Maximum waiting time in heavy-tailed fork-join queues

Dennis Schol, Maria Vlasiou, and Bert Zwart

November 2022

Abstract

In this paper, we study the maximum waiting time $\max_{i \leq N} W_i(\cdot)$ in an $N$-server fork-join queue with heavy-tailed services as $N \to \infty$. The service times are the product of two random variables. One random variable has a regularly varying tail probability and is the same among all $N$ servers, and one random variable is Weibull distributed and is independent and identically distributed among all servers. This setup has the physical interpretation that if a job has a large size, then all the subtasks have large sizes, with some variability described by the Weibull-distributed part. We prove that after a temporal and spatial scaling, the maximum waiting time process converges in $D[0,T]$ to the supremum of an extremal process with negative drift. The temporal and spatial scaling are of order $\tilde{L}(b_N) b_N^{\beta/(\beta-1)}$, where $\beta$ is the shape parameter in the regularly varying distribution, $\tilde{L}(x)$ is a slowly varying function, and $(b_N, N \geq 1)$ is a sequence for which holds that $\max_{i \leq N} A_i/b_N \to 1$, as $N \to \infty$, where $A_i$ are i.i.d. Weibull-distributed random variables. Finally, we prove steady-state convergence.

1 Introduction

The fork-join queue is a useful tool to model streams of jobs, consisting of subtasks, in parallel systems. A key quantity of interest is the behavior of the longest queue. In this paper, we investigate a fork-join queue with $N$ servers, where each of these servers has to complete a subtask of an incoming task. We assume that $N$ is large, and we investigate the longest waiting time among the $N$ subtasks. Moreover, we assume that service times are mutually dependent, and can be written as a product of two random variables, where one term is independent and identically distributed for all servers, and has a Weibull-like tail, while the other term is the same for all servers and has a regularly varying tail. This describes the situation that if a job has a large size, all the subtasks also have a large size, where the fluctuation is described by the Weibull-like distributed random variable.

We obtain a convergence result for the rescaled transient maximum waiting time $\max_{i \leq N} W_i(t_{cN})/c_N$ as $N \to \infty$, after choosing the proper temporal and spatial scaling $(c_N, N \geq 1)$. This maximum waiting time converges in distribution to a process which is the supremum of Fréchet-distributed random variables minus a drift term. The temporal and spatial scaling $c_N$ depends on the extreme-value scaling of $N$ independent Weibull-distributed random variables, a slowly varying function, and the parameter of regular variation. Hence, to obtain this result, a mixture of classic extreme value theory and analysis of heavy tails is needed. Furthermore, we show that this rescaled maximum waiting time process $(\max_{i \leq N} W_i(t_{cN})/c_N, t \in [0,T])$ converges as a process in $D[0,T]$ to an extremal process $(\sup_{s \in [0,t]} (X_{s,t} - \mu(t-s)), t \in [0,T])$ with Fréchet marginals, with $D[0,T]$ the space of cadlag functions on $[0,T]$, which we equip with the $d^0$ metric, cf. [2] Eq. (12.16). Finally, we prove steady-state convergence of $\max_{i \leq N} W_i(\infty)/c_N$ to $\lim_{t \to \infty} \sup_{s \in [0,t]} (X_{s,t} - \mu(t-s))$.

Applications of these heavy-tailed fork-join queues are usually found in parallel computing. Companies such as Google, Microsoft and Alibaba have datacenters with thousands of servers that are available for cloud computing, where there is often a form of parallel scheduling. Jobs in these systems have typically large sizes, and are often heavy tailed. However, most literature on parallel queueing theory assumes service times to be light tailed, cf. the survey [15]. This motivates the analysis of parallel queueing networks with heavy-tailed job sizes.

Our work relates to the literature on fork-join queues. Exact results on probability distributions of fork-join queues are only derived for fork-join queues with two service stations, cf. [2] [12] [17] [32]. Approximations and bounds for performance measure of the fork-join queue with an arbitrary but fixed number of servers can be found in [3] [18] [24]. In [31], a heavy-traffic analysis for fork-join queues is derived; see also [25] and [26]. More recent work in this direction may be found in [19] [20] [21] [22] [30].

Moreover, our work is connected to literature on heavy-tailed phenomena; cf. [23] for a summary. Specific results on the interplay between fork-join queues and heavy-tailed services can be found in [27] [33] [34]. In [27] Thm. 2, asymptotic lower and upper bounds for the tail probability of the maximum waiting time in steady state are given;
however these bounds are not sharp when $N$ is large. In [33] and [34], the authors investigate the fork-join queue with heavy-tailed services under a blocking mechanism. This paper contributes to the existing literature, as we give sharp convergence results for the maximum waiting time with heavy-tailed service times, where the number of servers $N$ grows large. We combine results from extreme value theory, the analysis of heavy-tailed random variables, and results on process convergence in $D[0,T]$. 

Furthermore, the limiting process in this paper has an interesting form; this process is an extremal process with negative drift. Several papers have been written on extremal processes, cf. [4, 5, 10, 11, 29]. These results are, among others, used and applied on the analysis of records in sport, cf. [4, 5]. For example, in [4], a model is used to analyze the times in the mile run.

This paper is organized as follows. We present our model in Section 2 and our main results in Theorems 1, 2, 3, and Proposition 1. We give a heuristic analysis of our results in Section 2.1. In Section 2.2, we discuss other modeling choices. In Section 3 we present some auxiliary results. We prove process convergence in Section 4. Finally, we prove our main results in Section 6.

2 Model and main results

In this paper, we analyze a fork-join queue with a common arrival process, and a service process that consists of a Weibull-like i.i.d. part and a regularly varying part that is the same among all the servers. This models the situation that if a job has a large size, then all the subtasks have a large size, with some variability. We show in Section 2.1 that the Weibull distribution has convenient properties that we exploit in this paper; in Section 2.2, we briefly discuss what happens when the i.i.d. part of the service process has a lighter tail. We write the random variable $A_{i,j}B_j$ as the representation of a service time at server $i$ of the subtask of the $j$-th job, while the random variable $T_j$ is the interarrival time between the $j$-th and $(j+1)$-st job. Now, by Lindley’s recursion, the waiting time at server $i$ upon arrival of the $(n+1)$-st job equals

$$W_i(n) = \sup_{0 \leq k \leq n} \sum_{j=k+1}^n (A_{i,j}B_j - T_j),$$

with $W_i(0) = 0$ and $\sum_{j=n+1}^\infty (A_{i,j}B_j - T_j) = 0$. Moreover, we write $W_i(t) = W_i(\lfloor t \rfloor)$. Furthermore, the maximum of the $N$ waiting times equals

$$\max_{i \leq N} W_i(n) = \max_{i \leq N} \sup_{0 \leq k \leq n} \sum_{j=k+1}^n (A_{i,j}B_j - T_j).$$

We assume that the sequence of random variables $(B_j, j \geq 1)$ are independent random variables that satisfy

$$P(B_j > x) = L(x)/x^\beta,$$

with $L(x)$ a slowly varying function and $\beta > 1$, indicating possible large job sizes. We let i.i.d. random variables $(A_{i,j}, i \geq 1, j \geq 1)$ satisfy

$$\log P(A_{i,j} > x) \sim -qx^\alpha,$$

as $x \to \infty$, with $0 < \alpha < 1$ and $q > 0$. Let $b_N = (\log N/q)^{1/\alpha}$. Then, we know from standard extreme value theory [14] Thm. 5.4.1, p. 188] that

$$\frac{\max_{i \leq N} A_{i,j}}{b_N} \overset{p}{\to} 1,$$

as $N \to \infty$. Thus, the number $b_N$ indicates the approximate size of the largest of $N$ independent Weibull-distributed random variables. Furthermore, we have independent and identically distributed distributed random variables $(T_j, j \geq 1)$, such that

$$\mathbb{E}[A_{i,j}B_j - T_j] = -\mu,$$

with $\mu > 0$.

In this paper, we prove process convergence of the scaled maximum waiting time over $N$ servers in Theorem 2 cf. [30] for a similar result for fork-join queues with light-tailed services. In order to achieve this result, we need to
scale the number of arriving jobs and the maximum waiting time with a sequence \((c_N, N \geq 1)\), where the sequence \((c_N, N \geq 1)\) satisfies
\[
c_N \sim \frac{(c_N/b_N)^\beta}{L(c_N/b_N)},
\]
as \(N \to \infty\), with \(c_N/b_N \xrightarrow{N \to \infty} \infty\). We explain in Section 2.4 in more detail why this sequence scales as given in (7).

Following standard arguments on generalized inverses of regularly varying functions; cf. [25] Prop. 2.6 (v,vii) and [8] Thm. 1.5.12], we can solve the right-hand side of (7) and get that \(c_N/b_N \sim c_N^{\beta/(\beta-1)}/L(c_N)\), with \(\hat{L}\) a slowly varying function. From this, it follows that \(b_N \sim \hat{L}(c_N)c_N^{\beta/(\beta-1)}\). Now, we define the sequence \((c_N, N \geq 1)\)
\[
eq 0 \phi(c_N/b_N)^{\beta/(\beta-1)}
\]
where \(\hat{L}\) is a slowly varying function that equals
\[
\hat{L}(x)^{\beta/(\beta-1)} = \left(\left(\frac{x}{(x^\beta/L(x))} \right)^{\star} \right)^{\star},
\]
with \(H(y)^\star = \inf\{s : H(s) \geq y\}\), and \(f(x)^*\) is a monotone function with the property that \(f(x)^* \sim f(x)\) as \(x \to \infty\). Thus, the sequence \((c_N, N \geq 1)\) satisfies the relationship described in (7). More precise properties of the function \(\hat{L}\) are given in Lemma 1.

As we have a proper scaling of the number of arriving jobs and the maximum waiting time by a sequence \((c_N, N \geq 1)\), the scaled maximum waiting time has the form
\[
\frac{\max_{i \leq N} W_i(tc_N)}{c_N} = \frac{\sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j)}{c_N}.
\]
Notice that
\[
\sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j) \xrightarrow{d} \sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j).
\]
Thus, to prove convergence of a single random variable \(\max_{i \leq N} W_i(tc_N)/c_N\) it suffices to prove convergence of the right-hand side in Equation (11).

However, the processes \((\sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j))/c_N, t \in [0, T]\) are not equal in distribution. For instance, \((\sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j))/c_N, t \in [0, T]\), which we will refer to as the auxiliary process, is non-decreasing in \(t\), and \((\max_{i \leq N} W_i(tc_N)/c_N, t \in [0, T]\) is not non-decreasing in \(t\). In Theorem 1 we show that this auxiliary process converges in distribution to a limiting process;
\[
\left(\sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j), t \in [0, T]\right) \xrightarrow{d} \left(\sup_{s \in [0,t]} (X_s - \mu s), t \in [0, T]\right),
\]
as \(N \to \infty\). The process \((X_t, t \in [0, T])\) is a stochastic process with Fréchet-distributed marginals. This process has cumulative distribution function \(P(X_t \leq x) = \exp(-t/x^\beta)\) for \(x > 0\). Furthermore, \(X_{t+s} = \max(X_t, X_s)\), where \(X_s\) is an independent copy of \(X_t\), because \(P(X_{t+s} < x) = P(X_t < x) P(X_s < x) + \exp(-t/x^\beta) \exp(-s/x^\beta) = \exp(-(t + s)/x^\beta)\). Thus, the process \((X_t, t \in [0, T])\) is a function in \(D[0, T]\) and is called an extremal process, cf. [29]. It is easy to see that \((\sup_{s \in [0,t]} (X_s - \mu s), t \in [0, T])\) is also non-decreasing in \(t\). The limiting process of \((\max_{i \leq N} W_i(tc_N)/c_N, t \in [0, T]\) has the same marginals as the process \((\sup_{s \in [0,t]} (X_s - \mu s), t \in [0, T])\), but is not non-decreasing. We write the limiting process of the maximum waiting time as \((\sup_{s \in [0,t]} (X_{s,t} - \mu (t-s)), t \in [0, T])\), with \(X_{s,t} \xrightarrow{d} X_{t,s}\). For \(r < s < t\), we have that \(X_{(r,t)} = \max(X_{r,s}, X_{s,t})\), and we have that \(X_{s,t}\) and \(X_{u,v}\) are independent if and only if the intervals \((s, t)\) and \((u, v)\) are disjoint. We write \(X_{t} = X_{0,t}\). In conclusion, the random variable \(X_t\) involves a single time parameter, while the random variable \(X_{(s,t)}\) is defined by two time parameters, which complicates the proof. There is a clear connection between the stochastic processes however, and in this paper, we first prove convergence of the non-decreasing process \((\max_{i \leq N} \sup_{s \in [0,t]} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j))/c_N, t \in [0, T]\) and we use this result with some additional steps to prove process convergence of the scaled maximum waiting time \((\max_{i \leq N} W_i(tc_N)/c_N, t \in [0, T]\).
Assumption 1 (Waiting time). Let the waiting time of customers in front of the i-th server be given in Equation 1, the i.i.d. random variables \((A_{i,j}, i \geq 1, j \geq 1)\) satisfy 4, and the i.i.d. random variables \((B_j, j \geq 1)\) satisfy 3 with \(L(x)\) a slowly varying function.

