Externalities in the M/G/1 queue: LCFS-PR versus FCFS

Royi Jacobovic[*]   Nikki Levering[*]   Onno Boxma[†]

February 1, 2023

Abstract

Consider a stable M/G/1 system in which, at time $t = 0$, there are exactly n customers with residual service times equal to $v_1, v_2, \ldots, v_n$. In addition, assume that there is an extra customer c who arrives at time $t = 0$ and has a service requirement of x. The *externalities* which are created by c are equal to the total waiting time that others will save if her service requirement is reduced to zero. In this work, we study the joint distribution (parameterized by $n, v_1, v_2, \ldots, v_n, x$) of the externalities created by c when the underlying service distribution is either last-come, first-served with preemption or first-come, first-served. We start by proving a decomposition of the externalities under the above-mentioned service disciplines. Then, this decomposition is used to derive several other results regarding the externalities: moments, asymptotic approximations as $x \to \infty$, asymptotics of the tail distribution, and a functional central limit theorem.

Keywords: Externalities, M/G/1 queue, LCFS-PR, FCFS, heavy-tailed distribution, Gaussian approximation.

AMS Subject Classification (MSC2010): 60K25, 60K30, 60K37.

---

[*]Korteweg-de Vries Institute; University of Amsterdam, the Netherlands. `royi.jacobovic@mail.huji.ac.il`, `n.a.c.levering@uva.nl`

[†]Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands. `o.j.boxma@tue.nl`

1 Introduction

Consider a stable M/G/1 queue with an arrival rate λ and a service distribution $G(\cdot)$. In addition, assume that at time $t = 0$, there are $n \geq 0$ customers in the system whose residual service times are respectively equal to $v_1, v_2, \ldots, v_n$. Now, assume that there is an arrival of an additional customer with a service requirement $x \geq 0$ at time $t = 0$. In general, we say that the *externalities* which are generated by this additional customer are equal to the total amount of waiting time that others would save if she reduced her service requirement from x to zero. Some motives for the research about the properties of the externalities are in the context of, e.g., choice of a management scheme [17], regulation of queues with discretionary services [21, 22] and server-allocation (scheduling) problems in multiclass queues (especially with dynamic class types) [3, 19, 20, 28]. For an extra elaboration regarding these aspects, see [23, Section 1.1].

Naturally, the externalities which are created by the additional customer are stochastic and the corresponding analysis relies heavily on the underlying service discipline. In previous work [23], there is an analysis of the externalities under the assumption that the service discipline is first-come, first-served (FCFS). The main motivation for the current research has been to examine the externalities under a different service discipline, viz., the last-come, first-served with preemption (LCFS-PR), and to obtain insight into the effect of the service discipline on externalities. This is triggered by the observation that FCFS and LCFS-PR are in some respect two extreme cases w.r.t. externalities: In FCFS, the additional customer only affects later arrivals, whereas in LCFS-PR she only affects the customers who are already present. To the best of our knowledge, the only existing result under the LCFS-PR discipline is about the expected externalities in a stationary queue [17, Theorem 2.2].

The analysis of the current work is based on a joint representation of the externalities under LCFS-PR and FCFS in terms of a bivariate compound Poisson process. The arrival rate of this process is λ and its bivariate jump distribution is determined uniquely by the primitives of the model (λ and $G(\cdot)$) in a way which is to be explained in the sequel. Notably, this decomposition is a generalization of the univariate decomposition which was introduced in [23, Corollary 2] for the externalities under the FCFS discipline. In order to give a formal statement of this decomposition, a precise definition of the externalities under both LCFS-PR and FCFS disciplines should be addressed. This is done in Section 2. Then, Section 3 is dedicated to the statement and the proof of the decomposition result. In other

sections of this work, there are various applications of the above-mentioned decomposition:

1. In Section 4, the decomposition is used to derive the first and second moment of the externalities under the LCFS-PR discipline, and the correlation with the externalities under the FCFS discipline. In particular, the results include: (i) a generalization of [17, Theorem 2.2] and (ii) an explicit expression for the variance of the externalities in a stationary M/M/1 LCFS-PR queue.

2. In Section 5, the decomposition is used to derive a first order approximation of the externalities under LCFS-PR when $x \to \infty$. In addition, it is applied to determine the rate at which the externalities under FCFS tend to infinity as $x \to \infty$.

3. In Section 6, utilizing the decomposition, we study the tail asymptotics of the externalities under both LCFS-PR and FCFS. In particular, we show that if the service times belong to a specific large subclass of the subexponential distributions, then so do the externalities.

4. In Section 7, a functional central limit theorem (CLT) for the externalities under the LCFS-PR service discipline is derived. This result is analogous to [23, Theorem 4] which states a functional CLT for the externalities under the FCFS discipline.

Finally, Section 8, and also Subsections 6.1 and 7.2, contain a discussion about open problems which arise from the current research and might be interesting to consider in the future.

2 Model description

Let $\lambda > 0$ and assume that $G(\cdot)$ is a distribution function such that $G(0-) = 0$. In addition, denote its LST (Laplace-Stieltjes transform) by

$$\widetilde{G}(s) = \int_0^\infty e^{-st} \mathrm{d}G(t), \qquad s \geq 0, \tag{1}$$

and for each $m \geq 1$ denote

$$\mu_m = \int_0^\infty t^m \mathrm{d}G(t). \tag{2}$$

In particular, assume that $\rho \equiv \lambda\mu_1 < 1$.

3

Let $\{J(t); t \geq 0\}$ be a compound Poisson process with rate λ and jump distribution $G(\cdot)$. Also, let $\boldsymbol{v} \equiv (v_1, \ldots, v_n)$ be a vector with positive coordinates and denote $v \equiv v_1 + v_2 + \ldots + v_n$. Then, define the process

$$X_v(t) \equiv v + J(t) - t, \tag{3}$$

and denote its reflection at the origin (for more details, Te.g., [7, Section 2.4]) by $W_v(t)$. Note that $W_v(t)$ represents the workload process of an M/G/1 system with strong service discipline (i.e., the order in which customers are served is not a function of their service requirements; cf. [17]), an arrival rate λ and a service distribution $G(\cdot)$ under the following assumption: At time $t = 0$, there are $n \geq 1$ customers $c_1, c_2, \ldots, c_n$ such that for every $1 \leq i \leq n$, the remaining service time of c_i is equal to v_i. In addition, without loss of generality we shall assume that for every $1 \leq i < j \leq n$, c_i arrived to the queue before c_j.

