Large Deviations and Queue Length Asymptotics for GI/GI/d Queues and Stochastic Fluid Networks with Log-Normal Tails

Zhe Su and Chang-Han Rhee

Industrial Engineering and Management Sciences, Northwestern University
Evanston, IL, 60613, USA

February 28, 2025

Abstract

We investigate rare events in two popular queueing models under the assumption of log-normal service requirements. The first model considers a stochastic fluid system with external input processes modeled as compound Poisson processes with log-normal jumps. For this system, we derive large deviation estimates for the buffer content process. The second model examines the GI/GI/d multi-server queue with log-normal-type service times, and we focus on the rare event of large queue lengths. In both cases, our analysis reveals that rare events are driven by a finite number of extreme events, confirming the *catastrophe* principle in the log-normal uncertainty regime. The proofs build on recent advances in sample-path large deviation theory for Lévy processes and random walks with log-normal-type increments.

1 Introduction

The multi-server queue, commonly known as the GI/GI/d queue, and the stochastic fluid network are two fundamental models in queueing theory, each with extensive real-world applications. These models have been widely studied over the past several decades, yet they continue to present significant analytical challenges. Early work on multi-server queues [23, 19] focused on exact analyses of the invariant waiting-time distribution, while research on queueing networks has largely centered on generalized Jackson networks [2], monotone separable networks [1], and max-plus networks [3]. Compared to discrete-customer queueing models, fluid networks capture the flow of "continuous customers," making them particularly well-suited for high-speed telecommunications and continuous production processes. Despite notable progress in both areas, finding tractable solutions remains difficult, leading to ongoing research in approximation methods and asymptotic analysis. For instance, heavy-traffic approximations for multi-server queues have been explored in [17, 22], while rare-event probabilities, such as long waiting times or large queue lengths under light-tailed service assumptions, have been analyzed in [25, 24]. In the context of fluid networks, [8] demonstrates that tail asymptotics for downstream nodes can be characterized through the busy periods of upstream nodes. A broader survey of results on fluid networks with light-tailed distributions can be found in [21].

Recent research has made significant strides in understanding queueing asymptotics under heavy-tailed assumptions. The study of multi-server queues with heavy-tailed service times traces back to [31], which proposed a conjecture on the tail behavior of the steady-state waiting time distribution under a sub-exponential service-time assumption. This foundational work has since led to further developments, including necessary and sufficient conditions for the existence of finite moments in the waiting-time distribution [26] and analyses of its tail asymptotics [14, 13].

In the context of fluid networks, tandem fluid queues—where the input to the first node is modeled as a Lévy process with regularly varying jump sizes—have been examined in [20]. A heavy-tailed large deviation framework introduced in [9] extends the analysis to multidimensional fluid queues, characterizing the asymptotic behavior of the vector of time-dependent buffer content. Another class of heavy-tailed distributions, namely Weibull-type distributions, has been considered in [5, 4], where the focus is on the asymptotics of extreme queue lengths in multi-server systems and buffer overflows in fluid networks.

Among the most commonly used models for heavy-tailed distributions are regularly varying, Weibull, and log-normal distributions. While queueing systems under the first two assumptions have been studied extensively, the log-normal case remains largely unexplored. This paper addresses this gap by analyzing queueing systems with log-normal-type uncertainties, revisiting the rare-event analyses conducted in [5, 4].

Specifically, in Section 3, we examine a stochastic fluid system with d nodes, where some nodes receive external fluid. The input processes are modeled as compound Poisson processes with log-normal-type jumps. Letting $\bar{\boldsymbol{Z}}_n(T)$ denote the time-dependent buffer content process and $\boldsymbol{h}$ represent the cost associated with buffer content, our main result, Proposition 3.2, establishes a large deviation principle for the rare event $\{\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) > y\}$, given by

$$-V(y) \leq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) \geq y\right)}{r_0(\log n)} \leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) \geq y\right)}{r_0(\log n)} \leq -\lim_{\epsilon\downarrow 0} V(y-\epsilon).$$

where $r_0(\cdot)$ is a regularly varying function that determines the decay rate of the rare-event probability. The function $V(\cdot)$ is derived by solving a family of linear systems, each corresponding to a transient state in which the system receives large fluid volumes at certain nodes. Intuitively, $V(y)$ represents the minimal number of large-volume arrivals required to incur a buffer cost of at least y. A detailed characterization of $V(\cdot)$ is provided in Proposition 3.3.

In Section 4, we analyze a GI/GI/d queue with d servers, where service times follow a log-normal-type distribution. Let $Q(t)$ denote the queue length at time t, and let p be a positive constant. Propositions 4.6 and 4.7 establish a large deviation principle for the rare event of a large queue length, given by

$$-W(p) \leq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(Q(n) \geq np\right)}{r_0(\log n)} \leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(Q(n) \geq np\right)}{r_0(\log n)} \leq -\lim_{x<p, x\uparrow p} W(x).$$

The function $W(y)$ is

$$W(y) = \begin{cases} \sum_{i=1}^{\lfloor y+d-\frac{1}{\mathbb{E}[A]}\rfloor+1} \lambda_{(i)} & y < \frac{1}{\mathbb{E}[A]} \\ \infty & \text{otherwise} \end{cases},$$

which, intuitively, quantifies the minimum number of large job size required for the queue to reach the level np, when $n \to \infty$.

In both studies, the functions $V(\cdot)$ and $W(\cdot)$ are piecewise increasing with discrete jumps, confirming that these rare events follow the catastrophe principle—they occur due to extreme individual events. This finding is closely aligned with the results in [9], where the uncertainty is modeled using regularly varying distributions. However, it contrasts with the Weibull-type case studied in [5, 4], where a non-trivial trade-off emerges between the number of large jobs and their sizes.

At the core of our proof is a novel result from [28], which establishes the *extended large deviation principle* for stochastic processes with log-normal-type jumps under the J_1 topology (see Theorems 2.1 and 2.2 for details). The entire proof is conducted within the J_1 topology and, when necessary, its product topology in d-dimensional space. This approach differs from [5], where the M_1' topology is used to facilitate the proof. The fundamental reason for this distinction is that the extended LDP established in [28] has a discrete rate function, which allows for the extension from a one-dimensional LDP to a multi-dimensional setting. Proposition 2.1 plays a key role in enabling this extension.

The organization of the paper is as follows: Section 2 introduces general notation and essential large deviation results. Sections 3 and 4 present the main arguments for the rare-event analyses in stochastic fluid networks and multi-server queues, respectively. The technical proofs supporting these results are collected in Section 5. Finally, we conclude with an appendix, where we develop several auxiliary large deviation results and other miscellaneous findings.

2 Notations and Preliminary Results

We start with recurring notations used in the paper. Let $(\mathcal{X}, d)$ denote a metric space $\mathcal{X}$ equipped with a metric d. For a set A, let A°, $\bar{A}$, A^c denote the interior, closure and complement of A. Also, let $d(x, A) \triangleq \inf_{y\in A} d(x, y)$, $B_r(x) \triangleq \{y \in \mathcal{X} : d(x, y) < r\}$, $A^\epsilon \triangleq \{x \in \mathcal{X} : d(x, A) \le \epsilon\}$, and $A^{-\epsilon} \triangleq \big((A^c)^\epsilon\big)^c$ denote the distance between x and A, the open ball with radius r centered at x, the closed ϵ-fattening of A, and the open ϵ-shrinking of A, respectively. In the paper, Bold letters are used to represent multi-valued elements, with coordinate indices indicated by subscripts for vectors in $\mathbb{R}^d$, (e.g., $\boldsymbol{x} = (x_1, \cdots, x_d) \in \mathbb{R}^d$) and by superscripts for multi-valued processes (e.g, $\boldsymbol{X} = (X^{(1)}, \ldots, X^{(d)})$). $[d]$ is a simplified notation for the index set $\{1, \cdots, d\}$. ζ_μ is the linear function with a rate of μ.

Many intermediate results are formulated using two large deviation principles, whose definitions are provided below.

Definition 2.1. Let $\{X_n\}_{n\ge1}$ be a sequence of random objects taking value in the metric space $(\mathcal{X}, d)$. Suppose there exists a lower semicontinuous function $I : \mathcal{X} \to \mathbb{R}_+$ and a sequence $\{a_n\}_{n\ge1}$ such that, for any Borel set A, the following inequalities hold:

$$-\inf_{x\in A^\circ} I(x) \le \liminf_{n\to\infty} \frac{\log \mathbb{P}(X_n \in A)}{a_n} \le \limsup_{n\to\infty} \frac{\log \mathbb{P}(X_n \in A)}{a_n} \le -\inf_{x\in\bar{A}} I(x) \tag{2.1}$$

Then I is called a *rate function*, and the sequence $\{X_n\}_{n\ge1}$ is said to satisfy the *full large deviation principle*(abbreviated as *full LDP*) on $(\mathcal{X}, d)$ with the rate function I and the speed sequence $\{a_n\}_{n\ge1}$.

Definition 2.2. Let $\{X_n\}_{n\ge1}$ be a sequence of random objects taking value in the metric space $(\mathcal{X}, d)$. Suppose there exists a lower semicontinuous function $I : \mathcal{X} \to \mathbb{R}_+$ and a sequence $\{a_n\}_{n\ge1}$ such that, for any Borel set A, the following inequalities hold:

$$-\inf_{x\in A^\circ} I(x) \le \liminf_{n\to\infty} \frac{\log \mathbb{P}(X_n \in A)}{a_n} \le \limsup_{n\to\infty} \frac{\log \mathbb{P}(X_n \in A)}{a_n} \le -\lim_{\epsilon\downarrow0}\inf_{x\in A^\epsilon} I(x) \tag{2.2}$$

Then I is called a *rate function*, and the sequence $\{X_n\}_{n\ge1}$ is said to satisfy the *extended large deviation principle*(abbreviated as *extended LDP*) on $(\mathcal{X}, d)$ with the rate function I and the speed sequence $\{a_n\}_{n\ge1}$.

The full LDP in Definition 2.1, although modified to suit our purposes, can be found in standard large deviations textbooks, such as [10, 11, 29, 27, 12, 16], which offer comprehensive discussions on large deviation principles and their applications. The extended LDP in Definition 2.2 was introduced in [7].

The leftmost and rightmost inequalities in (2.1) and (2.2) are referred to as the lower and upper bounds of the full LDP and extended LDP, respectively. The key difference between Definitions 2.1 and 2.2 lies in their upper bounds, where the following inequality holds:

$$\lim_{\epsilon\downarrow 0}\inf_{x\in A^\epsilon} I(x) \leq \inf_{x\in\bar{A}} I(x),$$

This distinction makes the full LDP a stronger principle than the extended LDP. In cases where the rate function I satisfies the additional requirement that the level sets $\{x \in \mathcal{X} : I(x) \leq \alpha\}$ are compact for all $\alpha \in \mathbb{R}_+$, it can be shown that the two bounds are equal, causing the two principles to coincide. However, some rate functions considered in this paper do not satisfy this compactness condition, necessitating a distinction between the full LDP and the extended LDP. In some scenarios, I may not be known to be lower semicontinuous, but the sequence $\{X_n\}_{n\geq 1}$ satisfies the inequalities in (2.1) and (2.2). In such cases, we say that the function I along with the speed sequence *control* the lower and upper bounds of the full/extended LDP for $\{X_n\}_{n\geq 1}$.

Several results in the domain of large deviations, which are instrumental in building the results presented in this paper, are collected in the Appendix .

At the core of our proof are recent advancements from [28], which establishes sample-path-level large deviations for Lévy processes and random walks with log-normal-type increments. These results are framed in the metric space $(\mathbb{D}[0,T], J_1)$, where $\mathbb{D}[0,T]$ denotes the Skorokhod space of real-valued càdlàg functions on $[0,T]$. The J_1 metric on $\mathbb{D}[0,T]$ is defined as

$$d_{J_1}(\xi,\zeta) = \inf_{\lambda\in\Lambda} \|\lambda - e\| \vee \|\xi\circ\lambda - \zeta\|,$$

where Λ is the collection of all non-decreasing homeomorphisms on $[0,T]$, e is the identity map on $[0,T]$, and $\|\cdot\|$ is the supremum norm of a function. Throughout the paper, we use the simplified notation $\mathbb{D}$ for $\mathbb{D}[0,1]$. These results assume uncertainties with log-normal-type tails, defined as follows:

Definition 2.3. A measure or a distribution μ supported on $\mathbb{R}_+$ is said to have the log-normal-type tail if

$$\mu[x,\infty) = \exp(-r(\log x)) \tag{2.3}$$

where $r(\cdot)$ is a regularly varying function with index $\gamma > 1$. Additionally, there exists some $\gamma' \in (0,\gamma)$ such that

$$\liminf_{x\to\infty} \frac{r(x+c) - r(x)}{\exp\{-x^{\gamma'}\}} \geq 1 \tag{2.4}$$

for each $c > 0$.

In Section 3, the external input processes of the stochastic fluid network are modeled as compound Poisson processes with log-normal-type increments. As a specific case of a Lévy process, the large deviation properties of the compound Poisson process, presented below, can be directly inferred from Theorem 3.3 in [28].

Theorem 2.1. Let $X(t) = \sum_{i=1}^{N(t)} Z_i$ be a 1-dimensional compound Poisson process where the increments Z_i's are independent and identically distributed, following the distribution defined in

Definition 2.3. $N(t), t \geq 0$ is the counting variable of the Poisson process with rate 1. Consider the centered and scaled sequence

$$\bar{X}_n(t) = \frac{1}{n}\big[X(nt) - t\mathbb{E}\left[Z_1\right]\big]$$

for $t \in [0, T]$ and $n \geq 1$. Then the sequence $\{\bar{X}_n\}_{n\geq 1}$ satisfies the extended LDP on $(\mathbb{D}[0,T], J_1)$ with the speed $\{r(\log n)\}_{n\geq 1}$ and rate function $I : \mathbb{D}[0,T] \to \mathbb{R}_+$ given by

$$I(\xi) = \begin{cases} \sum_{t\in(0,T]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \text{if } \xi \in \mathbb{D}_{<\infty}[0,T] \\ \infty & \text{otherwise} \end{cases} \tag{2.5}$$

In the above, $\mathbb{D}_{<\infty}[0,T]$ is the subspace of $\mathbb{D}[0,T]$, consisting of non-decreasing step functions with finite jumps, vanishing at the origin and continuous at T.

The following result is Theorem 4.1 from [28], which establishes the large deviation properties for random walks with log-normal-type increments.

Theorem 2.2. [Theorem 4.1 in [28]] Let $S(t) = \sum_{i=1}^{\lfloor t\rfloor} Z_i, t \geq 0$ be a 1-dimensional random walk where the increments Z_i's are independent and identically distributed, following the distribution defined in Definition 2.3. Consider the centered and scaled sequence

$$\bar{S}_n(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor}(Z_i - \mathbb{E}\left[Z_1\right])$$

for $t \in [0, 1]$ and $n \geq 1$. Then the sequence $\{\bar{S}_n\}_{n\geq 1}$ satisfies the extended large deviation on $(\mathbb{D}, J_1)$ with speed $\{r(\log n)\}_{n\geq 1}$ and the following rate function:

$$\hat{I}(\xi) = \begin{cases} \sum_{t\in[0,T]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \text{if } \xi \in \tilde{\mathbb{D}}_{<\infty} \\ \infty & \text{otherwise} \end{cases} \tag{2.6}$$

In the above, $\tilde{\mathbb{D}}_{<\infty}$ is the subspace of $\mathbb{D}$,consisting of non-decreasing step functions with finite jumps, vanishing at the origin.

The following general result, stated as Proposition 3.8 in [28], is applicable when the extend LDP for a multi-valued random sequence is desired.

Proposition 2.1. Suppose that $\boldsymbol{X}_n = (X_n^{(1)}, \cdots, X_n^{(k)})$ satisfies the following conditions:

(i) $X_n^{(i)}$ is a random object taking value in the metric space $(\mathcal{X}^{(i)}, d^{(i)})$.

(ii) There is independency between $\{X_n^{(i)}\}_{n\geq 1}$ and $\{X_n^{(j)}\}_{n\geq 1}$ for any $i \neq j$.

(iii) $\{X_n^{(i)}\}_{n\geq 1}$ satisfies the extended LDP with the rate function $I^{(i)}$ and the speed $\{a_n\}_{n\geq 1}$

(iv) $I^{(i)}$ takes at most countable distinct values in $\mathbb{R}_+$ with no limit point.

Then $\{\boldsymbol{X}_n\}_{n\geq 1}$ satisfies the extended LDP on the product space $\big(\prod_{i=1}^k \mathcal{X}^{(i)}, \sum_{i=1}^k d^{(i)}\big)$ with the rate function $\sum_{i=1}^k I^{(i)}$ and the speed $\{a_n\}_{n\geq 1}$.

In the context of this paper, the above proposition is instrumental in deriving the extended LDP for multivalued processes. Specifically, we consider the path space as the metric space $\left(\prod_{i=1}^{d}\mathbb{D}[0,T], d_p\right)$. The metric d_p is given by:

$$d_p(\boldsymbol{\xi},\boldsymbol{\eta}) = \sum_{i\in[d]} d_{J_1}(\xi_i,\eta_i). \tag{2.7}$$

Before concluding this section, we provide the following list of subsets of $\mathbb{D}[0,T]$ that appear throughout the paper, serving as a reference to facilitate reading.