Assumption 2 (Scaling). Let \(b_N = (\log N/q)^{1/\alpha}\), and \((c_N, N \geq 1)\) and \(\hat{L}\) satisfy 7, 8, and 9.

Assumption 3 (Limiting process). Let \((X(s,t), t \in [0,T])\) be a stochastic process with Fréchet-distributed marginals. For \(r < s < t\), we have that \(X(r,s) = \max(X(r,s), X(s,t))\), and we have that \(X(s,t)\) and \(X(u,v)\) are independent if and only if the intervals \((s,t)\) and \((u,v)\) are disjoint. We write \(X_t = X_{(0,t)}\) and we have that \(X(s,t) \xrightarrow{d} X_{t-s}\). Furthermore, \(P(X_t \leq x) = \exp(-t/x^\beta)\) for \(x > 0\).

Theorem 1. Given that Assumptions 7-9 hold, we have that
\[
\left(\sup_{s \in [0,t]} \max_{i \leq N} \sum_{j=1}^{\lfloor s/c_N \rfloor} (A_{i,j}B_j - T_j), t \in [0,T]\right) \xrightarrow{d} \left(\sup_{s \in [0,t]} (X_s - \mu s), t \in [0,T]\right),
\]
as \(N \to \infty\).

Now, the main result proven in this paper is given in Theorem 2.

Theorem 2. Given that Assumptions 7-9 hold, we have that
\[
\left(\frac{\max_{i \leq N} W_i(tc_N)}{c_N}, t \in [0,T]\right) \xrightarrow{d} \left(\sup_{s \in [0,t]} (X_{s,t} - \mu (t-s)), t \in [0,T]\right),
\]
as \(N \to \infty\).

Now, when letting \(t \to \infty\) on the left-hand and the right-hand side of (13), we expect from this convergence result that the maximum steady-state waiting time satisfies \(P(\max_{i \leq N} W_i(\infty) > xc_N) \xrightarrow{N \to \infty} P(\sup_{t>0}(X_t - \mu t) > x)\). Though this does not trivially follow from Theorem 2, it is indeed true, and we prove this in Theorem 3.

Theorem 3. Given that Assumptions 7-9 hold, we have that
\[
P\left(\max_{i \leq N} W_i(\infty) > xc_N\right) \xrightarrow{N \to \infty} P\left(\sup_{t>0}(X_t - \mu t) > x\right).
\]

We can write the limiting probabilities explicitly.

Proposition 1. Given that Assumptions 7-9 hold, we have that
\[
P\left(\sup_{t>0}(X_t - \mu t) > x\right) = 1 - \exp\left(-\frac{1}{\mu(\beta-1)x^{\beta-1}}\right),
\]
and
\[
P\left(\sup_{s \in [0,t]} (X_{s,t} - \mu (t-s)) > x\right) = 1 - \exp\left(-\frac{1}{\mu^\beta(\beta-1)}\left(\frac{1}{(x/\mu)^{\beta-1}} - \frac{1}{(x/\mu + t)^{\beta-1}}\right)\right).
\]

2.1 Main ideas for the proofs

To prove Theorem 2 directly is challenging, since the limiting random variable \(X_{s,t}\) depends on two parameters and cannot be written as a difference of the form \(Y_1 - Y_2\), as is the case in standard queueing theory. However, the marginal distributions of \(X_{s,t}\) and \(X_{t-s}\) are the same. Thus, we first prove Theorem 1 after which we prove Theorem 2 using some auxiliary results on bounds on tail probabilities, convergence rates of sums of Weibull-distributed random variables, and auxiliary results on process convergence in \(D[0,T]\); cf. Section 3. To get a better understanding of the convergence result in Theorem 1 it benefits to first examine the process
\[
\left(\max_{i \leq N} \sum_{j=1}^{\lfloor t/c_N \rfloor} (A_{i,j}B_j - T_j), t \in [0,T]\right),
\]
so we remove the supremum term from the expression on the left-hand side of (12) and we are left with a maximum of \( N \) random walks. We can however apply the continuous mapping theorem on this stochastic process and obtain the result in Theorem 1 because the supremum is a continuous functional. Obviously, the law of large numbers implies that

\[
\frac{\sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)}{c_N} \xrightarrow{p} -\mu t,
\]

as \( N \to \infty \). However, when we investigate the largest of \( N \) of these random variables, we obtain that

\[
\max_{j \leq [tc_N]} \frac{\sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)}{c_N} \xrightarrow{d} X_t - \mu t,
\]

as \( N \to \infty \). The fact that we see this limiting behavior has two main reasons; first of all, a standard result is that for i.i.d. regularly varying \((B_j, j \geq 1)\), the tail behavior of a finite sum is the same as the tail behavior of the largest regularly varying random variable. Second, for Weibull-distributed random variables and a deterministic sequence \((b_j, j \geq 1)\), we have that \( \max_{j \leq N} \sum_{j=1}^{n} A_{ij} b_j/b_N \xrightarrow{p} \max_{j \leq n} b_j \), as \( N \to \infty \), cf. [30] Lem. B.1 for a proof. Therefore, \( \max_{j \leq N} \sum_{j=1}^{n} A_{ij} B_j \approx \max_{j \leq N} A_i \cdot \max_{j \leq n} B_j + \mathbb{E}[A_{ij} B_j] (n-1) \) for \( N \) large. Thus, we can conclude that for \( N \) large,

\[
\max_{j \leq [tc_N]} \frac{\sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)}{c_N} \approx \max_{j \leq N} A_i \max_{j \leq [tc_N]} \frac{B_j}{c_N/b_N} + \sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)
\]

\[
\approx \max_{j \leq N} A_i \max_{j \leq [tc_N]} \frac{B_j}{c_N/b_N} - \mu t
\]

\[
\approx \max_{j \leq [tc_N]} \frac{B_j}{c_N/b_N} - \mu t.
\]

We see that the largest regularly varying random variable \( \max_{j \leq [tc_N]} B_j \) determines the stochastic part in the limit, and is of order \( c_N/b_N \). Now, it is easy to see that

\[
\mathbb{P}\left( \max_{j \leq [tc_N]} B_j \leq (x + \mu t)\frac{c_N}{b_N} \right) = \mathbb{P}\left( B_j \leq (x + \mu t)\frac{c_N}{b_N} \right)^{[tc_N]} \sim \left( 1 - \frac{L((x + \mu t)c_N/b_N)}{(x + \mu t)c_N/b_N)^\beta} \right)^{[tc_N]}
\]

Because we have defined \( c_N \) as having the relation \( c_N \sim (c_N/b_N)^\beta/L(c_N/b_N) \) as \( N \to \infty \), we get that,

\[
\left( 1 - \frac{L((x + \mu t)c_N/b_N)}{(x + \mu t)c_N/b_N)^\beta} \right)^{[tc_N]} \sim \left( 1 - \frac{1}{(x + \mu t)^\beta c_N} \right)^{[tc_N]} \xrightarrow{N \to \infty} \exp(-\frac{t}{(x + \mu t)^\beta})
\]

In conclusion, the limiting distribution of \( \max_{j \leq N} \sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)/c_N \) is a Fréchet-distributed random variable with a negative drift term. We also see that we can approximate \( \max_{j \leq [tc_N]} \sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)/c_N \) with \( \max_{j \leq [tc_N]} B_j/(c_N/b_N) - \mu t \) as \( N \) is large. This approximating random variable has convenient properties, since the stochastic term is non-decreasing in \( t \). Therefore, to prove process convergence of \( \max_{j \leq N} \sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)/c_N, t \in [0, T] \) to \( (X_t - \mu t, t \in [0, T]) \), we first prove that \( \max_{j \leq [tc_N]} B_j/(c_N/b_N) - \mu t, t \in [0, T] \) converges to \( (X_t - \mu t, t \in [0, T]) \). Furthermore, we prove in Lemma 9 that for all \( \epsilon > 0 \),

\[
\mathbb{P}\left( \sup_{t \in [0,T]} \left| \max_{j \leq [tc_N]} \frac{\sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)}{c_N} - \frac{\max_{j \leq [tc_N]} B_j}{c_N/b_N} - \mu t \right| > \epsilon \right) \xrightarrow{N \to \infty} 0
\]

After applying the triangle inequality, we obtain that

\[
\left( \max_{j \leq N} \sum_{j=1}^{[tc_N]} (A_{ij} B_j - T_j)/c_N, t \in [0, T] \right) \xrightarrow{d} (X_t - \mu t, t \in [0, T])
\]

as \( N \to \infty \). Now, by applying the continuous mapping theorem, we obtain the result of Theorem 1. This is still an auxiliary result, because the process on the left-hand side of (12) is not the maximum waiting time process. We can
however prove the process convergence of the maximum waiting time in Theorem 2 by using some additional results, as the marginals of the processes on the left side of the limit in Theorems 1 and 2 are the same, and the marginals of the limiting processes in Theorems 1 and 2 are the same. Thus, we already know that pointwise convergence holds.

In order to prove the convergence of the finite-dimensional distributions for the maximum waiting time process, we show that we can decompose the joint probabilities of both the maximum waiting time process and the limiting process into an operation of marginal probabilities, and thus, convergence of finite-dimensional distributions follows from pointwise convergence. For example, for \( x_2 + \mu t_2 > x_1 + \mu t_1 \),

\[
\mathbb{P}\left( \sup_{s \in [0,t_1]} (X(s,t_1) - \mu(t_1 - s)) < x_1 \cap \sup_{s \in [0,t_2]} (X(s,t_2) - \mu(t_2 - s)) < x_2 \right) = \frac{\mathbb{P}\left( \sup_{s \in [0,t_1]} (X(s,t_1) - \mu(t_1 - s)) < x_1 \right)}{\mathbb{P}\left( \sup_{s \in [0,t_1]} (X(s,t_1) - \mu(t_1 - s)) < x_1 + \mu(t_1 - t_1) \right)} \mathbb{P}\left( \sup_{s \in [0,t_2]} (X(s,t_2) - \mu(t_2 - s)) < x_2 \right).
\]

An analogous equation holds for the process \( \sup_{s \in [0,t]} \left( \max_{|s| \leq |t|} B_j (c_N/b_N) - \mu(t - s) \right), t \in [0,T] \).

In Lemma 10 we prove that the maximum waiting time in (10) satisfies

\[
\mathbb{P}\left( \sup_{t \in [0,T]} \sup_{s \in [0,t]} \frac{\max_{j \leq N} \sum_{j=1}^{|s|} (A_{i,j} B_j - T_j)}{c_N} - \sup_{s \in [0,t]} \frac{\max_{|s| \leq |t|} B_j (c_N/b_N) - \mu(t - s)}{c_N/b_N} \right) > \epsilon \xrightarrow{N \to \infty} 0,
\]

by using similar techniques as in Lemma 8. Finally, we show in the proof of Theorem 2 that

\[
\left( \sup_{s \in [0,t]} \left( \frac{\max_{|s| \leq |t|} B_j (c_N/b_N) - \mu(t - s)}{c_N/b_N} \right) \right) \xrightarrow{d} \left( \sup_{s \in [0,t]} (X(s,t) - \mu(t - s)), t \in [0,T] \right),
\]

as \( N \to \infty \), by using the earlier results together with Lemma 4 from [7, Thm. 13.3].

In summary, we prove process convergence of the maximum waiting time through three steps; first, pointwise convergence follows from Theorem 1, second, we show in Lemma 10 that the maximum waiting process is asymptotically equivalent to an extremal process that only depends on the regularly varying random variables, and finally, we prove process convergence for this latter process in Theorem 2.