2.1   The Externalities under LCFS-PR

Assume that the service discipline is LCFS-PR and observe that the preemption implies that an arrival of an additional customer c with a service requirement $x \geq 0$ at time $t = 0$ increases the waiting times of $c_1, c_2, \ldots, c_n$ while the waiting times of the customers who arrive after time $t = 0$ remain the same. Now, for every $0 \leq j \leq n$ denote $\bar{v}_j = \sum_{s=1}^{j} v_s$ (an empty sum being zero) and observe that the additional waiting time of c_k $(1 \leq k \leq n)$ due to the arrival of the additional customer at time $t = 0$ is equal to

$$E_k(x, \boldsymbol{v}) \equiv \min\{t \geq 0; W_{\bar{v}_n + x}(t) = \bar{v}_{k-1}\} - \min\{t \geq 0; W_{\bar{v}_n}(t) = \bar{v}_{k-1}\}. \tag{4}$$

It is common to refer to $E_k(x, \boldsymbol{v})$ as the *externalities* which are imposed on c_k and correspondingly to define the *total externalities* by

$$E(x, \boldsymbol{v}; \text{LCFS}) \equiv \sum_{k=1}^{n} E_k(x, \boldsymbol{v}). \tag{5}$$

Here and in the sequel we generally suppress n when denoting externalities.

2.2   The Externalities under FCFS

Assume that the service discipline is FCFS and for every $i \geq 1$, let τ_i be the time when $J(\cdot)$ has its i-th jump. Assume that an additional customer c with a service requirement $x \geq 0$ arrives at time $t = 0$. Then, her arrival can only affect the waiting times of those who arrive after time $t = 0$. Specifically,

for every $i \geq 1$, the externalities which are imposed on the customer who arrived at τ_i are equal to $W_{v+x}(\tau_i) - W_v(\tau_i)$. Thus, the total externalities are

$$E(x, \boldsymbol{v}; \text{FCFS}) \equiv \sum_{i=1}^{\infty} \left[W_{v+x}(\tau_i) - W_v(\tau_i) \right], \qquad (6)$$

and it is instructive to notice that they depend on $\boldsymbol{v}$ only through v. A thorough analysis of (6) is provided in [23].

3 Decomposition

Denote the externalities vector

$$\mathcal{E}(x, \boldsymbol{v}) \equiv (\mathcal{E}_L(x, \boldsymbol{v}, n), \mathcal{E}_F(x, \boldsymbol{v}, n)) \equiv (E(x, \boldsymbol{v}; \text{LCFS}), E(x, \boldsymbol{v}; \text{FCFS})) \quad (7)$$

and note that once the vector $\boldsymbol{v}$ is fixed, then $\{\mathcal{E}(x, \boldsymbol{v}); x \geq 0\}$ is a bivariate stochastic process. In Subsection 3.1 we present a decomposition of this process in terms of a compound Poisson process; cf. Theorem 1. In Subsection 3.2 we briefly discuss such a decomposition for *non-preemptive* LCFS.

3.1 LCFS-PR and FCFS

In order to phrase the externalities decomposition for LCFS-PR and FCFS, denote $w_j \equiv v - \overline{v}_{j-1}$ for every $1 \leq j \leq n$. In addition, consider a FCFS M/G/1 queue with an arrival rate λ and a service distribution $G(\cdot)$ which is empty at $t = 0$. Let B be the length of the first busy period and let N be the number of customers who received service during this period. Its joint distribution is given on p. 250 of [5]; see also Formula (28) below.

Theorem 1 *In the probability space of the model, there exists a bivariate compound Poisson process $S \equiv (S_1, S_2)$ with rate λ and jumps which are distributed like the vector (B, N) such that:*

1. For each $1 \leq k \leq n$,

$$E_k(x, \boldsymbol{v}) = x + S_1(x + w_k) - S_1(w_k) \;\;, \;\; x \geq 0 , \qquad (8)$$

and hence

$$\mathcal{E}_L(x, \boldsymbol{v}) = nx + \sum_{k=1}^{n} \left[S_1(x + w_k) - S_1(w_k) \right] \;\;, \;\; x \geq 0 . \qquad (9)$$

5

2.

$$\mathcal{E}_F(x, \boldsymbol{v}) = \int_0^x S_2(v+y)\mathrm{d}y \quad , \quad x \geq 0 \, . \tag{10}$$

Remark 1 Assume that the initial workload is $v + x$. In addition, consider the customers whose arrival times determine the descending ladder process (of the workload). Due to the strong Markov property of the workload process, each of these customers initiates a busy period which ends when the workload level returns to its original level (the level just before the first arrival epoch of that busy period). For each of these busy periods we may compute: (i) The length of the busy period and (ii) the number of customers who arrive until it ends. Since these busy periods are independent, the resulting bivariate random vectors are independent. Furthermore, standard properties of the Poisson process imply that the number of customers whose arrival times determine the descending ladder process has a Poisson distribution with rate $\lambda(v + x)$ and it is independent from the busy periods initialized by these customers. Thus, the sum of the bivariate random vectors has a bivariate compound Poisson distribution. This observation is the basis of the construction of $S = (S_1, S_2)$. In [23], it was shown that $\mathcal{E}_F(x, \boldsymbol{v})$ is determined uniquely by $S_2(\cdot)$ via the integral (10). Here, we complete the picture by expressing $\mathcal{E}_L(x, \boldsymbol{v})$ as a functional of $S_1(\cdot)$. Remarkably, the statement of Theorem 1 is about a *joint* decomposition of $\mathcal{E}(x, \boldsymbol{v})$ as a functional of $S(\cdot)$. We will demonstrate the additional value of having this joint decomposition by computing the cross covariance (and correlation) of $\mathcal{E}_L(x, \boldsymbol{v})$ and $\mathcal{E}_F(x, \boldsymbol{v})$ in Subsection 4.2. In particular, this computation reveals that the coordinates of $\mathcal{E}(x, \boldsymbol{v})$ are positively correlated, and quite strongly correlated in heavy traffic. This is not surprising in view of the fact that the number of customers and the length of each busy period in an M/G/1 queue are positively dependent. Furthermore, in Sections 6 and 7 we mention two open problems for which we suspect that the joint decomposition in Theorem 1 might be useful in their solutions.