- $\mathbb{D}^{\uparrow}[0,T] \triangleq \{\xi\in\mathbb{D}[0,T] : \xi \text{ is non-decreasing}, \xi(0)=0\}$
- $\mathbb{D}^{\beta}[0,T] \triangleq \{\zeta\in\mathbb{D}[0,T] : \zeta(t)=\xi(t)+\beta\cdot t, \xi\in\mathbb{D}^{\uparrow}[0,T]\}$ for some $\beta\in\mathbb{R}$
- $\mathbb{D}_{=j}[0,T] \triangleq \{\xi\in\mathbb{D}^{\uparrow}[0,T], \xi$ is a pure jump function with j jumps on $(0,T)$ and ξ is continuous at $T\}$
- $\mathbb{D}_{\leqslant j}[0,T] \triangleq \cup_{i=0}^{j}\mathbb{D}_{=i}[0,T]$ and $\mathbb{D}_{<\infty}[0,T] \triangleq \cup_{j=0}^{\infty}\mathbb{D}_{\leqslant j}[0,T]$
- $\mathbb{D}^{\beta}_{<\infty}[0,T] \triangleq \{\zeta\in\mathbb{D}[0,T] : \zeta(t)=\xi(t)+\beta\cdot t, \xi\in\mathbb{D}_{<\infty}[0,T]\}$ for some $\beta\in\mathbb{R}$

Note that the sets $\mathbb{D}_{=j}[0,T]$, $\mathbb{D}_{\leqslant j}[0,T]$, $\mathbb{D}_{<\infty}[0,T]$ and $\mathbb{D}^{\beta}_{<\infty}[0,T]$ require the paths to be continuous at the end time T. When relaxing this requirement, we define the corresponding sets as $\tilde{\mathbb{D}}_{=j}[0,T]$, $\tilde{\mathbb{D}}_{\leqslant j}[0,T]$, $\tilde{\mathbb{D}}_{<\infty}[0,T]$ and $\tilde{\mathbb{D}}^{\beta}_{<\infty}[0,T]$.

3 Stochastic Fluid Network and the Overflow Asymptotics

Throughout the section, we consider a mathematical model for continuous-flow between a network of nodes.

Definition 3.1. We define the fluid network(FN) as a network of d fully connected nodes with buffers. The fluid can flow out from each buffer at the maximum rate $\boldsymbol{r}=(r_1,\cdots,r_d)$, and the outflow is governed by an invertible matrix $Q=[q_{ij}]_{i,j=1,\cdots,d}$. Specifically, for fluid flowing out of node i, a fraction $q_{ij}\in[0,1]$ is redirected to the node j for all $j\in[d]$, while the remaining fraction,$1-\sum_{j=1}^{d}q_{ij}$, leaves the network. To prevent self-reinforcing cycles, we set $q_{ii}=0$.

A critical setting of the fluid network is how the exogenous fluid joins the network. Our study focuses on stochastic exogenous inflow to all nodes that can be modeled by a multidimensional compound Poisson process $\boldsymbol{H}=(H^{(1)},\ldots,H^{(d)})$. Under this set-up, we call the the fluid network a stochastic fluid network and denote it by FN($\boldsymbol{H}$). We impose the following assumption for $\boldsymbol{H}$ throughout the section.

Assumption 1. A subset of these nodes, $\mathcal{J}\subset\{1,\ldots,d\}$, receives random fluid from external sources over time.

- For $i\notin\mathcal{J}$, $H^{(i)}$ is the zero function. For $i\in\mathcal{J}$, $H^{(i)}$ is a compound Poisson process of 1 dimension whose increments are independent and have the same distribution as specified in Definition 2.3 with a regularly varying function r_i. The inter-arrival time between increments follows the exp(1) distribution.
- For $i\in\mathcal{J}$, the regularly varying functions r_i's associated with $H^{(i)}$'s share the same index $\beta>1$, and there exists a function r_0 such that $\lim_{x\to\infty}\frac{r_i(x)}{r_0(x)}=\lambda_i$. We make the convention that $\lambda_i=0$ if $i\notin\mathcal{J}$.

- The vector $\boldsymbol{\gamma}$ is the expectation of increments for all nodes, with the convention that $\gamma_i = 0$ if $i \notin \mathcal{J}$.

- We assume $\boldsymbol{r} > \mathcal{Q}^{-1}\boldsymbol{\gamma}$ where $\mathcal{Q}= \mathrm{I} - Q^{\intercal}$. This condition guarantees the stability of the fluid network (see [18]).

Our goal is to analyze the large deviation of buffer overflows in DF($\boldsymbol{H}$). More specifically, let the vector-valued process $\boldsymbol{Z}(t), t \geq 0$ stands for the buffer content at time t for all nodes, and consider the vector $\boldsymbol{h}$ with non-negative entries, the goal is to characterize the asymptotics of the probability

$$\mathbb{P}\left(\boldsymbol{h}^{\intercal} \cdot \boldsymbol{Z}(nT) \geq ny\right), \tag{3.1}$$

or

$$\mathbb{P}\left(\boldsymbol{h}^{\intercal} \cdot \bar{\boldsymbol{Z}}_n(T) \geq y\right) \tag{3.2}$$

if we define $\bar{\boldsymbol{Z}}_n(t) \triangleq \frac{1}{n}Z(nt)$ for $t \in [0,T]$, as $n \to \infty$ with given $T, y > 0$.

If the fluid in all buffers continues to flow out at the maximum rate $\boldsymbol{r}$ regardless of the buffer content, the resulting buffer content, which we call the potential buffer content and denote it as $\boldsymbol{B}(t), t > 0$, should satisfies:

$$\boldsymbol{B}(t) = \boldsymbol{H}(t) - \mathcal{Q}\boldsymbol{r} \cdot t \tag{3.3}$$

for $t \geq 0$. The actual buffer content $\boldsymbol{Z}(t)$ and the potential buffer content $\boldsymbol{B}(t)$ differ due to the imaginary fluid flowing out while the buffers are empty over time. If $\boldsymbol{Y}(t), t \geq 0$ represents the imaginary fluids accumulated over time, the relation $\boldsymbol{Z}(t) = \boldsymbol{B}(t) + \mathcal{Q}\boldsymbol{Y}(t)$ holds for $t \geq 0$. Let's define the following scaled process

$$\bar{\boldsymbol{B}}_n(t) \triangleq \frac{1}{n}\boldsymbol{B}(nt) \text{ and } \bar{\boldsymbol{Y}}_n(t) \triangleq \frac{1}{n}\boldsymbol{Y}(nt)$$

for $n \geq 1$ and $t \in [0,T]$. The following proposition 3.1 outlines the large deviations of the sequence $\{\bar{\boldsymbol{B}}_n\}_{n\geq 1}$.

Proposition 3.1. On $\left(\prod_{i=1}^{d} \mathbb{D}^{-(\mathcal{Q}\boldsymbol{r})_i}[0,T], d_p\right)$, the sequence $\{\bar{\boldsymbol{B}}_n\}_{n\geq 1}$ satisfies the extended LDP with the speed sequence $\{r_0(\log n)\}_{n\geq 1}$ and the rate function

$$I_{\boldsymbol{B}}(\boldsymbol{\xi}) \triangleq \begin{cases} \sum_{i=1}^{d} \lambda_i \cdot \left(\sum_{t\in(0,1]} \mathbb{1}\{\xi_i(t) \neq \xi_i(t-)\}\right) & \begin{array}{l}\xi_i \in \mathbb{D}_{<\infty}^{(\boldsymbol{\gamma}-\mathcal{Q}\boldsymbol{r})_i}[0,T] \text{ if } i \in \mathcal{J} \\ \xi_i(t) \equiv (-\mathcal{Q}\boldsymbol{r})_j \cdot t \text{ for } t \in [0,T] \text{ if } i \notin \mathcal{J}\end{array} \\ \\ \infty & \text{otherwise} \end{cases}. \tag{3.4}$$

Proof. We begin by deriving the large deviations for the sequence $\{\bar{\boldsymbol{H}}_n\}_{n\geq 1}$, where

$$\bar{\boldsymbol{H}}_n(t) = \frac{1}{n}[\boldsymbol{H}(t) - \boldsymbol{\gamma}t]$$

for $t \in [0,T]$ and $n \geq 1$. For a node $i \in \mathcal{J}$-that is, node i receives external inflow-$H^{(i)}$ is a special case of Lévy process with log-normal type Lévy measure (see Assumption 1). By Theorem 2.1, the sequence $\{\bar{H}_n^{(i)}\}_{n\geq 1}$ satisfies the extended LDP on $(\mathbb{D}[0,T], J_1)$ with the speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function

$$I(\xi, \lambda_i) = \begin{cases} \lambda_i \cdot \left(\sum_{t\in(0,T]} \mathbb{1}\{\xi(t) \neq \xi(t-)\}\right) & \xi \in \mathbb{D}_{<\infty}[0,T] \\ \infty & \text{otherwise} \end{cases}.$$

For a node i without external inflow, $H^{(i)}$ is the zero function. It is straightforward to verify that $\{\bar{H}_n^{(i)}\}_{n\geq 1}$ satisfies the extended LDP with the speed $\{r_0(\log n)\}_{n\geq 1}$ and a rate function that takes the value 0 at the zero path and ∞ elsewhere. Using the extended LDPs for the coordinates of $\bar{\boldsymbol{H}}_n$ and their independence, Theorem 2.1 implies that $\{\bar{\boldsymbol{H}}_n\}_{n\geq 1}$ satisfies the extended LDP with the speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function

$$I_{\boldsymbol{H}}(\xi) \triangleq \begin{cases} \sum_{i=1}^d \lambda_i \cdot \left(\sum_{t\in(0,T]} \mathbb{1}\{\xi_i(t) \neq \xi_i(t-)\}\right) & \begin{array}{l}\xi_i \in \mathbb{D}_{<\infty}[0,T] \text{ if } i \in \mathcal{J} \\ \xi_i \equiv 0 \text{ if } i \notin \mathcal{J}\end{array} \\ \\ \infty & \text{otherwise} \end{cases}. \tag{3.5}$$

Since $\bar{\boldsymbol{B}}_n = \Upsilon^{\gamma-\mathcal{Q}\boldsymbol{r}}(\bar{\boldsymbol{H}}_n)$ and $\Upsilon^{\gamma-\mathcal{Q}\boldsymbol{r}}$ (see Lemma 6.5) is a homeomorphism, the extended LDP of $\{\bar{\boldsymbol{B}}_n\}_{n\geq 1}$ follows directly from that of $\{\bar{\boldsymbol{H}}_n\}_{n\geq 1}$ by the contraction principle(see Lemma 6.3). More specifically, $\{\bar{\boldsymbol{B}}_n\}_{n\geq 1}$ satisfies the extended LDP on $(\prod_{i=1}^d \mathbb{D}[0,T], d_p)$ with speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function $I_{\boldsymbol{B}} \triangleq I_{\boldsymbol{H}} \circ (\Upsilon^{\gamma-\mathcal{Q}\boldsymbol{r}})^{-1}$. It's straightforward to verify that $I_{\boldsymbol{B}}$ takes the form in (3.4).

The effective domain $\mathcal{D}_{I_{\boldsymbol{B}}} = \{\xi, I_{\boldsymbol{B}}(\xi) < \infty\}$ is a subset of $\prod_{i=1}^d \mathbb{D}^{-(\mathcal{Q}\boldsymbol{r})_i}[0,T]$, which is closed as it is the image of the closed set $\mathbb{D}^{\uparrow}$ under the homeomorphism $\Upsilon^{-(\mathcal{Q}\boldsymbol{r})_i}$. Consequently, by Lemma 6.1 part (1), the extended LDP of $\{\bar{\boldsymbol{B}}_n\}_{n\geq 1}$ also holds on the smaller domain $\left(\prod_{i=1}^d \mathbb{D}^{-(\mathcal{Q}\boldsymbol{r})_i}[0,T], d_p\right)$. This completes the proof. □

Let $\pi : \prod_{i=1}^d \mathbb{D}[0,T] \to \mathbb{R}^d$ be a mapping such that $\pi(\boldsymbol{\xi}) = \boldsymbol{\xi}(T)$. The next proposition states the large deviations for the sequence $\{h^\intercal \cdot \bar{\boldsymbol{Z}}_n(T)\}_{n\geq 1}$.

Proposition 3.2. Let $\boldsymbol{R} = (\psi, \phi)$ be the reflection map in Definition 6.1(in Appendix 6.2). The function

$$I_{\boldsymbol{h}^\intercal\cdot\bar{\boldsymbol{Z}}}(x) = \inf\{I_{\boldsymbol{B}}(\boldsymbol{\xi}) : x = \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}), \boldsymbol{\xi} \in \prod_{i=1}^d \mathbb{D}^{-(\mathcal{Q}\gamma)_i}[0,T]\} \tag{3.6}$$

and the speed sequence $\{r_0(\log n)\}_{n\geq 1}$, control the extended LDP for the sequence $\{\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T)\}_{n\geq 1}$.

Proof. Based on the definitions of $\bar{\boldsymbol{B}}_n, \bar{\boldsymbol{Y}}_n$ and $\bar{\boldsymbol{Z}}_n$, it is true that

- All coordinates of $\bar{\boldsymbol{Y}}_n(t)$ are nonnegative and nondecreasing with respect to t.
- $\bar{\boldsymbol{Z}}_n(t) = \bar{\boldsymbol{B}}_n(t) + \mathcal{Q}\bar{\boldsymbol{Y}}_n(t)$, and all the coordinates of $\bar{\boldsymbol{Z}}_n(t)$ are nonnegative.
- For each coordinate i, $\int_0^T \bar{Z}_n^{(i)}(t) d\bar{Y}_n^{(i)}(t) = 0$. This means that $\bar{Y}_n^{(i)}(t)$ increases only when the buffer content $\bar{Z}_n^{(i)}(t)$ is zero.

The above three bullet points, along with the complementarity characterization of the reflection mapping(see Theorem 6.1 in Appendix 6.2) confirm

$$\boldsymbol{R}(\bar{\boldsymbol{B}}_n) = (\psi(\bar{\boldsymbol{B}}_n), \phi(\bar{\boldsymbol{B}}_n)) = (\bar{\boldsymbol{Y}}_n, \bar{\boldsymbol{Z}}_n).$$

This implies that $h^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) = \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\bar{\boldsymbol{B}}_n)$.

According to Theorem 6.2 in Appendix 6.2, ϕ is Lipschitz continuous on $\prod_{i=1}^d \mathbb{D}^{-(\mathcal{Q}\boldsymbol{r})_i}[0,T]$, so is $h^\intercal \cdot \pi \circ \phi$. Therefore, we can apply the contraction principle (Lemma 6.3 part (1)) to $\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\bar{\boldsymbol{B}}_n)$, and this gives the statement of the proposition. □

The conclusion of Proposition 3.2 gives the asymptotic of $\mathbb{P}\left(\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) \geq y\right)$:

$$-V(y) \leq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) \geq y\right)}{r_0(\log n)} \leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\boldsymbol{h}^\intercal \cdot \bar{\boldsymbol{Z}}_n(T) \geq y\right)}{r_0(\log n)} \leq -\lim_{\epsilon\downarrow 0} V(y-\epsilon), \quad (3.7)$$

where

$$V(y) = \inf_{x\in(y,\infty)} I_{\boldsymbol{h}^\intercal\cdot\bar{\boldsymbol{Z}}}(x) = \min\left\{I_{\boldsymbol{B}}(\boldsymbol{\xi}) : \boldsymbol{h}^\intercal \cdot \pi\circ\phi(\boldsymbol{\xi}) > y, \boldsymbol{\xi} \in \prod_{i=1}^{d} \mathbb{D}^{-(\mathcal{Q}\gamma)_i}[0,T]\right\}. \quad (3.8)$$

Note that determining the value of $V(\cdot)$ involves optimization problems in function space, therefore, for the remainder of the section, we dive deeper into the form of $V(\cdot)$, with the aim of providing numerical solutions for the bounds in (3.7).

The value of $V(\cdot)$ defined in (3.8) is closely related to the value of $\boldsymbol{h}^\intercal \cdot \pi\circ\phi(\boldsymbol{\xi})$ for $\boldsymbol{\xi}$ belonging to the effective domain $\mathcal{D}_{I_{\boldsymbol{B}}}$, which contains paths with jumps at coordinates that belong to $\mathcal{J}$ and are linear everywhere else (with the rate $(\boldsymbol{\gamma}-\mathcal{Q}\boldsymbol{r})_i$ at the coordinate i). To present $\mathcal{D}_{I_{\boldsymbol{B}}}$, we introduce the notation $\mathbb{S}_{\mathcal{K}}= \prod_{i=1}^d S_i$ and $\mathbb{S}^1_{\mathcal{K}}= \prod_{i=1}^d S^1_i$ for any $\mathcal{K}\subset\mathcal{J}$, where

$$S_i = \begin{cases} \cup_{n=1}^{\infty}\mathbb{D}^{(\boldsymbol{\gamma}-\mathcal{Q}\boldsymbol{r})_i}_{=n}[0,T] & i\in\mathcal{K} \\ \{(\boldsymbol{\gamma}-\mathcal{Q}\boldsymbol{r})_i\cdot t, t\in[0,T]\} & i\notin\mathcal{K}\end{cases} \text{ and } S^1_i = \begin{cases} \mathbb{D}^{(\boldsymbol{\gamma}-\mathcal{Q}\boldsymbol{r})_i}_{=1}[0,T] & i\in\mathcal{K} \\ \{(\boldsymbol{\gamma}-\mathcal{Q}\boldsymbol{r})_i\cdot t, t\in[0,T]\} & i\notin\mathcal{K}\end{cases} \quad (3.9)$$

Intuitively, $\mathbb{S}_{\mathcal{K}}$ is the subset of $\mathcal{D}_{I_{\boldsymbol{B}}}$ that restricts jumps to occur only at the coordinates in $\mathcal{K}$ and $\mathbb{S}^1_{\mathcal{K}}$ requires the jump number to be one. It is straightforward to verify

$$\mathcal{D}_{I_{\boldsymbol{B}}} = \bigcup_{\mathcal{K}\subset\mathcal{J}} \mathbb{S}_{\mathcal{K}}, \quad (3.10)$$

where the union is disjoint over all possible subset $\mathcal{K}\subset\mathcal{J}$.

