.v.s, each of which belongs to the class $\mathbb{D}(\cdot, \alpha_{r,i})$, then we use two quantities:

$$\alpha_r = \min(\alpha_{r,1}, \alpha_{r,2}, \dots, \alpha_{r,n}),$$

and

$$V = \max(V_1, V_2, \dots, V_n),$$

where V_i 's are the constants from the majorizing power functions. Clearly, that each X_i 's right tail then can be majorized by a power function $Vx^{-\alpha_r}$. In a similar way we define α_l and W , that is when all $X_i \sim \mathbb{D}(\alpha_{l,i}, \cdot)$, then

$$\alpha_l = \min(\alpha_{l,1}, \alpha_{l,2}, \dots, \alpha_{l,n}),$$

and

$$W = \max(W_1, W_2, \dots, W_n).$$

Additionally, when all X_i 's come from the $\mathbb{D}(\alpha_{l,i}, \alpha_{r,i})$ classes, then we use another important quantity α , which is the minimum among all $\alpha_{l,i}$'s and $\alpha_{r,i}$'s, i.e.

$$\alpha = \min(\alpha_l, \alpha_r).$$

In this paper we show how harnessing the values of tail exponents of variables X_i 's allows us to bound probabilities for the sum $S_n = \sum X_i$ to have large values or to deviate much from its expected value (given that its expectation exists).

3 Tail bounds of S_n , when $0 < \min(\alpha_l, \alpha_r) \leq 1$

Tail inequalities in this section are of the most general nature, meaning they can be applied to **any** variables with tail functions that can be majorized by some power-law function $Ck^{-\alpha}$ with $\alpha > 0$. Unfortunately, there exists a trade-off between the specificity of the inequalities over random variables and the tightness of the bounds they provide, i.e. the more general the inequalities are, the wider bounds on the values of S_n they assert.

Theorem 3. Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\cdot, \alpha_{r,i})$ are independent not necessary identically distributed integer-valued random variables, and $0 < \alpha_r \leq 1$. Then for any $\epsilon > 0$, we have

$$\Pr[S_n \geq n^{\frac{1}{\alpha_r} + \epsilon}] \leq (V + e^{2V}) n^{-\alpha_r \epsilon}, \quad (1)$$

when $n \rightarrow \infty$.

Proof. First, as was mentioned previously, when all $X_i \sim X_i \sim \mathbb{D}(\cdot, \alpha_{r,i})$, then

$$F_{X_{i+}}(k) \leq V k^{-\alpha_r}, \quad (2)$$

which is valid for all $k \geq 1$ and any $1 \leq i \leq n$.

Next, let us introduce events $B_i := \{X_i \leq x\}$, which indicate that the variable X_i does not exceed x , and also the event B that none of the variables X_i 's exceed x , i.e.

$$B := \bigcap_{i=1}^n B_i.$$

Then we have

$$\begin{aligned} \Pr[S_n \geq x] &= \Pr[\{S_n \geq x\} \wedge \bar{B}] + \Pr[\{S_n \geq x\} \wedge B] \\ &\leq \Pr[\bar{B}] + \Pr[\{S_n \geq x\} \wedge B] \\ &\leq \Pr\left[\bigcup_{i=1}^n \bar{B}_i\right] + \Pr[S_n \geq x | B] \cdot \Pr[B]. \end{aligned} \quad (3)$$

Next,

$$\begin{aligned} \Pr\left[\bigcup_{i=1}^n \bar{B}_i\right] &= \Pr\left[\bigcup_{i=1}^n \{X_i > x\}\right] \\ &\leq \sum_{i=1}^n \Pr[X_i > x] \quad (\text{by Union bound}) \\ &\leq \sum_{i=1}^n \Pr[X_i \geq x] \\ &= \sum_{i=1}^n F_{X_{i+}}(x) \\ &\leq \sum_{i=1}^n V x^{-\alpha_r}, \quad (\text{from (2)}) \\ &= n V x^{-\alpha_r}, \end{aligned} \quad (4)$$

and since X_i 's are independent random variables, we also have

$$\Pr[B] = \Pr\left[\bigcap_{i=1}^n B_i\right] = \prod_{i=1}^n \Pr[X_i \leq x]. \quad (5)$$

Therefore, inequality (3) can be simplified further to

$$\begin{aligned} \Pr[S_n \geq x] &\leq \Pr\left[\bigcup_{i=1}^n \bar{B}_i\right] + \Pr[S_n \geq x | B] \cdot \Pr[B] \\ &\leq n V x^{-\alpha_r} + \Pr[S_n \geq x | B] \cdot \prod_{i=1}^n \Pr[X_i \leq x]. \end{aligned} \quad (6)$$

Now we take a closer look at the probability $\Pr[S_n \geq x | B]$. This sum consists of n independent random variables having conditionally the same distribution as the original random variables X_i 's but truncated at x . More formally, let's introduce new truncated random variables Y_i with the following probability distribution function

$$\Pr[Y_i = k] = \begin{cases} \frac{\Pr[X_i = k]}{\Pr[X_i \leq x]}, & \text{when } k \leq x \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\Pr[S_n \geq x | B] = \Pr\left[\sum_{i=1}^n Y_i \geq x\right] =: \Pr[S_n^{(x)} \geq x],$$

where $S_n^{(x)}$ is the sum of the Y_i 's. After introducing Y_i 's and $S_n^{(x)}$, we can rewrite (6) as

$$\begin{aligned} \Pr[S_n \geq x] &\leq nVx^{-\alpha_r} + \Pr[S_n \geq x | B] \cdot \prod_{i=1}^n \Pr[X_i \leq x] \\ &\leq nVx^{-\alpha_r} + \Pr[S_n^{(x)} \geq x] \cdot \prod_{i=1}^n \Pr[X_i \leq x]. \end{aligned} \quad (7)$$

Next, from Markov's inequality we know that for any non-negative random variable A , positive x and non-negative non-decreasing function $\phi(x)$ we have

$$\Pr[A \geq x] \leq \frac{\mathbb{E}\phi(A)}{\phi(x)}.$$

Applying this inequality to the $S_n^{(x)}$ and letting $\phi(x) = e^{\mu x}$ for some positive μ , which will be defined later (in what follows, we will maintain $\mu \rightarrow 0$ when $n \rightarrow \infty$, but $\mu x \rightarrow \infty$). Moreover, we assume that n is large enough for $\mu \leq 1$), we obtain

$$\Pr[S_n^{(x)} \geq x] \leq \frac{\mathbb{E}e^{\mu S_n^{(x)}}}{e^{\mu x}} = e^{-\mu x} \mathbb{E}e^{\mu \sum_{i=1}^n Y_i} = e^{-\mu x} \mathbb{E} \prod_{i=1}^n e^{\mu Y_i}.$$

Exploiting the fact that all the Y_i 's are independent random variables, we can further simplify the above probability:

$$\begin{aligned} \Pr[S_n^{(x)} \geq x] &\leq e^{-\mu x} \mathbb{E} \prod_{i=1}^n e^{\mu Y_i} \\ &= e^{-\mu x} \prod_{i=1}^n \mathbb{E}e^{\mu Y_i} \\ &= e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu k} \Pr[Y_i = k] \\ &= e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu k} \frac{\Pr[X_i = k]}{\Pr[X_i \leq x]} \\ &= e^{-\mu x} \frac{\prod_{i=1}^n \sum_{k \leq x} e^{\mu k} \Pr[X_i = k]}{\prod_{i=1}^n \Pr[X_i \leq x]}. \end{aligned}$$

Plugging this inequality for $\Pr[S_n^{(x)} \geq x]$ back into (7), we get

$$\begin{aligned} \Pr[S_n \geq x] &\leq nVx^{-\alpha_r} + \Pr[S_n^{(x)} \geq x] \cdot \prod_{i=1}^n \Pr[X_i \leq x] \\ &\leq nVx^{-\alpha_r} + e^{-\mu x} \frac{\prod_{i=1}^n \sum_{k \leq x} e^{\mu k} \Pr[X_i = k]}{\prod_{i=1}^n \Pr[X_i \leq x]} \cdot \prod_{i=1}^n \Pr[X_i \leq x] \\ &= nVx^{-\alpha_r} + e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu k} \Pr[X_i = k] \\ &=: nVx^{-\alpha_r} + e^{-\mu x} \prod_{i=1}^n R_i(\mu, x), \end{aligned}$$

where $R_i(\mu, x) := \sum_{k \leq x} e^{\mu k} \Pr[X_i = k]$. Clearly, in order to obtain the final bound, we need to upper bound $R_i(\mu, x)$. For that purpose, we split sum in $R_i(\mu, x)$ into 3 disjoint intervals, and bound each interval separately:

$$\begin{aligned} R_i(\mu, x) &= \sum_{k \leq x} e^{\mu k} \Pr[X_i = k] \\ &\leq \left[\sum_{k \leq 0} e^{\mu k} \Pr[X_i = k] \right] + \left[\sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i = k] \right] + \left[\sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i = k] \right] \\ &=: I_{i,1} + I_{i,2} + I_{i,3}, \end{aligned}$$

where $M = \frac{2\alpha_r}{\mu}$ (now it should be clear why we require $\mu x \rightarrow \infty$, since we would like to have $0 \ll M < x$).

3.1 Upper bound of $I_{i,1}$, when $0 < \alpha_r \leq 1$

The first interval is the easiest to bound, as the next lemma shows this.

Lemma 1. *Let $X_i \sim \mathbb{D}(\cdot, \alpha_{r,i})$ be an integer valued r.v. with $0 < \alpha_r \leq 1$ and an arbitrary left tail function. Then*

$$I_{i,1} = \sum_{k \leq 0} e^{\mu k} \Pr[X_i = k]$$

is bounded from above by

$$I_{i,1} \leq \Pr[X_i \leq 0].$$

Proof. The proof is straightforward. For $I_{i,1}$ we have

$$\begin{aligned} I_{i,1} &= \sum_{k \leq 0} e^{\mu k} \Pr[X_i = k] \\ &\leq \sum_{k \leq 0} \Pr[X_i = k] \\ &= \Pr[X_i \leq 0]. \end{aligned}$$

□

3.2 Upper bound of $I_{i,2}$, when $0 < \alpha_r \leq 1$

In order to bound the second interval, we apply a slightly more sophisticated approach than that for $I_{i,1}$. Here we first apply summation by parts to express $I_{i,2}$ as a function of the right tail $F_{X_i+}(k)$, and then we use the integral bound of summation to upper bound the $I_{i,2}$ interval.

Lemma 2. *Let $X_i \sim \mathbb{D}(\cdot, \alpha_{r,i})$ be an integer valued r.v. with $0 < \alpha_r \leq 1$ and an arbitrary left-tail function. Then*

$$I_{i,2} = \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i = k],$$

where $M = \frac{2\alpha_r}{\mu}$, is bounded from above by

$$I_{i,2} \leq \Pr[X_i \geq 1] + C_1(\alpha_r) T(\alpha_r, \mu),$$

where $C_1(\alpha_r) = 2Ve^{2\alpha_r+1}$, and $T(\alpha_r, \mu)$ is defined as

$$T(\alpha_r, \mu) = \begin{cases} \frac{4-\alpha_r}{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{5}{2}\mu - \frac{3}{2}\mu \ln \mu, & \text{when } \alpha_r = 1. \end{cases} \quad (8)$$

Proof. As was mentioned, we first apply summation by parts to $I_{i,2}$. Hence, we obtain

$$\begin{aligned}
I_{i,2} &= \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i = k] \\
&= \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \left(\Pr[X_i \leq k] - \Pr[X_i \leq k-1] \right) \\
&= \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i \leq k] - \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i \leq k-1] \\
&= \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i \leq k] - \sum_{k=0}^{\lfloor M \rfloor - 1} e^{\mu(k+1)} \Pr[X_i \leq k] \\
&= \left[\sum_{k=0}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i \leq k] - \Pr[X_i \leq 0] \right] - \left[\sum_{k=0}^{\lfloor M \rfloor} e^{\mu(k+1)} \Pr[X_i \leq k] - e^{\mu(\lfloor M \rfloor + 1)} \Pr[X_i \leq \lfloor M \rfloor] \right] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] - \Pr[X_i \leq 0] + e^{\mu(\lfloor M \rfloor + 1)} \Pr[X_i \leq \lfloor M \rfloor] \\
&\leq \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] - \Pr[X_i \leq 0] + e^{\mu(\lfloor M \rfloor + 1)}.
\end{aligned}$$

Next, from the obvious equality

$$\Pr[X_i \leq k-1] + \Pr[X_i \geq k] = 1 \text{ for any } k, \quad (9)$$

it follows that

$$\begin{aligned}
I_{i,2} &= \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] - \Pr[X_i \leq 0] + e^{\mu(\lfloor M \rfloor + 1)} \\
&= \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] - 1 + \Pr[X_i \geq 1] + e^{\mu(\lfloor M \rfloor + 1)} \\
&= \Pr[X_i \geq 1] - 1 + \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] + e^{\mu(\lfloor M \rfloor + 1)}.
\end{aligned}$$

One more time we apply equality (9) to obtain

$$\begin{aligned}
I_{i,2} &= \Pr[X_i \geq 1] - 1 + \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] + e^{\mu(\lfloor M \rfloor + 1)} \\
&= \Pr[X_i \geq 1] - 1 + \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \left(1 - \Pr[X_i \geq k+1] \right) + e^{\mu(\lfloor M \rfloor + 1)} \\
&= \Pr[X_i \geq 1] - 1 + \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) - \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] + e^{\mu(\lfloor M \rfloor + 1)} \\
&= \Pr[X_i \geq 1] - 1 + e^{\mu 0} - e^{\mu(\lfloor M \rfloor + 1)} - \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] + e^{\mu(\lfloor M \rfloor + 1)} \\
&= \Pr[X_i \geq 1] - \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] \\
&\leq \Pr[X_i \geq 1] + \sum_{k=0}^{\lfloor M \rfloor} \left(e^{\mu} - 1 \right) e^{\mu k} \Pr[X_i \geq k+1],
\end{aligned}$$

and since $e^\mu \leq 1 + 2\mu$ when μ is close to 0, the above sum transforms into

$$\begin{aligned}
I_{i,2} &= \Pr[X_i \geq 1] + \sum_{k=0}^{\lfloor M \rfloor} \left(e^\mu - 1 \right) e^{\mu k} \Pr[X_i \geq k+1] \\
&\leq \Pr[X_i \geq 1] + 2\mu \sum_{k=0}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i \geq k+1] \\
&= \Pr[X_i \geq 1] + 2\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu(k-1)} \Pr[X_i \geq k] \\
&= \Pr[X_i \geq 1] + 2\mu e^{-\mu} \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} \Pr[X_i \geq k] \\
&\leq \Pr[X_i \geq 1] + 2\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} \Pr[X_i \geq k] \\
&= \Pr[X_i \geq 1] + 2\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} F_{X_i+}(k).
\end{aligned}$$