Remark 2 Importantly, the equations (8),(9) and (10) all hold for every realization of the probability space and for every $x \geq 0$. Thus, the decomposition in Theorem 1 is a strong one.

Remark 3 Note that the stationary increments of the compound Poisson process imply that for every $x \geq 0$ and $1 \leq k \leq n$,

$$E_k(x, \boldsymbol{v}) \stackrel{d}{=} x + S_1(x) \, . \tag{11}$$

6

Notably, the distribution of the right-hand side is invariant with respect to $\boldsymbol{v}$ and k, which is a bit surprising. In practice, this means that when the service discipline is LCFS-PR, then for every $1 \leq k \leq n$, the distribution of the externalities which are imposed on c_k is invariant with respect to: (i) the vector of the remaining service times $\boldsymbol{v}$ and (ii) the place k of c_k in the order of service.

Remark 4 Fix $\boldsymbol{v}$ and observe that $x \mapsto \mathcal{E}_F(x, \boldsymbol{v})$ is an integral of a non-decreasing right-continuous process and hence it is convex. On the other hand, the sample paths of $x \mapsto \mathcal{E}_L(x, \boldsymbol{v})$ are composed of a positive drift and positive jumps. Consequently, deduce that $x \mapsto \mathcal{E}_L(x, \boldsymbol{v})$ is neither convex nor concave in x.

Proof: [Theorem 1] Fix $\boldsymbol{v}$ and let T_1 equal τ_1, the time of the first jump of the process $J(\cdot)$. Under LCFS-PR, this arrival interrupts the service of one of the customers present at time 0, whose service is resumed once the workload process is back at level $X_{\bar{v}_n+x}(T_1)$, a time which we denote by U_1. Now, let T_2 be the time of the first jump of the process $J(\cdot)$ after U_1, and U_2 the first time after T_2 the process $X_{\bar{v}_n+x}(\cdot)$ is back at level $X_{\bar{v}_n+x}(T_2)$. Similarly, let T_3 be the time of the first jump of the process $J(\cdot)$ after U_2, and U_3 the first time after T_3 the process $X_{\bar{v}_n+x}(\cdot)$ hits $X_{\bar{v}_n+x}(T_3)$. We may continue recursively with this construction, yielding the sequences $(T_i)_{i \geq 1}$ and $(U_i)_{i \geq 1}$. An illustration of these sequences for a single sample path of $J(\cdot)$ is given in Figure 1.

Denote $U_0 \equiv 0$ and for each $i \geq 1$, let $I_i \equiv T_i - U_{i-1}$. Note that $I_1, I_2, \ldots$ is an iid sequence of exponentially distributed random variables with rate λ. Furthermore, define for each $i \geq 1$

$$W_i \equiv U_i - T_i \ , \quad N_i \equiv |\{T_i \leq t \leq U_i; J(t) - J(t-) > 0\}| \, , \qquad (12)$$

and observe that $(W_1, N_1), (W_2, N_2), \ldots$ are iid bivariate random vectors with the distribution of (B, N). Importantly, the sequences $I_1, I_2, \ldots$ and $(W_1, N_1), (W_2, N_2), \ldots$ are independent, and hence may be used to construct a bivariate compound Poisson process. Specifically, with

$$N_\lambda(t) = \min\left\{k \geq 1; \sum_{i=1}^{k} I_i > t\right\} - 1 \ , \quad t \geq 0$$

a Poisson process with rate λ, we consider the bivariate process

$$S(t) \equiv (S_1(t), S_2(t)) \equiv \sum_{i=1}^{N_\lambda(t)} (W_i, N_i) \ , \quad t \geq 0 \, . \qquad (13)$$
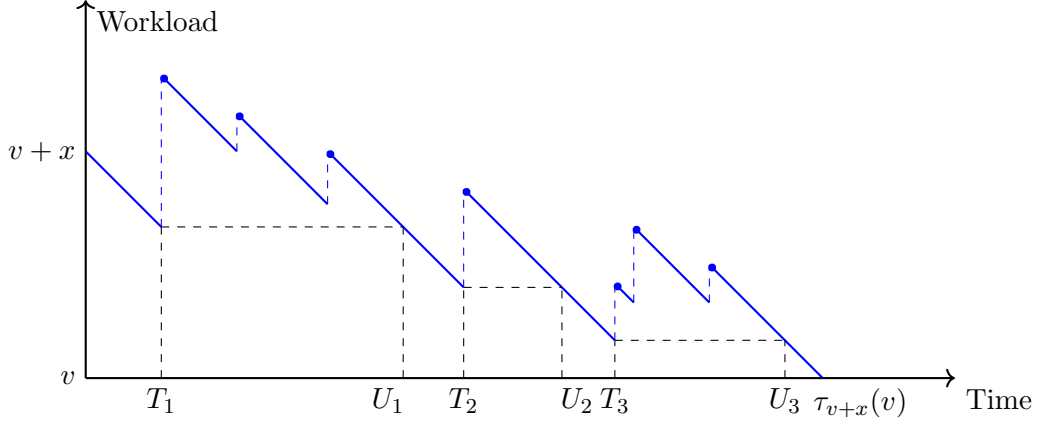
7

Figure 1: The blue graph illustrates a path of the workload process from level $v + x$ to the lower level v. Note that $U_1 - T_1, U_2 - T_2$ and $U_3 - T_3$ are three iid random variables which are distributed like B. In addition, observe that $x = (v + x) - v = T_1 + T_2 - U_1 + T_3 - U_2 + \tau_{v+x}(v) - U_3$.

Denote

$$\tau_v(y) \equiv \min\{t \geq 0; X_v(t) = y\}, \qquad y > 0, \tag{14}$$

and observe that

$$E_k(x, \boldsymbol{v}) = \tau_{\overline{v}_n + x}(\overline{v}_{k-1}) - \tau_{\overline{v}_n + x}(\overline{v}_{k-1} + x). \tag{15}$$

In addition, for every $0 \leq y \leq v$,

$$\tau_v(y) = \min\{t \geq 0; X_{v-y}(t) = 0\} = v - y + S_1(v - y), \tag{16}$$

and hence by combining (15) and (16), deduce (8). Finally, [23, Corollary 2] implies (10). ∎

3.2 Non-preemptive LCFS

One may wonder why we did not include the non-preemptive LCFS (LCFS-NP) in the decomposition which was stated in Theorem 1. The reason is that there exist some crucial differences between LCFS-PR and LCFS-NP. In order to explain these differences, we also provide a concrete example with a pictorial illustration in Figure 2.