In Section 6 we show that the cumulative distribution function of the limiting maximum steady-state waiting time converges to \( \mathbb{P}(\sup_{t \geq 0} (X_t - \mu t) < x) \). This means that the limiting cumulative distribution function of the maximum steady-state waiting time is the same as \( \lim_{T \to \infty} \mathbb{P}(\sup_{s \in [0,T]} (X(s,T) - \mu(T - s)) < x) \), thus the steady-state behavior of the limiting process of \( \sup_{s \in [0,t]} (X(s,t) - \mu(t - s)), t \in [0,T] \) is the same as the extreme-value limit of the maximum steady-state waiting time, which is not a trivial result.

2.2 Other choices for \( A_{i,j} \)

Our main result in Theorem 2 heavily relies on the fact that \( A_{i,j} \) is Weibull-like, and we are able to derive general results. Furthermore, when \( A_{i,j} \) has finite support, we are also able to derive general results. Under the assumption that \( A_{i,j} \) has a finite right endpoint \( b \), it follows that \( b_N = b \). Furthermore, in [30] Lem. B.1, we have shown that \( \max_{1 \leq j \leq N} \sum_{j=1}^n x_j A_{i,j}/b \xrightarrow{P} \sum_{j=1}^n x_j \), as \( N \to \infty \). Then,

\[
\mathbb{P}\left( \max_{1 \leq j \leq N} \sum_{j=1}^k (A_{i,j} B_j - T_j) > x \right) \xrightarrow{N \to \infty} \mathbb{P}\left( \sum_{j=1}^k (bB_j - T_j) > x \right).
\]

Furthermore, if \( \mathbb{E}[bB_j - T_j] < 0 \), we get that

\[
\mathbb{P}\left( \max_{1 \leq j \leq N} \sum_{j=1}^k (A_{i,j} B_j - T_j) > x \right) \xrightarrow{N \to \infty} \mathbb{P}\left( \sup_{k \geq 0} \sum_{j=1}^k (bB_j - T_j) > x \right).
\]

In general, when \( A_{i,j} \) has unbounded support, from [30] Lem. B.1 follows that

\[
\mathbb{P}\left( \max_{1 \leq j \leq N} \sum_{j=1}^k (A_{i,j} B_j - T_j) > x b_N \right) \xrightarrow{N \to \infty} \mathbb{P}\left( \max_{z \in \mathbb{R}} \left\{ \sum_{j=1}^l z_j B_j \right\} 1 \leq j \leq 1 \right) > x.
\]
For the heavy-tailed case, i.e., when $0 < \alpha \leq 1$, we have that \[ \max_j \left\{ \sum_{j=1}^l z_j B_j \left| \sum_{j=1}^l z_j^\alpha \leq 1, 0 \leq z_j \leq 1 \right. \right\} = \max_{j \leq l} B_j. \] In contrary, for the light-tailed case, i.e., when $\alpha > 1$, we get nontrivial results depending on $\alpha$. For instance for $\alpha = 2$, we obtain \[ \max_{j \leq l} \left\{ \sum_{j=1}^l z_j B_j \left| \sum_{j=1}^l z_j^\alpha \leq 1, 0 \leq z_j \leq 1 \right. \right\} = \sqrt{B_1^2 + \cdots + B_l^2}. \] In conclusion, when $\alpha > 1$, we cannot rely on the property described in (20)–(22) that follows from the behavior of Weibull-distributed random variables. In this paper, we limit ourselves to the case $0 < \alpha < 1$; the case $\alpha > 1$ falls outside the scope of this paper. The case $\alpha = 1$ lies on the boundary between these two regimes; this case needs a separate analysis.

3 Preliminary results

In Section 2.1 we gave heuristic ideas of our results. In order to be able to prove these, we need some auxiliary lemmas.

In (7), (8), and (9), we heuristically describe the behavior of the sequence $(c_N, N \geq 1)$ and the slowly varying function $\tilde{L}$ given a sequence $(b_N, N \geq 1)$ and a slowly varying function $L$. An unanswered question is whether this sequence $(c_N, N \geq 1)$ and this function $\tilde{L}$ exists. In Lemma 1 we show how, if this $\tilde{L}$ exists, $L$ and $\tilde{L}$ are asymptotically related. Their asymptotic relation resembles the asymptotic relation between a slowly varying function $l$ and its de Bruijn conjugate $l^\#$, cf. [8, Thm. 1.5.13]. The proof of the existence of $\tilde{L}$ is analogous to the proof of existence of $l^\#$ given in [8, Thm. 1.5.13], thus we omit it here.

Lemma 1 (Asymptotic behavior of $\tilde{L}(x)$). From (7), (8), and (9) follows that the function $\tilde{L}$ satisfies the relation

$$\tilde{L}(x) \sim L(\tilde{L}(x)x^{1/(\beta-1)})^{1/(\beta-1)}, \tag{26}$$

as $x \to \infty$.

Proof. We write $x = b_N$, then the relation in (7) can be rewritten to

$$\tilde{L}(x)x^{\beta/(\beta-1)} \sim \frac{(\tilde{L}(x)x^{\beta/(\beta-1)}/x)^\beta}{L(\tilde{L}(x)x^{\beta/(\beta-1)}/x)},$$

as $x \to \infty$. This simplifies to

$$\tilde{L}(x) \sim \frac{\tilde{L}(x)^\beta}{L(\tilde{L}(x)x^{1/(\beta-1)})},$$

as $x \to \infty$. The lemma follows.

Remark 1 (Asymptotic solutions of $\tilde{L}(x)$). It is not trivial to find functions $\tilde{L}(x)$ that have the asymptotic relation described in (26), since $\tilde{L}$ appears both on the left and the right side of the equation. However, we know that $\tilde{L}$ is slowly varying, thus the term $x^{1/(\beta-1)}$ is dominant in $L(\tilde{L}(x)x^{1/(\beta-1)})^{1/(\beta-1)}$, so we can remove $\tilde{L}$ from the right-hand side in (20) and look at the function $\tilde{L}^{(1)}$ that equals

$$\tilde{L}^{(1)}(x) = L(x^{1/(\beta-1)})^{1/(\beta-1)}.$$

For example, when $L(x) = \log x$, $\tilde{L}^{(1)}$ satisfies (25). However, there are also examples where $\tilde{L}^{(1)}$ does not satisfy the relation in (20), for example, when $L(x) = \exp(\sqrt{\log x})$. Still, we are able to find candidates that satisfy the relation in (20). First, we see that the relation in (20) is actually an iterative relation. Thus, we can rewrite (20) to

$$\tilde{L}(x) \sim L(L(\tilde{L}(x)x^{1/(\beta-1)})^{1/(\beta-1)})^{1/(\beta-1)},$$

as $x \to \infty$. Now, with the same reasoning as before, we define

$$\tilde{L}^{(2)}(x) = L(L(x^{1/(\beta-1)})^{1/(\beta-1)})^{1/(\beta-1)}.$$ 

The function $\tilde{L}^{(2)}$ satisfies the relation in (20) when $L(x) = \exp(\sqrt{\log x})$ and is a slowly varying function itself.

In order to prove that the heuristic approximations in Equations (20)–(22) are correct, we need to prove two things; first, that the largest regularly varying random variable determines the stochastic part of the limit, and second, that the other random variables satisfy the law of large numbers. To prove this second property, we use Bennett’s inequality as stated below. In Corollary 1 we state a simplified version of this inequality which we use in our proofs.
Corollary 1. Let $Y_1, \ldots, Y_n$ be independent random variables, $\mathbb{E}[Y_i] = 0$, $\mathbb{E}[Y_i^2] = \sigma_i^2$, and $|Y_i| < M \in \mathbb{R}$ almost surely. Then for $y > 0$,

$$
\mathbb{P}\left( \sum_{i=1}^{n} Y_i > y \right) \leq \exp\left( -\frac{n}{M^2} \sum_{i=1}^{n} \sigma_i^2 \log\left( 1 + \frac{yM}{\sum_{i=1}^{n} \sigma_i^2} \right) \right),
$$

with $h(x) = (1 + x)\log(1 + x) - x$.

For a proof, cf. [35].

**Proof.** Observe that for $x > 0$ we get that $h(x) > x\log(1 + x) - 1$. Now, the corollary follows from Lemma 2. \qed

Though in [30, Lem. B.1] it is proven that for Weibull-distributed random variables $\max_{i \leq N} \sum_{j=1}^{n} A_{i,j}b_j/b_N \xrightarrow{p} \max_{j \leq n} b_j$, as $N \to \infty$, which heuristically explains the nature of our main result, our approximations in Equations (20)–(22) suggest that we should take the sum of $[tCN]$ random variables. In [30, Lem. B.1] however, $n$ does not depend on $N$. Thus, we cannot resort to [30, Lem. B.1] in our proofs. However, in Lemma 3 a result is presented that we can use in this paper and proves the approximations in Equations (20)–(22).

**Lemma 3** ([9, Thm. 2]). Let $Y_1, \ldots, Y_n$ be independent random variables, with $\log \mathbb{P}(Y_i > x) \sim -qx^\alpha$, as $x \to \infty$, with $0 < \alpha < 1$ and $q > 0$. Let $(x_n, n \geq 1)$ be a sequence such that $\lim_{n \to \infty} x_n/n^{1/(2-\alpha)} = \infty$. Then

$$
\lim_{n \to \infty} \frac{1}{x_n^\alpha} \log \mathbb{P}\left( \sum_{i=1}^{n} Y_i > x_n \right) = -q
$$

We want to prove process convergence of the maximum waiting time to a limiting process, this limiting process is a function in $D[0,T]$. In [7, Thm. 13.3], a result is given that guarantees the convergence of a process in $D[0,T]$ when three conditions are satisfied, which we will apply in this paper.

**Lemma 4** ([7, Thm. 13.3]). Assume a sequence of processes $(Y^{(n)}(t), t \in [0,T])$ and a process $(Y(t), t \in [0,T])$ in $D[0,T]$, equipped with the $d^0$ metric, satisfy the following conditions:

1. For all $\{t_1, \ldots, t_k\} \subseteq [0,T]$: $(Y^{(n)}(t_1), \ldots, Y^{(n)}(t_k)) \xrightarrow{d} (Y(t_1), \ldots, Y(t_k))$ as $N \to \infty$.
2. $Y(T) - Y(T - \delta) \xrightarrow{p} 0$ as $\delta \downarrow 0$, and
3. For $0 < r < s < t < T$, $\epsilon, \eta > 0$ there exists $N_0 \geq 1$ and $\delta > 0$ such that

$$
\mathbb{P}\left( \sup_{s \in [r,t], t-r<\delta} \min \left( \left| Y^{(n)}(s) - Y^{(n)}(r) \right|, \left| Y^{(n)}(t) - Y^{(n)}(s) \right| \right) > \epsilon \right) \leq \eta
$$

for $N \geq N_0$. Then $(Y^{(n)}(t), t \in [0,T]) \xrightarrow{d} (Y(t), t \in [0,T])$ as $N \to \infty$.

Finally, to prove pointwise convergence of the maximum waiting time process in $\mathbb{B}^2$ to the limiting random variable, we need to pay special attention to the case that $B$ is a regularly varying random variable with $1 < \beta \leq 2$, since in this case the second moment of $B$ is not finite. In Lemma 5, we give a useful convergence result of the second moment of $B$ conditioned on $B$ being bounded.

**Lemma 5.** Let $B$ be a positive random variable that satisfies $\mathbb{P}(B > x) = L(x)/x^\beta$, with $L(x)$ a slowly varying function and $1 < \beta \leq 2$. Then,

$$
\frac{\mathbb{E}[B^2|B < r]}{r} \to 0,
$$

as $r \to \infty$.

**Proof.** Choose $0 < \epsilon < \beta - 1$. Because $\mathbb{P}(B > x) = L(x)/x^\beta$, we have that $\mathbb{E}[B^{\beta-\epsilon}] < \infty$. Therefore,

$$
\frac{\mathbb{E}[B^2|B < r]}{r} \leq \frac{r^{2-(\beta-\epsilon)}}{r} \mathbb{E}[B^{\beta-\epsilon}] \to 0,
$$

as $r \to \infty$. \qed
4 Convergence of the auxiliary process in $D[0, T]$

In this section, we prove Theorem 1. As explained in Section 2.1, we first remove the supremum functional from the random variable on the left-hand side in (12) and prove convergence of the process $(\max_{t \leq N} \sum_{j=1}^{[tCN]} (A_{i,j} B_j - T_j)/\epsilon N, t \in [0, T])$ to $(X_i - \mu, t \in [0, T])$. To do so, we first show pointwise convergence in Lemma 6. Afterwards we prove process convergence in Lemma 7. In order to prove Lemma 7, we need two auxiliary results, which are given in Lemmas 8 and 9. By using the continuous mapping theorem, Theorem 1 follows.