Now recall that $F_{X_i+}(k) \leq V k^{-\alpha_r}$ with $0 < \alpha_r \leq 1$. Then

$$\begin{aligned}
I_{i,2} &\leq \Pr[X_i \geq 1] + 2\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} F_{X_i+}(k) \\
&\leq \Pr[X_i \geq 1] + 2\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} V k^{-\alpha_r} \\
&= \Pr[X_i \geq 1] + 2V\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} k^{-\alpha_r}.
\end{aligned}$$

We have $M = \frac{2\alpha_r}{\mu}$, hence,

$$\begin{aligned}
I_{i,2} &\leq \Pr[X_i \geq 1] + 2V\mu \sum_{k=1}^{\lfloor M \rfloor + 1} e^{\mu k} k^{-\alpha_r} \\
&\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + \mu} \mu \sum_{k=1}^{\lfloor M \rfloor + 1} k^{-\alpha_r} \\
&\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu \sum_{k=1}^{\lfloor M \rfloor + 1} k^{-\alpha_r}.
\end{aligned}$$

Function $k^{-\alpha_r}$ is monotonically decreasing, hence, we can apply the integral upper bound of a sum of decreasing function, i.e. for some decreasing in $[a \dots b]$ function $\psi(k)$, we have

$$\sum_{k=a}^b \psi(k) \leq \int_{a-1}^b \psi(t) dt. \tag{10}$$

Then

$$\begin{aligned}
I_{i,2} &\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu \sum_{k=1}^{\lfloor M \rfloor + 1} k^{-\alpha_r} \\
&= \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu \left(1 + \sum_{k=2}^{\lfloor M \rfloor + 1} k^{-\alpha_r} \right) \\
&= \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu + 2Ve^{\alpha_r + 1} \mu \sum_{k=2}^{\lfloor M \rfloor + 1} k^{-\alpha_r} \\
&\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu + 2Ve^{2\alpha_r + 1} \mu \int_1^{\lfloor M \rfloor + 1} t^{-\alpha_r} dt \\
&\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu + 2Ve^{2\alpha_r + 1} \mu \int_1^{M+1} t^{-\alpha_r} dt \\
&=: \Pr[X_i \geq 1] + 2Ve^{2\alpha_r + 1} \mu + 2Ve^{2\alpha_r + 1} I_{i,2}^0,
\end{aligned} \tag{11}$$

where $I_{i,2}^0 = \mu \int_1^{M+1} t^{-\alpha_r} dt$. Next, let's bound the $I_{i,2}^0$ term:

$$\begin{aligned}
I_{i,2}^0 &= \mu \int_1^{M+1} t^{-\alpha_r} dt \\
&= \mu \begin{cases} \frac{(M+1)^{1-\alpha_r}}{1-\alpha_r} - \frac{1}{1-\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \ln(M+1), & \text{when } \alpha_r = 1 \end{cases} \\
&\leq \begin{cases} \frac{\mu}{1-\alpha_r} (M+1)^{1-\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \mu \ln(M+1), & \text{when } \alpha_r = 1 \end{cases} \\
&\leq \frac{3}{2} \begin{cases} \mu \frac{M^{1-\alpha_r}}{1-\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \mu \ln M, & \text{when } \alpha_r = 1. \end{cases}
\end{aligned}$$

Numerical constant in front of $I_{i,2}^0$ comes from two obvious inequalities:

$$(M+1)^{1-\alpha_r} = \left(1 + \frac{1}{M}\right)^{1-\alpha_r} M^{1-\alpha_r} \leq \left(\frac{3}{2}\right)^{1-\alpha_r} M^{1-\alpha_r} \leq \frac{3}{2} M^{1-\alpha_r},$$

which is valid for any $0 < \alpha_r < 1$ and $M \geq 2$; similarly, easy to verify that

$$\ln(M+1) \leq \frac{3}{2} \ln M,$$

for every $M \geq 3$. Therefore, when $M = \frac{2\alpha_r}{\mu} \rightarrow \infty$, we have

$$\begin{aligned} I_{i,2}^0 &\leq \frac{3}{2} \begin{cases} \mu \frac{M^{1-\alpha_r}}{1-\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \ln M, & \text{when } \alpha_r = 1 \end{cases} \\ &= \begin{cases} \frac{3}{1-\alpha_r} 2^{-\alpha_r} \alpha_r^{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{3}{2} \ln(2\alpha_r) \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1. \end{cases} \end{aligned}$$

Observe that

$$\max_{0 < \alpha_r < 1} 2^{-\alpha_r} \alpha_r^{1-\alpha_r} \approx 0.55380... \leq 1,$$

hence,

$$\begin{aligned} I_{i,2}^0 &\leq \begin{cases} \frac{3}{1-\alpha_r} 2^{-\alpha_r} \alpha_r^{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{3}{2} \ln(2\alpha_r) \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1, \end{cases} \\ &\leq \begin{cases} \frac{3}{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{3}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1. \end{cases} \end{aligned}$$

Thus, from (11) we obtain

$$\begin{aligned} I_{i,2} &\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r+1} \mu + 2Ve^{2\alpha_r+1} I_{i,2}^0 \\ &\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r+1} \mu + 2Ve^{2\alpha_r+1} \begin{cases} \frac{3}{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{3}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1 \end{cases} \\ &= \Pr[X_i \geq 1] + 2Ve^{2\alpha_r+1} \begin{cases} \mu + \frac{3}{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \mu + \frac{3}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1. \end{cases} \end{aligned}$$

Now, since we assumed that $\mu \leq 1$, we further have

$$\begin{aligned} I_{i,2} &\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r+1} \begin{cases} \mu + \frac{3}{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \mu + \frac{3}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1 \end{cases} \\ &\leq \Pr[X_i \geq 1] + 2Ve^{2\alpha_r+1} \begin{cases} \left(1 + \frac{3}{1-\alpha_r}\right) \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{5}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1 \end{cases} \\ &= \Pr[X_i \geq 1] + 2Ve^{2\alpha_r+1} \begin{cases} \left(1 + \frac{3}{1-\alpha_r}\right) \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{5}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1 \end{cases} \\ &=: \Pr[X_i \geq 1] + C_1(\alpha_r) T(\alpha_r, \mu), \end{aligned}$$

where $C_1(\alpha_r) = 2Ve^{2\alpha_r+1}$ and

$$T(\alpha_r, \mu) = \begin{cases} \frac{4-\alpha_r}{1-\alpha_r} \mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \frac{5}{2} \mu - \frac{3}{2} \mu \ln \mu, & \text{when } \alpha_r = 1. \end{cases}$$

And the lemma is proved. \square

3.3 Upper bound of $I_{i,3}$

Proof for this interval's bound closely resembles the proof of the bound of the second interval $I_{i,2}$ differing only in some details. However, unlike the previous interval, the bound for $I_{i,3}$ presented below is valid for any $\alpha_r > 0$. We emphasize this observation, since result of Lemma 3 will be re-used to prove a bound for a similar interval when $\alpha_r > 1$ (see Lemma 10).

Lemma 3. *Let $X_i \sim \mathbb{D}(\cdot, \alpha_{r,i})$ be an integer valued r.v. with $\alpha_{r,i} > 0$ and an arbitrary left tail function. Then*

$$I_{i,3} = \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i = k],$$

where $M = \frac{2\alpha_r}{\mu}$, is bounded from above by

$$I_{i,3} \leq C_2(\alpha_r)\mu^{\alpha_r} + Ve^{\mu x}x^{-\alpha_r},$$

where $C_2(\alpha_r) := Ve^{2\alpha_r+1}(2\alpha_r)^{-\alpha_r}$.

Proof. First, just like with the previous interval $I_{i,2}$, we apply summation by parts in order to express $I_{i,3}$

as a function of the right-tail function of X_i :

$$\begin{aligned}
I_{i,3} &= \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i = k] \\
&= \sum_{k=\lceil M \rceil}^x e^{\mu k} \left(\Pr[X_i \leq k] - \Pr[X_i \leq k-1] \right) \\
&= \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i \leq k] - \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i \leq k-1] \\
&= \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i \leq k] - \sum_{k=\lceil M \rceil-1}^{x-1} e^{\mu(k+1)} \Pr[X_i \leq k] \\
&= \left[\sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} \Pr[X_i \leq k] + e^{\mu x} \Pr[X_i \leq x] \right] - \left[\sum_{k=\lceil M \rceil}^{x-1} e^{\mu(k+1)} \Pr[X_i \leq k] + e^{\mu \lceil M \rceil} \Pr[X_i \leq \lceil M \rceil - 1] \right] \\
&\leq e^{\mu x} - e^{\mu \lceil M \rceil} \Pr[X_i \leq \lceil M \rceil - 1] + \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \leq k] \\
&= e^{\mu x} - e^{\mu \lceil M \rceil} \Pr[X_i \leq \lceil M \rceil - 1] + \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \left(1 - \Pr[X_i \geq k+1] \right) \\
&= e^{\mu x} - e^{\mu \lceil M \rceil} \Pr[X_i \leq \lceil M \rceil - 1] + \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) - \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} - e^{\mu \lceil M \rceil} \Pr[X_i \leq \lceil M \rceil - 1] - \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} \left(1 - \Pr[X_i \leq \lceil M \rceil - 1] \right) - \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} \Pr[X_i \geq \lceil M \rceil] - \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} F_{X_i+}(\lceil M \rceil) - \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu k} - e^{\mu(k+1)} \right) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} F_{X_i+}(\lceil M \rceil) + \sum_{k=\lceil M \rceil}^{x-1} \left(e^{\mu(k+1)} - e^{\mu k} \right) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} F_{X_i+}(\lceil M \rceil) + \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} \left(e^{\mu} - 1 \right) \Pr[X_i \geq k+1].
\end{aligned}$$

Again, recall that $e^\mu \leq 1 + 2\mu$, whenever $\mu \rightarrow 0$. Then

$$\begin{aligned}
I_{i,3} &\leq e^{\mu \lceil M \rceil} F_{X_{i+}}(\lceil M \rceil) + \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} (e^\mu - 1) \Pr[X_i \geq k+1] \\
&\leq e^{\mu \lceil M \rceil} F_{X_{i+}}(\lceil M \rceil) + \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} (1 + 2\mu - 1) \Pr[X_i \geq k+1] \\
&= e^{\mu \lceil M \rceil} F_{X_{i+}}(\lceil M \rceil) + 2\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} \Pr[X_i \geq k+1] \\
&\leq e^{\mu \lceil M \rceil} F_{X_{i+}}(\lceil M \rceil) + 2\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} \Pr[X_i \geq k] \\
&= e^{\mu \lceil M \rceil} F_{X_{i+}}(\lceil M \rceil) + 2\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} F_{X_{i+}}(k),
\end{aligned}$$

and since $F_{X_{i+}}(x) \leq Vx^{-\alpha_r}$, we further have that

$$\begin{aligned}
I_{i,3} &\leq e^{\mu \lceil M \rceil} F_{X_{i+}}(\lceil M \rceil) + 2\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} F_{X_{i+}}(k) \\
&\leq e^{\mu \lceil M \rceil} V \lceil M \rceil^{-\alpha_r} + 2V\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} k^{-\alpha_r} \\
&\leq e^{2\alpha_r+1} VM^{-\alpha_r} + 2V\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} k^{-\alpha_r}, \text{ since } \mu \lceil M \rceil \leq \mu \left(\frac{2\alpha_r}{\mu} + 1 \right) \leq 2\alpha_r + 1 \\
&= e^{2\alpha_r+1} V \left(\frac{2\alpha_r}{\mu} \right)^{-\alpha_r} + 2V\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} k^{-\alpha_r} \\
&= Ve^{2\alpha_r+1} (2\alpha_r)^{-\alpha_r} \mu^{\alpha_r} + 2V\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} k^{-\alpha_r} \\
&=: C_2(\alpha_r) \mu^{\alpha_r} + 2V\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} k^{-\alpha_r},
\end{aligned}$$

where $C_2(\alpha_r) = Ve^{2\alpha_r+1} (2\alpha_r)^{-\alpha_r}$. Next, let's investigate the monotonicity of the function under summation $e^{\mu k} k^{-\alpha_r}$. We have

$$\frac{d}{dk} \left[e^{\mu k} k^{-\alpha_r} \right] = \mu e^{\mu k} k^{-\alpha_r} - \alpha_r e^{\mu k} k^{-\alpha_r-1} = \mu e^{\mu k} k^{-\alpha_r-1} \left(k - \frac{\alpha_r}{\mu} \right),$$

which is clearly positive when $k \geq 2\alpha_r/\mu = M$. Thus, when $k \geq M$, the function $e^{\mu k} k^{-\alpha_r}$ is monotonically increasing. Therefore, we can apply the integral upper bound of summation of an increasing function, i.e. for some increasing in $[a \dots b]$ function $\psi(k)$, we have

$$\sum_{k=a}^b \psi(k) \leq \int_a^{b+1} \psi(t) dt.$$

So

$$\begin{aligned}
I_{i,3} &\leq C_2(\alpha_r)\mu^{\alpha_r} + 2V\mu \sum_{k=\lceil M \rceil}^{x-1} e^{\mu k} k^{-\alpha_r} \\
&\leq C_2(\alpha_r)\mu^{\alpha_r} + 2V\mu \int_{\lceil M \rceil}^x e^{\mu t} t^{-\alpha_r} dt \\
&\leq C_2(\alpha_r)\mu^{\alpha_r} + 2V\mu \int_M^x e^{\mu t} t^{-\alpha_r} dt.
\end{aligned}$$

Using substitution of variables $u := \mu(x - t)$ and $dt = -\frac{1}{\mu} du$, the above integral can be transformed into

$$\begin{aligned}
I_{i,3} &\leq C_2(\alpha_r)\mu^{\alpha_r} + 2V\mu \int_M^x e^{\mu t} t^{-\alpha_r} dt \\
&= C_2(\alpha_r)\mu^{\alpha_r} - 2V \int_{\mu(x-M)}^0 e^{\mu x - u} \left(x - \frac{u}{\mu}\right)^{-\alpha_r} du \\
&= C_2(\alpha_r)\mu^{\alpha_r} + 2V e^{\mu x} x^{-\alpha_r} \int_0^{\mu(x-M)} e^{-u} \left(1 - \frac{u}{\mu x}\right)^{-\alpha_r} du. \tag{12}
\end{aligned}$$

Now consider the function $f(u) = \left(1 - \frac{u}{\mu x}\right)^{-\alpha_r}$. We have for $u \in [0 \cdots \mu(x - M)]$

$$\begin{aligned}
\frac{d}{du} \ln f(u) &= \frac{d}{du} \ln \left(1 - \frac{u}{\mu x}\right)^{-\alpha_r} \\
&= -\alpha_r \frac{d}{du} \ln \left(1 - \frac{u}{\mu x}\right) \\
&= \frac{\alpha_r}{\mu x - u} \\
&\leq \frac{1}{2},
\end{aligned}$$

where the last inequality follows from the fact that function $\frac{\alpha_r}{\mu x - u}$ reaches its maximum at the rightmost point when $u = \mu(x - M) = \mu x - 2\alpha_r$. Furthermore, $f(0) = 1$, hence, $f(u) \leq e^{u/2}$, when $0 \leq u \leq \mu(x - M)$.