Consider an M/G/1 queue under the LCFS-NP discipline, and recall that for every $1 \leq i < j \leq n$, c_i arrived to the queue before c_j. Then, c_n
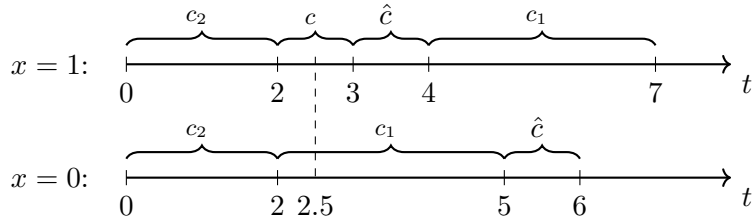
Figure 2: Consider a LCFS-NP M/G/1 queue with $n = 2$, $v_1 = 3$, $v_2 = 2$. In addition, assume that there is only one customer $\hat{c}$ who arrived at $t = 2.5$ and that besides him, no one else arrived during the time interval $(0, 5]$. If $x = 1$, the service of $\hat{c}$ would have been initiated at $t = 3$. If $x = 0$, the service of $\hat{c}$ would have been initiated at $t = 5$. This example demonstrates a possibility of having negative externalities.

was not preempted by the additional customer c who arrives at $t = 0$ with a service requirement $x \geq 0$. Thus, the externalities which were imposed on c_n by c had been equal to zero. In addition, for every $1 \leq k < n$, there were no interruptions during the service period of c_k. Therefore, the externalities which were imposed on c_k by c had been equal to

$$\min\{t \geq 0; W_{\bar{v}_n + x}(t) = \bar{v}_k\} - \min\{t \geq 0; W_{\bar{v}_n}(t) = \bar{v}_k\} = E_{k+1}(x, \boldsymbol{v}). \quad (17)$$

While this is in line with the decomposition (8), now a big difference is showing up: Unlike in the case of the LCFS-PR discipline, under LCFS-NP the additional customer might influence the waiting times of the customers who arrived *after* time $t = 0$ and the computation of the externalities for these customers is very complicated. This is mainly because their waiting times might be affected by the customer who was in the service position when they arrived. In fact, externalities could be negative for LCFS-NP.

4   The moments of $\mathcal{E}(x, \boldsymbol{v})$

When $\boldsymbol{v}$ is fixed, the moments of $\mathcal{E}_F(\cdot, \boldsymbol{v})$ are given in [23, Section 4]. If $\boldsymbol{v}$ is a random vector which is independent from the arrival process and the service demands of the customers, then the moments of $\mathcal{E}_F(\cdot, \boldsymbol{v})$ are given in [23, Section 7]. In Subsection 4.1 we analyse the first two moments of $\mathcal{E}_L(\cdot, \boldsymbol{v})$ and in Subsection 4.2 the joint moment of the coordinates of $\mathcal{E}(x, \boldsymbol{v})$. A mean-variance analysis for LCFS-PR is performed in Subsection 4.3. Initially, $\boldsymbol{v}$ is assumed to be fixed, but this assumption is relaxed in Subsection 4.4.

9

4.1 Mean and auto-covariance of $\mathcal{E}_L(\cdot, \boldsymbol{v})$

It is well-known (cf. [5], Section II.4.4) that

$$\mathbb{E}B = \frac{\mu_1}{1-\rho} \quad , \quad \mathbb{E}B^2 = \frac{\mu_2}{(1-\rho)^3} \, . \tag{18}$$

Thus, (11) implies that

$$\frac{\mathbb{E}\mathcal{E}_L(x, \boldsymbol{v})}{n} = \mathbb{E}E_1(x, \boldsymbol{v}) \tag{19}$$

$$= x + \mathbb{E}S(x)$$

$$= x(1 + \lambda\mathbb{E}B) = \frac{x}{1-\rho} \, .$$

Also, a standard calculation yields that

$$\mathrm{Cov}\,(S_1(t_1), S_1(t_2)) = \lambda\mathbb{E}B^2 \min\{t_1, t_2\} \tag{20}$$

$$= \frac{\lambda\mu_2 \min\{t_1, t_2\}}{(1-\rho)^3} \quad , \quad t_1, t_2 \ge 0,$$

and for simplicity of notations write

$$\sigma^2(u) \equiv \mathrm{Var}\,(S_1(u)) = \frac{\lambda\mu_2 u}{(1-\rho)^3} \quad , \quad u \ge 0 \, . \tag{21}$$

Therefore, for every $1 \le k, l \le n$ and $x_1, x_2 \ge 0$ deduce that

$$R_{k,l}(x_1, x_1 + x_2) \equiv \mathrm{Cov}\big(E_k(x_1, \boldsymbol{v}), E_l(x_1 + x_2, \boldsymbol{v})\big) \tag{22}$$

$$= \begin{cases} \sigma^2(x_1 + w_k) - \sigma^2(\min\{x_1 + w_k, w_l\}), & w_k \le w_l \\ \sigma^2(x_1 + \min\{w_k, x_2 + w_l\}) - \sigma^2(\min\{w_k, x_1 + x_2 + w_l\}), & w_k > w_l \end{cases},$$

$$= \frac{\lambda\mu_2}{(1-\rho)^3} \begin{cases} (x_1 + w_k - w_l)^+, & k \ge l \\ (x_1 - (x_2 + w_l - w_k)^-)^+, & k < l \end{cases},$$

where we have used the standard notations $u^+ \equiv \max\{u, 0\}$ and $u^- \equiv \max\{-u, 0\}$. As a result, this formula implies that

$$R_{\boldsymbol{v}}(x_1, x_1 + x_2) \equiv \text{Cov}\left(\mathcal{E}_L(x_1, \boldsymbol{v}), \mathcal{E}_L(x_1 + x_2, \boldsymbol{v})\right) \tag{23}$$

$$= \sum_{k=1}^{n} R_{k,k}(x_1, x_1 + x_2) + \sum_{k,l=1; k \neq l}^{n} R_{k,l}(x_1, x_1 + x_2)$$