**Lemma 6.** Given that Assumptions 4, 5 hold, $t > 0$, and $x > 0$, then

$$\mathbb{P}\left( \max_{t \leq N} \sum_{j=1}^{[tCN]} (A_{i,j} B_j - T_j) > xcN \right) \xrightarrow{N \to \infty} 1 - \exp\left( -\frac{t}{(x+\mu)^\beta} \right).$$  \hspace{1cm} (27)

**Proof.** The approach to prove this lemma is by analyzing upper and lower bounds of the probability given in (27) and by proving that these bounds are sharp as $N \to \infty$. Thus, first we see that

$$\mathbb{P}\left( \max_{t \leq N} \sum_{j=1}^{[tCN]} (A_{i,j} B_j - T_j) > xcN \right) \leq \mathbb{P}\left( \max_{t \leq N} \sum_{j=1}^{[tCN]} (A_{i,j} B_j - T_j) > xcN \right) \mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right) \mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right).$$ \hspace{1cm} (28)

We analyze this term as follows; first, we have that (28) yields

$$\mathbb{P}\left( \max_{t \leq N} \sum_{j=1}^{[tCN]} (A_{i,j} B_j - T_j) > xcN \right) \leq \mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right) \mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right).$$ \hspace{1cm} (29)

The first term in (29) yields

$$\mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right) \sim 1 - \left( 1 - \frac{L((x+\mu-\delta)\epsilon N/bN)}{(x+\mu-\delta)\epsilon N/bN)^\beta} \right)^{[tCN]} \xrightarrow{N \to \infty} 1 - \exp\left( -\frac{t}{(x+\mu)^\beta} \right).$$

Hence, in order to prove that the upper bound of (27) is asymptotically sharp, we are left with proving that the second term in (30) vanishes as $N \to \infty$. We analyze this term as follows; first, we have that $(x+\mu-\delta/2)/(x+\mu-\delta) > 1$ for $\delta$ small enough, thus we write $(x+\mu-\delta/2)/(x+\mu-\delta) = 1 + \epsilon$ with $\epsilon > 0$. Second, we can bound

\begin{align*}
\mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right) &\leq \mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right) \\
&= \mathbb{P}\left( \max_{j \leq [tCN]} B_j > (x+\mu-\delta)\frac{\epsilon N}{bN} \right)
\end{align*}
the second term in (31) as

\[
P\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j}B_j - T_j) > xCN \right) \leq \max_{j \leq [tc_N]} B_j \leq (x + \mu t - \delta)\frac{CN}{bN} \tag{32} \]

\[
\leq \mathbb{P}\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j}I(A_{i,j} < (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha})B_j - T_j) + \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j}I(A_{i,j} \geq (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha})B_j > xCN \right) \tag{33} \]

\[
\leq \mathbb{P}\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j}I(A_{i,j} < (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha})B_j > \left( \mathbb{E}[A_{i,j}B_j]t + \frac{\delta}{4} \right) cN \right) \left( \max_{j \leq [tc_N]} B_j \leq (x + \mu t - \delta)\frac{CN}{bN} \right) \tag{34} \]

\[
+ \mathbb{P}\left( \sum_{j=1}^{[tc_N]} -T_j > \left( -\mathbb{E}[T_j]t + \frac{\delta}{4} \right) cN \right) \tag{35} \]

\[
+ \mathbb{P}\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j}I(A_{i,j} \geq (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha})B_j > \left( x + \mu t - \frac{\delta}{2} \right) cN \right) \left( \max_{j \leq [tc_N]} B_j \leq (x + \mu t - \delta)\frac{CN}{bN} \right). \tag{36} \]

The upper bound in (33) holds because for a sequence of numbers \((a_{i,j}, i \geq 1, j \geq 1)\), we have that

\[
\max_{i \leq N} \left( \sum_{j=1}^{k} a_{ij} \right) \leq \max_{i \leq N} \sum_{j=1}^{k} a_{ij}I(a_{ij} \leq c) + \max_{i \leq N} \sum_{j=1}^{k} a_{ij}I(a_{ij} > c). \]

The upper bound from (33) to (34), (35), and (36) holds because of the union bound. The term in (35) converges to 0 due to the law of large numbers. For the term in (34), we know by the union bound that

\[
\mathbb{P}\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j}I(A_{i,j} < (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha})B_j > \left( \mathbb{E}[A_{i,j}B_j]t + \frac{\delta}{4} \right) cN \right) \left( \max_{j \leq [tc_N]} B_j \leq (x + \mu t - \delta)\frac{CN}{bN} \right)
\]

\[
\leq N \mathbb{P}\left( \sum_{j=1}^{[tc_N]} A_{i,j}I(A_{i,j} < (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha})B_j > \left( \mathbb{E}[A_{i,j}B_j]t + \frac{\delta}{4} \right) cN \right) \left( \max_{j \leq [tc_N]} B_j \leq (x + \mu t - \delta)\frac{CN}{bN} \right) \tag{37} \]

Now, since we have a probability of sums of almost surely bounded random variables, we can apply Bennett’s inequality with the setting given in Lemma 2 and Corollary 1. We see that \(\mathbb{E}[A_{i,j}B_j] < (1 + \epsilon)^{1-\alpha}b_1^{1-\alpha} \leq (x + \mu t - \delta)CN/bN\), \(B_j \leq (x + \mu t - \delta)CN/bN\), and \(y = \delta/4CN\). Thus,

\[
y/M = \frac{\delta}{4(x + \mu t - \delta)(1 + \epsilon)^{1-\alpha}b_1^{1-\alpha}} = \frac{\delta}{4(x + \mu t - \delta)(1 + \epsilon)^{1-\alpha}} \log N.
\]

It is important to note here, that \(y/M\) equals a constant times \(\log N\). We now add a subscript \(N\) to the variables \(y, M\), and \(\sigma^2\) to indicate sequences that change with \(N\). Now, for \(\beta > 2\), \(\lim \sup_{N \to \infty} \sigma^2_{i,N} \to \infty\). Thus,

\[
\frac{y_N M_N}{\sum_{j=1}^{[tc_N]} \sigma^2_{i,N}} \xrightarrow{N \to \infty} \infty.
\]

Therefore, using the information that \(y/M\) equals a constant times \(\log N\) and by using Corollary 1 we see that the exponent in Corollary 1 grows faster to infinity than \(\log N\). Thus, by applying Bennett’s inequality, we get that the expression in (37) converges to 0 as \(N \to \infty\). When \(1 < \beta \leq 2\), \(\sigma^2_{i,N} \xrightarrow{N \to \infty} \infty\), however, from Lemma 5 follows that \(\sigma^2_{i,N}/(CN/bN) \xrightarrow{N \to \infty} 0\). Therefore, \(y_N M_N/\sum_{j=1}^{[tc_N]} \sigma^2_{i,N} \xrightarrow{N \to \infty} \infty\). Concluding, from Corollary 1 we again get that the expression in (37) and therefore the expression in (34) converge to 0.
Furthermore, for the term in (36) we have that
\[
P\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}(A_{i,j} \geq (1 + \epsilon)1^{-\alpha}b_N^{1-\alpha})B_j > \left(x + \mu t - \frac{\delta}{2}\right)c_N \right) \leq P \left( \max_{j \leq [tc_N]} B_j \leq \left(x + \mu t - \delta \right)\frac{CN}{b_N} \right)
\]
\[
\leq P \left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}(A_{i,j} \geq (1 + \epsilon)1^{-\alpha}b_N^{1-\alpha}) > \frac{x + \mu t - \delta/2}{x + \mu t - \delta}b_N \right).
\]
We have \((x + \mu t - \delta/2)/(x + \mu t - \delta) = 1 + \epsilon\) with \(\epsilon > 0\), thus we can further simplify and bound this probability as follows:
\[
P\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}(A_{i,j} \geq (1 + \epsilon)1^{-\alpha}b_N^{1-\alpha}) > (1 + \epsilon)b_N \right)
\]
\[
\leq P \left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}(A_{i,j} \geq (1 + \epsilon)1^{-\alpha}b_N^{1-\alpha}) > (1 + \epsilon)b_N \cap \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} > (1 + \epsilon)b_N \right)
\]
\[
+ P \left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}(A_{i,j} \geq (1 + \epsilon)1^{-\alpha}b_N^{1-\alpha}) > (1 + \epsilon)b_N \cap \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} < (1 + \epsilon)b_N \right) \quad (38)
\]
Since
\[
\max_{i \leq N} \max_{j \leq [tc_N]} A_{i,j} \frac{b_N}{P} \rightarrow 1,
\]
as \(N \rightarrow \infty\), the term in (38) converges to 0 as \(N \rightarrow \infty\), and we only need to focus on the term in (39). Observe that by the union bound,
\[
P\left( \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}(A_{i,j} \geq (1 + \epsilon)1^{-\alpha}b_N^{1-\alpha}) > (1 + \epsilon)b_N \cap \max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} < (1 + \epsilon)b_N \right)
\]
\[
\leq N P \left( \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}((1 + \epsilon)1^{-\alpha}b_N^{1-\alpha} \leq A_{i,j} \leq (1 + \epsilon)b_N) > (1 + \epsilon)b_N \right) \quad (40)
\]
Following the proof given in [9] Lem. 8], we assume without loss of generality that \(q = 1\) and choose \(1/(1 + \epsilon) < q' < 1\) and \(q' < q'' < 1\). Now, we have by using Chernoff’s bound, that for \(\theta > 0\),
\[
N P \left( \sum_{j=1}^{[tc_N]} A_{i,j} \mathbb{1}((1 + \epsilon)1^{-\alpha}b_N^{1-\alpha} \leq A_{i,j} \leq (1 + \epsilon)b_N) > (1 + \epsilon)b_N \right)
\]
\[
\leq N \left( 1 + \mathbb{E} \left( \exp(\theta A_{i,j}) \mathbb{1}((1 + \epsilon)1^{-\alpha}b_N^{1-\alpha} \leq A_{i,j} \leq (1 + \epsilon)b_N) \right) \right)^{[tc_N]} \exp(-\theta(1 + \epsilon)b_N).
\]
Then, for \(\theta = q'(1 + \epsilon)^{\alpha-1}b_N^{\alpha-1}\), in [9] Lem. 8] it is proven that for \(N\) large enough
\[
\mathbb{E} \left( \exp(\theta A_{i,j}) \mathbb{1}((1 + \epsilon)1^{-\alpha}b_N^{1-\alpha} \leq A_{i,j} \leq (1 + \epsilon)b_N) \right) \leq (1 + q'(1 + \epsilon)\alpha\alpha b_N^{\alpha(1-\alpha)})^\alpha.
\]
Now, by using the fact that \(x > 0\) we have the simple bound \(1 + x \leq \exp(x)\), and that \(c_N = \tilde{L}(b_N)\beta^{(\beta-1)}\), it is easy to see that
\[
\left( 1 + (1 + q'(1 + \epsilon)\alpha\alpha b_N^{\alpha(1-\alpha)}) \right)^{[tc_N]} \rightarrow N \rightarrow \infty 1.
\]
Therefore, we know that Chernoff’s bound with \(\theta = q'(1 + \epsilon)^{\alpha-1}b_N^{\alpha-1}\) applied to the expression in (40) satisfies
\[
\limsup_{N \rightarrow \infty} N \left( 1 + \mathbb{E} \left( \exp(\theta A_{i,j}) \mathbb{1}((1 + \epsilon)1^{-\alpha}b_N^{1-\alpha} \leq A_{i,j} \leq (1 + \epsilon)b_N) \right) \right)^{[tc_N]} \cdot \exp(-q'(1 + \epsilon)\alpha\alpha b_N^{\alpha(1-\alpha)}) \leq \limsup_{N \rightarrow \infty} N \exp(-q'(1 + \epsilon)\alpha\alpha b_N^{\alpha}).
\]
Since $q' > 1/(1 + \epsilon)^{\alpha}$, we have that $q'(1 + \epsilon)^{\alpha} b_N^q > \log N$ and therefore that $N \exp(-q'(1 + \epsilon)^{\alpha} b_N^q) \xrightarrow{N \to \infty} 0$. Thus we can conclude that the expression in (40) converges to 0 as $N \to \infty$. From this, it follows that the term in (39) converges to 0 as $N \to \infty$ as well, and we can conclude that the upper bound proposed in (31) is asymptotically sharp.