Then (12) can be bounded as

$$\begin{aligned}
I_{i,3} &= C_2(\alpha_r)\mu^{\alpha_r} + 2Ve^{\mu x}x^{-\alpha_r} \int_0^{\mu(x-M)} e^{-u} \left(1 - \frac{u}{\mu x}\right)^{-\alpha_r} du \\
&= C_2(\alpha_r)\mu^{\alpha_r} + 2Ve^{\mu x}x^{-\alpha_r} \int_0^{\mu(x-M)} e^{-u} f(u) du \\
&\leq C_2(\alpha_r)\mu^{\alpha_r} + 2Ve^{\mu x}x^{-\alpha_r} \int_0^{\mu(x-M)} e^{-u} e^{u/2} du \\
&= C_2(\alpha_r)\mu^{\alpha_r} + 2Ve^{\mu x}x^{-\alpha_r} \int_0^{\mu(x-M)} e^{-u/2} du \\
&\leq C_2(\alpha_r)\mu^{\alpha_r} + 2Ve^{\mu x}x^{-\alpha_r} \int_0^{\infty} e^{-u/2} du \\
&= C_2(\alpha_r)\mu^{\alpha_r} + Ve^{\mu x}x^{-\alpha_r},
\end{aligned}$$

and, therefore, the lemma proved. \square

3.4 Final assembling steps, when $0 < \alpha_r \leq 1$

Recall, that the goal was to bound

$$\Pr[S_n \geq x] \leq nVx^{-\alpha_r} + e^{-\mu x} \prod_{i=1}^n R_i(\mu, x), \quad (13)$$

where $x = n^{\frac{1}{\alpha_r} + \epsilon}$, and

$$R_i(\mu, x) \leq I_{i,1} + I_{i,2} + I_{i,3}.$$

After proving Lemmas 1, 2, and 3, we can obtain final bounds of $\Pr[S_n \geq x]$. As was shown,

$$\begin{aligned}
I_{i,1} &\leq \Pr[X_i \leq 0], \\
I_{i,2} &\leq \Pr[X_i \geq 1] + C_1(\alpha_r)T(\alpha_r, \mu) \\
I_{i,3} &\leq C_2(\alpha_r)\mu^{\alpha_r} + Ve^{\mu x}x^{-\alpha_r}.
\end{aligned}$$

Hence,

$$R_i(\mu, x) = I_{i,1} + I_{i,2} + I_{i,3} \quad (14)$$

$$\begin{aligned}
&\leq \Pr[X_i \leq 0] + \Pr[X_i \geq 1] + C_1(\alpha_r)T(\alpha_r, \mu) + C_2(\alpha_r)\mu^{\alpha_r} + Ve^{\mu x}x^{-\alpha_r} \\
&= \left[\Pr[X_i \leq 0] + \Pr[X_i \geq 1] \right] + C_1(\alpha_r)T(\alpha_r, \mu) + C_2(\alpha_r)\mu^{\alpha_r} + Ve^{\mu x}x^{-\alpha_r} \\
&= 1 + C_1(\alpha_r)T(\alpha_r, \mu) + C_2(\alpha_r)\mu^{\alpha_r} + Ve^{\mu x}x^{-\alpha_r} \\
&=: 1 + T_0(\alpha_r, \mu) + Ve^{\mu x}x^{-\alpha_r}, \quad (15)
\end{aligned}$$

where we defined

$$\begin{aligned}
T_0(\alpha_r, \mu) &= C_1(\alpha_r)T(\alpha_r, \mu) + C_2(\alpha_r)\mu^{\alpha_r} \\
&= \begin{cases} C_2(\alpha_r)\mu^{\alpha_r} + C_1(\alpha_r)\frac{4-\alpha_r}{1-\alpha_r}\mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ C_2(\alpha_r)\mu^{\alpha_r} + \frac{5C_1(\alpha_r)}{2}\mu - \frac{3C_1(\alpha_r)}{2}\mu \ln \mu, & \text{when } \alpha_r = 1 \end{cases} \\
&= \begin{cases} \left(C_2(\alpha_r) + C_1(\alpha_r)\frac{4-\alpha_r}{1-\alpha_r}\right)\mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ \left(C_2(\alpha_r) + \frac{5C_1(\alpha_r)}{2}\right)\mu - \frac{3C_1(\alpha_r)}{2}\mu \ln \mu, & \text{when } \alpha_r = 1 \end{cases} \\
&=: \begin{cases} C'_1(\alpha_r)\mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ C'_2(\alpha_r)\mu - C'_3(\alpha_r)\mu \ln \mu, & \text{when } \alpha_r = 1, \end{cases}
\end{aligned}$$

where

$$C'_1(\alpha_r) = C_2(\alpha_r) + C_1(\alpha_r)\frac{4-\alpha_r}{1-\alpha_r}, \quad (16)$$

$$C'_2(\alpha_r) = C_2(\alpha_r) + \frac{5C_1(\alpha_r)}{2}, \quad (17)$$

$$C'_3(\alpha_r) = \frac{3C_1(\alpha_r)}{2}. \quad (18)$$

Since $1 + x \leq e^x$ for every x , we further have from (15)

$$\begin{aligned}
R_i(\mu, x) &\leq 1 + T_0(\alpha_r, \mu) + Ve^{\mu x}x^{-\alpha_r} \\
&\leq \exp\left(T_0(\alpha_r, \mu) + Ve^{\mu x}x^{-\alpha_r}\right),
\end{aligned}$$

and then (13) transforms into

$$\begin{aligned}
\Pr[S_n \geq x] &\leq nVx^{-\alpha_r} + e^{-\mu x} \prod_{i=1}^n R_i(\mu, x) \\
&\leq nVx^{-\alpha_r} + e^{-\mu x} \prod_{i=1}^n \exp\left(T_0(\alpha_r, \mu) + Ve^{\mu x}x^{-\alpha_r}\right) \\
&= nVx^{-\alpha_r} + e^{-\mu x} \exp\left(nT_0(\alpha_r, \mu) + Ve^{\mu x}nx^{-\alpha_r}\right) \\
&= nVx^{-\alpha_r} + \exp\left(-\mu x + nT_0(\alpha_r, \mu) + Ve^{\mu x}nx^{-\alpha_r}\right). \quad (19)
\end{aligned}$$

Next, we need to fix the value of μ , such that the above exponent is minimized as much as possible, while keeping $\mu \rightarrow 0$ but $\mu x \rightarrow \infty$ to make sure that $M = 2\alpha_r/\mu$ is much greater than 0, yet less than x . One such possible value is $\mu = \frac{1}{x} \ln \frac{x^{\alpha_r}}{n}$. The next lemma verifies that the chosen value of μ satisfies both constraints, when $0 < \alpha_r \leq 1$.

Lemma 4. *Let $x = n^{\frac{1}{\alpha_r} + \epsilon}$ with $0 < \alpha_r \leq 1$ and any $\epsilon > 0$. Then*

$$\mu = \frac{1}{x} \ln \frac{x^{\alpha_r}}{n} \rightarrow 0,$$

but

$$\mu x = \ln \frac{x^{\alpha_r}}{n} \rightarrow \infty,$$

when $n \rightarrow \infty$.

Proof. Simple calculation shows that

$$\begin{aligned}
\mu &= \frac{1}{x} \ln \frac{x^{\alpha_r}}{n} \\
&= x^{-1} (\alpha_r \ln x - \ln n) \\
&= n^{-\frac{1}{\alpha_r} - \epsilon} \left(\alpha_r \ln n^{\frac{1}{\alpha_r} + \epsilon} - \ln n \right) \\
&= n^{-\frac{1}{\alpha_r} - \epsilon} ((1 + \alpha_r \epsilon) \ln n - \ln n) \\
&\leq \alpha_r \epsilon n^{-\frac{1}{\alpha_r} - \epsilon} \ln n \\
&= o(1).
\end{aligned}$$

In a similar way we prove that $\mu x \rightarrow \infty$.

$$\begin{aligned}
\mu x &= \frac{x}{x} \ln \frac{x^{\alpha_r}}{n} \\
&= \alpha_r \ln x - \ln n \\
&= \alpha_r \ln n^{\frac{1}{\alpha_r} + \epsilon} - \ln n \\
&= (1 + \alpha_r \epsilon) \ln n - \ln n \\
&= \alpha_r \epsilon \ln n \rightarrow \infty,
\end{aligned}$$

when $n \rightarrow \infty$. □

Next, after fixing μ , we analyze the exponent in (19)

$$\begin{aligned}
\Pr[S_n \geq x] &\leq nVx^{-\alpha_r} + \exp\left(-\mu x + nT_0(\alpha_r, \mu) + Ve^{\mu x}nx^{-\alpha_r}\right) \\
&= nVx^{-\alpha_r} + \exp\left(nT_0(\alpha_r, \mu) + (Ve^{\mu x}nx^{-\alpha_r} - \mu x)\right)
\end{aligned}$$

by studying asymptotic behaviour of its two components, i.e. $nT_0(\alpha_r, \mu)$ and $Ve^{\mu x}nx^{-\alpha_r} - \mu x$.

Lemma 5. *Let $x = n^{\frac{1}{\alpha_r} + \epsilon}$ and $\mu = \frac{1}{x} \ln \frac{x^{\alpha_r}}{n}$ with $0 < \alpha_r \leq 1$ and any $\epsilon > 0$. Then*

$$nT_0(\alpha_r, \mu) = o(1),$$

where

$$T_0(\alpha_r, \mu) = \begin{cases} C'_1(\alpha_r)\mu^{\alpha_r}, & \text{when } 0 < \alpha_r < 1, \\ C'_2(\alpha_r)\mu - C'_3(\alpha_r)\mu \ln \mu, & \text{when } \alpha_r = 1. \end{cases}$$

Proof. Let us consider cases. When $0 < \alpha_r < 1$, we have $T_0(\alpha_r, \mu) = C'_1(\alpha_r)\mu^{\alpha_r}$, and so

$$\begin{aligned}
nT_0(\alpha_r, \mu) &= C'_1(\alpha_r)n\mu^{\alpha_r} \\
&= C'_1(\alpha_r)\frac{n}{x^{\alpha_r}} \ln^{\alpha_r} \frac{x^{\alpha_r}}{n} \\
&= C'_1(\alpha_r)\frac{n}{n^{1+\alpha_r\epsilon}} \ln^{\alpha_r} \frac{n^{1+\alpha_r\epsilon}}{n} \\
&= O\left(n^{-\alpha_r\epsilon} \ln^{\alpha_r} n^{\alpha_r\epsilon}\right) \\
&= o(1).
\end{aligned}$$

However, when $\alpha_r = 1$, then $T_0(\alpha_r, \mu) = C'_2(\alpha_r)\mu - C'_3(\alpha_r)\mu \ln \mu$, while

$$\begin{aligned}
nT_0(\alpha_r, \mu) &= C'_2(\alpha_r)n\mu - C'_3(\alpha_r)n\mu \ln \mu \\
&= C'_2(\alpha_r)\frac{n}{x} \ln \frac{x^{\alpha_r}}{n} - C'_3(\alpha_r)\frac{n}{x} \ln \frac{x^{\alpha_r}}{n} \cdot \ln \frac{1}{x} \ln \frac{x^{\alpha_r}}{n} \\
&= C'_2(\alpha_r)\frac{n}{n^{1/\alpha_r+\epsilon}} \ln \frac{n^{1+\alpha_r\epsilon}}{n} - C'_3(\alpha_r)\frac{n}{n^{1/\alpha_r+\epsilon}} \ln \frac{n^{1+\alpha_r\epsilon}}{n} \cdot \ln \left(\frac{1}{n^{1/\alpha_r+\epsilon}} \ln \frac{n^{1+\alpha_r\epsilon}}{n} \right) \\
&= C'_2(1)\frac{n}{n^{1+\epsilon}} \ln \frac{n^{1+\epsilon}}{n} - C'_3(1)\frac{n}{n^{1+\epsilon}} \ln \frac{n^{1+\epsilon}}{n} \cdot \ln \left(\frac{1}{n^{1+\epsilon}} \ln \frac{n^{1+\epsilon}}{n} \right), \text{ since } \alpha_r = 1 \\
&= C'_2(1)\epsilon n^{-\epsilon} \ln n - C'_3(1)\epsilon n^{-\epsilon} \ln n \cdot \ln (n^{-\epsilon} \ln n^\epsilon) \\
&= C'_2(1)\epsilon n^{-\epsilon} \ln n + C'_3(1)\epsilon^2 n^{-\epsilon} \ln^2 n - C'_3(1)\epsilon n^{-\epsilon} \ln n \cdot \ln \ln n^\epsilon \\
&\leq C'_2(1)\epsilon n^{-\epsilon} \ln n + C'_3(1)\epsilon^2 n^{-\epsilon} \ln^2 n, \text{ since } C'_3(1)\epsilon n^{-\epsilon} \ln n \cdot \ln \ln n^\epsilon > 0 \\
&= O(n^{-\epsilon} \ln^2 n) \\
&= o(1).
\end{aligned}$$

Thus, after combining two cases, we see that $nT_0(\alpha_r, \mu) \rightarrow 0$, and so the lemma follows. \square

However, unlike the $nT_0(\alpha_r, \mu)$ term, which approaches 0, when $n \rightarrow \infty$, the term $(Ve^{\mu x}nx^{-\alpha_r} - \mu x) \rightarrow -\infty$, and the next lemma states this fact in a more rigorous way.