$$= \frac{n x_1 \lambda \mu_2}{(1-\rho)^3} + \sum_{k=2}^{n} \sum_{l=1}^{k-1} \left(R_{k,l}(x_1, x_1+x_2) + R_{l,k}(x_1, x_1+x_2) \right)$$

$$= \frac{\lambda \mu_2}{(1-\rho)^3} \left[n x_1 + \sum_{k=2}^{n} \sum_{l=1}^{k-1} \left((x_1 + w_k - w_l)^+ + (x_1 - (x_2 + w_k - w_l)^-)^+ \right) \right]$$

$$= \frac{\lambda \mu_2}{(1-\rho)^3} \left[n x_1 + \sum_{k=2}^{n} \sum_{l=1}^{k-1} \left((x_1 - v_{l,k})^+ + (x_1 - (x_2 - v_{l,k})^-)^+ \right) \right],$$

where

$$v_{l,k} \equiv w_l - w_k = \sum_{i=1}^{k-1} v_i - \sum_{i=1}^{l-1} v_i , \quad 1 \leq l, k \leq n. \tag{24}$$

For future reference, we mention that taking $x_2 = 0$ in (23) gives

$$\text{Var}\left[\mathcal{E}_L(x, \boldsymbol{v})\right] = \frac{\lambda \mu_2}{(1-\rho)^3} \left[n x + 2 \sum_{1 \leq l < k \leq n} (x - v_{l,k})^+ \right]. \tag{25}$$

Remark 5 Observe that $R_{\boldsymbol{v}}(x_1, x_1 + x_2)$ is positive whenever $x_1 > 0$.

Remark 6 The auto-correlation function of $E_k(\cdot, \boldsymbol{v})$ is easily seen to equal

$$\frac{R_{k,k}(x_1, x_1 + x_2)}{\sqrt{R_{k,k}(x_1, x_1) R_{k,k}(x_1 + x_2, x_1 + x_2)}} = \sqrt{\frac{x_1}{x_1 + x_2}}.$$

Notably, this expression is invariant with respect to the parameters of the model, i.e., the arrival rate and the service distribution. The auto-correlation function of the total externalities satisfies the same invariance. The auto-correlation function of $\mathcal{E}_F(x, \boldsymbol{v})$ has the same property (see, [23, Eq. (33)]).

4.2 Cross-covariance of $\mathcal{E}_L(\cdot, \boldsymbol{v})$ and $\mathcal{E}_F(\cdot, \boldsymbol{v})$

We first mention two results from [23]:

$$\mathbb{E}\mathcal{E}_F(x, \boldsymbol{v}) = \lambda x \left(v + \frac{x}{2} \right) \mathbb{E}N = \lambda x \left(v + \frac{x}{2} \right) \frac{1}{1-\rho}, \tag{26}$$

11

$$\text{Var}[\mathcal{E}_F(x, \boldsymbol{v})] = \lambda x^2 \left(v + \frac{x}{3}\right) \mathbb{E}N^2 = \lambda x^2 \left(v + \frac{x}{3}\right) \frac{\lambda^2 \mu_2 + 1 - \rho^2}{(1 - \rho)^3}. \quad (27)$$

Next consider the joint distribution of the length B of an M/G/1 busy period and the number of customers N served in it. It is well-known (cf. p. 250 of [5]) that for every $\alpha, \beta \geq 0$,

$$\phi(\alpha, \beta) \equiv \mathbb{E}e^{-\alpha N - \beta B} = e^{-\alpha} \widetilde{G} \{\beta + \lambda [1 - \phi(\alpha, \beta)]\}. \quad (28)$$

Observe that

$$\frac{\mathrm{d}\widetilde{G}(\beta + \lambda(1 - y))}{\mathrm{d}y} = \lambda \int_0^\infty t e^{-\beta t - \lambda(1 - y)t} \mathrm{d}B(t). \quad (29)$$

Therefore, differentiating both sides of (28) once with respect to β, then with respect to α and taking both α and β equal to zero yields

$$\mathbb{E}BN = \frac{1}{(1 - \rho)^3} \big((1 - \rho)\mu_1 + \lambda\mu_2\big). \quad (30)$$

In addition, observe that for every $u, y > 0$,

$$c(u, y) \equiv \mathbb{E}S_1(u)S_2(y) \quad (31)$$

$$= \mathbb{E} \sum_{i=1}^{N_\lambda(u \wedge y)} B_i N_i + \mathbb{E} \sum_{i=1}^{N_\lambda(u)} \sum_{j=1; j \neq i}^{N_\lambda(y)} B_i N_j$$

$$= \lambda(u \wedge y)\mathbb{E}BN + \Big(\mathbb{E}N_\lambda(u \wedge y)\big(N_\lambda(u \vee y) - 1\big)\Big)\mathbb{E}B\mathbb{E}N$$

$$= \lambda(u \wedge y)\mathbb{E}BN + \lambda^2 uy\mathbb{E}B\mathbb{E}N.$$

Therefore, using (26) in the third step below,

$$\text{Cov}\left(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v})\right) \quad (32)$$

$$= \sum_{k=1}^n \left\{ \int_v^{v+x} \big(c(x + w_k, y) - c(w_k, y)\big)\mathrm{d}y \right\} - n\lambda x \mathbb{E}B\mathbb{E}\mathcal{E}_F(x, \boldsymbol{v})$$

$$= \sum_{k=1}^n \left\{ \lambda\mathbb{E}BN \int_v^{v+x} \big((y - w_k) \wedge x\big)\mathrm{d}y + \lambda^2 x^2 \left(v + \frac{x}{2}\right)\mathbb{E}B\mathbb{E}N \right\} - \frac{n\rho x}{1 - \rho}\mathbb{E}\mathcal{E}_F(x, \boldsymbol{v})$$

$$= \frac{(1 - \rho)\rho + \lambda^2\mu_2}{(1 - \rho)^3} \sum_{k=1}^n \int_v^{v+x} \big((y - w_k) \wedge x\big)\mathrm{d}y$$

$$= \frac{(1 - \rho)\rho + \lambda^2\mu_2}{(1 - \rho)^3} \left\{ nx^2 - \frac{1}{2}\sum_{k=1}^n ((x - (v - w_k))^+)^2 \right\}.$$