To prove a sharp lower bound for the probability in (27), observe that, for a sequence $(a_{i,j}, i \geq 1, j \geq 1)$, we have that $\max_{i \leq N} \sum_{j=1}^{k} a_{i,j} \geq \max_{j \leq N} \max_{i,j} a_{i,j} + \sum_{j=1, j \neq j^*}^{k} a_{i,j}$.

$$\liminf_{N \to \infty} \mathbb{P}(\max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j) > xc_N) \geq \liminf_{N \to \infty} \mathbb{P}(\max_{i \leq N} \max_{j \leq [tc_N]} A_{i,j} B_j - T_j^*(t) + \sum_{j=1, j \neq j^*}^{[tc_N]} (A_{i,j} B_j - T_j) > xc_N), \quad (41)$$

where $j^*(t) \in \arg\max\{j : B_j^*(t) = \max_{j \leq [tc_N]} B_j\}$ and $i^*(t) \in \arg\max\{i : A_{i,j}^*(t) = \max_{i \leq N} A_{i,j}^*(t)\}$. Because, $\max_{j \leq [tc_N]} B_j$ scales as $c_N/b_N$, we get that $\mathbb{E}[A_i] \max_{j \leq [tc_N]} B_j/c_N \xrightarrow{p} 0$, as $N \to \infty$, and thus we have that $\sum_{j=1, j \neq j^*(t)}^{[tc_N]} (A_{i,j}^*(t) B_j - T_j)/c_N \xrightarrow{p} -\mu t$, cf. [16, Thm. 1]. Furthermore,

$$\mathbb{P}\left(\max_{i \leq N} \max_{j \leq [tc_N]} (A_{i,j} B_j)/c_N > x + \mu t\right) \xrightarrow{N \to \infty} 1 - \exp(-t/(x + \mu t)^{\beta})$$

and $T_j^*(t)/c_N \xrightarrow{p} 0$ as $N \to \infty$. In conclusion, the lower bound in (41) is sharp, as the limit is the same as the limit in (27).

We have established pointwise convergence. In Lemma 7, we prove convergence in $D[0, T]$.

**Lemma 7.** Given that Assumptions [7][8] hold, and $T > 0$, then

$$\left(\frac{\max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j)}{c_N}, t \in [0, T]\right) \xrightarrow{d} (X_t - \mu t, t \in [0, T]), \quad (42)$$

as $N \to \infty$.

This lemma follows from the two results stated in Lemma 8 and 9.

**Lemma 8.** Given that Assumptions [7][8] hold, and $T > 0$, then

$$\left(\frac{\max_{j \leq [tc_N]} B_j}{c_N/b_N}, t \in [0, T]\right) \xrightarrow{d} (X_t, t \in [0, T]), \quad (43)$$

as $N \to \infty$.

**Lemma 9.** Given that Assumptions [7][8] hold, and $T > 0$, then we have that for all $\epsilon > 0$

$$\mathbb{P}\left(\sup_{t \in [0, T]} \left|\max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j) - \left(\frac{\max_{j \leq [tc_N]} B_j}{c_N/b_N} - \mu t\right)\right| > \epsilon\right) \xrightarrow{N \to \infty} 0. \quad (44)$$

Using the triangle inequality, we get that (42) follows from (43) and (44).

**Proof of Lemma 8.** In this proof, we use Lemma 4 thus we need to prove the three conditions stated in Lemma 4. First of all, we need to prove that

$$\left(\frac{\max_{j \leq [tc_N]} B_j}{c_N/b_N}, \ldots, \frac{\max_{j \leq [tm_c N]} B_j}{c_N/b_N}\right) \xrightarrow{d} (X_{t_1}, \ldots, X_{t_m}),$$

as $N \to \infty$. Let us assume that $m = 2$ and $t_2 > t_1$. If $x_2 \leq x_1$, because $\max_{j \leq k} B_j$ is increasing in $k$, we have that

$$\mathbb{P}\left(\frac{\max_{j \leq [t c_N]} B_j}{c_N/b_N} \leq x_1 \cap \frac{\max_{j \leq [t_2 c_N]} B_j}{c_N/b_N} \leq x_2\right) = \mathbb{P}\left(\frac{\max_{j \leq [t_2 c_N]} B_j}{c_N/b_N} \leq x_2\right) \xrightarrow{N \to \infty} \mathbb{P}(X_{t_2} \leq x_2) = \mathbb{P}(X_{t_1} \leq x_1 \cap X_{t_2} \leq x_2).$$
When \( x_2 > x_1 \), we have that

\[
P \left( \frac{\max_{j \leq \lfloor t_{2CN} \rfloor} B_j}{c_N/b_N} \leq x_1 \cap \frac{\max_{j \leq \lfloor t_{2CN} \rfloor} B_j}{c_N/b_N} \leq x_2 \right) = P \left( \frac{\max_{j \leq \lfloor t_{2CN} \rfloor} B_j}{c_N/b_N} \leq x_2 \left| \frac{\max_{j \leq \lfloor t_{1CN} \rfloor} B_j}{c_N/b_N} \leq x_1 \right. \right) P \left( \frac{\max_{j \leq \lfloor t_{1CN} \rfloor} B_j}{c_N/b_N} \leq x_1 \right)
\]

\[
= P \left( \frac{\max_{j \leq \lfloor t_{2CN} \rfloor} B_j}{c_N/b_N} \leq x_2 \right) P \left( \max_{j \leq \lfloor t_{1CN} \rfloor - \lfloor t_{1CN} \rfloor} B_j \leq x_1 \right)
\]

\[
N \to \infty \quad P(\tau_{t_2-t_1} \leq x_2) P(\tau_{t_1} \leq x_1) = P(\tau_{t_1} \leq x_1 \cap \tau_{t_2} \leq x_2).
\]

For \( m > 2 \) but finite, analogous derivations hold. Second, we need to prove that

\[
X_T - X_{T-\delta} \xrightarrow{p} 0,
\]

as \( \delta \downarrow 0 \). We can write \( X_T = \max(X_{T-\delta}, \hat{X}_\delta) \). Therefore, \( X_T - X_{T-\delta} \leq \hat{X}_\delta \). Let \( \epsilon > 0 \), then

\[
P(X_T - X_{T-\delta} > \epsilon) \leq P(\hat{X}_\delta > \epsilon) = 1 - \exp \left( -\frac{\delta}{\epsilon} \right) \frac{\delta \epsilon}{0}.
\]

Finally, we show that the process \( (\max_{j \leq \lfloor c_N \rfloor} B_j/c_N/b_N, t \in [0, T]) \) satisfies the third condition in Lemma 4. The random variable \( \max_{j \leq k} B_j \) is increasing with \( k \). Furthermore, the minimum of two numbers is bounded from above by the average. Also, because for \( k > l \), \( \max_{j \leq k} B_j - \max_{j \leq l} B_j = \max(\max_{j \leq k} B_j, \max_{j \leq k+1} B_j) - \max_{j \leq k} B_j \), we can bound

\[
\frac{\max_{j \leq \lfloor sCN \rfloor} B_j - \max_{j \leq \lfloor rCN \rfloor} B_j}{c_N/b_N} \leq \frac{\max_{j \leq \lfloor sCN \rfloor} B_j}{c_N/b_N} - \frac{\max_{j \leq \lfloor rCN \rfloor} B_j}{c_N/b_N} \leq \frac{\max_{j \leq \lfloor tCN \rfloor} B_j}{c_N/b_N} - \frac{\max_{j \leq \lfloor sCN \rfloor} B_j}{c_N/b_N}.
\]

Thus, using the expression in the third condition of Lemma 4, we obtain that

\[
P \left( \sup_{s \in [r,t], t-r<\delta} \min_{2\lfloor sCN/b_N \rfloor} \frac{\max_{j \leq \lfloor sCN \rfloor} B_j - \max_{j \leq \lfloor rCN \rfloor} B_j}{c_N/b_N} \right) > \epsilon
\]

\[
\leq P \left( \frac{\max_{j \leq \lfloor sCN \rfloor - \lfloor tCN \rfloor} B_j}{c_N/b_N} > 2\epsilon \right)
\]

\[
\leq \frac{\lfloor sCN \rfloor}{2\epsilon c_N/b_N}, (2\epsilon c_N/b_N)^3.
\]

We have that \( c_N L(2\epsilon c_N/b_N)/(c_N/b_N)^\beta \to 1 \) because \( c_N \sim (c_N/b_N)^\beta/L(c_N/b_N) \) as \( N \to \infty \) and \( L \) is a slowly varying function, so we choose \( N_0 > 1 \) such that \( \|c_N\| L(2\epsilon c_N/b_N)/(2\epsilon c_N/b_N)^\beta < (1+\epsilon)\delta/(2\epsilon)^\beta \) for all \( N > N_0 \).

Now, choose \( 0 < \delta < \eta(2\epsilon)^\beta/(1+\epsilon) \) and we get that for \( N > N_0 \)

\[
P \left( \sup_{s \in [r,t], t-r<\delta} \min_{2\lfloor sCN/b_N \rfloor} \frac{\max_{j \leq \lfloor sCN \rfloor} B_j - \max_{j \leq \lfloor rCN \rfloor} B_j}{c_N/b_N}, \frac{\max_{j \leq \lfloor tCN \rfloor} B_j - \max_{j \leq \lfloor sCN \rfloor} B_j}{c_N/b_N} \right) > \epsilon
\]

\[
\leq \eta.
\]

Hence, the process \( (\max_{j \leq \lfloor tCN \rfloor} B_j/c_N/b_N, t \in [0, T]) \) also satisfies the third condition in Lemma 4 and the result follows.
We have proven process convergence of \((\max_{j \leq [t\wedge N]} B_j/(cN/bN), t \in [0, T])\) to \((X_t, t \in [0, T])\), now in order to prove Lemma 9, we are left by proving that the convergence result in (44) holds. We do this in Lemma 9.

**Proof of Lemma 9.** The random variable in (44) has the form of a supremum of the absolute value of a stochastic process. We know that \(|X| = \max(X, -X)|. Then, by applying the union bound we get that \(P(|X| > x) \leq P(X > x) + P(-X > x)\). Thus, to prove the convergence result in (44) we can remove the absolute value and need to prove that the probability of a supremum of a stochastic process converges to 0 as \(N \to \infty\), cf. (45), and we need to prove that the probability of the supremum of the mirrored process converges to 0 as \(N \to \infty\), cf. (42). We first prove that

\[
P \left( \sup_{t \in [0, T]} \left( -\max_{i \leq N} \sum_{j=1}^{[t\wedge N]} (A_i, B_j - T_j) \right) + \left( \max_{j \leq [t\wedge N]} B_j \right) > \epsilon \right) > 0 \tag{45} \]

converges to 0 as \(N \to \infty\). We have, by using \(i^*(t)\) and \(j^*(t)\) as defined in Lemma 6 that

\[
P \left( \sup_{t \in [0, T]} \left( -\max_{i \leq N} \sum_{j=1}^{[t\wedge N]} (A_i, B_j - T_j) \right) + \left( \max_{j \leq [t\wedge N]} B_j \right) > \epsilon \right) \leq P \left( \sup_{t \in [0, T]} \left( -\mu t - \sum_{j=1}^{[t\wedge N]} (A_i^*(t), B_j - T_j) + \sum_{j=1}^{[t\wedge N]} \max_{i \leq N} A_i, j^*(t) \sum_{j=1}^{[t\wedge N]} B_j > \epsilon \right) \right. \]

\[
+ P \left( \sup_{t \in [0, T]} \left( \frac{\max_{j \leq [t\wedge N]} B_j}{cN/bN} - \max_{i \leq N} A_i, j^*(t) \sum_{j=1}^{[t\wedge N]} B_j \right) \right) > \epsilon \right). \tag{46} \]

For the term in (46), we use the union bound to obtain that

\[
P \left( \sup_{t \in [0, T]} \left( -\mu t - \sum_{j=1}^{[t\wedge N]} (A_i^*(t), B_j - T_j) + \sum_{j=1}^{[t\wedge N]} \max_{i \leq N} A_i, j^*(t) \sum_{j=1}^{[t\wedge N]} B_j \right) \right) > \epsilon \tag{47} \]

\[
P \left( \sup_{t \in [0, T]} \left( -\mu t - \sum_{j=1}^{[t\wedge N]} (A_i^*(t), B_j - T_j) + \sum_{j=1}^{[t\wedge N]} \max_{i \leq N} A_i, j^*(t) \sum_{j=1}^{[t\wedge N]} B_j \right) \right) > \epsilon \tag{48} \]