Let us consider the following system of linear equations:

$$\boldsymbol{f} = Q^\intercal\boldsymbol{o} + \boldsymbol{\gamma} \quad (3.11)$$

and

$$\boldsymbol{o} = \mathbb{1}_{\mathcal{B}} * \boldsymbol{r} + (\boldsymbol{1}-\mathbb{1}_{\mathcal{B}}) * \min\{Q^\intercal\boldsymbol{o}+\boldsymbol{\gamma}, \boldsymbol{r}\}. \quad (3.12)$$

In the above equations, $\boldsymbol{r}, \boldsymbol{\gamma}, Q$ are the parameters used in Definition 3.1 and Assumption 1 . $\mathcal{B}$ is a subset of $[d]$ and $\mathbb{1}_{\mathcal{B}}$ is the vector equal to 1 for the coordinates in $\mathcal{B}$ and 0 elsewhere. The unknowns $\boldsymbol{f}$ and $\boldsymbol{o}$ are vectors in $\mathbb{R}^d_+$, and the symbol $*$ denotes coordinate-wise multiplication. To motivate, equations (3.11) and (3.12) need to be satisfied by the transient states of $\mathrm{DF}(\boldsymbol{\gamma})$, a fluid network with linear inflow to all nodes with constant rate $\boldsymbol{\gamma}$. (see Definition 5.1 and more discussion in Appendix 5.1) The corollary 5.1 in Appendix 5.1 confirms that, given $\mathbb{1}_{\mathcal{B}}$, there exist unique solutions $\boldsymbol{o}^{\mathcal{B}}$ and $\boldsymbol{f}^{\mathcal{B}}$for equations (3.11) and (3.12).

The following lemma characterizes several properties of $\boldsymbol{h}^\intercal \cdot \pi\circ\phi(\boldsymbol{\xi})$ for $\boldsymbol{\xi}\in\mathcal{D}_{I_{\boldsymbol{B}}}$.

Lemma 3.1. For fixed $\boldsymbol{h}\in\mathbb{R}^d_+$ and $\mathcal{K}\subset\mathcal{J}$, the following statements hold:

(1) For $\boldsymbol{\xi}\in\mathbb{S}_{\mathcal{K}}$, there exists $\boldsymbol{\eta}\in\mathbb{S}^1_{\mathcal{K}}$ such that $\boldsymbol{h}^\intercal \cdot \pi\circ\phi(\boldsymbol{\xi}) \leq \boldsymbol{h}^\intercal \cdot \pi\circ\phi(\boldsymbol{\eta})$.

(2) If $\mathcal{K}\cap\{i\in[d] : h_i > 0\} \neq \emptyset$, then $\sup\{\boldsymbol{h}^\intercal \cdot \pi\circ\phi(\boldsymbol{\xi}), \boldsymbol{\xi}\in\mathbb{S}^1_{\mathcal{K}}\} = \infty$.

(3) If $\mathcal{K} \cap \{i \in [d] : h_i > 0\} = \emptyset$. Let $\boldsymbol{o}^{\mathcal{K}}$ and $\boldsymbol{f}^{\mathcal{K}}$ be the unique solution of (3.11) and (3.12) with respect to $\mathcal{B} = \mathcal{K}$. Then

$$\sup\{\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}), \boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}\} = \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T. \tag{3.13}$$

Furthermore, the value in (3.13) cannot be reached by any $\boldsymbol{\xi} \in \mathbb{S}_{\mathcal{K}}$ if this value is strictly greater than zero.

The proof of Lemma is put in With the help of Lemma 3.1, the following proposition gives a way to evaluate the value of $V(\cdot)$ function.

Proposition 3.3. For fixed $\boldsymbol{h} \in \mathbb{R}^d_+$, let $c^* = \min\{\lambda_i, i \in \mathcal{J}, h_i > 0\}$. The $V(\cdot)$ defined in (3.8) satisfies

$$V(y) = \Big(\min_{\mathcal{K} \subset \mathcal{J}} \Big\{ \sum_{i \in \mathcal{K}} \lambda_i : \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T > y, \mathcal{K} \cap \{i \in [d] : h_i > 0\} = \emptyset \Big\}\Big) \wedge c^* \tag{3.14}$$

where $\boldsymbol{o}^{\mathcal{K}}$ and $\boldsymbol{f}^{\mathcal{K}}$ are the unique solutions of (3.11) and (3.12) with $\mathcal{B} = \mathcal{K}$.

Proof. According to (3.8), the $V(y)$ is the minimum value of $I_{\boldsymbol{B}}(\boldsymbol{\xi})$ such that $\boldsymbol{\xi}$ belongs to the effective domain $\mathcal{D}_{I_{\boldsymbol{B}}}$ and $\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) > y$. Due to Lemma 3.1 part (1), we can only consider $\boldsymbol{\xi} \in \bigcup_{\mathcal{K} \subset \mathcal{J}} \mathbb{S}^1_{\mathcal{K}}$, and since $I_{\boldsymbol{B}}(\boldsymbol{\xi}) = \sum_{i \in \mathcal{K}} \lambda_i$ for $\boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}$, we have

$$V(y) = \inf\{\sum_{i \in \mathcal{K}} \lambda_i : \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) > y, \boldsymbol{\xi} \in \bigcup_{\mathcal{K} \subset \mathcal{J}} \mathbb{S}^1_{\mathcal{K}}\}.$$

Due to the above equation, we can represent $V(y) = (I) \wedge (II)$, where

$$\begin{aligned}
(\mathrm{I}) &= \min_{\mathcal{K} \subset \mathcal{J}, \mathcal{K} \cap \{i \in [d]: h_i > 0\} \neq \emptyset} \big\{ \inf\{I_{\boldsymbol{B}}(\boldsymbol{\xi}) : \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) > y, \boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}\} \big\} \\
(\mathrm{II}) &= \min_{\mathcal{K} \subset \mathcal{J}, \mathcal{K} \cap \{i \in [d]: h_i > 0\} = \emptyset} \big\{ \inf\{I_{\boldsymbol{B}}(\boldsymbol{\xi}) : \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) > y, \boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}\} \big\}
\end{aligned}$$

For term (I), part (2) of Lemma 3.1 implies $\sup\{\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}), \boldsymbol{\xi} \in \mathbb{S}_{\mathcal{K}}\} = \infty$. Hence

$$(\mathrm{II}) = \min_{\mathcal{K} \subset \mathcal{J}, \mathcal{K} \cap \{i \in [d]: h_i > 0\} \neq \emptyset} \big\{ \sum_{i \in \mathcal{K}} \lambda_i \big\} = c^*.$$

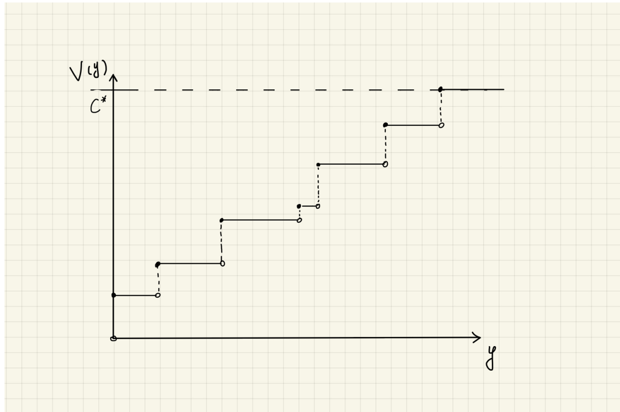
For term (II), part (3) of Lemma 3.1 implies that we can find some $\boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}$ such that $\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) > y$ if and only if $\boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T > y$. Therefore

$$(\mathrm{I}) = \min_{\mathcal{K} \subset \mathcal{J}} \Big\{ \sum_{i \in \mathcal{K}} \lambda_i : \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T > y, \mathcal{K} \cap \{i \in [d] : h_i > 0\} = \emptyset \Big\}.$$

Combining (I) and (II) completes the proof.

□

To illustrate the conclusion in Proposition 3.3, it is possible to depict the graph of $V(\cdot)$ for not too large d and $|\mathcal{J}|$. This boils down to compute $\boldsymbol{f}^{\mathcal{K}}$ and $\boldsymbol{o}^{\mathcal{K}}$ as the solution of equations (3.11) and (3.12) for all subsets $\mathcal{K}$ of $\mathcal{J}$. By the $V(\cdot)$'s formulation, its graph is a piecewise constant and nondecreasing function (shown as an instant in the figure below). The order of the values of $\boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T$ for all $\mathcal{K} \subset \mathcal{J}$ such that $\mathcal{K} \cap \{i \in [d] : h_i > 0\} \neq \emptyset$. will enable us to locate the jump location and size of the V' s graph.


V(y)
C*
y

4 Queue Length Tail Asymptotics in a GI/GI/d System

The $GI/GI/d$ is a system composed of d servers. Jobs arrive one at a time and are served by one of the d servers. If there is an available server, an arrived job will immediately start service. Otherwise, the job will join the queue and wait to be served in a first-come-first-served (FCFS) manner. The interarrival times of the jobs are independent and identically distributed (i.i.d.), while the service times for all jobs are also i.i.d. and independent of the arrival process.

Throughout the section, we have the following distribution assumptions in terms of the inter-arrival times and services times.

Assumption 2. Let A and $S^{(i)}$ be the generic random variables that have the distribution of inter-arrival time and job service time at station i, respectively.

- There exists $\theta_+ > 0$ such that $\mathbb{E}\left[e^{\theta A}\right] < \infty$ for every $\theta < \theta_+$
- For $i \in [d]$, $S^{(i)}$ has the distribution as specified in Definition 2.3 with a regularly varying function r_i. Without loss of generality, we assume $\mathbb{E}\left[S^{(i)}\right] = 1$.
- For $i \in [d]$, the regularly varying functions r_i' associated with $S^{(i)}$' share the same index $\beta > 1$, and there exists a function r_0 such that $\lim\limits_{x\to\infty} \frac{r_i(x)}{r_0(x)} = \lambda_i$.
- $\frac{1}{\mathbb{E}[A]} < \frac{d}{\mathbb{E}[S]} = d$. This is the requirement that guarantees the stability of the system.

Consider a FCFS $GI/GI/d$ system with no previous job arrivals at $t = 0$. Let $Q(t), t \geq 0$ denote the number of jobs in the system at time t, including those being served and waiting in the queue. This section aims at characterizing the asymptotic of

$$\mathbb{P}\left(Q(n) \geq pn\right)$$

as $n \to \infty$ for any $p > 0$.

Let A_j's and $S_j^{(i)}$'s be i.i.d.copies of A and S representing the sequence of inter-arrival times and the sequence of service times at server i, respectively. We define the following pure jump processes on the non-negative time interval: for $s \in [0, \infty)$

$$A(s) = A_1 + \cdots A_{\lfloor s \rfloor} \text{ and } S^{(i)}(s) = S_1^{(i)} + \cdots S_{\lfloor s \rfloor}^{(i)}, \tag{4.1}$$

and their renewal process

$$M(t) = \inf\{s : A(s) > t\} \text{ and } N^{(i)}(t) = \inf\{s : S^{(i)}(s) > t\}.$$

Corollary 1 in [15] states that $Q(\cdot)$, $M(\cdot)$ and $N^{(i)}(\cdot)$ satisfy the following inequality:

$$\mathbb{P}\left(Q(t) > x\right) \leq \mathbb{P}\left(\sup_{s\in[0,t]} \left\{\big(M(t) - M(t-s)\big) - \sum_{i=1}^{d} \left(N^{(i)}(t) - N^{(i)}(t-s)\right)\right\} > x\right).$$

If we define $\bar{M}(t)$= $\frac{M(nt)}{n}$ and $\bar{N}_n^{(i)}(t)$= $\frac{N^{(i)}(nt)}{n}$, the above result translates to

$$\mathbb{P}\left(Q(n) \geq pn\right) \leq \mathbb{P}\left(\sup_{s\in[0,1]} \left\{\big(\bar{M}_n(1) - \bar{M}_n(s)\big) - \sum_{i=1}^{d} \left(\bar{N}_n^{(i)}(1) - \bar{N}_n^{(i)}(s)\right)\right\} \geq p\right). \tag{4.2}$$

In light of (4.2), we will derive an asymptotic upper bound for $\mathbb{P}(Q(n) \geq pn)$ by analyzing the large deviations of the sequences $\{\bar{M}_n\}_{n\geq 1}$ and $\{\bar{N}_n^{(i)}\}_{n\geq 1}$. We will also establish a matching lower bound for $\mathbb{P}(Q(n) \geq pn)$, confirming that the same asymptotic rate applies in both directions.

We start with the following Lemma 4.1, stating the large deviations of $\{\bar{A}_n\}_{n\geq 1}$ and $\{\bar{S}_n^{(i)}\}_{n\geq 1}$, where $\bar{A}_n(t) \triangleq A(nt)/n$ for $t \in [0, 1/\mathbb{E}[A]]$ and $\bar{S}_n^{(i)}(t) \triangleq S^{(i)}(nt)/n$ for $t \in [0,1]$.

Lemma 4.1. (1) The sequence $\{\bar{A}_n\}_{n\geq 1}$ on $\big(\mathbb{D}[0, \frac{1}{\mathbb{E}[A]}], J_1\big)$ satisfies full LDP with speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function

$$I_0(\xi) = \begin{cases} 0 & \xi = \zeta_{\mathbb{E}[A]} \\ \infty & \text{otherwise} \end{cases} \tag{4.3}$$

(2) For $i \in [d]$, the sequence $\{\bar{S}_n^{(i)}\}_{n\geq 1}$ on $\big(\mathbb{D}[0,1], J_1\big)$ satisfies extended LDP with speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function

$$I_i(\xi) = \begin{cases} \lambda_i \cdot \sum_{t\in[0,1]} \mathbb{1}\{\xi(-t) \neq \xi(t)\} & \xi \in \tilde{\mathbb{D}}^1_{<\infty}[0,1] \\ \infty & \text{otherwise} \end{cases} \tag{4.4}$$

Proof. For (1), consider the random walk process $\bar{A}'_n$ defined by

$$\bar{A}'_n(t) = \frac{1}{n}\sum_{j=1}^{\lfloor nt \rfloor}(A_j - \mathbb{E}[A]) \text{ for } t \in [0, \frac{1}{\mathbb{E}[A]}].$$

By Lemma 3.2 in [24], the sequence $\{\bar{A}'_n\}_{n\geq 1}$ satisfies a full LDP on $\big(\mathbb{D}[0, \frac{1}{\mathbb{E}[A]}], J_1\big)$ with speed $\{r_0(\log n)\}_{n\geq 1}$ and the good rate function

$$I_A(\xi) = \begin{cases} 0 & \xi = \mathbf{0} \\ \infty & \text{otherwise} \end{cases}.$$

Next, define $\bar{A}''_n$ as

$$\bar{A}''_n(t) = \frac{1}{n}\sum_{j=1}^{\lfloor nt \rfloor} A_j - \mathbb{E}[A] \cdot t$$

for $t \in [0, \frac{1}{\mathbb{E}[A]}]$. The uniform gap between $\bar{A}'_n$ and $\bar{A}''_n$'s vanishes as $n \to \infty$, making the two processes exponentially equivalent. By part 1 of Lemma 6.4, the sequence $\{\bar{A}''_n\}_{n\geq 1}$ satisfies the same full LDP as $\{\bar{A}'_n\}_{n\geq 1}$, with the same speed and rate function.

Now, consider $\bar{A}_n = \Upsilon^{\mathbb{E}[A]}(\bar{A}''_n)$. By Lemma 6.5, $\Upsilon^{\mathbb{E}[A]}$ is continuous under the J_1 topology. Applying the contraction principle (part 2 of Lemma 6.2), the sequence $\{\bar{A}_n\}_{n\geq 1}$ satisfies a full LDP on $\big(\mathbb{D}[0, \frac{1}{\mathbb{E}[A]}], J_1\big)$ with the same speed $\{r_0(\log n)\}_{n\geq 1}$ and rate function given by (4.3). This establishes the conclusion for part (1).