Lemma 6. *Let $x = n^{\frac{1}{\alpha_r}+\epsilon}$ and $\mu = \frac{1}{x} \ln \frac{x^{\alpha_r}}{n}$ with $0 < \alpha_r \leq 1$ and any $\epsilon > 0$. Then*

$$Ve^{\mu x}nx^{-\alpha_r} - \mu x = V - \alpha_r \epsilon \ln n.$$

Proof. Clearly,

$$\begin{aligned}
Ve^{\mu x}nx^{-\alpha_r} - \mu x &= Ve^{\frac{x}{x} \ln \frac{x^{\alpha_r}}{n}} nx^{-\alpha_r} - \frac{x}{x} \ln \frac{x^{\alpha_r}}{n} \\
&= V - \ln \frac{x^{\alpha_r}}{n} \\
&= V - \ln \frac{n^{1+\alpha_r\epsilon}}{n} \\
&= V - \alpha_r \epsilon \ln n.
\end{aligned}$$

\square

Hence, after collecting results of Lemmas 5 and 6, we can finally show that $\Pr[S_n \geq x]$ from (19) is bounded from above by

And so inequality (19) gets its final form

$$\begin{aligned}
\Pr[S_n \geq x] &= nVx^{-\alpha_r} + \exp \left(nT_0(\alpha_r, \mu) + (Ve^{\mu x}nx^{-\alpha_r} - \mu x) \right) \\
&\leq nVx^{-\alpha_r} + \exp \left(o(1) + V - \alpha_r \epsilon \ln n \right) \\
&\leq nVx^{-\alpha_r} + \exp \left(2V - \alpha_r \epsilon \ln n \right) \\
&= nVx^{-\alpha_r} + e^{2V} n^{-\alpha_r \epsilon},
\end{aligned}$$

and after recalling that $x = n^{\frac{1}{\alpha_r}+\epsilon}$, we obtain the final form of the above inequality

$$\Pr[S_n \geq n^{\frac{1}{\alpha_r}+\epsilon}] \leq (V + e^{2V}) n^{-\alpha_r \epsilon},$$

which proves Theorem 3. \square

Theorem 3 implies an obvious corollary that asserts an upper bound of values for S_n , when every X_i in the aforementioned sum has a right-tail function that can be bounded by $V_i x^{-\alpha_{r,i}}$ with $0 < \alpha_{r,i} \leq 1$:

Corollary 1. Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\cdot, \alpha_{r,i})$ are independent not necessary identically distributed integer-valued random variables, and $0 < \alpha_r \leq 1$. Then w.h.p.

$$S_n \leq C n^{1/\alpha_r},$$

where $C > 0$ is some constant.

Note, that in Theorem 3, where we were interested in S_n exceeding some *positive* $x = n^{\frac{1}{\alpha_r} + \epsilon}$, we completely ignored the left tails of X_i 's. However, if we were trying to bound S_n from below, then the left tails of the random variables X_i 's play a vital role, and the following theorem verifies this fact.

Theorem 4. Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\alpha_{l,i}, \cdot)$ are independent not necessary identically distributed integer-valued random variables, and $0 < \alpha_{l,i} \leq 1$. Then for any $\epsilon > 0$, we have

$$\Pr \left[S_n \leq -n^{\frac{1}{\alpha_l} + \epsilon} \right] \leq (W + e^{2W}) n^{-\alpha_l \epsilon},$$

when $n \rightarrow \infty$.

Proof. After establishing Theorem 3, proof of the left tail bound becomes easy. First, let's introduce random variables X'_i that have the same distributions as $-X_i$, i.e. $X'_i \stackrel{d}{=} -X_i$. Clearly $X'_i \sim \mathbb{D}(\cdot, \alpha_l)$, hence, from Definition 3, it follows that $F_{X'_i+}(k) \leq W X_i k^{-\alpha_{l,i}}$.

Then for $x = n^{1/\alpha_r + \epsilon}$ we have

$$\begin{aligned} \Pr [S_n \leq -x] &= \Pr \left[\sum_{i=1}^n X_i \leq -x \right] \\ &= \Pr \left[-\sum_{i=1}^n X_i \geq x \right] \\ &= \Pr \left[\sum_{i=1}^n -X_i \geq x \right] \\ &= \Pr \left[\sum_{i=1}^n X'_i \geq x \right] \\ &= \Pr \left[S'_n \geq x \right], \text{ where } S'_n := \sum_{i=1}^n X'_i \\ &\leq (W + e^{2W}) n^{-\alpha_l \epsilon}, \end{aligned}$$

where the last inequality follows after applying Theorem 3 to the sum S'_n , which consists of random variables with the right tail functions that can be bounded by some $W x^{-\alpha_l}$. What is left is to recall that $x = n^{1/\alpha_r + \epsilon}$, and the left-tail bound is established. \square

Similarly to the right-tail bound of S_n , we can state a corollary that restricts the lower range of values of S_n .

Corollary 2. Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\alpha_{l,i}, \cdot)$ are independent not necessary identically distributed integer-valued random variables, and $0 < \alpha_l \leq 1$. Then w.h.p.

$$-C n^{1/\alpha_l} \leq S_n,$$

where $C > 0$ is some constant.

4 Tail bounds of $S_n - \mathbb{E}S_n$, when $\alpha_l, \alpha_r > 1$

As was mentioned previously, Theorem 3 and Theorem 4 can be used to upper and lower bound the sum of integer-valued random variables, when the variables have left- or right-tail function that can be bounded in

the "best" case by $Ck^{-\alpha}$ with constants $C > 0$ and $0 < \alpha \leq 1$. However, if the variables, which the sum consists of, have distributions from $\mathbb{D}(\alpha_l, \alpha_r)$ with both $\alpha_l > 1$ and $\alpha_r > 1$, then, these variables must have finite expectations, and therefore, the sum itself has finite expectation; moreover, as we show in Theorems 5 and 6, the sum does not deviate much from its expected value. This result is summarized in Corollaries 4 and ??.

Note, that results in this section require that both tails of every variable X_i can be bounded by a power function with power $-\alpha$, where $\alpha > 1$.

First, let's verify that the variable $X \sim \mathbb{D}(> 1, > 1)$ has finite expectation. For that we provide an alternative way to compute expectation of random variables, which is a generalization of the tail sum of expectation [?]:

Lemma 7 (Generalized Tails Sum Formula). *Let X be a random variable with support on \mathbb{Z} , such that $\mathbb{E}X$ exists. Then*

$$\mathbb{E}X = \sum_{j=1}^{\infty} F_{X^+}(j) - \sum_{j=1}^{\infty} F_{X^-}(j).$$

Proof. From the definition of expectation, it follows that

$$\begin{aligned} \mathbb{E}X &= \sum_k k \Pr[X = k] \\ &= \sum_{k \leq -1} k \Pr[X = k] + \sum_{k \geq 1} k \Pr[X = k] \\ &= \sum_{k=1}^{\infty} (-k) \Pr[X = -k] + \sum_{k=1}^{\infty} k \Pr[X = k] \\ &= \sum_{k=1}^{\infty} k \Pr[X = k] - \sum_{k=1}^{\infty} k \Pr[X = -k] \\ &= \sum_{k=1}^{\infty} \sum_{j=1}^k \Pr[X = k] - \sum_{k=1}^{\infty} \sum_{j=1}^k \Pr[X = -k] \\ &= \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} \Pr[X = k] - \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} \Pr[X = -k] \\ &= \sum_{j=1}^{\infty} \Pr[X \geq j] - \sum_{j=1}^{\infty} \Pr[X \leq -j]. \end{aligned}$$

Recall that $F_{X^+}(j) := \Pr[X_i \geq j]$ and $F_{X^-}(j) := \Pr[X_i \leq -j]$ for every $j \geq 1$, and, therefore,

$$\mathbb{E}X = \sum_{j=1}^{\infty} \Pr[X \geq j] - \sum_{j=1}^{\infty} \Pr[X \leq -j] \tag{20}$$

$$= \sum_{j=1}^{\infty} F_{X^+}(j) - \sum_{j=1}^{\infty} F_{X^-}(j). \tag{21}$$

The lemma is proved. □

Now, after establishing an alternative way for calculating the expected value of a random variable, we can state a simple corollary, which verifies that random variables from $\mathbb{D}(> 1, > 1)$ have finite expectations.

Corollary 3. *An integer-valued r.v. $X \sim \mathbb{D}(> 1, > 1)$ has finite expectation.*

Proof. Since $X \sim \mathbb{D}(> 1, > 1)$, we have that

$$F_{X^+}(k) \leq V k^{-\alpha_r}, \quad \text{and} \quad F_{X^-}(k) \leq W k^{-\alpha_l} \text{ for any } k > 0,$$

where $V, W > 0$ and $\alpha_l, \alpha_r > 1$.

Next, introduce quantities

$$R = \sum_{k=1}^{\infty} F_{X_+}(k) \quad \text{and} \quad L = \sum_{k=1}^{\infty} F_{X_-}(k).$$

Since $0 \leq F_{X_+}(k) \leq V k^{-\alpha_r}$ and $\alpha_r > 1$, we have that

$$R = \sum_{k=1}^{\infty} F_{X_+}(k) \leq V \sum_{k=1}^{\infty} k^{-\alpha_r} < \infty.$$

Moreover, since R is the sum of non-negative terms, we obtain that $0 \leq R < \infty$. The same idea we apply to $|L|$.

Since both $0 \leq R, |L| < \infty$ are finite, we obtain from 7 that

$$|\mathbb{E}X| = \sum_{j=1}^{\infty} F_{X_+}(j) - \sum_{j=1}^{\infty} F_{X_-}(j) \leq \left| \sum_{j=1}^{\infty} F_{X_+}(j) - \sum_{j=1}^{\infty} F_{X_-}(j) \right| = |R - L| \leq |R| + |L| < \infty.$$

Hence, when $X \sim \mathbb{D}(> 1, > 1)$, then $|\mathbb{E}X| < \infty$. □

Therefore, every random variable $X \sim \mathbb{D}(\alpha_l, \alpha_r)$ with $\alpha_l, \alpha_r > 1$ has finite expectation. We can exploit this knowledge in order to obtain tighter bounds for the sum S_n . We show that whenever random variables have tail functions that can be bounded by $Vx^{-\alpha}$ with $\alpha > 1$, then the sum of such variables is concentrated around its mean.

The next two theorems assert this fact, where first we show that S_n does not deviate much to the right from $\mathbb{E}S_n$, while the second theorem states a similar result but for the deviation to the left from the expectation.

Theorem 5. *Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\alpha_{l,i}, \alpha_{r,i})$ are independent not necessary identically distributed integer-valued random variables with $\alpha_{l,i}, \alpha_{r,i} > 1$. Then for any $\epsilon > 0$, we have*

$$\Pr \left[S_n - \mathbb{E}S_n \geq n^{\max(1/\alpha, 1/2) + \epsilon} \right] \leq V n^{1 - \max(1, \alpha/2) - \alpha\epsilon} + e^{2V} n^{-\alpha\epsilon},$$

when $n \rightarrow \infty$.

Proof. The basic idea of the proof is similar to that of Theorem 3, differing only in some details. First, from Lemma 3, it follows that $|\mathbb{E}X_i| < \infty$, and so

$$\mathbb{E}S_n = \mathbb{E} \sum_{i=1}^n X_i = \sum_{i=1}^n \mathbb{E}X_i$$

is finite as well. Next, since each $X_i \sim \mathbb{D}(\alpha_{l,i}, \alpha_{r,i})$ and $\alpha = \min(\alpha_l, \alpha_r)$, then, obviously,

$$F_{X_{i+}}(k) \leq V k^{-\alpha} \quad \text{and} \quad F_{X_{i-}}(k) \leq W k^{-\alpha}$$

for any $k \geq 1$.

Also we introduce a set of events $B_i := \{X_i \leq x\}$ and

$$B := \bigcap_{i=1}^n B_i.$$

Then, like in the previous proof, we have

$$\begin{aligned} \Pr[S_n \geq \mathbb{E}S_n + x] &= \Pr[\{S_n \geq \mathbb{E}S_n + x\} \wedge \bar{B}] + \Pr[\{S_n \geq \mathbb{E}S_n + x\} \wedge B] \\ &\leq \Pr[\bar{B}] + \Pr[\{S_n \geq \mathbb{E}S_n + x\} \wedge B] \\ &\leq \Pr[\bar{B}] + \Pr[S_n \geq \mathbb{E}S_n + x \mid B] \cdot \Pr[B]. \end{aligned}$$

Since event B is the intersection of independent events $\{X_i \leq x\}$, we have

$$\Pr[B] = \prod_{i=1}^n \Pr[X_i \leq x];$$

and to bound the probability of event \bar{B} we apply Union bound (like we did in (4))

$$\Pr[\bar{B}] \leq nVx^{-\alpha}.$$

Hence, the above inequality of the probability $\Pr[S_n \geq \mathbb{E}S_n + x]$ can be further simplified

$$\begin{aligned} \Pr[S_n \geq \mathbb{E}S_n + x] &\leq \Pr[\bar{B}] + \Pr[S_n \geq \mathbb{E}S_n + x | B] \cdot \Pr[B] \\ &\leq nVx^{-\alpha} + \Pr[S_n \geq \mathbb{E}S_n + x | B] \cdot \prod_{i=1}^n \Pr[X_i \leq x]. \end{aligned}$$

Next, consider the sum S_n in probability $\Pr[S_n \geq \mathbb{E}S_n + x | B]$. This sum consists of n independent random variables having conditionally the same distribution as the original variable X_i but truncated at x . Formally speaking, let's introduce new "nearly-centered"¹ truncated random variables Z_i with the following probability distribution function

$$\Pr[Z_i = k - \mathbb{E}X_i] = \begin{cases} \frac{\Pr[X_i=k]}{\Pr[X_i \leq x]}, & \text{when } k \leq x \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

Then

$$\Pr[S_n \geq \mathbb{E}S_n + x | B] = \Pr\left[\sum_{i=1}^n Z_i \geq x\right] =: \Pr\left[S_n^{(x)} \geq x\right],$$

where Z_i 's are the truncated versions of the respective r.v. X_i 's, and $S_n^{(x)} = \sum_{i=1}^n Z_i$. So now we have

$$\begin{aligned} \Pr[S_n \geq \mathbb{E}S_n + x] &\leq nVx^{-\alpha} + \Pr[S_n \geq \mathbb{E}S_n + x | B] \cdot \prod_{i=1}^n \Pr[X_i \leq x] \\ &= nVx^{-\alpha} + \Pr\left[S_n^{(x)} \geq x\right] \cdot \prod_{i=1}^n \Pr[X_i \leq x]. \end{aligned} \quad (23)$$