Because all random variables \(A_i, B_j,\) and \(T_j\) are positive, it is easy to see that the term in (48) can be upper bounded by

\[
P \left( \sup_{t \in [0, T]} \left( \sum_{j=1}^{[t\wedge N]} \frac{T_j}{cN} \right) > \epsilon \right) = \mathbb{P} \left( \sum_{j=1}^{[t\wedge N]/cN} \frac{T_j}{cN} > \epsilon \right) \to 0 \text{ as } N \to \infty \tag{49} \]

as we can conclude from the law of large numbers that \(\sum_{j=1}^{[t\wedge N]/cN} T_j/cN \to \epsilon/4\) as \(N \to \infty\). For the term in
conclude that the number of different instances of converging to 1, and therefore, we can use the union bound to bound \[ P \left( \sup_{t \in [\epsilon/(4E[T_j]), T]} \left( -\mu t - \sum_{j=1,j \neq j^*(t)}^{[tcN]} c_N \right) \right) \geq \epsilon \]

\[ \leq \sup_{t \in [\epsilon/(4E[T_j]), T]} \frac{1}{\delta} \left( \sup_{s \in [t,t+\delta]} \left( -\mu s - \sum_{j=1,j \neq j^*(s)}^{[scN]} c_N \right) \right) > \epsilon \]

\[ \leq \sup_{t \in [\epsilon/(4E[T_j]), T]} \frac{1}{\delta} \left( \frac{-\mu \inf_{s \in [t,t+\delta]} \sum_{j=1,j \neq j^*(s)}^{[tcN]} c_N}{c_N} + \sum_{j=1}^{[tcN]} \frac{T_j}{c_N} \right) > \epsilon . \] (50)

To bound the term in (50), we use the result from [13, Eq. (6)] that the expected number of new extremes of the process \( \max_{j \leq [scN]} B_j, s \geq 0 \) in the interval \([t, t+\delta]\) equals \( \sum_{j=1}^{[t+\delta]cN} 1/j \rightarrow \infty \) \( N \rightarrow \infty \log((t+\delta)/t) \). Therefore, we can conclude that the number of different instances of \( i^*(s) \) when \( s \in [t, t+\delta] \) is asymptotically finite, with probability converging to 1, and therefore, we can use the union bound to bound

\[ \sup_{t \in [\epsilon/(4E[T_j]), T]} \frac{1}{\delta} \left( -\mu t - \sum_{j=1,j \neq j^*(t)}^{[tcN]} A_{i,j} B_j + \sum_{j=1}^{[tcN]} T_j \right) > \epsilon \]

\[ \leq \sup_{t \in [\epsilon/(4E[T_j]), T]} \sum_{k=0}^{k} (k+1) \left( \# \text{ new extremes of } \left( \max_{j \leq [scN]} B_j, s \geq 0 \right) \text{ in } [t, t+\delta] = k \right) \]

\[ \cdot \frac{1}{\delta} \left( -\mu t - \sum_{j=1,j \neq j^*(t)}^{[tcN]} A_{i,j} B_j + \sum_{j=1}^{[tcN]} T_j \right) > \epsilon \]

\[ = \sup_{t \in [\epsilon/(4E[T_j]), T]} \left( 1 + \sum_{j=1}^{[tcN]} \frac{1}{j} \right) \frac{1}{\delta} \left( -\mu t - \sum_{j=1,j \neq j^*(t)}^{[tcN]} A_{i,j} B_j + \sum_{j=1}^{[tcN]} T_j \right) > \epsilon . \]

This last expression converges to 0 as \( N \to \infty \), when \( \delta > 0 \) is small enough; that is \( 0 < \delta < \epsilon / (2E[T_j]) \), because by the law of large numbers we obtain that \( -\mu t - \sum_{j=1,j \neq j^*(t)}^{[tcN]} A_{i,j} B_j/c_N + \sum_{j=1}^{[tcN]} T_j/c_N \rightarrow -\mu t - \mathbb{E}[A_{i,j} B_j] t + (t+\delta)E[T_j] = \mathbb{E}[T_j] \delta \) as \( N \to \infty \). Thus, we can conclude that the expression in (50), and therefore also in (49), converge to 0, as \( N \to \infty \). For the term in (47), we have that

\[ P \left( \sup_{t \in [0,T]} \left( \frac{\max_{j \leq [tcN]} B_j}{c_N/b_N} - \max_{i \leq N} A_{i,j} (t) \max_{j \leq [tcN]} B_j \right) > \epsilon \right) \]

\[ \leq P \left( \sup_{t \in [0,T]} \left( 1 - \frac{\max_{i \leq N} A_{i,j} (t)}{b_N} \right) \frac{\max_{j \leq [tcN]} B_j}{c_N/b_N} > \epsilon / 2 \right) . \] (51)

This tail probability converges to 0 as \( N \to \infty \), since we know that \( \max_{j \leq [tcN]} B_j/c_N \) converges in distribution to a Fréchet random variable as \( N \to \infty \), and \( \sup_{t \in [0,T]} \left( 1 - \max_{i \leq N} A_{i,j (t)} / b_N \right) \rightarrow 0 \), as \( N \to \infty \). To see this, we first bound

\[ P \left( \sup_{t \in [0,T]} \left( 1 - \frac{\max_{i \leq N} A_{i,j (t)}}{b_N} \right) > \epsilon / 2 \right) = P \left( \inf_{t \in [0,T]} \frac{\max_{i \leq N} A_{i,j (t)}}{b_N} < 1 - \epsilon / 2 \right) \]

\[ \leq [TcN] P \left( \frac{\max_{i \leq N} A_{i,1}}{b_N} < 1 - \epsilon / 2 \right) . \]

Now, we have that \( P(\max_{i \leq N} A_{i,1}/b_N < 1 - \epsilon / 2) \leq \exp(-N P(A_{i,1}/b_N > 1 - \epsilon / 2)) \), cf. the proof of [14, Thm. 5.4.4].
conclude that by applying the continuous mapping theorem, Theorem 1 follows.

Hence, the upper bound in (31) converges to 0 as \( N \to \infty \). These results together give that the tail probability in (15) converges to 0 as \( N \to \infty \).

To prove the convergence result in (44), we are left with proving that the probability

\[
\mathbb{P} \left( \sup_{t \in [0,T]} \left( \frac{\max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j)}{\epsilon N} - \left( \frac{\max_{j \leq [tc_N]} B_j}{\epsilon N / b_N} - \mu t \right) \right) > \epsilon \right)
\]

(52)

converges to 0 as \( N \to \infty \). In order to do so, we use the upper bound

\[
\mathbb{P} \left( \sup_{t \in [0,T]} \left( \frac{\max_{i \leq N} \sum_{j=1}^{[tc_N]} (A_{i,j} B_j - T_j)}{\epsilon N} - \left( \frac{\max_{j \leq [tc_N]} B_j}{\epsilon N / b_N} - \mu t \right) \right) > \epsilon \right) \leq \sup_{t \in [0,T]} \frac{1}{\delta} \mathbb{P} \left( \sup_{s \in [t, t+\delta]} \left( \frac{\max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j)}{\epsilon N} - \left( \frac{\max_{j \leq [sc_N]} B_j}{\epsilon N / b_N} - \mu s \right) \right) > \epsilon \right)
\]

(53)

with \( 0 < \delta < 1 \). Now, we bound this further as follows;

\[
\frac{1}{\delta} \mathbb{P} \left( \sup_{s \in [t, t+\delta]} \left( \frac{\max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j)}{\epsilon N} - \left( \frac{\max_{j \leq [sc_N]} B_j}{\epsilon N / b_N} - \mu s \right) \right) > \epsilon \right) \leq \frac{1}{\delta} \mathbb{P} \left( \sup_{s \in [t, t+\delta]} \frac{\max_{i \leq N} \sum_{j=1}^{[sc_N]} (A_{i,j} B_j - T_j)}{\epsilon N} - \left( \frac{\max_{j \leq [sc_N]} B_j}{\epsilon N / b_N} + \mu s \right) > \frac{\epsilon}{2} \right)
\]

+ \frac{1}{\delta} \mathbb{P} \left( \sup_{s \in [t, t+\delta]} \left( \frac{\max_{i \leq N} \sum_{j=1}^{[sc_N]} A_{i,j} 1(A_{i,j} > b_N^{1-\alpha}) B_j}{\epsilon N} - \frac{\max_{j \leq [sc_N]} B_j}{\epsilon N / b_N} \right) > \frac{\epsilon}{2} \right).

The first term vanishes asymptotically, by taking \( \delta \) small enough compared to \( \epsilon \), by using a similar argument as for bounding (50) and by using the same argument as in Lemma 6. For the second term, we see that from Lemma 6 that

\[
\left( \frac{\max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} 1(A_{i,j} > b_N^{1-\alpha})}{b_N} - 1 \right) \rightarrow 0,
\]

as \( N \to \infty \). We also know that \( \max_{j \leq [(t+\delta)c_N]} B_j / (\epsilon N / b_N) \) converges in distribution as \( N \to \infty \). Now, we can conclude that

\[
\frac{1}{\delta} \mathbb{P} \left( \sup_{s \in [t, t+\delta]} \left( \frac{\max_{i \leq N} \sum_{j=1}^{[sc_N]} A_{i,j} 1(A_{i,j} > b_N^{1-\alpha}) B_j}{\epsilon N} - \frac{\max_{j \leq [sc_N]} B_j}{\epsilon N / b_N} \right) > \frac{\epsilon}{2} \right) \leq \frac{1}{\delta} \mathbb{P} \left( \left( \frac{\max_{i \leq N} \sum_{j=1}^{[tc_N]} A_{i,j} 1(A_{i,j} > b_N^{1-\alpha})}{b_N} - 1 \right) \rightarrow 0 \right) \rightarrow 0.
\]

Now we have established that the probabilities in (15) and (52) converge to 0 as \( N \to \infty \), the result follows.

From the results in Lemmas 7 and 8, we can conclude that the convergence result in (12) holds. Furthermore, by applying the continuous mapping theorem, Theorem 1 follows.
5 Process convergence of the maximum waiting time in $D[0, T]$

At this point, we have proven the convergence result of an auxiliary process whose marginals are the same as the marginals of the maximum waiting time. Now, we can extend these results to prove convergence of the maximum waiting time ($\max_{i \leq N} W_i(t; c_N) / c_N, t \in [0, T]$) to the process $(\sup_{s \in [0, t]} (X(s, t) - \mu(t - s)), t \in [0, T])$ as $N \to \infty$. We first show in Lemma 10 that the maximum waiting time can be approximated by an auxiliary process, as we did in Lemma 9 and then we prove the main result described in Theorem 2.

**Lemma 10.** Given that Assumptions 7–9 hold, and $T > 0$, then we have that for all $\epsilon > 0$

$$\mathbb{P} \left( \sup_{t \in [0, T]} \left| \sup_{s \in [0, t]} \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} - \sup_{s \in [0, t]} \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right| > \epsilon \right) \to 0. \quad (54)$$

**Proof.** As in Lemma 9, we first use that $|X| = \max(X, -X)$. Then, by applying the union bound we get that $\mathbb{P}(|X| > x) \leq \mathbb{P}(X > x) + \mathbb{P}(-X > x)$. Now, we have that

$$\sup_{t \in [0, T]} \left( \sup_{s \in [0, t]} \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) - \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} \right) \leq \sup_{t \in [0, T]} \sup_{s \in [0, t]} \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) - \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N}.$$

Similarly,

$$\sup_{t \in [0, T]} \left( \sup_{s \in [0, t]} \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) - \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} \right) \leq \sup_{t \in [0, T]} \sup_{s \in [0, t]} \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) - \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N}.$$

Therefore,

$$\mathbb{P} \left( \sup_{t \in [0, T]} \left| \sup_{s \in [0, t]} \left( \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} - \sup_{s \in [0, t]} \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right) \right| > \epsilon \right) \leq 2 \mathbb{P} \left( \sup_{t \in [0, T]} \sup_{s \in [0, t]} \left| \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} - \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right| > \epsilon \right). \quad (55)$$

Now, we use the same approach as in Lemma 9 with the somewhat different upper bound:

$$\mathbb{P} \left( \sup_{t \in [0, T]} \sup_{s \in [0, t]} \left( \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} - \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right) > \epsilon \right) \leq \frac{1}{\delta^2} \mathbb{P} \left( \sup_{r \in [t, t + \delta]} \sup_{q \in [s - \delta, s]} \left( \frac{\max_{i \leq N} \sum_{j=\lfloor q c_N \rfloor}^{\lfloor r c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} - \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right) > \epsilon \right). \quad (56)$$

Also,

$$\mathbb{P} \left( \sup_{t \in [0, T]} \sup_{s \in [0, t]} \left( \frac{\max_{i \leq N} \sum_{j=\lfloor s c_N \rfloor}^{\lfloor t c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} + \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right) > \epsilon \right) \leq \frac{1}{\delta^2} \mathbb{P} \left( \sup_{r \in [t, t + \delta]} \sup_{q \in [s - \delta, s]} \left( \frac{\max_{i \leq N} \sum_{j=\lfloor q c_N \rfloor}^{\lfloor r c_N \rfloor} (A_{i,j} B_j - T_j)}{c_N} + \left( \frac{\max_{s \leq j \leq |t c_N|} B_j}{c_N/b_N} - \mu(t - s) \right) \right) > \epsilon \right). \quad (57)$$

The proof that these upper bounds converge to 0 as $N \to \infty$ is analogous to the proof of Lemma 9. \qed
Proof of Theorem 2. We have proven in Lemma 10 that the maximum waiting time can be approximated with the process \( \sup_{s \in [0,t]} \left( \max_{j \geq i} \left\{ \frac{B_j}{C_N} - \mu(t-s) \right\}, t \in [0,T] \right) \). Therefore, in order to prove convergence of the maximum waiting time to the process \( \sup_{s \in [0,t]} (X(s,t) - \mu(t-s)), t \in [0,T] \) in \( D[0,T] \), it suffices to prove convergence of the process \( \sup_{s \in [0,t]} (max_{j \leq i} B_j/(C_N/b_N) - \mu(t-s)), t \in [0,T] \) to the process \( \sup_{s \in [0,t]} (X(s,t) - \mu(t-s)), t \in [0,T] \) in \( D[0,T] \). As in Lemma 3, we again check the conditions given in Lemma 8.