We follow a similar approach to prove part (2). Fix $i \in [d]$, define the random walk process $\bar{S}'_n$ as

$$\bar{S}'_n(t) = \frac{1}{n}\sum_{j=1}^{\lfloor nt \rfloor}(S_j^{(i)} - \mathbb{E}[S]) \text{ for } t \in [0,1].$$

By Theorem 2.2, the sequence $\{\bar{S}'_n\}_{n\geq 1}$ satisfies an extended LDP on $\big(\mathbb{D}[0,1], J_1\big)$ with speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function defined in (2.6). Define $\bar{S}''_n$ as

$$\bar{S}''_n(t) = \frac{1}{n}\sum_{j=1}^{\lfloor nt \rfloor} S_j^{(i)} - \mathbb{E}[S] \cdot t$$

for $t \in [0,1]$. Since $\bar{S}'_n$ and $\bar{S}''_n$ are exponentially equivalent, part (2) of Lemma 6.4 ensures that $\{\bar{S}''_n\}_{n\geq 1}$ satisfies the same extended LDP with the same speed and rate function. Finallly, consider $\bar{S}^{(i)}_n = \Upsilon^1(\bar{S}''_n)$. By Lemma 6.5, Υ^1 is Lipschitz continuous. Applying the contraction principle(part (1) of Lemma 6.3), the rate function I_i defined in (4.4) governs the upper and lower bounds of the extended LDP for $\{\bar{S}^{(i)}_n\}_{n\geq 1}$. Since I_i is a lower semicontinuous, part (2) of Lemma 6.3 confirms the conclusion for part (2). □

Define $\Psi(\xi)(t)$ as the running supremum of the path ξ up to time t, that is, $\Psi(\xi)(t) = \sup_{s\in[0,t]} \xi(s)$. Also, for any $\mu \in \mathbb{R}_+$, define $\Phi_\mu : \mathbb{D}[0, \frac{1}{\mu}] \to \mathbb{D}[0,1]$ by

$$\Phi_\mu(\xi)(t) \triangleq \varphi_\mu(\xi)(t) \wedge \psi_\mu(\xi)(t), \tag{4.5}$$

where

$$\varphi_\mu(\xi)(t) \triangleq \inf\{s \in [0, 1/\mu] : \xi(s) > t\}, \tag{4.6}$$

$$\psi_\mu(\xi)(t) \triangleq \frac{1}{\mu}\big[1 + [t - \Psi(\xi)(\frac{1}{\mu})]_+\big]. \tag{4.7}$$

An important property o Φ_μ, which plays a key role in establishing the conclusions of Propositions 4.1 and 4.2 below, is summarized in the following Lemma 4.2. The proof of Lemma 4.2 is provided in Appendix 5.2.

Lemma 4.2. For any $\xi \in \tilde{\mathbb{D}}^{\mu}_{<\infty}[0, \frac{1}{\mu}]$ and $\eta \in \mathbb{D}[0, \frac{1}{\mu}]$,

$$d_{J_1}(\Phi_\mu(\xi), \Phi_\mu(\eta)) \leq (2 + \frac{4}{\mu}) d_{J_1}(\xi, \eta). \tag{4.8}$$

Proposition 4.1. The sequence $\{\Phi_{\mathbb{E}[A]}(\bar{A}_n)\}_{n\geq 1}$ satisfies the full LDP on $\big(\mathbb{D}[0,1], J_1\big)$ with speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function given by

$$I'_0(\xi) = \begin{cases} 0 & \xi = \zeta_{\frac{1}{\mathbb{E}[A]}} \\ \infty & \text{otherwise} \end{cases} \tag{4.9}$$

Proof. By part (1) of Lemma 4.1, the sequence $\{\bar{A}_n\}_{n\geq 1}$ satisfies a full LDP with the rate function I_0. Furthermore, Lemma 4.2 confirms that the mapping $\Phi_{\mathbb{E}[A]}$ is continuous with respect to the J_1 topology at the path $\zeta_{\mathbb{E}[A]}$, , which corresponds to the effective domain of I_0. Applying the contraction principle(parts (1) and (2) of Lemma 6.2) to $\{\bar{A}_n\}_{n\geq 1}$ via the mapping $\Phi_{\mathbb{E}[A]}$ yields the desired conclusion of the proposition. □

Let $\mathbb{C}[0,1]$ be the subspace of $\mathbb{D}[0,1]$ that contains continuous functions. Define

$$\check{\mathbb{C}}^1[0,1] = \big\{\xi \in \mathbb{C}[0,1] : \xi = \Phi_1(\eta), \eta \in \tilde{\mathbb{D}}^1_{<\infty}[0,1]\big\} \tag{4.10}$$

Proposition 4.2. For each $i \in [d]$, the sequence $\{\Phi_1(\bar{S}^{(i)}_n)\}_{n\geq 1}$ satisfies the extended LDP on $\big(\mathbb{D}[0,1], J_1\big)$ with speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function given by

$$I'_i(\xi) = \begin{cases} \lambda_i \cdot \sum_{s\in(0,1]} \tau_s(\xi) & \text{if } \xi \in \check{\mathbb{C}}^1[0,1] \\ \infty & \text{otherwise} \end{cases} \tag{4.11}$$

where

$$\tau_s(\xi) = \mathbb{1}\{\sup\{t \in (0,1] : \xi(t) = s\} \neq \inf\{t \in (0,1] : \xi(t) = s\}\} \tag{4.12}$$

Proof. By Lemma 4.2, Φ_1 is Lipschitz continuous at the effective domain of I_i(defined in (4.4)). Therefore, the contraction principle (part (1) of Lemma 6.3) applies to the sequence $\{\Phi_{\mathbb{E}[A]}(\bar{S}_n^{(i)})\}_{n\geq 1}$. Since part (2) of Lemma 4.1 confirms that $\{\bar{S}_n\}_{n\geq 1}$ satisfies an extended LDP, the function $\tilde{I}_i$, defined below, controls the upper and lower bounds of the extended LDP for $\{\Phi_1(\bar{S}_n^{(i)})\}_{n\geq 1}$:

$$\tilde{I}_i(\xi) = \inf\{I_i(\eta) : \xi = \Phi_1(\eta), \eta \in \tilde{\mathbb{D}}^1_{<\infty}[0,1]\}.$$

We now argue that $\tilde{I}_i = I'_i$. By definition, the effective domain of $\tilde{I}_i$ and I'_i are identical, both being $\check{\mathbb{C}}^1[0,1]$, where paths alternate between flat segments and increasing segments with rate 1 over subintervals of $[0,1]$.

For any $\xi \in \check{\mathbb{C}}^1[0,1]$, there exists $\eta \in \tilde{\mathbb{D}}^1_{<\infty}[0,1]$ such that $\xi = \Phi_1(\eta)$. Since $\eta(1) \geq 1$, it follows that $\xi(1) \leq 1$. Construct $\eta^* \in \tilde{\mathbb{D}}^1_{<\infty}[0,1]$ by letting η^* equal η on $[0, \xi(1))$ and increase with rate 1 on $[\xi(1), 1]$. Then, $\Phi_1(\eta^*) = \xi$, and $I_i(\eta^*) = I'_i(\xi)$, which shows that $I'_i(\xi) \geq \tilde{I}_i(\xi)$. To prove the reverse inequality, assume there exists $\eta' \in \tilde{\mathbb{D}}^1_{<\infty}[0,1]$ such that $\xi = \Phi_1(\eta')$ and $I'_i(\xi) > I_i(\eta')$. The number of flat regions of ξ corresponds 1-to-1 to the jumps of η' on $[0, \varphi_1(\eta')(1)]$. Since $I'_i(\xi)$ counts the flat regions of ξ', it cannot exceed the total number of jumps of η', which is equal to $I_i(\eta')$. This contradiction establishes $I'_i(\xi) \leq \tilde{I}_i(\xi)$. Thus ,$\tilde{I}_i = I'_i$

Finally, Lemma 5.2 confirms that the level sets of I'_i, defined as $L_{\leq k} = \{\xi \in \check{\mathbb{C}}^1[0,1], I'_i(\xi) \leq \lambda \cdot k\}$, are closed. Therefore, I'_i is lower semicontinuous. Applying the contraction principle (part (2) of the lemma 6.3) completes the proof of the proposition. □

Proposition 4.3. On $(\mathbb{D}[0,1], J_1)$, the pair $\{\bar{M}_n\}_{n\geq 1}$ and $\{\Phi_{\mathbb{E}[A]}(\bar{A}_n)\}_{n\geq 1}$, and the pair $\{\bar{N}_n^{(i)}\}_{n\geq 1}$ and $\{\Phi_1(\bar{S}_n^{(i)})\}_{n\geq 1}$, are both exponentially equivalent w.r.t. the speed sequence $\{r_0(\log n)\}_{n\geq 1}$

Proof. The proof of the exponential equivalence for the two pairs are similar; hence we focus on the pair $\{\bar{N}_n^{(i)}\}_{n\geq 1}$ and $\{\Phi_1(\bar{S}_n^{(i)})\}_{n\geq 1}$.

We claim for $\epsilon > 0$,

$$\Big\{d_{J_1}\big(\Phi_1(\bar{S}_n^{(i)}), \bar{N}_n^{(i)}\big) \geq \epsilon\Big\} \subset \Big\{\big\|\Phi_1(\bar{S}_n^{(i)}) - \bar{N}_n^{(i)}\big\| \geq \epsilon\Big\} \tag{4.13}$$

$$=\Big\{\sup_{t\in[\Psi(\bar{S}_n^{(i)}),1]} \big|\Phi_1(\bar{S}_n^{(i)})(t) - \bar{N}_n^{(i)}(t)\big| \geq \epsilon\Big\} \tag{4.14}$$

$$\subset\Big\{\sup_{t\in[\Psi(\bar{S}_n^{(i)}),1]} \big|\Phi_1(\bar{S}_n^{(i)})(t) - 1\big| \geq \frac{\epsilon}{2}\Big\} \cup \Big\{\sup_{t\in[\Psi(\bar{S}_n^{(i)}),1]} \big|1 - \bar{N}_n^{(i)}(t)\big| \geq \frac{\epsilon}{2}\Big\} \tag{4.15}$$

$$\subset\Big\{\Phi_1(\bar{S}_n^{(i)})(1) - 1 \geq \frac{\epsilon}{2}\Big\} \cup \Big\{\bar{N}_n^{(i)}(1) - 1 \geq \frac{\epsilon}{2}\Big\} \tag{4.16}$$

$$=\Big\{\bar{S}_n^{(i)}(1) \leq 1 - \frac{\epsilon}{2}\Big\} \cup \Big\{\bar{S}_n^{(i)}(1 + \frac{\epsilon}{2}) \leq 1\Big\} \tag{4.17}$$

In the above, the transition from (4.13) to (4.14) holds because the paths $\bar{N}_n^{(i)}$ and $\Phi_1(\bar{S}_n^{(i)})$ differ only on $[\Psi(\bar{S}_n^{(i)}), 1]$, where $\Psi(\bar{S}_n^{(i)}) \leq 1$. If $\Psi(\bar{S}_n^{(i)}) > 1$, the paths are identical. The transition from (4.15) to (4.16) follows from the nondecreasing nature of both $\Phi_1(\bar{S}_n^{(i)})$ and $\bar{N}_n^{(i)}$, ensuring that they take values greater than or equal to 1 in$[\Psi(\bar{S}_n^{(i)}), 1]$. The transition from (4.16) to (4.17) uses the fact that when $\Phi_1(\bar{S}_n^{(i)}) \geq 1 + \epsilon/2$, $\bar{S}_n^{(i)}(1) \leq 1 - \frac{\epsilon}{2}$ due to the linear increase of Φ_1 at rate 1. Similarly, $\bar{N}_n^{(i)}(1) - 1 \geq \frac{\epsilon}{2}$ implies $\bar{S}_n^{(i)}(1 + \frac{\epsilon}{2}) \leq 1$.

The above inclusion imply:

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(d_{J_1}\left(\Phi_1(\bar{S}_n^{(i)}), \bar{N}_n^{(i)}\right) \geq \epsilon\right)}{r_0(\log n)}$$
$$\leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\left\{\bar{S}_n^{(i)}(1) \leq 1 - \frac{\epsilon}{2}\right\}\right)}{r_0(\log n)} \vee \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\left\{\bar{S}_n^{(i)}(1 + \frac{\epsilon}{2}) \leq 1\right\}\right)}{r_0(\log n)} \tag{4.18}$$

Any path ξ in the effective domain of I_i(defined in (4.4)) satisfies $\xi(1) \geq 1$. By the convergence of J_1, events$\left\{\bar{S}_n^{(i)}(1) \leq 1 - \frac{\epsilon}{2}\right\}$ and $\left\{\bar{S}_n^{(i)}(1 + \frac{\epsilon}{2}) \leq 1\right\}$ fall outside the effective domain of I_i' with respect to the J_1 metric. Using the extended LDP of $\{\bar{S}_n^{(i)}\}_{n\geq 1}$ established in part (2) of Lemma 4.1, both terms in (4.18) are equal to $-\infty$, and this finishes the proof. □

Proposition 4.4. (1) The sequence $\{\bar{M}_n\}_{n\geq 1}$ satisfies the full LDP on $\big(\mathbb{D}[0,1], J_1\big)$ with the speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function I_0' defined in (4.9).

(2) For $i = 1, \cdots, d$, the sequence $\{\bar{N}_n^{(i)}\}_{n\geq 1}$ satisfies the extended LDP on $\big(\mathbb{D}[0,1], J_1\big)$ with the speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function I_i' defined in (4.11).

Proof. By Proposition 4.3, the pairs $\{\bar{M}_n\}_{n\geq 1}$ and $\{\Phi_{\mathbb{E}[A]}(\bar{A}_n)\}_{n\geq 1}$, as well as $\{\bar{N}_n^{(i)}\}_{n\geq 1}$ and $\{\Phi_1(\bar{S}_n^{(i)})\}_{n\geq 1}$, are two exponential equivalent pairs. Applying Lemma 6.4 part (1) to the first pair, and part (2) to the second, we conclude that both sequences in each pair satisfy the same large deviation principle, sharing the same rate function and speed sequence. Consequently, Proposition 4.1 and Proposition 4.2 directly establish conclusions (1) and (2). □

Proposition 4.5. The multi-valued processes $\{(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)})\}_{n\geq 1}$ satisfies the extended LDP on $(\Pi_{i=1}^d \mathbb{D}[0,1], d_p)$ with the speed $\{r_0(\log n)\}_{n\geq 1}$ and the rate function given by

$$I'(\xi_1, \cdots, \xi_d) = \begin{cases} \sum_{i=1}^d \left(\lambda_i \cdot \sum_{s\in(0,1]} \tau_s(\xi_i)\right) & \text{if } \xi_i \in \check{\mathbb{C}}^1[0,1] \text{ for } i \in [d] \\ \infty & \text{otherwise} \end{cases} \tag{4.19}$$

Proof. Since the rate function I_i' (see (4.11)) corresponding to the extended LDP of $\{\bar{N}_n^{(i)}\}_{n\geq 1}$ take only non-negative multiples of λ_i, and $\bar{N}_n^{(i)}$'s are independent processes across $i \in [d]$, the conclusion of the proposition directly follows from Theorem 2.1. □

The large deviations of $\{\bar{M}_n\}_{n\geq 1}$ and $\{(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)})\}_{n\geq 1}$ derived in Propositions 4.4 and 4.5, will help determine the asymptotic upper and lower bound of $\mathbb{P}(Q(n) \geq pn)$, which will be presented in the following Propositions 4.6 and 4.7. Recall λ_i's to be the parameters in Assumption 2 and let $\lambda_{(1)}, \lambda_{(2)}, \cdots$ be their increasing order statistics. Define

$$W(y) = \begin{cases} \sum_{i=1}^{\lfloor y+d-\frac{1}{\mathbb{E}[A]}\rfloor+1} \lambda_{(i)} & y < \frac{1}{\mathbb{E}[A]} \\ \infty & \text{otherwise} \end{cases}. \tag{4.20}$$

Proposition 4.6. For $p > 0$,

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(Q(n) \geq np\right)}{r_0(\log n)} \leq -\lim_{x<p, x\uparrow p} W(x). \tag{4.21}$$

Proof. Define the function $f : \prod_{i=1}^{d} \mathbb{D}[0,1] \to \mathbb{R}$ as

$$f(\xi_1, \cdots, \xi_d) = \sup_{s\in[0,1]} \left\{ \frac{1-s}{\mathbb{E}[A]} - \sum_{i=1}^{d} \left(\xi_i(1) - \xi_i(s)\right) \right\}. \tag{4.22}$$

Starting from (4.2), for any $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}(Q(n) \geq pn) \leq \mathbb{P}\left(\sup_{s\in[0,1]} \left\{ \left(\bar{M}_n(1) - \bar{M}_n(s)\right) - \sum_{i=1}^{d} \left(\bar{N}_n^{(i)}(1) - \bar{N}_n^{(i)}(s)\right) \right\} \geq p \right) \\
\leq &\mathbb{P}\left(\bar{M}_n(1) - \frac{1}{\mathbb{E}[A]} \geq \epsilon \right) + \mathbb{P}\left(\inf_{s\in[0,1]} \left\{ \bar{M}_n(s) - \frac{s}{\mathbb{E}[A]} \right\} \leq -\epsilon \right) + \mathbb{P}\left(f(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)}) \geq p - 2\epsilon \right)
\end{aligned} \tag{4.23}$$

By the logrithmic function maximum principle, (4.23) implies

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}(Q(n) \geq pn)}{r_0(\log n)} \leq \max\left\{ (\mathrm{I})\ ,\ (\mathrm{II})\ ,\ (\mathrm{III}) \right\} \tag{4.24}$$

where

$$\begin{aligned}
(\mathrm{I}) &= \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{M}_n(1) - \frac{1}{\mathbb{E}[A]} \geq \epsilon \right)}{r_0(\log n)} \\
(\mathrm{II}) &= \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\inf_{s\in[0,1]} \left\{ \bar{M}_n(s) - \frac{s}{\mathbb{E}[A]} \right\} \leq -\epsilon \right)}{r_0(\log n)} \\
(\mathrm{III}) &= \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(f(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)}) \geq p - 2\epsilon \right)}{r_0(\log n)}
\end{aligned}$$

The events associated with (I) and (II) are bounded away from $\zeta_{1/\mathbb{E}[A]}$(w.r.t the J_1 topology). Thus, by the full LDP of $\{\bar{M}_n\}_{n\geq 1}$(Proposition 4.4 part (1)), both terms are equal to $-\infty$.