As we did before, we apply generalized Markov's inequality to the probability of the sum of truncated variables to obtain the following inequality

$$\Pr\left[S_n^{(x)} \geq x\right] \leq \frac{\mathbb{E}e^{\mu S_n^{(x)}}}{e^{\mu x}},$$

which holds for any $\mu \geq 0$, however, in what follows, we will require that $\mu \rightarrow 0$ when $n \rightarrow \infty$ (moreover, we will assume that n is large enough for $\mu \leq 1$). Hence,

$$\begin{aligned} \Pr\left[S_n^{(x)} \geq x\right] &\leq \frac{\mathbb{E}e^{\mu S_n^{(x)}}}{e^{\mu x}} \\ &= e^{-\mu x} \mathbb{E}e^{\mu \sum_{i=1}^n Z_i} \\ &= e^{-\mu x} \prod_{i=1}^n \mathbb{E}e^{\mu Z_i}, \text{ since all } Z_i \text{'s are independent} \\ &= e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu(k - \mathbb{E}X_i)} \Pr[Z_i = k - \mathbb{E}X_i]. \end{aligned}$$

¹The new variable is not fully centred, since $\mathbb{E}Z_i < \mathbb{E}X_i$

Next, we use the definition of the variable Z_i from (22)

$$\begin{aligned}
\Pr \left[S_n^{(x)} \geq x \right] &\leq e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu(k - \mathbb{E}X_i)} \Pr[Z_i = k - \mathbb{E}X_i] \\
&\leq e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu(k - \mathbb{E}X_i)} \frac{\Pr[X_i = k]}{\Pr[X_i \leq x]} \\
&= e^{-\mu x} \frac{\prod_{i=1}^n \sum_{k \leq x} e^{\mu(k - \mathbb{E}X_i)} \Pr[X_i = k]}{\prod_{i=1}^n \Pr[X_i \leq x]}
\end{aligned}$$

Thus, we can substitute $\Pr[S_n^{(x)} \geq x]$ in (23) with the above calculated inequality to obtain

$$\begin{aligned}
\Pr[S_n \geq \mathbb{E}S_n + x] &\leq nVx^{-\alpha} + \Pr \left[S_n^{(x)} \geq x \right] \cdot \prod_{i=1}^n \Pr[X_i \leq x] \\
&\leq nVx^{-\alpha} + e^{-\mu x} \frac{\prod_{i=1}^n \sum_{k \leq x} e^{\mu(k - \mathbb{E}X_i)} \Pr[X_i = k]}{\prod_{i=1}^n \Pr[X_i \leq x]} \cdot \prod_{i=1}^n \Pr[X_i \leq x] \\
&= nVx^{-\alpha} + e^{-\mu x} \prod_{i=1}^n \sum_{k \leq x} e^{\mu(k - \mathbb{E}X_i)} \Pr[X_i = k] \\
&= nVx^{-\alpha} + e^{-\mu x} \prod_{i=1}^n e^{-\mu \mathbb{E}X_i} \sum_{k \leq x} e^{\mu k} \Pr[X_i = k] \\
&=: nVx^{-\alpha} + e^{-\mu x} \prod_{i=1}^n P_i(\mu, x). \tag{24}
\end{aligned}$$

where

$$P_i(\mu, x) := \sum_{k \leq x} e^{\mu k} \Pr[X_i = k].$$

To obtain a bound on $P_i(\mu, x)$, we split its summation into 3 disjoint intervals, and bound each interval separately

$$\begin{aligned}
P_i(\mu, x) &= \sum_{k \leq x} e^{\mu k} \Pr[X_i = k] \\
&\leq \sum_{k \leq 0} e^{\mu k} \Pr[X_i = k] + \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i = k] + \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i = k] \\
&=: J_{i,1} + J_{i,2} + J_{i,3}, \tag{25}
\end{aligned}$$

where $M = \frac{2\alpha}{\mu}$ (again, since we would like to have $0 \ll M < x$, we require $\mu x \rightarrow \infty$), and the next three sections identify bounds for each interval, and the fourth section assembles upper bounds of $J_{i,1}$, $J_{i,2}$, and $J_{i,3}$ to provide the final bound for the probability $\Pr[S_n - \mathbb{E}S_n \geq x]$.

4.1 Upper bound of $J_{i,1}$, when $\alpha > 1$

Observe, that $P_i(\mu, x)$ is multiplied by $e^{-\mu \mathbb{E}X_i}$ term in (24). When $\mathbb{E}X_i \geq 0$, this shouldn't cause any troubles, however, if $\mathbb{E}X_i < 0$, then clearly $e^{-\mu \mathbb{E}X_i} = 1 + \epsilon'$ with $\epsilon' > 0$. Taking into account that $e^{-\mu \mathbb{E}X_i} P_i(\mu, x)$ stands under product operator, this extra $(1 + \epsilon')$ term, when raised to the power of $n \rightarrow \infty$ may cause troubles if not dealt with properly.

So our goal, beside obtaining bounds of $P_i(\mu, x)$ in terms of tail functions, is to "extract" and include into the bound a term that will eventually diminish the $e^{-\mu \mathbb{E}X_i}$ multiplier.

The next lemma demonstrates how we achieve this goal by bounding the $J_{i,1}$ interval with the tail functions *and* partial expectation of X_i (the other part of $\mathbb{E}X_i$ will be included in the second interval).

Lemma 8. Let $X_i \sim \mathbb{D}(\alpha, \alpha)$ be an integer valued r.v. with $\alpha > 1$. Then the term

$$J_{i,1} = \sum_{k \leq 0} e^{\mu k} \Pr[X_i = k]$$

is bounded from above by

$$J_{i,1} \leq \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + W\mathcal{T}_0(\alpha, \mu),$$

where

$$\mathcal{T}_0(\alpha, \mu) = \begin{cases} C_4(\alpha)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 3\mu^2 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_5(\alpha)\mu^2, & \text{when } \alpha > 2 \end{cases}$$

with constants $C_4(\alpha) = \left(2 + \frac{1}{(\alpha-1)(2-\alpha)}\right)$ and $C_5(\alpha) = \left(\frac{\alpha-1}{\alpha-2} + 2\right)$.

Proof. We have

$$\begin{aligned} J_{i,1} &= \sum_{k \leq 0} e^{\mu k} \Pr[X_i = k] \\ &= \sum_{k \leq 0} \left((1 + \mu k) + e^{\mu k} - (1 + \mu k) \right) \Pr[X_i = k] \\ &= \sum_{k \leq 0} \Pr[X_i = k] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + \sum_{k \leq 0} \left(e^{\mu k} - 1 - \mu k \right) \Pr[X_i = k] \\ &= \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + \sum_{k \leq 0} \left(e^{\mu k} - 1 - \mu k \right) \Pr[X_i = k]. \end{aligned} \quad (26)$$

Consider the rightmost sum. First, let's denote by $\phi(k)$ the function

$$\phi(k) := e^{\mu k} - 1 - \mu k, \quad (27)$$

and then apply summation by parts

$$\begin{aligned} \sum_{k \leq 0} \phi(k) \Pr[X_i = k] &= \sum_{k \leq 0} \phi(k) \left(\Pr[X_i \leq k] - \Pr[X_i \leq k-1] \right) \\ &= \sum_{k \leq 0} \phi(k) \Pr[X_i \leq k] - \sum_{k \leq 0} \phi(k) \Pr[X_i \leq k-1] \\ &= \sum_{k \leq 0} \phi(k) \Pr[X_i \leq k] - \sum_{k \leq -1} \phi(k+1) \Pr[X_i \leq k] \\ &= \phi(0) \Pr[X_i \leq 0] + \sum_{k \leq -1} \phi(k) \Pr[X_i \leq k] - \sum_{k \leq -1} \phi(k+1) \Pr[X_i \leq k]. \end{aligned}$$

However, $\phi(0) = 0$, and so the sum simplifies to

$$\begin{aligned} \sum_{k \leq 0} \phi(k) \Pr[X_i = k] &= \phi(0) \Pr[X_i \leq 0] + \sum_{k \leq -1} \phi(k) \Pr[X_i \leq k] - \sum_{k \leq -1} \phi(k+1) \Pr[X_i \leq k] \\ &= \sum_{k \leq -1} \left(\phi(k) - \phi(k+1) \right) \Pr[X_i \leq k] \\ &= \sum_{k=1}^{\infty} \left(\phi(-k) - \phi(1-k) \right) \Pr[X_i \leq -k] \\ &= \sum_{k=1}^{\infty} \left(\phi(-k) - \phi(1-k) \right) F_{X_i-}(k). \end{aligned}$$

Next, for every X_i we have

$$F_{X_i-}(k) \leq Wk^{-\alpha} \text{ for } k \geq 1,$$

and, therefore,

$$\begin{aligned} \sum_{k \leq 0} \phi(k) \Pr[X_i = k] &= \sum_{k=1}^{\infty} \left(\phi(-k) - \phi(1-k) \right) F_{X_i-}(k) \\ &\leq W \sum_{k=1}^{\infty} \left(\phi(-k) - \phi(1-k) \right) k^{-\alpha}. \end{aligned}$$

After substituting back the function we denoted by ϕ (27), we obtain that

$$\begin{aligned} \phi(-k) - \phi(1-k) &= e^{-\mu k} - 1 + \mu k - e^{\mu(1-k)} + 1 + \mu(1-k) \\ &= e^{-\mu k} - e^{\mu(1-k)} + \mu \\ &= e^{-\mu k} (1 - e^{\mu}) + \mu \\ &\leq -\mu e^{-\mu k} + \mu, \quad \text{since } e^{\mu} \geq 1 + \mu \\ &= \mu (1 - e^{-\mu k}). \end{aligned}$$

Thus,

$$\begin{aligned} \sum_{k \leq 0} \phi(k) \Pr[X_i = k] &\leq W \sum_{k=1}^{\infty} \left(\phi(-k) - \phi(1-k) \right) k^{-\alpha} \\ &\leq W\mu \sum_{k=1}^{\infty} (1 - e^{-\mu k}) k^{-\alpha} \\ &\leq W\mu \left(\sum_{k=1}^{\lfloor 1/\mu \rfloor} (1 - e^{-\mu k}) k^{-\alpha} + \sum_{k=\lceil 1/\mu \rceil}^{\infty} (1 - e^{-\mu k}) k^{-\alpha} \right). \end{aligned}$$

Here we can split the sum into two intervals, since $\mu \leq 1$. Then

$$\begin{aligned} \sum_{k \leq 0} \phi(k) \Pr[X_i = k] &\leq W\mu \left(\sum_{k=1}^{\lfloor 1/\mu \rfloor} (1 - e^{-\mu k}) k^{-\alpha} + \sum_{k=\lceil 1/\mu \rceil}^{\infty} (1 - e^{-\mu k}) k^{-\alpha} \right) \\ &\leq W\mu \left(\sum_{k=1}^{\lfloor 1/\mu \rfloor} (1 - e^{-\mu k}) k^{-\alpha} + \sum_{k=\lceil 1/\mu \rceil}^{\infty} k^{-\alpha} \right) \\ &\leq W\mu \left(\mu \sum_{k=1}^{\lfloor 1/\mu \rfloor} k^{1-\alpha} + \sum_{k=\lceil 1/\mu \rceil}^{\infty} k^{-\alpha} \right), \text{ since } e^{-\mu k} \geq 1 - \mu k. \end{aligned}$$

Now we have two sums of strictly decreasing functions, which we bound using the integral upper bound of summation:

$$\begin{aligned}
\sum_{k \leq 0} \phi(k) \Pr[X_i = k] &\leq W\mu \left(\mu \sum_{k=1}^{\lfloor 1/\mu \rfloor} k^{1-\alpha} + \sum_{k=\lceil 1/\mu \rceil}^{\infty} k^{-\alpha} \right) \\
&\leq W\mu \left(\mu + \mu \sum_{k=2}^{\lfloor 1/\mu \rfloor} k^{1-\alpha} + (\lceil 1/\mu \rceil)^{-\alpha} + \sum_{k=\lceil 1/\mu \rceil+1}^{\infty} k^{-\alpha} \right) \\
&\leq W\mu \left(\mu + \mu \int_1^{\lfloor 1/\mu \rfloor} x^{1-\alpha} dx + \mu^\alpha + \int_{\lceil 1/\mu \rceil}^{\infty} x^{-\alpha} dx \right) \\
&\leq W\mu \left(\mu + \mu \int_1^{1/\mu} x^{1-\alpha} dx + \mu^\alpha + \int_{1/\mu}^{\infty} x^{-\alpha} dx \right) \\
&\leq W\mu \left(\mu + \mu^\alpha + \frac{1}{\alpha-1} \mu^{\alpha-1} + \mu \int_1^{1/\mu} x^{1-\alpha} dx \right) \\
&\leq W \left(\mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha + \mu^2 \int_1^{1/\mu} x^{1-\alpha} dx \right) \\
&=: W \left(\mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha + \mu^2 \sigma(\alpha, \mu) \right),
\end{aligned}$$

where we define $\sigma(\alpha, \mu) = \int_1^{1/\mu} x^{1-\alpha} dx$, which is upper bounded by

$$\begin{aligned}
\sigma(\alpha, \mu) &= \int_1^{1/\mu} x^{1-\alpha} dx \\
&= \begin{cases} \frac{\mu^{\alpha-2}}{2-\alpha} - \frac{1}{2-\alpha}, & \text{when } 1 < \alpha < 2, \\ -\ln \mu, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2} - \frac{\mu^{\alpha-2}}{\alpha-2}, & \text{when } \alpha > 2 \end{cases} \\
&\leq \begin{cases} \frac{1}{2-\alpha} \mu^{\alpha-2}, & \text{when } 1 < \alpha < 2, \\ -\ln \mu, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2}, & \text{when } \alpha > 2. \end{cases}
\end{aligned}$$

Thus, we have that

$$\begin{aligned}
\sum_{k \leq 0} \phi(k) \Pr[X_i = k] &\leq W \left(\mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha + \mu^2 \sigma(\alpha, \mu) \right) \\
&\leq W \begin{cases} \mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha + \frac{1}{2-\alpha} \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ \mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ \mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha + \frac{1}{\alpha-2} \mu^2, & \text{when } \alpha > 2. \end{cases} \\
&= W \begin{cases} \mu^2 + \mu^{\alpha+1} + \frac{1}{(\alpha-1)(2-\alpha)} \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 2\mu^2 + \mu^3 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ \frac{\alpha-1}{\alpha-2} \mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha, & \text{when } \alpha > 2. \end{cases}
\end{aligned}$$

Recall, that we have $\mu \leq 1$, therefore, we can further simplify the above relation

$$\begin{aligned} \sum_{k \leq 0} \phi(k) \Pr[X_i = k] &\leq W \begin{cases} \mu^2 + \mu^{\alpha+1} + \frac{1}{(\alpha-1)(2-\alpha)} \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 2\mu^2 + \mu^3 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ \frac{\alpha-1}{\alpha-2} \mu^2 + \mu^{\alpha+1} + \frac{1}{\alpha-1} \mu^\alpha, & \text{when } \alpha > 2 \end{cases} \\ &\leq W \begin{cases} \left(2 + \frac{1}{(\alpha-1)(2-\alpha)}\right) \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 3\mu^2 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ \left(\frac{\alpha-1}{\alpha-2} + 2\right) \mu^2, & \text{when } \alpha > 2 \end{cases} \\ &=: W \begin{cases} C_4(\alpha) \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 3\mu^2 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_5(\alpha) \mu^2, & \text{when } \alpha > 2, \end{cases} \end{aligned}$$

where $C_4(\alpha) = \left(2 + \frac{1}{(\alpha-1)(2-\alpha)}\right)$, and $C_5(\alpha) = \left(\frac{\alpha-1}{\alpha-2} + 2\right)$.