Now, we can further rewrite the event \( \{ \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_2 + \mu t_2 \} \) and relate this to the random variable \( X(s,t) \); namely:

\[
\mathbb{P} \left( \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_2 + \mu t_2 \right) = \mathbb{P} \left( \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_2 + \mu t_2 \right) \cap \left\{ X(t_1,t_2) + \mu t_1 < x_2 + \mu t_2 \right\} \cap \left\{ \sup_{s \in [t_1,t_2]} (X(s,t) + \mu s) < x_2 + \mu t_2 \right\}.
\]

Thus, when \( x_2 + \mu t_2 \leq x_1 + \mu t_1 \), then

\[
\mathbb{P} \left( \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_1 + \mu t_1 \right) \leq \mathbb{P} \left( \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_2 + \mu t_2 \right) = 1,
\]

and when \( x_2 + \mu t_2 > x_1 + \mu t_1 \),

\[
\mathbb{P} \left( \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_1 + \mu t_1 \right) \leq \mathbb{P} \left( \sup_{s \in [0,t]} (X(s,t) + \mu s) < x_2 + \mu t_2 \right) = \frac{\mathbb{P} \left( \sup_{s \in [0,t_1]} (X(s,t_1) + \mu s) < x_1 + \mu t_1 \right)}{\mathbb{P} \left( \sup_{s \in [0,t_1]} (X(s,t_1) + \mu s) < x_2 + \mu t_2 \right)}.
\]

From now on, we focus on the case \( x_2 + \mu t_2 > x_1 + \mu t_1 \); the proof of the case \( x_2 + \mu t_2 \leq x_1 + \mu t_1 \) is analogous. In conclusion,

\[
\mathbb{P} \left( \sup_{s \in [0,t_1]} (X(s,t_1) - \mu(t_1-s)) < x_1 \cap \sup_{s \in [0,t_2]} (X(s,t_2) - \mu(t_2-s)) < x_2 \right) = \frac{\mathbb{P} \left( \sup_{s \in [0,t_1]} (X(s,t_1) - \mu(t_1-s)) < x_1 \right)}{\mathbb{P} \left( \sup_{s \in [0,t_1]} (X(s,t_1) - \mu(t_1-s)) < x_2 + \mu(t_2-t_1) \right)} \cdot \mathbb{P} \left( \sup_{s \in [0,t_2]} (X(s,t_2) - \mu(t_2-s)) < x_2 \right).
\]

Thus, we can write the joint probability in (55) as an operation of marginal probabilities. We can do the same for
that the following bound holds:

\[
\mathbb{P}\left( \sup_{s \in [0,t]} \left( \max_{\lfloor s \leq t \leq \lfloor cN \rfloor} - \mu(t) \right) < x_1 \cup \sup_{s \in [0,t]} \left( \max_{\lfloor s \leq t \leq \lfloor cN \rfloor} - \mu(t) \right) < x_2 \right) < x_1
\]

Using Lemma 8 and the decomposition of a joint probability into marginal probabilities, we establish that the probability in (56) converges to the probability in (54) as \( N \to \infty \). Analogous extensions hold for higher dimensional distributions. Hence, convergence of finite-dimensional distributions follows. To prove process convergence, we show that the second and third condition of Lemma 4 also hold. To establish that the second condition holds, we observe that the following bound holds:

\[
\mathbb{P}\left( \sup_{s \in [0,t]} (X(s,T) - \mu(T) - s) - \sup_{s \in [0,T]} (X(s,T) - \mu(T - s)) > \epsilon \right) \leq \mathbb{P}\left( \sup_{s \in [0,t]} (X(s,T) + \mu s) - \sup_{s \in [0,T]} (X(s,T) + \mu s) + \mu \epsilon > \epsilon \right).
\]

Now, we can further simplify this as follows:

\[
\sup_{s \in [0,T]} (X(s,T) + \mu s) - \sup_{s \in [0,T]} (X(s,T) + \mu s) + \mu \epsilon = \max_{s \in [0,T]} (X(s,T) + \mu s) - \sup_{s \in [0,T]} (X(s,T) + \mu s) + \mu \epsilon \leq \max_{s \in [0,T]} (X(s,T) - X(s,T) - \mu T - \sup_{s \in [0,T]} (X(s,T) + \mu s) + \mu) \leq \max_{s \in [0,T]} (X(s,T) - X(s,T) - \mu T - \mu) + \mu \epsilon = X(T,\mu, T) + 2\mu \epsilon.
\]

We have that

\[
\mathbb{P}(X(T,\mu, T) + 2\mu \epsilon > \epsilon) = 1 - \exp\left( -\frac{\epsilon}{(\epsilon - 2\mu \epsilon)^3} \right) \to 0.
\]

To establish that the third condition holds, we first observe that for \( r \leq t \)

\[
\left| \sup_{s \in [0,t]} \left( \max_{\lfloor s \leq t \leq \lfloor cN \rfloor} B_j / cN/bN - \mu(t) - \mu(s) \right) \right| \leq \sup_{s \in [0,t]} \left( \max_{\lfloor s \leq t \leq \lfloor cN \rfloor} B_j / cN/bN + \mu \right) - \sup_{s \in [0,r]} \left( \max_{\lfloor s \leq t \leq \lfloor cN \rfloor} B_j / cN/bN + \mu \right) + \mu(t - r).
\]

With an analogous derivation as in the proof of Lemma 8, we see that the third condition holds. Thus, we have process convergence.

\[
\square
\]

6 Steady-state convergence of the maximum waiting time

Finally, we prove steady-state convergence of the longest of the \( N \) waiting times. We give lower and upper bounds of \( \mathbb{P}(\max_{i \leq N} W_i(\infty) > x_{cN}) \) and show that these are asymptotically tight.
Proof of Theorem 3. First of all, to prove a sharp lower bound, we first notice that

$$\max_{i \leq N} W_i(\infty) = \max_{i \leq N} \sup_{k \geq 0} \sum_{j=1}^{k} (A_{i,j}B_j - T_j).$$

Thus, the maximum steady-state waiting time is lower bounded by random variables of the form

$$\max_{i \leq N} \sup_{0 \leq k \leq l} \sum_{j=1}^{k} (A_{i,j}B_j - T_j)$$

with \( l > 0 \). Thus, by using the convergence result in (12) in Theorem 1 we know that

$$\lim_{N \to \infty} \inf P\left( \max_{i \leq N} W_i(\infty) > xc_N \right) \geq \lim_{N \to \infty} P\left( \max_{i \leq N} \sup_{t \in [0,M]} \sum_{j=1}^{\lfloor tc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) = P\left( \sup_{t \in [0,M]} (X_t - \mu t) > x \right).$$

Now, following Equations (15) and (16) in Proposition 1, it is easy to see that

$$P\left( \sup_{(t \in [0,M]} (X_t - \mu t) > x \right) \to P\left( \sup_{t > 0} (X_t - \mu t) > x \right),$$

as \( M \to \infty \). Thus, we have a tight lower bound.

Second, we want to find a tight upper bound for the tail probability of the steady-state maximum queue length. We have that

$$P\left( \max_{i \leq N} W_i(\infty) > xc_N \right) = P\left( \max_{i \leq N} \sup_{k \geq 0} \sum_{j=1}^{k} (A_{i,j}B_j - T_j) > xc_N \right) = P\left( \max_{i \leq N} \sup_{t \in [0,M]} \sum_{j=1}^{\lfloor tc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) \leq P\left( \max_{i \leq N} \sup_{t > M} \sum_{j=1}^{\lfloor tc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) + P\left( \max_{i \leq N} \sup_{t > M} \sum_{j=1}^{\lfloor tc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) .$$

For the first term in (59) we obtain that

$$P\left( \max_{i \leq N} \sup_{t \in [0,M]} \sum_{j=1}^{\lfloor tc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) \to 1 - \exp\left( \frac{1}{\mu(\beta - 1)x^{\beta - 1}} \right),$$

as \( M \to \infty \). Thus, we need to prove that the second term in (59) asymptotically vanishes when \( N, M \to \infty \). Let \( \hat{A}_{i,j} \) and \( \hat{B}_j \) be independent copies of \( A_{i,j} \) and \( B_j \) respectively. Then, we can bound the second term in (59) as follows:

$$P\left( \max_{i \leq N} \sup_{t > M} \sum_{j=1}^{\lfloor tc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) = P\left( \max_{i \leq N} \sum_{j=1}^{\lfloor Mtc_N \rfloor} (A_{i,j}B_j - T_j) + \sup_{k \geq 0} \sum_{j=1}^{k} (\hat{A}_{i,j}\hat{B}_j - \hat{T}_j) > xc_N \right) \leq P\left( \max_{i \leq N} \sum_{j=1}^{\lfloor Mtc_N \rfloor} (A_{i,j}B_j - T_j) > xc_N \right) + P\left( \sum_{j=1}^{\lfloor Mtc_N \rfloor} (\hat{A}_{i,j}\hat{B}_j - \hat{T}_j) > xc_N \right) \leq P\left( \sum_{j=1}^{\lfloor Mtc_N \rfloor} (A_{i,j}B_j - T_j) > \frac{\mu}{2}Mtc_N \right) + P\left( \sum_{j=1}^{\lfloor Mtc_N \rfloor} (\hat{A}_{i,j}\hat{B}_j - \hat{T}_j) > \left( x + \frac{\mu}{2}M \right)tc_N \right).$$

20
The bound in (61) holds because \( \max_{i \leq N} (a_i + b_i) \leq \max_{i \leq N} a_i + \max_{i \leq N} b_i \). The bound in (61) follows from the union bound. For the first term in (61) we have that

\[
P \left( \max_{i \leq N} \sum_{j=1}^{M_{CN}} (A_{i,j}B_j - T_j) > -\frac{\mu}{2} M_{CN} \right) \xrightarrow{N \to \infty} 1 - \exp \left( -\frac{M}{\mu M/2} \right) \to 0,
\]
as \( M \to \infty \). In order to analyze the second term in (61), we use the fact that \( \mathbb{E}[A_{i,j}B_j - T_j] = -\mu < 0 \), from this, it follows that there exists a \( \gamma > 1 \), such that \( \mathbb{E}[A_{i,j}B_j] < \mathbb{E}[T_j]/\gamma \). We write \( \gamma \mathbb{E}[A_{i,j}B_j] - \mathbb{E}[T_j] = -\mu \gamma < 0 \). Then,

\[
P \left( \max_{i \leq N} \sup_{k \geq 0} \sum_{j=1}^{k} (A_{i,j}B_j - \hat{T}_j) > \left( x + \frac{\mu}{2} M \right) c_N \right)
\leq P \left( \max_{i \leq N} \sup_{k \in \{0,\ldots,c_N\}} \sum_{j=1}^{k} (\hat{A}_{i,j}\hat{B}_j - \hat{T}_j) > \left( x + \frac{\mu}{2} M \right) c_N \right)
\]

\[+
\sum_{n=0}^{\infty} P \left( \max_{i \leq N} \sup_{k \in \left[\gamma^n c_N, \gamma^{n+1} c_N\right]} \sum_{j=1}^{k} (\hat{A}_{i,j}\hat{B}_j - \hat{T}_j) > \left( x + \frac{\mu}{2} M \right) c_N \right)
\]

\[+
\sum_{n=0}^{\infty} \left( \sum_{i=1}^{\infty} \sum_{j=1}^{\gamma^{n+1} c_N} \hat{A}_{i,j}\hat{B}_j - \sum_{j=1}^{\gamma^n c_N} \hat{T}_j > \left( x + \frac{\mu}{2} M \right) c_N \right)
\]

\[\xrightarrow{N \to \infty} P \left( \sup_{t \in [0,1]} (X_t - \mu t) > x + \frac{\mu}{2} M \right) + \sum_{n=0}^{\infty} \left( 1 - \exp \left( -\frac{\gamma^{n+1}}{(x + \mu M/2 + \gamma^n \mu \gamma)^\beta} \right) \right). \tag{62}
\]