For (III), we apply the contraction principle (Lemma 6.3) for the sequence $\{f(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)})\}_{n\geq 1}$. This is valid due to the extended LDP of $\{(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)})\}_{n\geq 1}$(Corollary 4.5) and the Lipschitz continuity of the function f on the effective domain of I'(Lemma 5.5 in appendix section 5.2). Consequently, the extended LDP bounds for $\{f(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)})\}_{n\geq 1}$'s are governed by the function

$$I^{''}(x) = \inf\{I'(\boldsymbol{\xi}) : f(\boldsymbol{\xi}) = x, \boldsymbol{\xi} \in \Pi_{i=1}^{d} \check{\mathbb{C}}^1[0,1]\} \tag{4.25}$$

Combining this result with (4.24), and letting $\epsilon \to 0$, we obtain:

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\log \mathbb{P}(Q(n) \geq pn)}{r_0(\log n)} &\leq -\lim_{\epsilon\to 0} \left(\lim_{\epsilon'\to 0} \inf_{x\in[p-2\epsilon-\epsilon',\infty)} I^{''}(x) \right) \\
&\leq -\lim_{\epsilon\to 0} \inf_{x\in[p-\epsilon,\infty)} I^{''}(x) \\
&= -\lim_{x<p, x\uparrow p} W'(x)
\end{aligned} \tag{4.26}$$

where $W'(y) \triangleq \inf\{I^{''}(x) : x \geq y\}$.

For the rest of the proof, we argue $W' = W$. From (4.25) and (4.22), $W'(y)$ is equivalent to the optimal value of

$$\begin{aligned}
\inf \quad & \sum_{i=1}^{d} \big(\lambda_i \cdot \sum_{s\in[0,1]} \tau_s(\xi_i)\big) \\
\text{s.t.} \quad & \sup_{s\in[0,1]} \big\{\frac{1-s}{\mathbb{E}\,[A]} - [\sum_{i=1}^{d} \xi_i(1) - \xi_i(s)]\big\} \geq y \\
& \xi_i \in \check{\mathbb{C}}^1[0,1], \ \text{for } i \in [d]
\end{aligned} \tag{4.27}$$

Introducing the change of variables $\eta_i(1-\cdot) = \xi_i(1) - \xi_i(\cdot)$, the feasible domain becomes

$$\hat{\mathbb{C}}^1[0,1] \triangleq \big\{\eta \in \mathbb{D}[0,1] : \eta(1-\cdot) = \xi(1) - \xi(\cdot), \xi \in \check{\mathbb{C}}^1[0,1]\big\},$$

The reformulated problem becomes:

$$\begin{aligned}
\inf \quad & \sum_{i=1}^{d} \lambda_i \cdot \sum_{s\in[0,1]} \tau_s(\eta_i) \\
\text{s.t.} \quad & \sup_{s\in[0,1]} \{\frac{s}{\mathbb{E}[A]} - \sum_{i=1}^{d} \eta_i(s)\} \geq y, \\
& \eta_i \in \hat{\mathbb{C}}^1[0,1], \ \forall i \in [d].
\end{aligned} \tag{4.28}$$

Each η_i in the feasible domain $\hat{\mathbb{C}}^1[0,1]$ is a continuous function that is flat (constant) over a finite number of subintervals of [0,1) and increases linearly with slope 1 elsewhere. The term $\sum_{s\in[0,1]} \tau_s(\eta_i)$ counts the number of flat subintervals in η_i, which directly contributes to the objective.

We rule out $\boldsymbol{\eta}$ with any coordinates having more than one flat region as optimal solutions of (4.28), since concatenating two or more flat regions remains feasible while reducing the objective. Similarly, all coordinates of $\boldsymbol{\eta}$ cannot be the linear function with rate 1 on $[0,1]$, as the stability condition of the GI/GI/d($1/\mathbb{E}[A] - d < 0$) makes such $\boldsymbol{\eta}$ infeasilbe. Thus, the solution of (4.28) can be constructed by introducing single flat regions to coordinates of $\boldsymbol{\eta}$ incrementally, starting with the smallest $\lambda_{(1)}$. For $\boldsymbol{\eta}$ to be feasible, the number of flat regions introduced, denoted by k, must satisfy $y < 1/\mathbb{E}[A] - (d-k)$. This gives the smallest k as $k^* = \lfloor y + d - 1/\mathbb{E}[A] \rfloor + 1$, at which the optimal value is $\sum_{i=1}^{k^*} \lambda_{(i)}$. If $y \geq 1/\mathbb{E}[A]$, the optimization is infeasible, and $W'(y) = \infty$. This confirms that the form of $W'(\cdot)$ matches (4.20), completing the proof.

□

Proposition 4.7. For $p > 0$,

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}\,(Q(n) \geq np)}{r_0(\log n)} \geq -W(p). \tag{4.29}$$

Proof. Considering the obvious coupling between Q and $(M, N^{(1)}, \cdots, N^{(d)})$, the term $M(s) - \sum_{i=1}^{d} N^{(i)}(s)$ represents a lower bound on the length of an imaginary queue at time s, where servers can work on jobs that have not yet arrived. Based on this observation, for any $s \in (0,1)$, we have $\mathbb{P}\,(Q(n) \geq pn) \geq \mathbb{P}\,(Q(n) \geq pn | Q(ns) = 0) \geq \mathbb{P}\left(\bar{M}_N(s) - \sum_{i=1}^{d} \bar{N}^{(i)}(s) \geq p\right)$. For any fixed $s \in (0,1)$ and any $\epsilon > 0$, this implies

$$\mathbb{P}\,(Q(n) \geq pn) \geq \mathbb{P}\left(\bar{M}_n(s) - \frac{s}{\mathbb{E}\,[A]} \geq -\epsilon, \frac{s}{\mathbb{E}\,[A]} - \sum_{i=1}^{d} \bar{N}_n^{(i)}(s) \geq p + \epsilon\right).$$

Using independence between $\bar{M}_n$ and $\bar{N}_n^{(i)}$, this can be written as

$$\mathbb{P}\left(\bar{M}_n(s) - \frac{s}{\mathbb{E}[A]} \geq -\epsilon\right) \cdot \mathbb{P}\left(\frac{s}{\mathbb{E}[A]} - \sum_{i=1}^{d} \bar{N}_n^{(i)}(s) \geq p+\epsilon\right). \tag{4.30}$$

Taking logarithmic limits, we obtain:

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}(Q(n) \geq pn)}{r_0(\log n)} \geq \max\left\{(\mathrm{I}), (\mathrm{II})\right\} \tag{4.31}$$

where

$$(\mathrm{I}) = \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{M}_n(s) - \frac{s}{\mathbb{E}[A]} \geq -\epsilon\right)}{r_0(\log n)}, (\mathrm{II}) = \liminf_{n\to\infty} \frac{\log \mathbb{P}\left((\bar{N}^{(i)}, \cdots, \bar{N}^{(d)}) \in A(s,\epsilon)\right)}{r_0(\log n)}.$$

and $A(s,\epsilon) = \{\boldsymbol{\xi} \in \Pi_{i=1}^{d} \mathbb{D}[0,1], s/\mathbb{E}[A] - \sum_{i=1}^{d} \xi_i(s) \geq p+\epsilon\}$.

For any $\epsilon > 0$ and $s \in (0,1)$, the event associated with (I) is bounded away from $\zeta_{1/\mathbb{E}[A]}$. By the large deviations principle $\{\bar{M}_n\}_{n\geq 1}$ (Proposition 4.4), we conclude that (I)$= -\infty$.

For (II), the large deviations of $\{(\bar{N}_n^{(1)}, \cdots, \bar{N}_n^{(d)})\}_{n\geq 1}$(Proposition 4.5) implies that

$$(\mathrm{II}) \geq -\inf_{\boldsymbol{\xi}\in\left(A(s,\epsilon)\right)^{\circ}} I'(\boldsymbol{\xi}) \tag{4.32}$$

with I' defined in (4.19). The compute the infimum, let $k = \lfloor \frac{p+\epsilon}{s} + d - \frac{1}{\mathbb{E}[A]} \rfloor + 1$. If $k \leq d$, consider the construction $\boldsymbol{\eta}$ where η_i increases at a rate of 1 on $[0,\delta]$ and remains flat on $[\delta, 1]$ for indices corresponding to the k-smallest λ_i, while the remaining η_i's increase linearly $[0,1]$. Choose $\delta > 0$ small enough ensures that $s/\mathbb{E}[A] - \sum_{i=1}^{d} \eta_i(s) > p+\epsilon$. Moreover, by continuity argument, this construction of $\boldsymbol{\eta}$ satisfies $\boldsymbol{\eta} \in \left(A(s,\epsilon)\right)^{\circ}$ and $I'(\boldsymbol{\eta}) = \sum_{i=1}^{k} \lambda_{(i)}$. If $k > d$, the value of k implies $\frac{s}{\mathbb{E}[A]} \leq p+\epsilon$. Since paths in the effective domain of I' must satisfy $\sum_{i=1}^{d} \eta_i(s) > 0$, this violates the constraint $s/\mathbb{E}[A] - \sum_{i=1}^{d} \eta_i(s) \geq p+\epsilon$, i.e $\boldsymbol{\eta} \notin A(s,\epsilon)$. Thus, $\inf_{\boldsymbol{\xi}\in\left(A(s,\epsilon)\right)^{\circ}} I'(\boldsymbol{\xi}) = \infty$. Combining those two cases, we have, for any $s \in (0,1)$,

$$\inf_{\boldsymbol{\xi}\in\left(A(s,\epsilon)\right)^{\circ}} I'(\boldsymbol{\xi}) = \begin{cases} \sum_{i=1}^{\lfloor \frac{p+\epsilon}{s} + d - \frac{1}{\mathbb{E}[A]} \rfloor + 1} \lambda_{(i)} & \lfloor \frac{p+\epsilon}{s} + d - \frac{1}{\mathbb{E}[A]} \rfloor + 1 \leq d \\ \infty & \lfloor \frac{p+\epsilon}{s} + d - \frac{1}{\mathbb{E}[A]} \rfloor + 1 > d \end{cases}. \tag{4.33}$$

Finally, Returning to (4.31), with (I)$= -\infty$ and (II)$\geq$-(4.33), by taking the infimum over $\epsilon > 0$ and $s \in (0,1)$, we conclude

$$\begin{aligned} \liminf_{n\to\infty} \frac{\log \mathbb{P}(Q(n) \geq pn)}{r_0(\log n)} &\geq -\inf_{\epsilon>0, s\in(0,1)} \inf_{\boldsymbol{\xi}\in\left(A(s,\epsilon)\right)^{\circ}} I'(\boldsymbol{\xi}) \\ &= \begin{cases} -\sum_{i=1}^{\lfloor p+d-\frac{1}{\mathbb{E}[A]} \rfloor + 1} \lambda_{(i)} & \lfloor p+d-\frac{1}{\mathbb{E}[A]} \rfloor + 1 \leq d \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} -\sum_{i=1}^{\lfloor p+d-\frac{1}{\mathbb{E}[A]} \rfloor + 1} \lambda_{(i)} & p < \frac{1}{\mathbb{E}[A]} \\ -\infty & \text{otherwise} \end{cases} \\ &= -W(p). \end{aligned}$$

This finishes the proof. $\square$

5 Proofs

5.1 Proofs in section 3

This section provides the proof of Lemma 3.1 and other supporting lemmas introduced in Section 3. To motivate the lemmas in this section, let us consider two types of fluid network with deterministic fluid inflow.

Definition 5.1. DF($\boldsymbol{\gamma}$) is a fluid network (as in Definition 3.1) with exogenous inflow to all nodes modeled by a linear function with constant rate $\boldsymbol{\gamma}$, where $\boldsymbol{\gamma}$ is the parameter in Assumption 1.

Definition 5.2. Given $\boldsymbol{\xi} \in \mathbb{S}_{\mathcal{K}}$ that takes the following form

$$\xi_i = \begin{cases} \sum_{j=1}^{n(i)} x_{ij} \mathbb{1}_{[t_{ij},T]}(t) + (\boldsymbol{\gamma} - \mathcal{Q}\boldsymbol{r})_i \cdot t & i \in \mathcal{K} \\ (\boldsymbol{\gamma} - \mathcal{Q}\boldsymbol{r})_i \cdot t & i \notin \mathcal{K} \end{cases}, \tag{5.1}$$

DF($\boldsymbol{\xi}$) is a fluid network (as in Definition 3.1) that has exogenous inflow to all nodes with the constant rate $\boldsymbol{\gamma}$ over time. In addition to that , for all $i \in \mathcal{K}$ and $1 \le j \le n(i)$, a large flow of fluid enters the buffer i at time t_{ij} with volume x_{ij}.

The transient state of both FN($\boldsymbol{\gamma}$) and FN($\boldsymbol{\xi}$), if we use $\boldsymbol{o}, \boldsymbol{f}$ and $\boldsymbol{b}$ to stand for their transient outflow rate, inflow rate and buffer content at certain time respectively, should satisfy equations (3.11) and (3.12), which are

$$\boldsymbol{f} = Q^{\intercal}\boldsymbol{o} + \boldsymbol{\gamma} \tag{5.2}$$

and

$$\boldsymbol{o} = \mathbb{1}_{\{i \in [d]: b_i > 0\}} * \boldsymbol{r} + (\boldsymbol{1} - \mathbb{1}_{\{i \in [d]: b_i > 0\}}) * \min\{Q^{\intercal}\boldsymbol{o} + \boldsymbol{\gamma}, \boldsymbol{r}\}. \tag{5.3}$$

The following lemma compares two transient states, stating that if one state has dominant buffer content, then so is the outflow rate.

Lemma 5.1. Assume both triples $(\boldsymbol{o}, \boldsymbol{f}, \boldsymbol{b})$ and $(\boldsymbol{o}', \boldsymbol{f}', \boldsymbol{b}')$ satisfy the equations (5.2) and (5.3). If $\mathcal{G}_1 = \{i \in [d] : b_i > 0\} \cup \{i \in [d] : f_i - r_i \ge 0\}$ and $\mathcal{G}_2 = \{i \in [d] : b'_i > 0\}$ are such that $\mathcal{G}_2 \subset \mathcal{G}_1$, then $\boldsymbol{o}' \le \boldsymbol{o}$ coordinately.

Proof of Lemma 5.1. We firstly prove the statement (1). Let $\boldsymbol{o}' = \boldsymbol{o} + \Delta\boldsymbol{o}$. If the conclusion of the lemma is not true, then $\mathcal{G} = \{i \in [d] : \Delta o_i > 0\}$ is not an empty set. For nodes $i \in \mathcal{G}$, since $o'_i \le r_i$ and $\Delta o_i > 0$, we have $o_i < r_i$. By equation (5.3), $o_i = r_i$ for $i \in \mathcal{G}_2$, hence $\mathcal{G} \cap \mathcal{G}_2 = \emptyset$, which further implies $\mathcal{G} \cap \mathcal{G}_1 = \emptyset$ as $\mathcal{G}_1 \subset \mathcal{G}_2$. Since $\mathcal{G} \cap \mathcal{G}_2 = \emptyset$, the equation (5.2) for $i \in \mathcal{G}$ simplifies to

$$\sum_{j \in [d]} q_{ji} o_j + \gamma_i = o_i \text{ for } i \in \mathcal{G} \tag{5.4}$$

Also, since $\mathcal{G} \cap \mathcal{G}_1 = \emptyset$, we have $o'_i = \min\{f_i, r_i\}$, which further implies

$$\sum_{j \in [d]} q_{ji} o'_j + \gamma_i \ge o'_i \text{ for } i \in \mathcal{G} \tag{5.5}$$

Subtracting the above two relation, with $o'_i = o_i + \Delta o_i$, yields:

$$\sum_{j \in [d]} q_{ji} \Delta o_j \ge \Delta o_i \text{ for } i \in \mathcal{G} \tag{5.6}$$

Sum the above inequality over $i \in \mathcal{G}$ and by some elementary algebra, we reach

$$\sum_{j \notin \mathcal{G}} \big(\sum_{i \in \mathcal{G}} q_{ji}\big)\Delta o_j \geq \sum_{j \in \mathcal{G}} \big(1 - \sum_{i \in \mathcal{G}} q_{ji}\big)\Delta o_j \tag{5.7}$$

Note that the Δo_j's on RHS are strictly greater zero, and Δo_j's on LHS are non-positive, so the only way the above inequality holds if

- $\sum_{i \in \mathcal{G}} q_{ji} = 1$ for $j \in \mathcal{G}$,
- Each term $q_{ji}\Delta o_j = 0$ for $j \notin \mathcal{G}$ and $i \in \mathcal{G}$,
- All inequality in (5.6) is an equality

This further implies all inequality in (5.6) could be rewritten as

$$\sum_{j \in [d]} q_{ji}\Delta o_j = \sum_{j \in \mathcal{G}} q_{ji}\Delta o_j + \sum_{j \notin \mathcal{G}} q_{ji}\Delta o_j = \sum_{j \in \mathcal{G}} q_{ji}\Delta o_j = \Delta o_i \text{ for } i \in \mathcal{G} \tag{5.8}$$

Let $Q_1 = [q_{ij}]_{i,j \in \mathcal{G}}$ and $Q_2 = [q_{ij}]_{i,j \notin \mathcal{G}}, Q_3 = [q_{ij}]_{i \in \mathcal{G}, j \notin \mathcal{G}}$ and $Q_4 = [q_{ij}]_{i \notin \mathcal{G}, j \in \mathcal{G}}$. On the one hand, the last equality above is equivalent to the homogeneous linear system $(I - Q_1^\intercal)x = 0$, which has a nonzero solution. On the other hand, $\sum_{i \in \mathcal{G}} q_{ji} = 1$ for $j \in \mathcal{G}$ implies that Q_3 is a zero matrix. So putting entries in Q in a proper way, we have

$$\mathcal{Q} = I - Q^\intercal = \begin{bmatrix} I - Q_1^\intercal & Q_4^\intercal \\ 0 & I - Q_2^\intercal \end{bmatrix}$$

Since $\mathcal{Q}$ is invertible, so is $I - Q_1^\intercal$, and we reach a contradiction. $\square$

Corollary 5.1. Given $\boldsymbol{b}$, there exists a unique pair $(\boldsymbol{o}, \boldsymbol{f})$ that satisfies equations (5.2) and (5.3).