Next, recall that $\phi(k) = e^{\mu k} - 1 - \mu k$ (27), and

$$\sum_{k \leq 0} \phi(k) \Pr[X_i = k] = \sum_{k \leq 0} (e^{\mu k} - 1 - \mu k) \Pr[X_i = k]$$

is an additive term in the upper bound of $J_{i,1}$ (26). Therefore,

$$\begin{aligned} J_{i,1} &= \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + \sum_{k \leq 0} (e^{\mu k} - 1 - \mu k) \Pr[X_i = k] \\ &= \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + \sum_{k \leq 0} \phi(k) \Pr[X_i = k] \\ &\leq \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + W \begin{cases} C_4(\alpha) \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 3\mu^2 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_5(\alpha) \mu^2, & \text{when } \alpha > 2, \end{cases} \\ &=: \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + W \mathcal{T}_0(\alpha, \mu), \end{aligned}$$

where

$$\mathcal{T}_0(\alpha, \mu) = \begin{cases} C_4(\alpha) \mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 3\mu^2 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_5(\alpha) \mu^2, & \text{when } \alpha > 2. \end{cases}$$

and lemma is proved. \square

4.2 Upper bound of $J_{i,2}$, when $\alpha > 1$

The basic idea of the proof in this section resembles the one presented in Lemma 8. First and foremost, we add and subtract the $(1 + \mu k)$ term to obtain the positive part of the $\mathbb{E}X_i$, and after that we apply summation by parts followed with simple integral bound.

Lemma 9. *Let $X_i \sim \mathbb{D}(\alpha, \alpha)$ be an integer valued r.v. with $\alpha > 1$. When $M = \frac{2\alpha}{\mu}$, then the term*

$$J_{i,2} = \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i = k],$$

is bounded from above by

$$J_{i,2} \leq \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + 2V e^{2\alpha} \mu^2 + V \mathcal{T}_1(\alpha, \mu),$$

where

$$\mathcal{T}_1(\alpha, \mu) = \begin{cases} C_6(\alpha)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ C_7(\alpha)\mu^2 - C_8(\alpha)\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_9(\alpha)\mu^2, & \text{when } \alpha > 2, \end{cases}$$

with

$$\begin{aligned} C_6(\alpha) &= 2V + Ve^{1+e/2} + Ve^{2\alpha} \frac{(2\alpha)^{2-\alpha}}{2-\alpha}, \\ C_7(\alpha) &= 2V + Ve^{1+e/2} + Ve^4 \ln 2, \\ C_8(\alpha) &= Ve^4, \\ C_9(\alpha) &= 2V + Ve^{1+e/2} + \frac{Ve^{2\alpha}}{\alpha-2}. \end{aligned}$$

Proof. For $J_{i,2}$ we have

$$\begin{aligned} J_{i,2} &= \sum_{k=1}^{\lfloor M \rfloor} e^{\mu k} \Pr[X_i = k] \\ &= \sum_{k=1}^{\lfloor M \rfloor} \left(1 + \mu k + e^{\mu k} - 1 - \mu k\right) \Pr[X_i = k] \\ &= \sum_{k=1}^{\lfloor M \rfloor} \Pr[X_i = k] + \mu \sum_{k=1}^{\lfloor M \rfloor} k \Pr[X_i = k] + \sum_{k=1}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 - \mu k\right) \Pr[X_i = k] \\ &\leq \sum_{k=1}^{\infty} \Pr[X_i = k] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + \sum_{k=1}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 - \mu k\right) \Pr[X_i = k] \\ &= \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + \sum_{k=1}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 - \mu k\right) \Pr[X_i = k] \\ &=: \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + \mathcal{T}_1(\alpha, \mu), \end{aligned} \tag{28}$$

where $J_{i,2}^0 = \sum_{k=1}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 - \mu k\right) \Pr[X_i = k]$. To bound the $J_{i,2}^0$ term, we apply Abel's transformation to perform summation by parts, however, first let's denote as before

$$\phi(k) := e^{\mu k} - 1 - \mu k,$$

and then

$$\begin{aligned} J_{i,2}^0 &= \sum_{k=1}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 - \mu k\right) \Pr[X_i = k] \\ &= \sum_{k=1}^{\lfloor M \rfloor} \phi(k) \Pr[X_i = k] \\ &= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i = k], \end{aligned}$$

since $\phi(0) = 0$. Next,

$$\begin{aligned}
J_{i,2}^0 &= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i = k] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \left(\Pr[X_i \leq k] - \Pr[X_i \leq k-1] \right) \\
&= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i \leq k] - \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i \leq k-1] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i \leq k] - \sum_{k=-1}^{\lfloor M \rfloor-1} \phi(k+1) \Pr[X_i \leq k] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i \leq k] - \sum_{k=0}^{\lfloor M \rfloor-1} \phi(k+1) \Pr[X_i \leq k],
\end{aligned}$$

again, because $\phi(0) = 0$, and therefore,

$$\begin{aligned}
J_{i,2}^0 &= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i \leq k] - \sum_{k=0}^{\lfloor M \rfloor-1} \phi(k+1) \Pr[X_i \leq k] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \phi(k) \Pr[X_i \leq k] - \sum_{k=0}^{\lfloor M \rfloor} \phi(k+1) \Pr[X_i \leq k] + \phi(\lfloor M \rfloor + 1) \Pr[X_i \leq \lfloor M \rfloor] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k) - \phi(k+1) \right) \Pr[X_i \leq k] + \phi(\lfloor M \rfloor + 1) \Pr[X_i \leq \lfloor M \rfloor] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k) - \phi(k+1) \right) \left(1 - \Pr[X_i > k] \right) + \phi(\lfloor M \rfloor + 1) \Pr[X_i \leq \lfloor M \rfloor] \\
&= \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k) - \phi(k+1) \right) + \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k+1) - \phi(k) \right) \Pr[X_i > k] + \phi(\lfloor M \rfloor + 1) \Pr[X_i \leq \lfloor M \rfloor] \\
&= \phi(0) - \phi(\lfloor M \rfloor + 1) + \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k+1) - \phi(k) \right) \Pr[X_i > k] + \phi(\lfloor M \rfloor + 1) \Pr[X_i \leq \lfloor M \rfloor] \\
&= \phi(\lfloor M \rfloor + 1) \Pr[X_i \leq \lfloor M \rfloor] - \phi(\lfloor M \rfloor + 1) + \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k+1) - \phi(k) \right) \Pr[X_i > k] \\
&\leq \sum_{k=0}^{\lfloor M \rfloor} \left(\phi(k+1) - \phi(k) \right) \Pr[X_i > k] \\
&\leq \sum_{k=1}^{\lfloor M \rfloor+1} \left(\phi(k) - \phi(k-1) \right) \Pr[X_i \geq k].
\end{aligned}$$

Now, since $\phi(k) = e^{\mu k} - 1 - \mu k$, we further have

$$\begin{aligned}
J_{i,2}^0 &\leq \sum_{k=1}^{\lfloor M \rfloor + 1} \left(\phi(k) - \phi(k-1) \right) \Pr[X_i \geq k] \\
&= \sum_{k=1}^{\lfloor M \rfloor + 1} \left(e^{\mu k} - 1 - \mu k - e^{\mu(k-1)} + 1 + \mu(k-1) \right) \Pr[X_i \geq k] \\
&= \sum_{k=1}^{\lfloor M \rfloor + 1} \left(e^{\mu k} - e^{\mu(k-1)} - \mu \right) \Pr[X_i \geq k] \\
&= \sum_{k=1}^{\lfloor M \rfloor + 1} \left(e^{\mu k} (1 - e^{-\mu}) - \mu \right) \Pr[X_i \geq k] \\
&\leq \sum_{k=1}^{\lfloor M \rfloor + 1} \left(\mu e^{\mu k} - \mu \right) \Pr[X_i \geq k], \text{ since } e^{-\mu} \geq 1 - \mu \\
&\leq \mu \sum_{k=1}^{\lfloor M \rfloor + 1} \left(e^{\mu k} - 1 \right) \Pr[X_i \geq k].
\end{aligned}$$

Next, from the definition of the right-tail function and the way we defined V and α , it follows that

$$\Pr[X_i \geq k] =: F_{X_i+}(k) \leq V k^{-\alpha} \text{ for every } k \geq 1.$$

Then

$$\begin{aligned}
J_{i,2}^0 &\leq \mu \sum_{k=1}^{\lfloor M \rfloor + 1} \left(e^{\mu k} - 1 \right) \Pr[X_i \geq k] \\
&\leq V \mu \sum_{k=1}^{\lfloor M \rfloor + 1} \left(e^{\mu k} - 1 \right) k^{-\alpha} \\
&\leq V \mu \left(e^\mu - 1 \right) + V \mu \left(e^{\mu(\lfloor M \rfloor + 1)} - 1 \right) (\lfloor M \rfloor + 1)^{-\alpha} + V \mu \sum_{k=2}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 \right) k^{-\alpha} \\
&\leq V \mu \left(e^\mu - 1 \right) + V \mu \left(e^{\mu(M+1)} - 1 \right) M^{-\alpha} + V \mu \sum_{k=2}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 \right) k^{-\alpha}.
\end{aligned}$$

Since $M = 2\alpha/\mu$, we further have

$$\begin{aligned}
J_{i,2}^0 &\leq V \mu \left(e^\mu - 1 \right) + V \mu \left(e^{\mu(M+1)} - 1 \right) M^{-\alpha} + V \mu \sum_{k=2}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 \right) k^{-\alpha} \\
&\leq V \mu \left(e^\mu - 1 \right) + V \mu^{\alpha+1} \left(e^{2\alpha+\mu} - 1 \right) (2\alpha)^{-\alpha} + V \mu \sum_{k=2}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 \right) k^{-\alpha}.
\end{aligned}$$

Now, since $\mu \leq 1$, we have that $e^\mu \leq 1 + 2\mu$, and so

$$\begin{aligned}
J_{i,2}^0 &\leq V \mu \left(e^\mu - 1 \right) + V \mu^{\alpha+1} \left(e^{2\alpha+\mu} - 1 \right) (2\alpha)^{-\alpha} + V \mu \sum_{k=2}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 \right) k^{-\alpha} \\
&\leq 2V \mu^2 + V \mu^{\alpha+1} e^{2\alpha+1} (2\alpha)^{-\alpha} + V \mu \sum_{k=2}^{\lfloor M \rfloor} \left(e^{\mu k} - 1 \right) k^{-\alpha}.
\end{aligned}$$

Furthermore, it is easy to verify that

$$\max_{\alpha > 1} [e^{2\alpha+1}(2\alpha)^{-\alpha}] = e^{1+e/2},$$

when $\alpha = \frac{e}{2}$. Thus, we obtain that

$$\begin{aligned} J_{i,2}^0 &\leq 2V\mu^2 + V\mu^{\alpha+1}e^{2\alpha+1}(2\alpha)^{-\alpha} + V\mu \sum_{k=2}^{\lfloor M \rfloor} (e^{\mu k} - 1)k^{-\alpha} \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + V\mu \sum_{k=2}^{\lfloor M \rfloor} (e^{\mu k} - 1)k^{-\alpha}. \end{aligned}$$

Next, observe that the function $f(k) = e^{\mu k} - 1$ is convex, hence, we can bound it from above with a straight line $l(k) = \frac{k}{M}(e^{2\alpha} - 1)$ when $k \in [0 \dots M]$. Therefore,

$$\begin{aligned} J_{i,2}^0 &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + V\mu \sum_{k=2}^{\lfloor M \rfloor} (e^{\mu k} - 1)k^{-\alpha} \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + V\mu \frac{e^{2\alpha} - 1}{M} \sum_{k=2}^{\lfloor M \rfloor} k^{1-\alpha} \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + V\mu^2 \frac{e^{2\alpha} - 1}{2\alpha} \sum_{k=2}^{\lfloor M \rfloor} k^{1-\alpha} \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \sum_{k=2}^{\lfloor M \rfloor} k^{1-\alpha}. \end{aligned}$$

Next, we apply the integral upper bound of summation of a decreasing function (10). So for $\alpha > 1$ we have

$$\begin{aligned} J_{i,2}^0 &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \sum_{k=2}^{\lfloor M \rfloor} k^{1-\alpha} \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \int_1^{\lfloor M \rfloor} x^{1-\alpha} dx \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \int_1^M x^{1-\alpha} dx \\ &= 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \begin{cases} \frac{M^{2-\alpha}}{2-\alpha} - \frac{1}{2-\alpha}, & \text{when } 1 < \alpha < 2, \\ \ln M, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2} - \frac{M^{2-\alpha}}{\alpha-2}, & \text{when } \alpha > 2 \end{cases} \\ &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \begin{cases} \frac{M^{2-\alpha}}{2-\alpha}, & \text{when } 1 < \alpha < 2, \\ \ln M, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2}, & \text{when } \alpha > 2. \end{cases} \end{aligned}$$