It is clear that \( P(\sup_{t \in [0,1]} (X_t - \mu t) > x + \mu M/2) \to 0 \) as \( M \to \infty \). The sum in (62) is finite and also converges to 0 as \( M \to \infty \), as the ratio test gives us that

\[
\lim_{n \to \infty} \frac{\left( 1 - \exp \left( -\frac{\gamma^{n+1}}{(x + \mu M/2 + \gamma^n \mu \gamma)^\beta} \right) \right)}{\left( 1 - \exp \left( -\frac{\gamma^{n+1}}{(x + \mu M/2 + \gamma^n \mu \gamma)^\beta} \right) \right)} = \frac{1}{\gamma^{\beta-1}} < 1.
\]

Hence, we can choose for all \( \epsilon > 0 \) a \( K \) large enough such that

\[
\sum_{n=K}^{\infty} \left( 1 - \exp \left( -\frac{\gamma^{n+1}}{(x + \mu M/2 + \gamma^n \mu \gamma)^\beta} \right) \right) < \epsilon,
\]

and it is obvious that

\[
\sum_{n=0}^{K} \left( 1 - \exp \left( -\frac{\gamma^{n+1}}{(x + \mu M/2 + \gamma^n \mu \gamma)^\beta} \right) \right) \to 0,
\]
as \( M \to \infty \). Thus, we can conclude that the both terms in (62) converge to 0 as \( M \to \infty \), and consequently, both terms in (61) asymptotically vanish. Returning to the upper bound for the steady-state tail probability of the maximum waiting time given in (59), we can conclude that

\[
\limsup_{N \to \infty} P \left( \max_{i \leq N} W_i(\infty) > xc_N \right) \leq P \left( \sup_{t \geq 0} (X_t - \mu t) > x \right).
\]

We have proven process convergence of the maximum transient waiting time and we have proven steady state convergence. The limiting processes have the form of a supremum of Fréchet-distributed random variables with a negative drift. We now give an explicit expression of the cumulative distribution function.
Proof of Proposition 7. To prove Equation (15), we provide sharp lower and upper bounds of \( P(\sup_{t>0}(X_t - \mu t) < x) \). First of all, let \( \delta > 0 \). We have that
\[
P(X_\delta - \mu \delta < x) = \exp \left( -\frac{\delta}{(x + \mu \delta)^\beta} \right).
\]
Obviously, we can bound \( P(\sup_{t>0}(X_t - \mu t) < x) \) from above as
\[
P\left( \sup_{t>0}(X_t - \mu t) < x \right) < P(\cap_{i=1}^\infty X_{i\delta} - \mu \delta < x).
\]
We can write \( X_{2\delta} = \max(\bar{X}_\delta, X_\delta) \). From this relation, we know that, if \( X_\delta - \delta < x \), then \( X_{2\delta} - 2\delta < x \) and only if \( \bar{X}_\delta - 2\delta < x \). Therefore,
\[
P(X_\delta - \delta < x \cap X_{2\delta} - 2\delta < x) = P(X_\delta - \delta < x \cap \bar{X}_\delta - 2\delta < x) = P(X_\delta - \delta < x) P(\bar{X}_\delta - 2\delta < x).
\]
Thus, in general, the cumulative distribution function of \( \sup_{t>0}(X_t - \mu t) \) is bounded from above as
\[
P\left( \sup_{t>0}(X_t - \mu t) < x \right) < P(\cap_{i=1}^\infty X_{i\delta} - \mu i\delta < x) = \prod_{i=1}^\infty \exp \left( -\frac{\delta}{(x + \mu i\delta)^\beta} \right).
\]
We can find a lower bound as well, because both \( X_t \) and \( \mu t \) are non-decreasing in \( t \) we know that \( \sup_{s \in ((i-1)\delta, i\delta]} (X_s - \mu s) \leq X_{i\delta} - \mu (i-1)\delta \). Therefore,
\[
P\left( \sup_{t>0}(X_t - \mu t) < x \right) = P\left( \cap_{i=1}^\infty \sup_{s \in ((i-1)\delta, i\delta]} (X_s - \mu s) < x \right) > P(\cap_{i=1}^\infty X_{i\delta} - \mu (i-1)\delta < x).
\]
With a similar derivation as before, we have that
\[
P(\cap_{i=1}^\infty X_{i\delta} - \mu (i-1)\delta < x) = \prod_{i=1}^\infty \exp \left( -\frac{\delta}{(x + \mu (i-1)\delta)^\beta} \right).
\]
Now, we can rewrite this expression as
\[
\prod_{i=0}^\infty \exp \left( -\frac{\delta}{(x + \mu i\delta)^\beta} \right) = \exp \left( -\frac{\delta}{(\mu \delta)^\beta} \sum_{i=0}^\infty \frac{1}{(x/\mu + i)^\beta} \right) = \exp \left( -\frac{\delta}{(\mu \delta)^\beta} \zeta \left( \beta, \frac{x}{\mu \delta} \right) \right),
\]
where \( \zeta(\beta, x) \) is the Hurwitz zeta function, cf. [1, Eq. (1.10)]. We have that
\[
\lim_{\delta \to 0} \frac{\delta}{(\mu \delta)^\beta} \zeta \left( \beta, \frac{x}{\mu \delta} \right) = \frac{1}{\mu (\beta - 1) x^{\beta - 1}}.
\]
The same limit holds for the upper bound in [64], thus Equation (15) follows. The proof of Equation (15) is analogous, and follows from the fact that
\[
\lim_{\delta \to 0} \frac{\delta}{(\mu \delta)^\beta} \sum_{i=0}^{[t/\delta]} \frac{1}{(x/(\mu \delta) + i)^\beta} = \frac{1}{\mu^\beta (\beta - 1)} \left( \frac{1}{(x/\mu)^{\beta - 1}} - \frac{1}{(x/\mu + t)^{\beta - 1}} \right).
\]

Acknowledgments

This work is part of the research program Complexity in high-tech manufacturing, (partly) financed by the Dutch Research Council (NWO) through contract 438.16.121.
References

[1] Victor S Adamchik and HM Srivastava. Some series of the zeta and related functions. *Analysis*, 18(2):131–144, 1998.

[2] François Baccelli. Two parallel queues created by arrivals with two demands: The $M/G/2$ symmetrical case. *Technical report RR–0426, INRIA*, 1985.

[3] François Baccelli and Armand M Makowski. Queueing models for systems with synchronization constraints. *Proceedings of the IEEE*, 77(1):138–161, 1989.

[4] Rocco Ballerini and Sidney I Resnick. Records from improving populations. *Journal of Applied probability*, 22(3):487–502, 1985.

[5] Rocco Ballerini and Sidney I Resnick. Records in the presence of a linear trend. *Advances in Applied Probability*, 19(4):801–828, 1987.

[6] George Bennett. Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association*, 57(297):33–45, 1962.

[7] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[8] Nicholas H Bingham, Charles M Goldie, and Jozef L Teugels. *Regular variation*. Number 27. Cambridge University Press, 1989.

[9] Fabien Brosset, Thierry Klein, Agnès Lagnoux, and Pierre Petit. Large deviations at the transition for sums of Weibull-like random variables. In *Séminaire de Probabilités LI*, pages 239–257. Springer, 2022.

[10] Meyer Dwass. Extremal processes. *The Annals of Mathematical Statistics*, 35(4):1718–1725, 1964.

[11] Meyer Dwass. Extremal processes, II. *Illinois Journal of Mathematics*, 10(3):381–391, 1966.

[12] Leopold Flatto and Sam Hahn. Two parallel queues created by arrivals with two demands I. *SIAM Journal on Applied Mathematics*, 44(5):1041–1053, 1984.

[13] Claude Godreche, Satya N Majumdar, and Gregory Schehr. Record statistics of a strongly correlated time series: random walks and lévy flights. *Journal of Physics A: Mathematical and Theoretical*, 50(33):333001, 2017.

[14] Laurens de Haan and Ana Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2006.

[15] Mor Harchol-Balter. Open problems in queueing theory inspired by datacenter computing. *Queueing Systems*, 97(1):3–37, 2021.

[16] Harry Kesten. Convergence in distribution of lightly trimmed and untrimmed sums are equivalent. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 113, pages 615–638. Cambridge University Press, 1993.

[17] Stephanus J de Klein. *Fredholm integral equations in queueing analysis*. PhD thesis, Rijksuniversiteit Utrecht, 1988.

[18] Sung-Seok Ko and Richard F. Serfozo. Response times in $M/M/s$ fork-join networks. *Advances in Applied Probability*, 36(3):854–871, 2004.

[19] Hongyuan Lu and Guodong Pang. Gaussian limits for a fork-join network with nonexchangeable synchronization in heavy traffic. *Mathematics of Operations Research*, 41(2):560–595, 2015.

[20] Hongyuan Lu and Guodong Pang. Heavy-traffic limits for a fork-join network in the Halfin-Whitt regime. *Stochastic Systems*, 6(2):519–600, 2017.

[21] Hongyuan Lu and Guodong Pang. Heavy-traffic limits for an infinite-server fork-join queueing system with dependent and disruptive services. *Queueing Systems*, 85(1-2):67–115, 2017.
[22] Mirjam Meijer, Dennis Schol, Willem van Jaarsveld, Maria Vlasiou, and Bert Zwart. Extreme-value theory for large fork-join queues, with applications to high-tech supply chains. https://arxiv.org/abs/2105.09189, 2021.

[23] Jayakrishnan Nair, Adam Wierman, and Bert Zwart. The fundamentals of heavy tails: Properties, emergence, and estimation, volume 53. Cambridge University Press, 2022.

[24] Randolph Nelson and Asser N Tantawi. Approximate analysis of fork/join synchronization in parallel queues. IEEE Transactions on Computers, 37(6):739–743, 1988.

[25] Viên Nguyen. Processing networks with parallel and sequential tasks: Heavy traffic analysis and Brownian limits. The Annals of Applied Probability, pages 28–55, 1993.

[26] Viên Nguyen. The trouble with diversity: Fork-join networks with heterogeneous customer population. The Annals of Applied Probability, pages 1–25, 1994.

[27] Youri Raaijmakers, Sem Borst, and Onno Boxma. Fork-join and redundancy systems with heavy-tailed job sizes. arXiv preprint arXiv:2105.13738, 2021.

[28] Sidney I Resnick. Heavy-tail phenomena: probabilistic and statistical modeling. Springer Science & Business Media, 2007.

[29] Sidney I Resnick and Michael Rubinovitch. The structure of extremal processes. Advances in Applied Probability, 5(2):287–307, 1973.

[30] Dennis Schol, Maria Vlasiou, and Bert Zwart. Large fork-join queues with nearly deterministic arrival and service times. Mathematics of Operations Research, 47(2):1335–1364, 2021.

[31] Subir Varma. Heavy and light traffic approximations for queues with synchronization constraints. PhD thesis, University of Maryland, 1990.

[32] Paul E Wright. Two parallel processors with coupled inputs. Advances in Applied Probability, 24(4):986–1007, 1992.

[33] Cathy H Xia, Zhen Liu, Don Towsley, and Marc Lelarge. Scalability of fork/join queueing networks with blocking. ACM SIGMETRICS Performance Evaluation Review, 35(1):133–144, 2007.

[34] Yun Zeng, Jian Tan, and Cathy H Xia. Fork and join queueing networks with heavy tails: Scaling dimension and throughput limit. Journal of the ACM (JACM), 68(3):1–30, 2021.

[35] Songfeng Zheng. An improved Bennett’s inequality. Communications in Statistics-Theory and Methods, 47(17):4152–4159, 2018.