Together with the expression (27) for the variance of $\mathcal{E}_F(x, \boldsymbol{v})$ this directly yields that

$$
\text{Corr}(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v})) \tag{33}
$$

$$
= \frac{\left((1-\rho)\mu_1 + \lambda\mu_2\right)\left\{nx^2 - \frac{1}{2}\sum_{k=1}^{n}((x - (v - w_k))^+)^2\right\}}{x\sqrt{\mu_2(1 + \lambda^2\mu_2 - \rho^2)(v + \frac{x}{3})(nx + 2\sum_{1 \le l < k \le n}(x - v_{l,k})^+)}},
$$

which is easily seen to be positive. In particular, denote the variance of the service distribution by σ^2 and observe that

$$
\lim_{x \to \infty} \text{Corr}(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v})) = \frac{\sqrt{3}}{2} \frac{(1-\rho)\mu_1 + \lambda\mu_2}{\sqrt{\mu_2(1 + \lambda^2\mu_2 - \rho^2)}} \tag{34}
$$

$$
= \frac{\sqrt{3}}{2} \frac{\mu_1 + \lambda\sigma^2}{\sqrt{(\mu_1 + \lambda\sigma^2)^2 + (1-\rho)^2\sigma^2}},
$$

which is bounded from above by $\sqrt{3}/2$ (the value that is attained in heavy traffic).

4.3 Mean-variance analysis

Fix $x > 0$ and some positive integer n. Define the function $V : [0, \infty)^n \to (0, \infty)$ as follows:

$$
V(\boldsymbol{v}) \equiv V(\boldsymbol{v}; x, n) \equiv \text{Var}\left[\mathcal{E}_L(x, \boldsymbol{v})\right],
$$

and notice that it is given by (25). Recall that (19) implies that the mean of $\mathcal{E}_L(x, \boldsymbol{v})$ is determined uniquely by the parameters x and n (which is the length of $\boldsymbol{v}$). This motivates the following natural question: Given fixed values of x and n, what are the extreme values that the variance of $\mathcal{E}_L(x, \boldsymbol{v})$ may get?

Corollary 1 *For every x and n,*

$$
\min_{\boldsymbol{v}} V(\boldsymbol{v}; x, n) = \frac{nx\lambda\mu_2}{(1 - \rho)^3} \quad , \quad \max_{\boldsymbol{v}} V(\boldsymbol{v}; x, n) = \frac{n^2 x\lambda\mu_2}{(1 - \rho)^3}. \tag{35}
$$

Proof: Observe that $V(\boldsymbol{v})$ is determined uniquely by $v_1, v_2, \ldots, v_{n-1}$, i.e., it is invariant with respect to the value of v_n. Furthermore, $V(\cdot)$ is decreasing in each of the coordinates $v_1, v_2, \ldots, v_{n-1}$. This immediately yields the result. ∎

13

Remark 7 The current discussion resembles the situation in the classical mean-variance analysis in which the aim is to decompose a portfolio with minimal variance under the constraint that the expected return is fixed (see, e.g., [29]).

Remark 8 Note that for every $n > 1$ and $v > 0$,

$$V(v; x, 1) = \frac{x\lambda\mu_2}{(1-\rho)^3} < \frac{n^2 x\lambda\mu_2}{(1-\rho)^3} = V(0, 0, \ldots, 0, v; x, n). \qquad (36)$$

This means that gradual reduction of the remaining service times of the customers towards zero does not have the same effect as removing these customers from the system. In particular, Corollary 1 implies that by reducing the remaining service times of $c_1, c_2, \ldots, c_{n-1}$ to be arbitrarily close to zero, the variance of the total externalities becomes arbitrarily close to its maximum. However, it is important to notice that once the remaining service times of the customers are elapsed, they leave the system. Thus, in that sense the maximum which appears in the statement of Corollary 1 should be considered as a supremum.

4.4 Random initial condition

Assume that the vector $\boldsymbol{v}$ is a random vector with a random size n such that $\boldsymbol{v}$ is independent from all other stochastic objects of the model. An application of (19) with proper conditioning implies that

$$\mathbb{E}\mathcal{E}_L(x, \boldsymbol{v}) = \mathbb{E}\frac{nx}{1-\rho} = \frac{x\mathbb{E}n}{1-\rho}. \qquad (37)$$

In addition, observe that for every $x_1, x_2 \geq 0$:

$$\mathrm{Cov}\Big(\mathbb{E}\big[\mathcal{E}_L(x_1, \boldsymbol{v})\,\big|\,\boldsymbol{v}\big], \mathbb{E}\big[\mathcal{E}_L(x_1 + x_2, \boldsymbol{v})\,\big|\,\boldsymbol{v}\big]\Big) \qquad (38)$$

$$= \mathrm{Cov}\Big(\frac{nx_1}{1-\rho}, \frac{n(x_1 + x_2)}{1-\rho}\Big)$$

$$= \frac{x_1(x_1 + x_2)\mathrm{Var}(n)}{(1-\rho)^2}.$$

Therefore, a proper conditioning with the help of (23) and the law of total covariance yields that

$$\mathrm{Cov}\Big(\mathcal{E}_L(x_1, \boldsymbol{v}), \mathcal{E}_L(x_1 + x_2, \boldsymbol{v})\Big) = \frac{x_1(x_1 + x_2)\mathrm{Var}(n)}{(1-\rho)^2}$$

$$+ \frac{\lambda\mu_2}{(1-\rho)^3}\left\{x_1\mathbb{E}n + \mathbb{E}\Big[\sum_{k=2}^{n}\sum_{l=1}^{k-1}\Big((x_1 - v_{l,k})^+ + (x_1 - (x_2 - v_{l,k})^-)^+\Big)\Big]\right\}.$$

14

Especially, by setting $x_1 = x \geq 0$ and $x_2 = 0$, deduce that:

$$\text{Var}\big[\mathcal{E}_L(x, \boldsymbol{v})\big] = \frac{x^2 \text{Var}(n)}{(1-\rho)^2} + \frac{\lambda \mu_2}{(1-\rho)^3} \left\{ x \mathbb{E}n + 2\mathbb{E} \sum_{1 \leq l < k \leq n} (x - v_{l,k})^+ \right\}.$$
(39)