Proof. The existence of the solution follows the dynamic of DF$(\boldsymbol{\gamma})$. For uniqueness, suppose $\boldsymbol{o}$ and $\boldsymbol{o}'$ are two solutions of (5.2) and (5.3), Lemma 5.1 implies $\boldsymbol{o} \geq \boldsymbol{o}'$ and $\boldsymbol{o} \leq \boldsymbol{o}'$ coordinately, hence $\boldsymbol{o} = \boldsymbol{o}'$. This further implies $\boldsymbol{f} = \boldsymbol{f}'$ according to (5.2). $\square$

We introduce the notation $\boldsymbol{o}^{\boldsymbol{\xi}}(t), \boldsymbol{f}^{\boldsymbol{\xi}}(t), \boldsymbol{b}^{\boldsymbol{\xi}}(t)$ with $t \in [0, T]$ to represent the transient outflow, the inflow rate, and the buffer content of FN$(\boldsymbol{\xi})$ over time, respectively. If $\boldsymbol{\xi}$ takes the form in (5.1), then the potential buffer content at node i, which is formed by imagining the fluid flow out by the maximum rate $\boldsymbol{r}$ at all nodes over time, satisfies

$$(\tilde{\boldsymbol{b}}^{\boldsymbol{\xi}})^{(i)}(t) = \gamma_i t + \sum_{j=1}^{n(i)} x_{ij} \mathbb{1}_{\{t_{ij} \leq t\}} - \big(\mathcal{Q}\boldsymbol{r}\big)_i t.$$

Hence $\tilde{\boldsymbol{b}}^{\boldsymbol{\xi}}(t) = \boldsymbol{\xi}(t)$. On the one hand, a similar argument as in Proposition 3.2 implies that the actual buffer content $\boldsymbol{b}^{\boldsymbol{\xi}}$ equals $\phi(\boldsymbol{\xi})$, where $\phi(\cdot)$ is the content component of the reflection mapping (see Definition 6.1). On the other hand, the buffer content $\boldsymbol{b}^{\boldsymbol{\xi}}(T)$ is the accumulation of $\boldsymbol{f}^{\boldsymbol{\xi}}(t) - \boldsymbol{o}^{\boldsymbol{\xi}}(t)$ over time. Therefore,

$$\pi \circ \phi(\boldsymbol{\xi}) = \boldsymbol{b}^{\boldsymbol{\xi}}(T) = \int_0^T (\boldsymbol{f}^{\boldsymbol{\xi}}(t) - \boldsymbol{o}^{\boldsymbol{\xi}}(t))dt. \tag{5.9}$$

With all the preparations above, we can end this section with the proof of Lemma 3.1.

Proof of Lemma 3.1. We prove the three statements of the lemma sequentially.

(1) Consider $\boldsymbol{\xi} \in \mathbb{S}_{\mathcal{K}}$ that takes the form (5.1). Let t^* be a real number such that $0 < t^* < \min_{i\in\mathcal{K}}\big(\min_{1\le j\le n(i)} t_{ij}\big)$ and y^* be a positive real number. Define $\boldsymbol{\eta} \in \mathbb{S}^1_{\mathcal{K}}$ with

$$\eta_i = \begin{cases} y^* \mathbb{1}_{[t^*,T]}(t) + (\boldsymbol{\gamma} - \mathcal{Q}\boldsymbol{r})_i \cdot t & i \in \mathcal{K} \\ (\boldsymbol{\gamma} - \mathcal{Q}\boldsymbol{r})_i \cdot t & i \notin \mathcal{K} \end{cases}. \tag{5.10}$$

We prove this part by showing that $\pi \circ \phi(\boldsymbol{\xi}) \le \pi \circ \phi(\boldsymbol{\eta})$ coordinate-wise for large enough y^*. Since both $\mathrm{DF}(\boldsymbol{\xi})$ and $\mathrm{DF}(\boldsymbol{\eta})$ are defined on finite time horizon $[0,T]$, and by the definition of t^*, we can select y^* large enough such that the buffer content of $\mathrm{DF}(\boldsymbol{\eta})$ is no smaller than that of $\mathrm{DF}(\boldsymbol{\xi})$ at at nodes $i \in \mathcal{K}$ at any time $t \in [0,T]$. By relation (5.9),

$$\big(\pi \circ \phi(\boldsymbol{\xi})\big)_i \le \big(\pi \circ \phi(\boldsymbol{\eta})\big)_i \text{ for } i \in \mathcal{K}. \tag{5.11}$$

The moments that can potentially lead to changes in the transient inflow and outflow rate of $\mathrm{DF}(\boldsymbol{\xi})$ and $\mathrm{DF}(\boldsymbol{\eta})$ occur when large bulks of fluid join them. For $\mathrm{DF}(\boldsymbol{\xi})$, those moments are given by $\bigcup_{i\in\mathcal{K}}\{t_{ij}, 1 \le j \le n(i)\}$ and we can order those timestamps in increasing order as $t_{(0)}(= t^*), t_{(1)}, t_{(2)}, \cdots, t_{(K)}$. On each $[t_{(k-1)}, t_{(k)})$, the values of $\boldsymbol{f}^{\boldsymbol{\xi}}(t)$, $\boldsymbol{o}^{\boldsymbol{\xi}}(t)$, $\boldsymbol{f}^{\boldsymbol{\eta}}(t)$ and $\boldsymbol{o}^{\boldsymbol{\eta}}(t)$ remain constant. Therefore, for all $i \in [d]$,

$$\big(\pi \circ \phi(\boldsymbol{\xi})\big)_i - \big(\pi \circ \phi(\boldsymbol{\eta})\big)_i = \sum_{k=1}^{K} \int_{t_{(k-1)}}^{t_{(k)}} \big(f_i^{\boldsymbol{\xi}}(t) - \boldsymbol{o}_i^{\boldsymbol{\xi}}(t)\big) - \big(f_i^{\boldsymbol{\eta}}(t) - \boldsymbol{o}_i^{\boldsymbol{\eta}}(t)\big) dt. \tag{5.12}$$

For $\mathrm{DF}(\boldsymbol{\eta})$, the only moment with such change is at $t_{(0)}$. Combining with the relation (5.3), it is clear that

$$f_i^{\boldsymbol{\eta}}(t) - \boldsymbol{o}_i^{\boldsymbol{\eta}}(t) \ge 0 \text{ for } i \notin \mathcal{K} \text{ and } t \in [t_{(0)}, T]. \tag{5.13}$$

We claim that

$$\boldsymbol{b}_i^{\boldsymbol{\xi}}(t_{(j)}) \le \boldsymbol{b}_i^{\boldsymbol{\eta}}(t_{(j)}) \text{ for } i \notin \mathcal{K} \text{ and } j = 0, 1, \cdots, K. \tag{5.14}$$

This claim is trivial when $j = 0$. Now, we assume the claim holds for some $0 < j < K$. By Lemma 5.1, $o^{\boldsymbol{\xi}}(t_{(j)}) \le \boldsymbol{o}^{\boldsymbol{\eta}}(t_{(j)})$ coordinate-wise. Consider a node $i \notin \mathcal{K}$. If $o_i^{\boldsymbol{\xi}}(t_{(j)})$, then $o_i^{\boldsymbol{\eta}}(t_{(j)}) = r_i$ as well. Thus,

$$f_i^{\boldsymbol{\xi}}(t_{(j)}) - \boldsymbol{o}_i^{\boldsymbol{\xi}}(t_{(j)}) = (Q^\top \boldsymbol{o}^{\boldsymbol{\xi}}(t_{(j)}))_i + \gamma_i - r_i \le (Q^\top \boldsymbol{o}^{\boldsymbol{\eta}}(t_{(j)}))_i + \gamma_i - r_i = f_i^{\boldsymbol{\eta}}(t_{(0)}) - \boldsymbol{o}_i^{\boldsymbol{\eta}}(t_{(j)});$$

If $o_i^{\boldsymbol{\xi}}(t_{(j)}) < r_i$, then $f_i^{\boldsymbol{\xi}}(t_{(j)}) - \boldsymbol{o}_i^{\boldsymbol{\xi}}(t_{(j)}) = 0$ by (5.3). Combined with (5.13), this confirms $f_i^{\boldsymbol{\xi}}(t_{(j)}) - \boldsymbol{o}_i^{\boldsymbol{\xi}}(t_{(j)}) \le f_i^{\boldsymbol{\eta}}(t_{(j)}) - \boldsymbol{o}_i^{\boldsymbol{\eta}}(t_{(j)})$.

Since $\boldsymbol{f}^{\boldsymbol{\xi}}(\cdot), \boldsymbol{o}^{\boldsymbol{\xi}}(\cdot), \boldsymbol{f}^{\boldsymbol{\eta}}(\cdot), \boldsymbol{o}^{\boldsymbol{\eta}}(\cdot)$ remain constant on $t \in [t_{(j)}, t_{(j+1)})$, it follows that

$$\begin{aligned} \boldsymbol{b}_i^{\boldsymbol{\xi}}(t_{(j+1)}) =& \boldsymbol{b}_i^{\boldsymbol{\xi}}(t_{(j)}) + \int_{t_{(j)}}^{t_{(j+1)}} \big(f_i^{\boldsymbol{\xi}}(t) - \boldsymbol{o}_i^{\boldsymbol{\xi}}(t)\big) dt \\ \le& \boldsymbol{b}_i^{\boldsymbol{\eta}}(t_{(j)}) + \int_{t_{(j)}}^{t_{(j+1)}} \big(f_i^{\boldsymbol{\eta}}(t) - \boldsymbol{o}_i^{\boldsymbol{\eta}}(t)\big) dt = \boldsymbol{b}_i^{\boldsymbol{\eta}}(t_{(j+1)}) \end{aligned} \tag{5.15}$$

for $i \notin \mathcal{K}$. Thus, the claim in (5.14) also holds for $j + 1$.

By mathematical induction, the claim (5.14) is proven. This implies

$$\big(\pi \circ \phi(\boldsymbol{\xi})\big)_i = \boldsymbol{b}_i^{\boldsymbol{\xi}}(t_{(K)}) \le \boldsymbol{b}_i^{\boldsymbol{\eta}}(t_{(K)}) = \big(\pi \circ \phi(\boldsymbol{\eta})\big)_i \text{ for } i \notin \mathcal{K}. \tag{5.16}$$

Combining this with (5.11) completes the proof of this part.

(2) Let i^* be some nodes in $\mathcal{K} \cap \{i \in [d] : h_i > 0\}$ and consider $\boldsymbol{\eta} \in \mathbb{S}^1_{\mathcal{K}}$ in the form of (5.10). Since $\pi \circ \phi(\boldsymbol{\eta}) = \boldsymbol{b}^{\boldsymbol{\eta}}(T)$, $\big(\pi \circ \phi(\boldsymbol{\xi})\big)_{i^*} \geq x_{i^*} - r_{i^*}T$, we have

$$\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) \geq h_{i^*} \cdot (y_{i^*} - r_{i^*}T).$$

As $h_{i^*} > 0$, we can choose y_{i^*} to make $\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi})$ arbitrarily large.

(3) According to the argument in part (1), for any $\boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}$, we can find $\boldsymbol{\eta} \in \mathbb{S}^1_{\mathcal{K}}$ taking the form (5.10) with t^* small enough and y^* large enough such that $\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) \leq \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\eta})$ and

$$\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\eta}) = \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\boldsymbol{\eta}}(t^*) - \boldsymbol{o}^{\boldsymbol{\eta}}(t^*)) \cdot (T - t^*).$$

By definition, $\boldsymbol{f}^{\boldsymbol{\eta}}(t^*)$ and $\boldsymbol{o}^{\boldsymbol{\eta}}(t^*)$ is the solution of (3.11) and (3.12) with respect to $\mathcal{B} = \{i \in [d] : b_i^{\boldsymbol{\eta}}(t^*) > 0\} = \mathcal{K}$. Hence $\boldsymbol{f}^{\boldsymbol{\eta}}(t^*) = \boldsymbol{f}^{\mathcal{K}}$ and $\boldsymbol{o}^{\boldsymbol{\eta}}(t^*) = \boldsymbol{o}^{\mathcal{K}}$. Therefore,

$$\sup\{\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}), \boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}\} \geq \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot (T - t^*).$$

Taking $t^* \to 0$ yields $\sup\{\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}), \boldsymbol{\xi} \in \mathbb{S}^1_{\mathcal{K}}\} \geq \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T$.

However, there is no path in $\mathbb{S}^1_{\mathcal{K}}$ that can have $\boldsymbol{h}^\intercal \cdot \pi \circ \phi(\cdot)$ greater than or equal to $\boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T$. Suppose such path $\boldsymbol{\xi}$ exists, then there exists $\boldsymbol{\eta} \in \mathbb{S}^1_{\mathcal{K}}$ taking the form (5.10) such that

$$\boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot T \leq \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\xi}) \leq \boldsymbol{h}^\intercal \cdot \pi \circ \phi(\boldsymbol{\eta}) = \boldsymbol{h}^\intercal \cdot (\boldsymbol{f}^{\mathcal{K}} - \boldsymbol{o}^{\mathcal{K}}) \cdot (T - t^*).$$

This requires $t^* = 0$, which makes $\boldsymbol{\eta} \notin \mathbb{S}^1_{\mathcal{K}}$, a contradiction. Therefore, the conclusion of this part is proved.

□

5.2 Proofs in section 4

Lemma 5.2. For $k \in \mathbb{N}$, the set $L_{\leq k} = \{\xi \in \check{\mathbb{C}}^1[0,1], I_i'(\xi) \leq \lambda \cdot k\}$ is closed w.r.t the J_1 topology.

Proof. The paths in $\check{\mathbb{C}}^1[0,1]$ are flat on disjoint intervals as subset of $(0,1]$, and increasing with rate 1 between those flat intervals. $L_{leqslantk}$ contains such path with the number of flat intervals smaller or equal to k. Apparently, $L_{\leqslant 0}$ has a single element ζ_1, hence it is trivially closed.

Now let's assume for $k \leq K$, the sets $L_{\leqslant k}$'s are closed. Below we will show $L_{\leqslant K+1}$ is closed as well, and the lemma is proved by mathematical induction.

Let $\xi_n, n \in \mathbb{N}$ be paths in $L_{\leqslant K+1}$, and for some $\xi \in \mathbb{D}[0,1]$ we have $d_{J_1}(\xi_n, \xi) \to 0$ as $n \to \infty$. We will show $\xi \in L_{\leqslant K+1}$. WLOG, we can assume $\xi_n \in L_{=K+1}$ for $n \in \mathbb{N}$ ($L_{=K+1} = L_{\leqslant K+1} \backslash L_{\leqslant k}$). This is because if there are infinite ξ_n's belonging to $L_{\leqslant K}$, then the closeness of $L_{\leqslant K}$, as we have assumed, implies that $\xi \in L_{\leqslant K}$, and the proof is finished.

Now for $n \in \mathbb{N}$, let $j_1^n, j_2^n, \cdots, j_{2K+1}^n, j_{2K+2}^n$ be the ordered consecutive endpoints of intervals on which ξ_n is flat. In other words, ξ_n is flat on the intervals $[j_{2k-1}^n, j_{2k}^n]$ for $k = 1, \cdots K+1$ and is increasing with rate 1 between those intervals. Since all the j_k^n's fall into the compact interval $[0,1]$, by applying the Bolzano–Weierstrass theorem multiple times, we can find index set $\{n_m, m \in \mathbb{N}\}$ as a subset of $\mathbb{N}$, and for each $k \in [2K+2]$, there exists j_k such that $\lim_{m\to\infty} j_k^{n_m} = j_k$. Let ξ' be the path in $\check{\mathbb{C}}^1[0,1]$ such that it is flat on $[j_{2k-1}, j_{2k}]$ for $k \in [K+1]$. Note that ξ' belongs $L_{\leqslant K+1}$ but not necessarily in $L_{=K+1}$. This is because it is likely that $j_{2k-1} = j_{2k}$ for some k, then the flat interval $[j_{2k-1}, j_{2k}]$ degenerates to a point. Or $j_{2k} = j_{2k+1}$ for some k, then the two consecutive intervals with endpoints j_{2k} and j_{2k+1} merge to one flat interval.

For any fixed $\epsilon > 0$, we can find M large enough such that for $m > M$ and $k \in [2K+2]$,

$$|j_k^{n_m} - j_k| \leq \epsilon$$

Based on the form of ξ_{n_m} and ξ', the discrepancy of $j_k^{n_m}$ and j_k would at most incur $|j_k^{n_m} - j_k|$ uniform gap between ξ^{n_m} and ξ', hence

$$\|\xi^{n_m} - \xi'\| \leq (2K+2) \cdot \epsilon$$

Since ϵ is arbitrary, we have shown $\|\xi^{n_m} - \xi'\| \to 0$ as $m \to 0$, and this further implies ξ^{n_m} converge to ξ' w.r.t the J_1 topology. As the limit of the subsequence $\{\xi^{n_m}, m \in \mathbb{N}\}$, ξ' is also the limit of the sequence $\{\xi^n, m \in \mathbb{N}\}$, thus $\xi = \xi'$, and this proves $\xi \in L_{\leqslant K+1}$.