Recall that $M = 2\alpha/\mu$. Then

$$\begin{aligned}
J_{i,2}^0 &\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \begin{cases} \frac{M^{2-\alpha}}{2-\alpha}, & \text{when } 1 < \alpha < 2, \\ \ln M, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2}, & \text{when } \alpha > 2. \end{cases} \\
&\leq 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha}\mu^2 \begin{cases} \frac{(2\alpha)^{2-\alpha}}{2-\alpha}\mu^{\alpha-2}, & \text{when } 1 < \alpha < 2, \\ \ln 2 - \ln \mu, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2}, & \text{when } \alpha > 2 \end{cases} \\
&= 2V\mu^2 + Ve^{1+e/2}\mu^{\alpha+1} + Ve^{2\alpha} \begin{cases} \frac{(2\alpha)^{2-\alpha}}{2-\alpha}\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ \mu^2 \ln 2 - \mu^2 \ln \mu, & \text{when } \alpha = 2, \\ \frac{1}{\alpha-2}\mu^2, & \text{when } \alpha > 2 \end{cases} \\
&= V \begin{cases} 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + e^{2\alpha}\frac{(2\alpha)^{2-\alpha}}{2-\alpha}\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + e^{2\alpha}\mu^2 \ln 2 - e^{2\alpha}\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + \frac{e^{2\alpha}}{\alpha-2}\mu^2, & \text{when } \alpha > 2 \end{cases} \\
&= V \begin{cases} 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + e^{2\alpha}\frac{(2\alpha)^{2-\alpha}}{2-\alpha}\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 2\mu^2 + e^{1+e/2}\mu^3 + e^4\mu^2 \ln 2 - e^4\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + \frac{e^{2\alpha}}{\alpha-2}\mu^2, & \text{when } \alpha > 2. \end{cases}
\end{aligned}$$

Also, since $\mu \leq 1$, we can further bound the $J_{i,2}^0$ as follows:

$$\begin{aligned}
J_{i,2}^0 &\leq V \begin{cases} 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + e^{2\alpha}\frac{(2\alpha)^{2-\alpha}}{2-\alpha}\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ 2\mu^2 + e^{1+e/2}\mu^3 + e^4\mu^2 \ln 2 - e^4\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ 2\mu^2 + e^{1+e/2}\mu^{\alpha+1} + \frac{e^{2\alpha}}{\alpha-2}\mu^2, & \text{when } \alpha > 2 \end{cases} \\
&\leq V \begin{cases} \left(2 + e^{1+e/2} + e^{2\alpha}\frac{(2\alpha)^{2-\alpha}}{2-\alpha}\right)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ (2 + e^{1+e/2} + e^4 \ln 2)\mu^2 - e^4\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ \left(2 + e^{1+e/2} + \frac{e^{2\alpha}}{\alpha-2}\right)\mu^2, & \text{when } \alpha > 2 \end{cases} \tag{29}
\end{aligned}$$

Let us denote the term to the right from V in (29) by $\mathcal{T}_1(\alpha, \mu)$, that is

$$J_{i,2}^0 = V \mathcal{T}_1(\alpha, \mu), \tag{30}$$

where

$$\mathcal{T}_1(\alpha, \mu) =: V \begin{cases} C_6(\alpha)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ C_7(\alpha)\mu^2 - C_8(\alpha)\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_9(\alpha)\mu^2, & \text{when } \alpha > 2, \end{cases} \tag{31}$$

with

$$\begin{aligned}
C_6(\alpha) &= 2 + e^{1+e/2} + e^{2\alpha}\frac{(2\alpha)^{2-\alpha}}{2-\alpha}, \\
C_7(\alpha) &= 2 + e^{1+e/2} + e^4 \ln 2, \\
C_8(\alpha) &= e^4, \\
C_9(\alpha) &= 2 + e^{1+e/2} + \frac{e^{2\alpha}}{\alpha-2}.
\end{aligned}$$

Thus, from (28), (30), and (31), it follows that the second interval is bounded by

$$\begin{aligned} J_{i,2} &\leq \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + J_{i,2}^0 \\ &\leq \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + V \mathcal{T}_1(\alpha, \mu), \end{aligned}$$

where

$$\mathcal{T}(\alpha, \mu) = \begin{cases} C_6(\alpha)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ C_7(\alpha)\mu^2 - C_8(\alpha)\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_9(\alpha)\mu^2, & \text{when } \alpha > 2, \end{cases}$$

which proves the lemma. \square

4.3 Upper bound of $J_{i,3}$, when $\alpha > 1$

This interval is the easiest to deal with. Note that $J_{i,3}$ is identical to the $I_{i,2}$ term from Lemma 2, and thus, we can re-use its result to bound the interval.

Lemma 10. *Let $X_i \sim \mathbb{D}(\alpha, \alpha)$ be an integer valued r.v. with $\alpha > 1$. When $M = \frac{2\alpha}{\mu}$, then the term*

$$J_{i,3} = \sum_{k=\lceil M \rceil}^x e^{\mu k} \Pr[X_i = k],$$

is bounded from above by

$$J_{i,3} \leq C_{10}(\alpha)\mu^\alpha + V e^{\mu x} x^{-\alpha},$$

where $C_{10}(\alpha) = V e^{2\alpha+1} (2\alpha)^{-\alpha}$ is a constant.

Proof. Proof follows from Lemma 3. \square

4.4 Final assembling steps, when $\alpha > 1$

Recall, that our goal was to bound $P_i(\mu, x) \leq J_{i,1} + J_{i,2} + J_{i,3}$ in (24). Thus, after collecting results of Lemmas 8, 9, and 10:

$$\begin{aligned} J_{i,1} &\leq \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + W \mathcal{T}_0(\alpha, \mu), \\ J_{i,2} &\leq \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + 2V e^{2\alpha} \mu^2 + V \mathcal{T}_1(\alpha, \mu), \\ J_{i,3} &\leq C_{10}(\alpha)\mu^\alpha + V e^{\mu x} x^{-\alpha}, \end{aligned}$$

we obtain

$$\begin{aligned} P_i(\mu, x) &\leq J_{i,1} + J_{i,2} + J_{i,3} \\ &\leq \Pr[X_i \leq 0] + \mu \sum_{k \leq 0} k \Pr[X_i = k] + W \mathcal{T}_0(\alpha, \mu) \\ &\quad + \Pr[X_i \geq 1] + \mu \sum_{k=1}^{\infty} k \Pr[X_i = k] + 2V e^{2\alpha} \mu^2 + V \mathcal{T}_1(\alpha, \mu) \\ &\quad + C_{10}(\alpha)\mu^\alpha + V e^{\mu x} x^{-\alpha} \\ &= \sum_k \Pr[X_i = k] + \mu \sum_k k \Pr[X_i = k] + V e^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu), \end{aligned}$$

where

$$\begin{aligned} \mathcal{T}(\alpha, \mu) &= W \mathcal{T}_0(\alpha, \mu) + V \mathcal{T}_1(\alpha, \mu) + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2 \\ &= \begin{cases} WC_4(\alpha)\mu^\alpha + VC_6(\alpha)\mu^\alpha + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2, & \text{when } 1 < \alpha < 2, \\ 3W\mu^2 - W\mu^2 \ln \mu + VC_7(\alpha)\mu^2 - VC_8(\alpha)\mu^2 \ln \mu + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2, & \text{when } \alpha = 2, \\ WC_5(\alpha)\mu^2 + VC_9(\alpha)\mu^2 + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2, & \text{when } \alpha > 2. \end{cases} \end{aligned}$$

Again, in order to simplify the above term, we recall that $\mu \leq 1$. Then

$$\begin{aligned} \mathcal{T}(\alpha, \mu) &= \begin{cases} WC_4(\alpha)\mu^\alpha + VC_6(\alpha)\mu^\alpha + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2, & \text{when } 1 < \alpha < 2, \\ 3W\mu^2 - W\mu^2 \ln \mu + VC_7(\alpha)\mu^2 - VC_8(\alpha)\mu^2 \ln \mu + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2, & \text{when } \alpha = 2, \\ WC_5(\alpha)\mu^2 + VC_9(\alpha)\mu^2 + C_{10}(\alpha)\mu^\alpha + 2Ve^{2\alpha}\mu^2, & \text{when } \alpha > 2 \end{cases} \\ &\leq \begin{cases} (WC_4(\alpha) + VC_6(\alpha) + C_{10}(\alpha) + 2Ve^{2\alpha})\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ (3W + VC_7(\alpha) + C_{10}(\alpha) + 2Ve^{2\alpha})\mu^2 - (W + VC_8(\alpha))\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ (WC_5(\alpha) + VC_9(\alpha) + C_{10}(\alpha) + 2Ve^{2\alpha})\mu^2, & \text{when } \alpha > 2 \end{cases} \\ &=: \begin{cases} C_{11}(\alpha)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ C_{12}(\alpha)\mu^2 - C_{13}(\alpha)\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_{14}(\alpha)\mu^2, & \text{when } \alpha > 2, \end{cases} \end{aligned}$$

where

$$\begin{aligned} C_{11}(\alpha) &= WC_4(\alpha) + VC_6(\alpha) + C_{10}(\alpha) + 2Ve^{2\alpha}, \\ C_{12}(\alpha) &= 3W + VC_7(\alpha) + C_{10}(\alpha) + 2Ve^{2\alpha}, \\ C_{13}(\alpha) &= W + VC_8(\alpha), \\ C_{14}(\alpha) &= WC_5(\alpha) + VC_9(\alpha) + C_{10}(\alpha) + 2Ve^{2\alpha}. \end{aligned}$$

Hence,

$$\begin{aligned} P_i(\mu, x) &\leq \sum_k \Pr[X_i = k] + \mu \sum_k k \Pr[X_i = k] + Ve^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu) \\ &= 1 + \mu \mathbb{E}X_i + Ve^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu). \end{aligned}$$

Then, after applying the well-known relation $1 + x \leq e^x$, we obtain

$$P_i(\mu, x) \leq 1 + \mu \mathbb{E}X_i + Ve^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu) \leq \exp\left(\mu \mathbb{E}X_i + Ve^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu)\right). \quad (32)$$

Now, let's recall inequality (24) and denote its right-hand side by $\mathcal{K}(\mu, x)$, that is

$$\begin{aligned} \Pr[S_n \geq \mathbb{E}S_n + x] &\leq nVx^{-\alpha} + e^{-\mu x} \prod_{i=1}^n e^{-\mu \mathbb{E}X_i} P_i(\mu, x) \\ &=: nVx^{-\alpha} + \mathcal{K}(\mu, x), \end{aligned} \quad (33)$$

where $\mathcal{K}(\mu, x) := e^{-\mu x} \prod_{i=1}^n e^{-\mu \mathbb{E}X_i} P_i(\mu, x)$.

Next, after having obtained bound for $P_i(\mu, x)$ (32), we can simplify $\mathcal{K}(\mu, x)$

$$\begin{aligned}
\mathcal{K}(\mu, x) &= e^{-\mu x} \prod_{i=1}^n e^{-\mu \mathbb{E}X_i} P_i(\mu, x) \\
&\leq e^{-\mu x} \prod_{i=1}^n e^{-\mu \mathbb{E}X_i} \exp\left(\mu \mathbb{E}X_i + V e^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu)\right) \\
&= e^{-\mu x} \prod_{i=1}^n \exp\left(V e^{\mu x} x^{-\alpha} + \mathcal{T}(\alpha, \mu)\right) \\
&= e^{-\mu x} \exp\left(V n e^{\mu x} x^{-\alpha} + n \mathcal{T}(\alpha, \mu)\right) \\
&= \exp\left(-\mu x + V n e^{\mu x} x^{-\alpha} + n \mathcal{T}(\alpha, \mu)\right). \tag{34}
\end{aligned}$$

Next, we need to fix the value of μ that minimizes the above exponent, while keeping $\mu \rightarrow 0$, but $\mu x \rightarrow \infty$. Similar to the case when $0 < \alpha_r \leq 1$, we take $\mu = \frac{1}{x} \ln \frac{x^\alpha}{n}$.

The next lemma verifies that the chosen value of μ satisfies both constraints.

Lemma 11. *Let $\mu = \frac{1}{x} \ln \frac{x^\alpha}{n}$ and $x = n^{\max(1/\alpha, 1/2)+\epsilon}$ for $\epsilon > 0$. Then $\mu \rightarrow 0$ and $\mu x \rightarrow \infty$ when $n \rightarrow \infty$.*

Proof. First, let's check that $\mu \rightarrow 0$. We have for μ

$$\begin{aligned}
\mu &= \frac{1}{x} \ln \frac{x^\alpha}{n} \\
&= \frac{1}{n^{\max(1/\alpha, 1/2)+\epsilon}} \ln \frac{n^{\alpha \max(1/\alpha, 1/2)+\alpha\epsilon}}{n} \\
&\leq \frac{1}{n^{1/2}} \ln \frac{n^{\alpha/\min(\alpha, 2)+\alpha\epsilon}}{n}.
\end{aligned}$$

However, $\alpha/\min(\alpha, 2) \geq 1$, for any $\alpha > 0$, and so it is safe to claim that $\alpha/\min(\alpha, 2) = 1 + \delta$, where $\delta \geq 0$. Then

$$\begin{aligned}
\mu &\leq \frac{1}{n^{1/2}} \ln \frac{n^{\alpha/\min(\alpha, 2)+\alpha\epsilon}}{n} \\
&\leq \frac{1}{n^{1/2}} \ln \frac{n^{1+\delta+\alpha\epsilon}}{n} \\
&\leq \frac{1}{n^{1/2}} O(\ln n) \\
&= O\left(\frac{\ln n}{n^{1/2}}\right) \\
&= o(1),
\end{aligned}$$

when $n \rightarrow \infty$.

Now, in a similar way, we prove that $\mu x \rightarrow \infty$. From the defined values of μ and x , it follows that

$$\begin{aligned}
\mu x &= \frac{x}{x} \ln \frac{x^\alpha}{n} \\
&= \ln \frac{x^\alpha}{n} \\
&= \ln \frac{n^{\alpha \max(1/\alpha, 1/2)+\alpha\epsilon}}{n} \\
&= \ln \frac{n^{\alpha/\min(\alpha, 2)+\alpha\epsilon}}{n} \\
&= \ln \frac{n^{1+\delta+\alpha\epsilon}}{n}, \text{ since } \frac{\alpha}{\min(\alpha, 2)} = 1 + \delta, \text{ where } \delta \geq 0 \\
&\geq \ln n^{\delta+\alpha\epsilon} \\
&= (\delta + \alpha\epsilon) \ln n \rightarrow \infty,
\end{aligned}$$

when $n \rightarrow \infty$. □

Now, after fixing μ , we are going to show that

$$\mathcal{K}(\mu, x) \leq \exp\left(n\mathcal{T}(\alpha, \mu) + (Vne^{\mu x}x^{-\alpha} - \mu x)\right) \leq \exp(o(1) + V - \alpha\epsilon \ln n), \quad (35)$$

when $n \rightarrow \infty$, by analyzing each term separately. And so we have

Lemma 12. *Let*

$$\mathcal{T}(\alpha, \mu) = \begin{cases} C_{11}(\alpha)\mu^\alpha, & \text{when } 1 < \alpha < 2, \\ C_{12}(\alpha)\mu^2 - C_{13}(\alpha)\mu^2 \ln \mu, & \text{when } \alpha = 2, \\ C_{14}(\alpha)\mu^2, & \text{when } \alpha > 2, \end{cases}$$

where $\mu = \frac{1}{x} \ln \frac{x^\alpha}{n}$, $x = n^{\max(1/\alpha, 1/2)+\epsilon}$ and any $\epsilon > 0$. Then

$$n\mathcal{T}(\alpha, \mu) = o(1),$$

when $n \rightarrow \infty$.