Moreover, observe that for every $x \geq 0$,

$$\text{Cov}\Big(\mathbb{E}\big[\mathcal{E}_L(x, \boldsymbol{v}) \,\big|\, \boldsymbol{v}\big], \mathbb{E}\big[\mathcal{E}_F(x, \boldsymbol{v}) \,\big|\, \boldsymbol{v}\big]\Big) = \text{Cov}\Big(\frac{nx}{1-\rho}, \frac{\lambda x}{1-\rho}\big(v + \frac{x}{2}\big)\Big)$$

$$= \frac{\lambda x^2}{(1-\rho)^2} \text{Cov}(n, v),$$

and hence the law of total variance can be used to find that

$$\text{Cov}\left(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v})\right)$$
$$= \frac{\lambda x^2}{(1-\rho)^2} \text{Cov}(n, v) + \frac{(1-\rho)\rho + \lambda^2 \mu_2}{(1-\rho)^3} \left\{ x^2 \mathbb{E}n - \frac{1}{2}\mathbb{E} \sum_{k=1}^{n} ((w_k + x - v)^+)^2 \right\}.$$

4.4.1 The stationary distribution as an initial condition

Assume that $\boldsymbol{v}$ is distributed according to the stationary distribution of the remaining service times and the queue length under the LCFS-PR. By applying [25, Corollary 1] and PASTA, this distribution is as follows:

1. n is a geometric random variable, i.e.,

$$\mathbb{P}\{n = s\} = (1 - \rho)\rho^s \ , \quad s \geq 0 \,.$$
(40)

2. For every $s \geq 1$, $(v_1, v_2, \ldots, v_s)|\{n = s\}$ is a random vector with iid coordinates such that v_1 is distributed according to the residual time distribution of $G(\cdot)$.

Consequently, (37) implies that

$$\mathbb{E}\mathcal{E}_L(x, \boldsymbol{v}) = \frac{x \mathbb{E}n}{1 - \rho} = \frac{x\rho}{(1-\rho)^2}$$
(41)

which yields a new derivation for [17, Theorem 2.2]. Also, (39) yields the following result.

15

Proposition 1 *If the service distribution is exponential with rate ν such that $\frac{\lambda}{\nu} < 1$, then the variance of the externalities in the stationary LCFS-PR M/G/1 queue is equal to*

$$\frac{x\rho}{(1-\rho)^4}\left(x+\lambda\mu_2\right) + \frac{2\mu_2\lambda\rho^2}{(1-\rho)^5}\left[x + \frac{\mu_1}{1-\rho}\left(e^{-(1-\rho)x/\mu_1}-1\right)\right]. \qquad (42)$$

Proof: For each $i \geq 1$, let X_i be a random variable which is distributed according to

$$F_i(u) \equiv 1 - \sum_{j=0}^{i-1}\frac{e^{-\nu u}(\nu u)^j}{j!} \quad, \quad u \geq 0, \qquad (43)$$

i.e., an Erlang distribution with shape i and rate ν. Then, conditioning on n yields:

$$\begin{aligned}
\mathbb{E}\sum_{1\leq l<k\leq n}(x-v_{l,k})^+ &= \sum_{s=2}^{\infty}(1-\rho)\rho^s\sum_{1\leq l<k\leq n}\mathbb{E}\left[(x-v_{l,k})^+|n=s\right]\\
&= \sum_{s=2}^{\infty}(1-\rho)\rho^s\sum_{1\leq l<k\leq s}\left\{x\mathbb{P}\left\{X_{k-l}<x\right\} - \mathbb{E}X_{k-l}1\{X_{k-l}<x\}\right\}\\
&= \sum_{s=2}^{\infty}(1-\rho)\rho^s\sum_{i=1}^{s-1}(s-i)\left\{x\mathbb{P}\left\{X_i<x\right\} - \mathbb{E}X_i1\{X_i<x\}\right\}\\
&= \sum_{s=2}^{\infty}(1-\rho)\rho^s\sum_{i=1}^{s-1}(s-i)\left\{xF_i(x) - \int_0^x\frac{\nu^i u^i e^{-\nu u}}{(i-1)!}\mathrm{d}u\right\}\\
&= \sum_{s=2}^{\infty}(1-\rho)\rho^s\sum_{i=1}^{s-1}(s-i)\left\{xF_i(x) - \frac{i}{\nu}F_{i+1}(x)\right\}\\
&= (1-\rho)\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}j\rho^{i+j}\left\{xF_i(x) - \frac{i}{\nu}F_{i+1}(x)\right\}\\
&= \frac{1}{(1-\rho)}\sum_{i=1}^{\infty}\rho^{i+1}\left\{xF_i(x) - \frac{i}{\nu}F_{i+1}(x)\right\}\\
&= \frac{1}{(1-\rho)}\sum_{i=1}^{\infty}\rho^i\left\{\rho xF_i(x) - \frac{(i-1)}{\nu}F_i(x)\right\}. \qquad (44)
\end{aligned}$$

It is well-known that a compound geometric distribution with a probability of success p and an exponential distribution with rate ν is an exponential

distribution with rate $p\nu$, and hence

$$\sum_{i=1}^{\infty} \rho^i F_i(x) = \frac{\rho}{1-\rho} \sum_{i=1}^{\infty} (1-\rho)\rho^{i-1} F_i(x) = \frac{\rho}{1-\rho}\left(1 - e^{-(1-\rho)\nu x}\right).$$

This identity may also be used to obtain that

$$\sum_{i=1}^{\infty} i\rho^i F_i(x) = \rho \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\sum_{i=1}^{\infty} \rho^i F_i(x)\right) = \rho \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\frac{\rho}{1-\rho}\left(1 - e^{-(1-\rho)\nu x}\right)\right)$$

$$= \frac{\rho}{(1-\rho)^2}\left(1 - e^{-(1-\rho)\nu x}\right) - \frac{\rho^2 \nu x}{1-\rho} e^{-(1-\rho)\nu x}.$$

Thus, the expression in (44) equals

$$\frac{\rho^2}{\nu(1-\rho)^3}\left\{e^{-(1-\rho)\nu x} - 1 + x\nu(1-\rho)\right\}.$$

Insertion of this formula into (39) yields the desired result. ∎

Remark 9 The variance of the externalities is very sensitive to a high traffic load. Concretely, notice that

$$x + \frac{\mu_1}{1-\rho}\left(e^{-(1-\rho)x/\mu_1} - 1\right) = \sum_{k=1}^{\infty} \frac{(-x)^{k+1}}{(k+1)!\mu_1^k}(1-\rho)^k$$

and hence $\mathrm{Var}\big[\mathcal{E}_L(x, \boldsymbol{v})\big]$ is of the order $(1-\rho)^{-4}$ as $\rho \uparrow 1$.