□

Lemma 5.3. Let Φ_μ be the mapping defined in (4.5). If there are two paths $\eta, \eta' \in \mathbb{D}[0, \frac{1}{\mu}]$, and there is a $\rho \in \Lambda[0, \frac{1}{\mu}]$ such that $\|\rho - e\| < \epsilon$ and $\eta = \eta' \circ \rho$, then $\|\Phi_\mu(\eta) - \Phi_\mu(\eta')\| \leq \epsilon$.

Proof. By the property of ρ, η' and η have the same range, hence $\Psi(\eta)(\frac{1}{\mu}) = \Psi(\eta')(\frac{1}{\mu})$. Let us denote this common value as t^*.

For $t < t^* \wedge 1$, it is clear that both $\varphi_\mu(\eta)(t)$ and $\varphi_\mu(\eta')(t)$ are smaller than or equal to $\frac{1}{\mu}$, hence $\Phi_\mu(\eta)(t) = \varphi_\mu(\eta)(t)$ and $\Phi_\mu(\eta')(t) = \varphi_\mu(\eta')(t)$. For such t,

$$\begin{aligned}\varphi_\mu(\eta')(t) =& \inf\{s \in [0, \frac{1}{\mu}] : \eta'(s) > t\} \\ =& \inf\{s \in [0, \frac{1}{\mu}] : \eta \circ \rho^{-1}(s) > t\} \\ =& \inf\{\rho(s) \in [0, \frac{1}{\mu}] : \eta(s) > t\}\end{aligned}$$

while

$$\varphi_\mu(\eta)(t) = \inf\{s \in [0, \frac{1}{\mu}] : \eta(s) > t\}$$

since $\|\rho - e\| \leq \epsilon$, by standard analysis, $|\varphi_\mu(\eta')(t) - \varphi_\mu(\eta')(t)| \leq \epsilon$. This yields $|\Phi\mu(\eta')(t) - \Phi\mu(\eta')(t)| \leq \epsilon$ for $t < t^* \wedge 1$.

For $t^* \leq t \leq 1$, we have $\varphi_\mu(\eta)(t) = \varphi_\mu(\eta')(t) = \infty$ while $\psi_\mu(\eta)(t) = \psi_\mu(\eta')(t) = \frac{1}{\mu}[1 + (t - t^*)]$. Hence $\Phi\mu(\eta')(t) = \Phi\mu(\eta')(t)$ for $t^* \leq t \leq 1$.

Combining the above two cases for t reaches the conclusion of the lemma. □

Proof of Lemma 4.2. We would prove a stronger statement: $\|\Phi_\mu(\xi) - \Phi_\mu(\eta)\| \leq (2 + 4/\mu) \cdot d_{J_1}(\xi, \eta)$. Suppose $d_{J_1}(\xi, \eta) = \epsilon$, then exists $\delta < 1$, $\eta' \in \mathbb{D}[0, \frac{1}{\mu}]$, and a $\rho \in \Lambda[0, \frac{1}{\mu}]$ such that

$$\|\rho\| \leq \epsilon(1+\delta) < 2\epsilon, \quad \eta = \eta' \circ \rho, \quad \|\xi - \eta'\| \leq (1+\delta)\epsilon \tag{5.17}$$

By lemma 5.3, we have $\|\Phi_\mu(\eta) - \Phi_\mu(\eta')\| \leq 2\epsilon$, and it would be enough to prove $\|\Phi_\mu(\xi) - \Phi_\mu(\eta')\| \leq 4\epsilon/\mu$.

For $t \in [0, 1]$, $u = \varphi_\mu(\xi)(t)$ can take value in $[0, 1/\mu]$ and ∞(when the infimum is taken w.r.t the empty set). Note that when $u \in [0, \frac{1}{\mu}]$, t and u should satisfy the following relation

$$\xi(u-) < t \leq \xi(u) \tag{5.18}$$

and

$$\Phi_\mu(\xi)(t) = \varphi_\mu(\xi)(t) \tag{5.19}$$

We will separate the discussion to the following three cases based on the value of u:

- **Case 1**: $u = \varphi_\mu(\xi)(t) < \frac{1}{\mu} - \frac{2\epsilon}{\mu}$. Due to the relation (5.18), along with the special form of $\xi \in \tilde{\mathbb{D}}^{\mu}_{<\infty}$, and the assumption of $\|\xi - \eta'\| \leq (1+\delta)\epsilon$, we have

$$\eta'(u') \leq \xi(u-) - \mu \cdot \frac{2\epsilon}{\mu} + (1+\delta)\epsilon < \xi(u-) < t \text{ for } u' \in [0, u - \frac{2\epsilon}{\mu}] \tag{5.20}$$

and

$$\eta'(u') \geq \xi(u) + \mu \cdot \frac{2\epsilon}{\mu} - (1+\delta)\epsilon > \xi(u) \geq t \text{ for } u' \in [u + \frac{2\epsilon}{\mu}, \frac{1}{\mu}] \tag{5.21}$$

The above inequalities (5.21) and (5.20) show $\varphi_\mu(\eta')(t) \geq u - \frac{2\epsilon}{\mu}$ and $\varphi_\mu(\eta')(t) \leq u + \frac{2\epsilon}{\mu} < \frac{1}{\mu}$, respectively. Since $\psi_\mu(\eta')$ takes value no smaller than $\frac{1}{\mu}$, and $\Phi_\mu(\eta') = \varphi_\mu(\eta') \wedge \psi_\mu(\eta')$, we have

$$u - \frac{2\epsilon}{\mu} \leq \varphi_\mu(\eta') = \Phi_\mu(\eta') \leq u + \frac{2\epsilon}{\mu}$$

Combine with (5.19), we conclude that for t in this case,

$$|\Phi_\mu(\xi)(t) - \Phi_\mu(\eta')(t)| \leq \frac{2\epsilon}{\mu} \tag{5.22}$$

Note that in (5.20), it is possible that $u - \frac{2\epsilon}{\mu} \leq 0$, hence the interval for u' degenerates. However, $\varphi_\mu(\eta')(t) \geq 0 \geq u - \frac{2\epsilon}{\mu}$ would be trivial, hence the same argument holds.

- **Case 2**: $\frac{1}{\mu} - \frac{2\epsilon}{\mu} \leq u = \varphi_\mu(\xi)(t) \leq \frac{1}{\mu}$. We begin with a lower bound for $\Phi_\mu(\eta')(t)$:

$$\Phi_\mu(\eta')(t) \geq \varphi_\mu(\eta')(t) \geq u - \frac{2\epsilon}{\mu} = \Phi_\mu(\xi)(t) - \frac{2\epsilon}{\mu} \tag{5.23}$$

In the above, the first inequality is a property of Φ_μ. The second inequality is a consequence of (5.20), which also holds in case 2 following the same argument as in case 1. The third equality is due to (5.19).

Then we seek an upper bound for $\Phi_\mu(\eta')(t)$: with ξ being non-decreasing and $\|\xi - \eta'\| \leq (1+\delta)\epsilon$, and the relation (5.18), the term $\Psi(\eta')(\frac{1}{\mu})$ satisfies the following

$$\Psi(\eta')(\frac{1}{\mu}) \geq \eta'(\frac{1}{\mu}) \geq \xi(\frac{1}{\mu}) - (1+\delta)\epsilon \geq \xi(u) - (1+\delta)\epsilon \geq t - (1+\delta)\epsilon \tag{5.24}$$

The above lower bound for $\Psi(\eta')(\frac{1}{\mu})$ implies

$$\Phi_\mu(\eta')(t) \leq \psi_\mu(\eta')(t) = \frac{1}{\mu}\big[1 + [t - \Psi(\eta')(\frac{1}{\mu})]_+\big] \leq \frac{1}{\mu} + \frac{(1+\delta)\epsilon}{\mu} \tag{5.25}$$

With the case assumption $\Phi_\mu(\xi)(t) = \varphi_\mu(\xi)(t) = u \geq \frac{1}{\mu} - \frac{2\epsilon}{\mu}$, the above inequality can further yield

$$\Phi_\mu(\eta')(t) - \Phi_\mu(\xi)(t) \leq \frac{2+1+\delta}{\mu}\epsilon \tag{5.26}$$

In view of the bounds (5.26) and (5.23) together, we conclude that for t in this case,

$$|\Phi_\mu(\xi)(t) - \Phi_\mu(\eta')(t)| \leq \frac{2+1+\delta}{\mu}\epsilon \leq \frac{4\epsilon}{\mu} \tag{5.27}$$

- **Case 3**: $u = \varphi_\mu(\xi)(t) = \infty$. Indeed, since $\xi \in \tilde{\mathbb{D}}^{\mu}_{<\infty}[0, \frac{1}{\mu}]$, this last case is only possible when $\xi = \zeta_\mu$ and $t = 1$. On one hand, following the same argument for (5.24) and (5.25) in case 2, we can yield the same upper bound for $\Phi_\mu(\eta')(t)$ and $t = 1$:

$$\Phi_\mu(\eta')(1) \le \frac{1}{\mu} + \frac{(1+\delta)\epsilon}{\mu} \le \frac{1}{\mu} + \frac{2\epsilon}{\mu} \tag{5.28}$$

On the other hand, similar to the inequality (5.20) in case 1, one can show for $u' \in [0, \frac{1}{\mu} - \frac{2\epsilon}{\mu}]$

$$\eta'(u') \le \xi(u') + (1+\delta)\epsilon \le \xi(\frac{1}{\mu}) - \mu \cdot \frac{2\epsilon}{\mu} + (1+\delta)\epsilon = 1 - (2 - 1 - \delta)\epsilon < 1 \tag{5.29}$$

which implies $\varphi_\mu(\eta')(1) \ge \frac{1}{\mu} - \frac{2\epsilon}{\mu}$, thus

$$\Phi_\mu(\eta')(1) \ge \frac{1}{\mu} - \frac{2\epsilon}{\mu} \tag{5.30}$$

Combining the bounds (5.28) and (5.30), and the fact $\Phi_\mu(\xi)(1) = \frac{1}{\mu}$, we have

$$|\Phi_\mu(\xi)(1) - \Phi_\mu(\eta')(1)| \le \frac{2\epsilon}{\mu} \tag{5.31}$$

To summarize, we have obtained inequalities (5.22)(5.27)(5.31) for case 1, case 2 and case 3, respectively, from which we conclude that for $t \in [0,1]$, $|\Phi_\mu(\xi)(t) - \Phi_\mu(\eta')(t)| \le 4\epsilon/\mu$. This finishes the proof. $\square$

Lemma 5.4. For $\xi \in \check{\mathbb{C}}^1[0,1]$ and $\eta \in \mathbb{D}[0,1]$, we have

$$\|\xi - \eta\| \le 2 \cdot d_{J_1}(\xi, \eta)$$

Proof. Assume $d_{J_1}(\xi, \eta) = \epsilon$, then arbitrary $\delta > 0$, there exists $\rho \in \Lambda[0,1]$ such that

$$\|\rho - e\| \vee \|\xi - \eta \circ \rho\| \le (1+\delta)\epsilon \tag{5.32}$$

For each $t \in [0,1]$,

$$|\xi(t) - \eta(t)| \le |\xi(t) - \xi(\rho(t))| + |\xi(\rho(t)) - \eta(t)| \le 2(1+\delta)\epsilon$$

In the above deduction, we use $\xi \in \check{\mathbb{C}}^1[0,1]$, hence the gap $|\xi(t) - \xi(\rho(t))|$ cannot exceed $(1+\delta)\epsilon$. Since δ is arbitrarily small, we reach the conclusion $\|\xi - \eta\| \le 2\epsilon$.

$\square$

Lemma 5.5. Let $f : \Pi_{i=1}^d \mathbb{D}[0,1] \to \mathbb{R}$ be defined as in (4.22). Then for $\boldsymbol{\xi} \in \Pi_{i=1}^d \check{\mathbb{C}}^1[0,1]$ and $\boldsymbol{\eta} \in \Pi_{i=1}^d \mathbb{D}[0,1]$, the following inequality holds:

$$\big|f(\boldsymbol{\xi}) - f(\boldsymbol{\eta})\big| \le 4d \cdot d_p\big(\boldsymbol{\xi}, \boldsymbol{\eta}\big)$$

Proof. Let $d_p(\boldsymbol{\xi}, \boldsymbol{\eta}) = \epsilon$, then for each $i \in [d]$, $d_{J_1}(\xi_i, \eta_i) \le \epsilon$. Since $\xi_i \in \check{\mathbb{C}}^1[0,1]$, Lemma 5.4 implies that $\|\xi_i - \eta_i\| \le 2\epsilon$. Hence

$$\begin{aligned} |f(\boldsymbol{\xi}) - f(\boldsymbol{\eta})| &\le \sup_{s \in [0,1]} \Big\{ \Big| \frac{1-s}{\mathbb{E}[A]} - \sum_{i=1}^d \big(\xi_i(1) - \xi_i(s)\big) - \frac{1-s}{\mathbb{E}[A]} + \sum_{i=1}^d \big(\eta_i(1) - \eta_i(s)\big) \Big| \Big\} \\ &\le \sum_{i=1}^d |\xi_i(1) - \eta_i(1)| + \sum_{i=1}^d \|\xi_i - \eta_i\| \le 4d\epsilon \end{aligned}$$

This finishes the proof.

$\square$

6 Appendix

6.1 Some large-deviations theory results

In this appendix, we include some important results in the field of large deviations that facilitate the establishment of full/extended LDPs.

Lemma 6.1. Let $\{X_n\}_{n\geq 1}$ be a sequence of stochastic processes on $(\mathcal{X}, d)$. Let $\mathcal{E} \subset \mathcal{X}$ be a closed set such that $\mathbb{P}(X_n \in \mathcal{E}) = 1$ for all $n \in \mathbb{N}$. Then

(1) If $\{X_n\}_{n\geq 1}$ on $(\mathcal{X}, d)$ satisfies an extended LDP with and rate function I (and some speed) and if $\{x \in \mathcal{X} : I(x) < \infty\} \subseteq \mathcal{E}$, then $\{X_n\}_{n\geq 1}$ on $(\mathcal{E}, d)$ satisfies the extended LDP with rate function $I' = I|_{\mathcal{E}}$ (and the same speed). Here $I|_{\mathcal{E}}$ is restriction of I on set $\mathcal{E}$.

(2) If $\{X_n\}_{n\geq 1}$ on $(\mathcal{E}, d)$ satisfies the extended LDP with the rate function I (and some speed). Then $\{X_n\}_{n\geq 1}$ on $(\mathcal{X}, d)$ satisfies the extended LDP with the rate function

$$I'(x) = \begin{cases} I(x) & x \in \mathcal{E} \\ \infty & x \notin \mathcal{E} \end{cases}$$

(and the same speed).

Proof. Lemma B.1 of [4] gives the proof of the first part. We omit the second part's proof as it is straightforward by checking that extended LDP's upper and lower bounds are satisfied. □

The Contraction Principles are a set of results that aim to infer the large deviation properties of $\{f(X_n)\}_{n\geq 1}$, given those of$\{X_n\}_{n\geq 1}$. Theorem 4.2.1 (along with the accompanying comments) in [10] addresses the full LDP case. We present it below as a lemma, adapted for random objects:

Lemma 6.2. Let $(\mathcal{X}, d)$ and $(\mathcal{X}', d')$ be two metric spaces, and let $\{X_n\}_{n\geq 1}$ be a sequence of random objects taking value in $(\mathcal{X}, d)$. Consider a good rate function $I : \mathcal{X} \to \mathbb{R}_+$ and a mapping $f : (\mathcal{X}, d) \to (\mathcal{X}', d')$ that is continuous at the *effective domain* $\mathcal{D}_I \triangleq \{x \in \mathcal{X} : I(x) < \infty\}$. For each $y \in \mathcal{X}'$, define

$$I'(\eta) = \inf\{I(\xi) : \eta = f(\xi), \xi \in \mathcal{X}\}, \tag{6.1}$$

where the infimum over an empty set is taken as ∞. Then the following statements hold:

(1) I' is a good rate function.

(2) If I controls the full LDP of $\{X_n\}_{n\geq 1}$, then the function I' controls the full LDP of $\{f(X_n)\}_{n\geq 1}$.

The following is a variation of the contraction principle, tailored for the extended LDP case.

Lemma 6.3. Let $(\mathcal{X}, d)$ and $(\mathcal{X}', d')$ be two metric spaces, and let $\{X_n\}_{n\geq 1}$ be a sequence of random objects taking value in $(\mathcal{X}, d)$. Additionally, let $I : \mathcal{X} \to \mathbb{R}_+$ be a rate function. Assume that $f : (\mathcal{X}, d) \to (\mathcal{X}', d')$ satisfies

$$d'\big(f(x), f(y)\big) \leq C \cdot d\big(x, y\big)$$

for some $C > 0$, for all point $x \in \mathcal{D}_I$ and $y \in \mathcal{X}$. Define

$$I'(\eta) = \inf\{I(\xi) : \eta = f(\xi), \xi \in \mathcal{X}\}, \tag{6.2}$$

where the infimum over an empty set is taken as ∞. Then the following statements hold:

(1) If I controls the extended LDP of $\{X_n\}_{n\geq 1}$, then the function I' controls the extended LDP of $\{f(X_n)\}_{n\geq 1}$.

(2) If, addtionally, I' is a rate function(i.e., lower semicontinuous), then $\{f(X_n)\}_{n\geq 1}$ satisfies the extended LDP with rate function I'.