Proof. First, let us consider the case when $1 < \alpha < 2$. Then $x = n^{1/\alpha+\epsilon}$ and $\mathcal{T}(\alpha, \mu) = C_{11}(\alpha)\mu^\alpha$. Therefore, we have

$$\begin{aligned} n\mathcal{T}(\alpha, \mu) &= C_{11}(\alpha)n\mu^\alpha \\ &= O\left(\frac{n}{x^\alpha} \ln^\alpha \frac{x^\alpha}{n}\right) \\ &= O\left(\frac{n}{n^{1+\alpha\epsilon}} \ln^\alpha \frac{n^{1+\alpha\epsilon}}{n}\right) \\ &= O(n^{-\alpha\epsilon} \ln^\alpha n^{\alpha\epsilon}) \\ &= o(1). \end{aligned}$$

Next, when $\alpha = 2$. Then $x = n^{1/2+\epsilon}$ and $\mathcal{T}(\alpha, \mu) = C_{12}(\alpha)\mu^2 - C_{13}(\alpha)\mu^2 \ln \mu$. Hence,

$$\begin{aligned} n\mathcal{T}(\alpha, \mu) &= C_{12}(\alpha)n\mu^2 - C_{13}(\alpha)n\mu^2 \ln \mu \\ &= O(n\mu^2 - n\mu^2 \ln \mu) \\ &= O\left(\frac{n}{x^2} \ln^2 \frac{x^2}{n} - \frac{n}{x^2} \ln^2 \frac{x^2}{n} \cdot \ln\left(\frac{1}{x} \ln \frac{x^2}{n}\right)\right) \\ &= O\left(\frac{n}{n^{1+2\epsilon}} \ln^2 \frac{n^{1+2\epsilon}}{n} - \frac{n}{n^{1+2\epsilon}} \ln^2 \frac{n^{1+2\epsilon}}{n} \cdot \ln\left(\frac{1}{n^{1/2+\epsilon}} \ln \frac{n^{1+2\epsilon}}{n}\right)\right) \\ &= O\left(n^{-2\epsilon} \ln^2 n^{2\epsilon} - n^{-2\epsilon} \ln^2 n^{2\epsilon} \cdot \ln\left(n^{-1/2-\epsilon} \ln n^{2\epsilon}\right)\right) \\ &= O\left(n^{-2\epsilon} \ln^2 n^{2\epsilon} - n^{-2\epsilon} \ln^2 n^{2\epsilon} \ln n^{-1/2-\epsilon} - n^{-2\epsilon} \ln^2 n^{2\epsilon} \ln \ln n^{2\epsilon}\right) \\ &= O\left(n^{-2\epsilon} \ln^2 n^{2\epsilon} - n^{-2\epsilon} \ln^2 n^{2\epsilon} \ln n^{-1/2-\epsilon}\right), \text{ since } n^{-2\epsilon} \ln^2 n^{2\epsilon} \ln \ln n^{2\epsilon} > 0 \\ &= O\left(n^{-2\epsilon} \ln^2 n^{2\epsilon} + (1/2 + \epsilon)n^{-2\epsilon} \ln^2 n^{2\epsilon} \ln n\right) \\ &= O\left(n^{-2\epsilon} \ln^2 n^{2\epsilon} \ln n\right) \\ &= o(1). \end{aligned}$$

Finally, when $\alpha > 2$, then $x = n^{1/2+\epsilon}$, $\mathcal{T}(\alpha, \mu) = C_{14}(\alpha)\mu^2$, while

$$\begin{aligned}
n\mathcal{T}(\alpha, \mu) &= C_{14}(\alpha)n\mu^2 \\
&= O(n\mu^2) \\
&= O\left(\frac{n}{x^2} \ln^2 \frac{x^2}{n}\right) \\
&= O\left(\frac{n}{n^{1+2\epsilon}} \ln^2 \frac{n^{1+2\epsilon}}{n}\right) \\
&= O(n^{-2\epsilon} \ln^2 n^{2\epsilon}) \\
&= o(1).
\end{aligned}$$

And this completes the proof of the lemma. \square

However, unlike the $n\mathcal{T}(\alpha, \mu)$ term in $\mathcal{K}(\mu, x)$ (34), term $Vne^{\mu x}x^{-\alpha} - \mu x \rightarrow -\infty$, when $n \rightarrow \infty$, and the next lemma verifies this.

Lemma 13. *Let $\alpha > 1$ and $x = n^{\max(1/\alpha, 1/2)+\epsilon}$. Then*

$$Vne^{\mu x}x^{-\alpha} - \mu x \leq V - \alpha\epsilon \ln n.$$

Proof. We have

$$\begin{aligned}
Vne^{\mu x}x^{-\alpha} - \mu x &= Vne^{\frac{x}{x} \ln \frac{x^\alpha}{n}} x^{-\alpha} - \frac{x}{x} \ln \frac{x^\alpha}{n} \\
&= V - \ln \frac{x^\alpha}{n}.
\end{aligned}$$

When $1 < \alpha \leq 2$, then $x = n^{1/\alpha+\epsilon}$, and so

$$\begin{aligned}
V - \ln \frac{x^\alpha}{n} &= V - \ln \frac{n^{1+\alpha\epsilon}}{n} \\
&= V - \ln n^{\alpha\epsilon} \\
&= V - \alpha\epsilon \ln n.
\end{aligned}$$

When $\alpha > 2$, then $x = n^{1/2+\epsilon}$, and

$$\begin{aligned}
V - \ln \frac{x^\alpha}{n} &= V - \ln \frac{n^{\alpha/2+\alpha\epsilon}}{n} \\
&\leq V - \ln \frac{n^{1+\alpha\epsilon}}{n} \\
&= V - \ln n^{\alpha\epsilon} \\
&= V - \alpha\epsilon \ln n.
\end{aligned}$$

Thus, after combining both cases, we obtain that $Vne^{\mu x}x^{-\alpha} - \mu x \leq V - \alpha\epsilon \ln n$, and the lemma follows. \square

Next, after collecting results of Lemmas 12, and 13, we obtain from (34) that

$$\mathcal{K}(\mu, x) \leq \exp\left(n\mathcal{T}(\alpha, \mu) + (Vne^{\mu x}x^{-\alpha} - \mu x)\right) \leq \exp\left(o(1) + V - \alpha\epsilon \ln n\right) \leq e^{2V}n^{-\alpha\epsilon},$$

and, therefore, (33) simplifies to

$$\begin{aligned}
\Pr[S_n \geq \mathbb{E}S_n + x] &\leq nVx^{-\alpha} + \mathcal{K}(\mu, x) \\
&\leq nVx^{-\alpha} + \exp\left(2V - \alpha\epsilon \ln n\right) \\
&= nVx^{-\alpha} + e^{2V}n^{-\alpha\epsilon}.
\end{aligned}$$

And finally, since $x = n^{\max(1/\alpha, 1/2)+\epsilon}$, we obtain that

$$\begin{aligned}\Pr[S_n \geq \mathbb{E}S_n + x] &= \Pr[S_n - \mathbb{E}S_n \geq n^{\max(1/\alpha, 1/2)+\epsilon}] \\ &\leq nVx^{-\alpha} + e^{2V}n^{-\alpha\epsilon} \\ &= Vn^{1-\max(1, \alpha/2)-\alpha\epsilon} + e^{2V}n^{-\alpha\epsilon},\end{aligned}$$

which proves Theorem 5. \square

Theorem 5 implies a simple corollary:

Corollary 4. *Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\alpha, \alpha)$ are independent not necessary identically distributed integer-valued random variables, with $\alpha > 1$. Then w.h.p.*

$$S_n - \mathbb{E}S_n \leq Cn^{\max(1/\alpha, 1/2)},$$

where $C > 0$ is some constant.

To find the left tail bounds of the r.v. $S_n - \mathbb{E}S_n$, we will apply the method we used while proving Theorem 4, i.e. we introduce "inverted" random variables to which we apply the right-tail bound from Theorem 5:

Theorem 6. *Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\alpha_{l,i}, \alpha_{r,i})$ are independent not necessary identically distributed integer-valued random variables with $\alpha_{l,i}, \alpha_{r,i} > 1$. Then for any $\epsilon > 0$, we have*

$$\Pr[S_n - \mathbb{E}S_n \leq -n^{\max(1/\alpha, 1/2)+\epsilon}] \leq Wn^{1-\max(1, \alpha/2)-\alpha\epsilon} + e^{2W}n^{-\alpha\epsilon},$$

when $n \rightarrow \infty$.

Proof. Let's introduce random variables X'_i that have the same distributions as $-X_i$, i.e. $X'_i \stackrel{d}{=} -X_i$. Clearly, $X'_i \sim \mathbb{D}(\alpha_{r,i}, \alpha_{l,i})$ with $\alpha_{l,i}, \alpha_{r,i} > 1$, and so, by Definition 3, $F_{X'_i}(k) \leq W_{X'_i}k^{-\alpha_{l,i}}$.

Since $\alpha_{l,i}, \alpha_{r,i} > 1$, from Lemma 3 it follows that $|\mathbb{E}X'_i| < \infty$. Then

$$\begin{aligned}\Pr[S_n - \mathbb{E}S_n \leq -x] &= \Pr\left[\sum_{i=1}^n (X_i - \mathbb{E}X_i) \leq -x\right] \\ &= \Pr\left[\sum_{i=1}^n (\mathbb{E}X_i - X_i) \geq x\right] \\ &= \Pr\left[\sum_{i=1}^n (-X_i - \mathbb{E}[-X_i]) \geq x\right] \\ &= \Pr\left[\sum_{i=1}^n (X'_i - \mathbb{E}X'_i) \geq x\right] \\ &= \Pr\left[\sum_{i=1}^n X'_i - \sum_{i=1}^n \mathbb{E}X'_i \geq x\right] \\ &= \Pr\left[S'_n - \mathbb{E}S'_n \geq x\right], \text{ where } S'_n := \sum_{i=1}^n X'_i \\ &\leq nWx^{-\alpha} + e^{2W}n^{-\alpha\epsilon}.\end{aligned}$$

The last inequality here follows after applying Theorem 5 to the sum $S'_n - \mathbb{E}S'_n$, which consists of random variables each of which has tail distribution functions that can be bounded by some $Ck^{-\alpha}$ with constants $C > 0$ and $\alpha > 1$.

Finally, since $x = n^{\max(1/\alpha, 1/2)+\epsilon}$, we obtain that

$$\Pr[S_n - \mathbb{E}S_n \leq -x] = \Pr[S_n - \mathbb{E}S_n \leq -n^{\max(1/\alpha, 1/2)+\epsilon}] \leq Wn^{1-\max(1, \alpha/2)-\alpha\epsilon} + e^{2W}n^{-\alpha\epsilon},$$

and the theorem follows. □

As a result from Theorem 6 and Corollary 4 another useful corollary follows:

Corollary 5. *Let $S_n = \sum_{i=1}^n X_i$, where $X_i \sim \mathbb{D}(\alpha, \alpha)$ are independent not necessary identically distributed integer-valued random variables, with $\alpha > 1$. Then w.h.p.*

$$|S_n - \mathbb{E}S_n| \leq Cn^{\max(1/\alpha, 1/2)},$$

where $C > 0$ is some constant.

Hence, when $S_n = \sum_{i=1}^n X_i$ consists of variables, whose tail functions can be bounded by function $Ck^{-\alpha}$ with $\alpha > 1$ and some constant $C > 0$, then we do not expect S_n to deviate much from its expected value $\mathbb{E}S_n$.

References

- [1] Carlos Ansótegui, Maria Luisa Bonet, and Jordi Levy. On the structure of industrial SAT instances. In *Principles and Practice of Constraint Programming - CP 2009, 15th International Conference, CP 2009, Lisbon, Portugal, September 20-24, 2009, Proceedings*, pages 127–141, 2009.
- [2] Carlos Ansótegui, Maria Luisa Bonet, and Jordi Levy. Towards industrial-like random SAT instances. In *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*, pages 387–392, 2009.
- [3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509512, 1999.
- [4] A.A. Borovkov and K.A. Borovkov, editors. *Asymptotic analysis of random walks: heavy-tailed distributions*, volume 118 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 2008.
- [5] Andrei Z. Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet L. Wiener. Graph structure in the web. *Computer Networks*, 33(1-6):309–320, 2000.
- [6] Maurice C. Bryson. Heavy-tailed distributions: Properties and tests. *Technometrics*, 16(1):61–68, 1974.
- [7] A. Clauset, C.R. Shalizi, and M.E. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661703, 2009.
- [8] Colin Cooper, Alan M. Frieze, and Pawel Pralat. Some typical properties of the spatial preferred attachment model. *Internet Mathematics*, 10(1-2):116–136, 2014.
- [9] Colin Cooper, Alan M. Frieze, and Gregory B. Sorkin. Random 2-sat with prescribed literal degrees. *Algorithmica*, 48(3):249–265, 2007.
- [10] E. Friedgut. Sharp thresholds of graph properties, and the k -sat problem. *J. ACM*, 12(4):10171054, 1999.
- [11] Tobias Friedrich, Anton Krohmer, Ralf Rothenberger, Thomas Sauerwald, and Andrew M. Sutton. Bounds on the satisfiability threshold for power law distributed random SAT. In *25th Annual European Symposium on Algorithms, ESA 2017, September 4-6, 2017, Vienna, Austria*, pages 37:1–37:15, 2017.
- [12] John Hopcroft and Ravindran Kannan. Foundations of data science. available at <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/book-No-Solutions-Aug-21-2014.pdf> 2014.

- [13] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D. Sivakumar, Andrew Tomkins, and Eli Upfal. Random graph models for the web graph. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 57–65, 2000.
- [14] Michael Molloy and Bruce A. Reed. A critical point for random graphs with a given degree sequence. *Random Struct. Algorithms*, 6(2/3):161–180, 1995.
- [15] Michael Molloy and Bruce A. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability & Computing*, 7(3):295–305, 1998.
- [16] M.E. Newman. Power laws, pareto distributions and zipfs law. *Contemporary Physics*, 46(5):323351, 2005.