Proof of Lemma 6.3. The statement (2) is direct by definition of extended LDP. To show statement (1), we firstly show the lower bound. For $x' \in A^\circ$ and any $x \in f^{-1}(x') \cap \mathcal{D}_I$ it satisfies

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(X_n \in f^{-1}(A)\right)}{a_n} \geq - \inf_{y \in \left(f^{-1}(A)\right)^\circ} I(y) \geq -I(x) \tag{6.3}$$

The first inequality above holds because I controls the extended LDP of $\{X_n\}_{n\geq 1}$ lower bound. The second inequality above holds because $x \in \left(f^{-1}(A)\right)^\circ$. To see this, since $x' \in A^\circ$, there exists δ such that $B(x', \delta) \subset A$. By assumption of f, we have $\{f(y), y \in B(x, \delta/C)\} \subset B(x', \delta) \subset A$, hence $B(x, \delta/C) \subset f^{-1}(A)$, which verifies x is an interior point of $f^{-1}(A)$. Now we can take supremum on the RHS of (6.3) among all $x \in f^{-1}(x') \cap \mathcal{D}_I$ and yield

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(f(X_n) \in A\right)}{a_n} \geq -\inf\{I(x); f(x) = x'\} = -I'(x')$$

and the lower bound of extended LDP is obtained by taking the supremum across $x' \in A^\circ$.

For the upper bound: We claim $\mathcal{D}_I \cap \left(f^{-1}(A)\right)^\epsilon \subset f^{-1}\left(A^{C\cdot\epsilon}\right)$. To see this, suppose $A \subset \mathcal{X}'$ and $x \in \mathcal{D}_I \cap \left(f^{-1}(A)\right)^\epsilon$. This means for arbitrary small $\delta > 0$, there exists $y \in f^{-1}(A)$ such that $d(x, y) < \epsilon + \delta$. By the assumption on f, we have

$$d'\big(f(x), f(y)\big) \leq C \cdot d(x, y) \leq C(\epsilon + \delta)$$

Since δ is arbitrarily small and $f(y) \in A$, we have $f(x) \in A^{C\cdot\epsilon}$, or equivalently, $x \in f^{-1}\left(A^{C\cdot\epsilon}\right)$, and this implies the claim. With this claim, we have

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(f(X_n) \in A\right)}{a_n} &= \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(X_n \in f^{-1}(A)\right)}{a_n} \leq -\lim_{\epsilon\downarrow 0} \inf_{x \in \left(f^{-1}(A)\right)^\epsilon} I(x) \\
&= -\lim_{\epsilon\downarrow 0} \inf_{x \in \mathcal{D}_I \cap \left(f^{-1}(A)\right)^\epsilon} I(x) \leq -\lim_{\epsilon\downarrow 0} \inf_{x \in f^{-1}\left(A^{C\cdot\epsilon}\right)} I(x) \\
&= -\lim_{\epsilon\downarrow 0} \inf_{y \in A^{C\cdot\epsilon}} \inf\{I(x) : x = f^{-1}(y)\} = -\lim_{\epsilon\downarrow 0} \inf_{y \in A^\epsilon} I'(y)
\end{aligned}$$

In the first line above, the set $f^{-1}(A)$ is a Borel set, as we have assumed f is a measurable function. The inequality in the first line holds due to X_n's satisfies the extended LDP upper bound with I. □

Two sequences $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ on $(\mathcal{X}, d)$ are said to be exponential equivalent(with respect to the speed sequence $\{a_n\}_{n\geq 1}$) if, for any $\epsilon > 0$,

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(d(X_n, Y_n) \geqslant \epsilon\right)}{a_n} = -\infty$$

If $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ are exponential equivalent, they share the same full/extended LDP. This result is formalized in the following lemma: the first part corresponds to Theorem 4.2.13 in [10], and the second part corresponds to Lemma 2.1 in [6].

Lemma 6.4. Suppose two sequences $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ taking value in $(\mathcal{X}, d)$ are exponential equivalent, then

(1) If the sequences $\{X_n\}_{n\geq 1}$ satisfies full LDP with a good rate function I and speed sequence $\{a_n\}_{n\geq 1}$, then $\{Y_n\}_{n\geq 1}$ satisfies the full LDP with the same rate function I and speed sequence $\{a_n\}_{n\geq 1}$.

(2) If the sequences $\{X_n\}_{n\geq 1}$ satisfies extended LDP with rate function I and speed sequence $\{a_n\}_{n\geq 1}$, then $\{Y_n\}_{n\geq 1}$ satisfies the extended LDP with the same rate function I and speed sequence $\{a_n\}_{n\geq 1}$.

6.2 Some Mappings and their Properties

Recall $\mathbb{D}^{\uparrow}[0,T] \triangleq \{\xi \in \mathbb{D}[0,T] : \xi \text{ is non-decreasing}, \xi(0) = 0\}$:

Definition 6.1. [Definition 14.2.1 of [30]] Fix a reflection matrix $\mathcal{Q}$. For any $\boldsymbol{\xi} \in \prod_{i=1}^{d} \mathbb{D}[0,T]$, let the feasible regulator set be

$$\Xi(\boldsymbol{\xi}) \triangleq \left\{ \boldsymbol{\zeta} \in \prod_{i=1}^{d} \mathbb{D}^{\uparrow}[0,T] : \boldsymbol{\xi} + \mathcal{Q}\boldsymbol{\zeta} \geq 0 \right\}, \tag{6.4}$$

and let the reflection map be $\boldsymbol{R} \triangleq (\boldsymbol{\psi}, \boldsymbol{\phi}) : \prod_{i=1}^{d} \mathbb{D}[0,T] \to \prod_{i=1}^{d} \mathbb{D}^{\uparrow}[0,T] \times \prod_{i=1}^{d} \mathbb{D}[0,T]$. The regulator component $\boldsymbol{\psi}(\boldsymbol{\xi}) = (\psi_1(\boldsymbol{\xi}), \cdots, \psi_d(\boldsymbol{\xi}))$ is coordinately defined by

$$\psi_i(\boldsymbol{\xi})(t) = \inf\{\omega_i(t) : \omega \in \Xi(\boldsymbol{\xi})\} \quad \text{for } t \in [0,T],$$

and its content component $\boldsymbol{\phi}(\boldsymbol{\xi})$ is defined by

$$\boldsymbol{\phi}(\boldsymbol{\xi}) \triangleq \boldsymbol{\xi} + \mathcal{Q}\boldsymbol{\psi}(\boldsymbol{\xi}).$$

The reflection map can be characterized by the following complementarity property:

Theorem 6.1. [Theorem 14.2.3 of [30]] Given a function $\boldsymbol{x} \in \prod_{i=1}^{d} \mathbb{D}[0,T]$ and $\boldsymbol{y} \in \Psi(\boldsymbol{x})$. Let x_i stands for the i-th coordinate of $\boldsymbol{x}$ (similarly for $\boldsymbol{y}$ and $\boldsymbol{z}$), then $\boldsymbol{R}(\boldsymbol{x}) = (\boldsymbol{y}, \boldsymbol{z})$ if and only if $\boldsymbol{z} = \boldsymbol{x} + \mathcal{Q}\boldsymbol{y}$ and

$$\int_0^T z_i dy_i = 0 \text{ for } i = 1, \cdots, d.$$

Recall that for fixed $T > 0$ and some $\beta \in \mathbb{R}$, $\mathbb{D}^{\beta}[0,T] \triangleq \{\zeta \in \mathbb{D}[0,T] : \zeta(t) = \xi(t) + \beta \cdot t, \xi \in \mathbb{D}^{\uparrow}[0,T]\}$. The following theorem summarizes a property of the reflection map that will play an important role in our proof.

Theorem 6.2. [Theorem 2.1 of [4]] Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d) \in \mathbb{R}^d$. The reflection map $\boldsymbol{R} = (\psi, \phi)$ is Lipschitz continuous w.r.t. d_p on $\prod_{i=1}^{d} \mathbb{D}^{\beta_i}[0,T]$.

For some $\boldsymbol{\kappa} \in \mathbb{R}^d$, define the mapping $\Upsilon^{\boldsymbol{\kappa}}$: $\prod_{i=1}^{d} \mathbb{D}[0,T] \to \prod_{i=1}^{d} \mathbb{D}[0,T]$ such that $\Upsilon^{\boldsymbol{\kappa}}(\boldsymbol{\xi})(t) = \boldsymbol{\xi}(t) + \boldsymbol{\kappa} \cdot t$. We conclude this section with the following lemma regarding this mapping:

Lemma 6.5. [Lemma 2.2 of [4]] For any $\boldsymbol{\kappa} \in \mathbb{R}^d$,

i) $\Upsilon^{\boldsymbol{\kappa}}$ is Lipschitz continuous w.r.t. d_p,

ii) $\Upsilon^{\boldsymbol{\kappa}}$ is a homeomorphism.

References

[1] François Baccelli and Serguei Foss. "Moments and tails in monotone-separable stochastic networks". In: *Ann. Appl. Probab.* 14.2 (2004), pp. 612–650. ISSN: 1050-5164,2168-8737. DOI: 10.1214/105051604000000044. URL: https://doi.org/10.1214/105051604000000044.

[2] François Baccelli, Serguei Foss, and Marc Lelarge. "Tails in generalized Jackson networks with subexponential service-time distributions". In: *J. Appl. Probab.* 42.2 (2005), pp. 513–530. ISSN: 0021-9002,1475-6072. DOI: 10.1239/jap/1118777185. URL: https://doi.org/10.1239/jap/1118777185.

[3] François Baccelli, Marc Lelarge, and Serguei Foss. "Asymptotics of subexponential max plus networks: the stochastic event graph case". In: *Queueing Syst.* 46.1-2 (2004), pp. 75–96. ISSN: 0257-0130,1572-9443. DOI: 10.1023/B:QUES.0000021142.51241.76. URL: https://doi.org/10.1023/B:QUES.0000021142.51241.76.

[4] Mihail Bazhba, Chang-Han Rhee, and Bert Zwart. "Large deviations for stochastic fluid networks with Weibullian tails". In: *Queueing Syst.* 102.1-2 (2022), pp. 25–52. ISSN: 0257-0130. DOI: 10.1007/s11134-022-09865-5. URL: https://doi.org/10.1007/s11134-022-09865-5.

[5] Mihail Bazhba et al. "Queue length asymptotics for the multiple-server queue with heavy-tailed Weibull service times". In: *Queueing Syst.* 93.3-4 (2019), pp. 195–226. ISSN: 0257-0130. DOI: 10.1007/s11134-019-09640-z. URL: https://doi.org/10.1007/s11134-019-09640-z.

[6] Mihail Bazhba et al. "Sample path large deviations for Lévy processes and random walks with Weibull increments". In: *Ann. Appl. Probab.* 30.6 (2020), pp. 2695–2739. ISSN: 1050-5164. DOI: 10.1214/20-AAP1570. URL: https://doi.org/10.1214/20-AAP1570.

[7] A. A. Borovkov and A. A. Mogul'skiı̆. "On large deviation principles in metric spaces". In: *Sibirsk. Mat. Zh.* 51.6 (2010), pp. 1251–1269. ISSN: 0037-4474. DOI: 10.1007/s11202-010-0098-0. URL: https://doi.org/10.1007/s11202-010-0098-0.

[8] Onno J Boxma and Vincent Dumas. "Fluid queues with long-tailed activity period distributions". In: *Computer Communications* 21.17 (1998), pp. 1509–1529.

[9] Bohan Chen et al. "Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound Poisson processes". In: *Math. Oper. Res.* 44.3 (2019), pp. 919–942. ISSN: 0364-765X,1526-5471. DOI: 10.1287/moor.2018.0950. URL: https://doi.org/10.1287/moor.2018.0950.

[10] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications.* Vol. 38. Stochastic Modelling and Applied Probability. Corrected reprint of the second (1998) edition. Springer-Verlag, Berlin, 2010, pp. xvi+396. ISBN: 978-3-642-03310-0. DOI: 10.1007/978-3-642-03311-7. URL: https://doi.org/10.1007/978-3-642-03311-7.

[11] Jean-Dominique Deuschel and Daniel W. Stroock. *Large deviations.* Vol. 137. Pure and Applied Mathematics. Academic Press, Inc., Boston, MA, 1989, pp. xiv+307. ISBN: 0-12-213150-9.

[12] Paul Dupuis and Richard S. Ellis. *A weak convergence approach to the theory of large deviations.* Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1997, pp. xviii+479. ISBN: 0-471-07672-4. DOI: 10.1002/9781118165904. URL: https://doi.org/10.1002/9781118165904.

[13] Sergey Foss and Dmitry Korshunov. “On large delays in multi-server queues with heavy tails”. In: *Math. Oper. Res.* 37.2 (2012), pp. 201–218. ISSN: 0364-765X,1526-5471. DOI: 10.1287/moor.1120.0539. URL: https://doi.org/10.1287/moor.1120.0539.

[14] Serguei Foss and Dmitry Korshunov. “Heavy tails in multi-server queue”. In: *Queueing Syst.* 52.1 (2006), pp. 31–48. ISSN: 0257-0130,1572-9443. DOI: 10.1007/s11134-006-3613-z. URL: https://doi.org/10.1007/s11134-006-3613-z.

[15] David Gamarnik and David A. Goldberg. “Steady-state $GI/GI/n$ queue in the Halfin-Whitt regime”. In: *Ann. Appl. Probab.* 23.6 (2013), pp. 2382–2419. ISSN: 1050-5164. DOI: 10.1214/12-AAP905. URL: https://doi.org/10.1214/12-AAP905.

[16] Frank den Hollander. *Large deviations*. Vol. 14. Fields Institute Monographs. American Mathematical Society, Providence, RI, 2000, pp. x+143. ISBN: 0-8218-1989-5. DOI: 10.1007/s00440-009-0235-5. URL: https://doi.org/10.1007/s00440-009-0235-5.

[17] Donald L. Iglehart and Ward Whitt. “Multiple channel queues in heavy traffic. I”. In: *Advances in Appl. Probability* 2 (1970), pp. 150–177. ISSN: 0001-8678,1475-6064. DOI: 10.1017/s0001867800037241. URL: https://doi.org/10.1017/s0001867800037241.

[18] Offer Kella. “Stability and nonproduct form of stochastic fluid networks with Lévy inputs”. In: *The Annals of Applied Probability* 6.1 (1996), pp. 186–199.

[19] J. Kiefer and J. Wolfowitz. “On the theory of queues with many servers”. In: *Trans. Amer. Math. Soc.* 78 (1955), pp. 1–18. ISSN: 0002-9947,1088-6850. DOI: 10.2307/1992945. URL: https://doi.org/10.2307/1992945.

[20] Pascal Lieshout and Michel Mandjes. “Asymptotic analysis of Lévy-driven tandem queues”. In: *Queueing Syst.* 60.3-4 (2008), pp. 203–226. ISSN: 0257-0130,1572-9443. DOI: 10.1007/s11134-008-9094-5. URL: https://doi.org/10.1007/s11134-008-9094-5.

[21] Masakiyo Miyazawa. “Light tail asymptotics in multidimensional reflecting processes for queueing networks”. In: *TOP* 19.2 (2011), pp. 233–299. ISSN: 1134-5764,1863-8279. DOI: 10.1007/s11750-011-0179-7. URL: https://doi.org/10.1007/s11750-011-0179-7.

[22] Guodong Pang, Rishi Talreja, and Ward Whitt. “Martingale proofs of many-server heavy-traffic limits for Markovian queues”. In: *Probab. Surv.* 4 (2007), pp. 193–267. ISSN: 1549-5787. DOI: 10.1214/06-PS091. URL: https://doi.org/10.1214/06-PS091.

[23] Félix Pollaczek. “Théorie analytique des problèmes stochastiques relatifs à un groupe de lignes téléphoniques avec dispositif d’attente”. In: (1961).

[24] A. Puhalskii. “Large deviation analysis of the single server queue”. In: *Queueing Systems Theory Appl.* 21.1-2 (1995), pp. 5–66. ISSN: 0257-0130. DOI: 10.1007/BF01158574. URL: https://doi.org/10.1007/BF01158574.

[25] John S. Sadowsky. “The probability of large queue lengths and waiting times in a heterogeneous multiserver queue. II. Positive recurrence and logarithmic limits”. In: *Adv. in Appl. Probab.* 27.2 (1995), pp. 567–583. ISSN: 0001-8678,1475-6064. DOI: 10.2307/1427839. URL: https://doi.org/10.2307/1427839.

[26] Alan Scheller-Wolf and Rein Vesilo. “Sink or swim together: necessary and sufficient conditions for finite moments of workload components in FIFO multiserver queues”. In: *Queueing Syst.* 67.1 (2011), pp. 47–61. ISSN: 0257-0130,1572-9443. DOI: 10.1007/s11134-010-9198-6. URL: https://doi.org/10.1007/s11134-010-9198-6.

[27] D. W. Stroock. *An introduction to the theory of large deviations*. Universitext. Springer-Verlag, New York, 1984, pp. vii+196. ISBN: 0-387-96021-X. DOI: 10.1007/978-1-4613-8514-1. URL: https://doi.org/10.1007/978-1-4613-8514-1.

[28] Zhe Su and Chang-Han Rhee. "Sample path large deviations for Lévy processes and random walks with log-normal type jumps". In: (2024).

[29] S. R. S. Varadhan. *Large deviations and applications*. Vol. 46. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984, pp. v+75. ISBN: 0-89871-189-4. DOI: 10.1137/1.9781611970241.bm. URL: https://doi.org/10.1137/1.9781611970241.bm.

[30] Ward Whitt. *Stochastic-process limits*. Springer Series in Operations Research. An introduction to stochastic-process limits and their application to queues. Springer-Verlag, New York, 2002, pp. xxiv+602. ISBN: 0-387-95358-2.

[31] Ward Whitt. "The impact of a heavy-tailed service-time distribution upon the $M/GI/s$ waiting-time distribution". In: *Queueing Systems Theory Appl.* 36.1-3 (2000), pp. 71–87. ISSN: 0257-0130,1572-9443. DOI: 10.1023/A:1019143505968. URL: https://doi.org/10.1023/A:1019143505968.