5 Rates of increase

By definition $x \mapsto \mathcal{E}_L(x, \boldsymbol{v})$ and $x \mapsto \mathcal{E}_F(x, \boldsymbol{v})$ are both nondecreasing stochastic processes. The following theorem gives the rates of increase.

Theorem 2 *For every $\boldsymbol{v}$ (with nonnegative coordinates whose sum is positive), the following statements hold with probability one:*

1.
$$\mathcal{E}_L(x, \boldsymbol{v}) \sim \frac{x}{1-\rho} \quad \text{as} \ \ x \to \infty. \tag{45}$$

2.
$$\frac{\lambda}{4(1-\rho)} \leq \liminf_{x \to \infty} \frac{\mathcal{E}_F(x; \boldsymbol{v})}{x^2} \leq \limsup_{x \to \infty} \frac{\mathcal{E}_F(x; \boldsymbol{v})}{x^2} \leq \frac{\lambda}{(1-\rho)}. \tag{46}$$

17

Remark 10 Note that Theorem 2 states that $x \mapsto \mathcal{E}_L(x, \boldsymbol{v})$ (resp. $x \mapsto \mathcal{E}_F(x, \boldsymbol{v})$) increases at the same rate in which $x \mapsto \mathbb{E}\mathcal{E}_L(x, \boldsymbol{v})$ (resp. $x \mapsto \mathbb{E}\mathcal{E}_F(x, \boldsymbol{v})$) increases (see (19) and also [23, Eq. (28)]).

Remark 11 It is important to notice that the rate which is stated in (45) and the bounds on the rate which are given in (46) are all uniform in $\boldsymbol{v}$. Since $\mathbb{E}\mathcal{E}_L(x, \boldsymbol{v})$ is invariant with respect to $\boldsymbol{v}$ (see (19)), it is not surprising that $\boldsymbol{v}$ has no effect on the first order approximation of $x \mapsto \mathcal{E}_L(x, \boldsymbol{v})$ as $x \to \infty$. However, [23, Eq. (28)] implies that $\boldsymbol{v}$ influences $\mathbb{E}\mathcal{E}_F(x, \boldsymbol{v})$ and hence we suspect that the bounds in (46) are sub-optimal.

Remark 12 Theorem 2 implies that

$$\sup \{x \geq 0; \mathcal{E}_L(x, \boldsymbol{v}) \geq \mathcal{E}_F(x, \boldsymbol{v})\} \tag{47}$$

is finite with probability one.

Proof: [Theorem 2] Consider some $x > 0$ and observe that for each $1 \leq k \leq n$, Theorem 1 and the renewal reward theorem imply that with probability one:

$$\frac{E_k(x, \boldsymbol{v})}{x} = 1 + \frac{S_1(x + w_k) - S_1(w_k)}{x} \tag{48}$$

$$= 1 + x^{-1} \sum_{i=N_\lambda(w_k)+1}^{N_\lambda(x+w_k)} B_i$$

$$= 1 + x^{-1} \sum_{i=1}^{N_\lambda(x+w_k)-N_\lambda(w_k)} B_{i+N_\lambda(w_k)}$$

$$\xrightarrow{x \to \infty} 1 + \frac{\frac{\mu_1}{1-\rho}}{1/\lambda} = \frac{1}{1-\rho}. \tag{49}$$

This yields (45) via summation over k.

In addition, observe that Theorem 1 and the renewal reward theorem might be applied once again in order to deduce that with probability one:

$$\frac{\mathcal{E}_F(x, \boldsymbol{v})}{x^2} = x^{-2} \int_0^x S_2(v + y)\mathrm{d}y \tag{50}$$

$$\leq x^{-1} S_2(v + x)$$

$$= \frac{x + v}{x} \cdot \frac{1}{x + v} \sum_{i=1}^{N_\lambda(v+x)} N_i \xrightarrow{x \to \infty} \frac{\frac{1}{1-\rho}}{1/\lambda} = \frac{\lambda}{1-\rho}.$$

18

In a similar fashion, deduce that for every $\alpha \in (0,1)$ with probability one:

$$
\frac{\mathcal{E}_F(x, \boldsymbol{v})}{x^2} = x^{-2} \int_0^x S_2(v+y)\mathrm{d}y \tag{51}
$$

$$
\geq x^{-2} \int_{\alpha x}^x S_2(v+y)\mathrm{d}y
$$

$$
\geq \frac{(1-\alpha)(v+\alpha x)}{x} \cdot \frac{1}{v+\alpha x} S_2(v+\alpha x)
$$

$$
\xrightarrow{x \to \infty} \alpha(1-\alpha) \cdot \frac{\frac{1}{1-\rho}}{1/\lambda} = \frac{\alpha(1-\alpha)\lambda}{1-\rho}.
$$

Finally, an insertion of $\alpha = \frac{1}{2}$ completes the proof. ∎

6 Tail asymptotics

Our goal in this section is to determine the tail behavior of $\mathcal{E}_L(x, \boldsymbol{v})$ and $\mathcal{E}_F(x, \boldsymbol{v})$ in the case of heavy-tailed service times. Consider $x > 0$ and for every distribution function $F(\cdot)$, denote $\overline{F} \equiv 1 - F$. In addition, let $\star$ denote a convolution between two distribution functions. In this section, we assume that $G(\cdot)$ is subexponential, i.e.,

$$
\overline{G \star G}(u) \sim 2\overline{G}(u) \quad \text{as} \quad u \to \infty. \tag{52}
$$

We have already seen that the M/G/1 busy period length B plays a key role in $\mathcal{E}_L(x, \boldsymbol{v})$, and that the corresponding number of customers N in that busy period plays a key role in $\mathcal{E}_F(x, \boldsymbol{v})$. In the literature it was first shown [8] for the M/G/1 queue with regularly varying service times (a subclass of the class of subexponential distributions) that

$$
\mathbb{P}\{B > u\} \sim \frac{1}{1-\rho}\overline{G}\left[u(1-\rho)\right] \quad \text{as} \quad u \to \infty. \tag{53}
$$

Exactly the same result was subsequently derived for the larger classes of intermediate regularly varying service times [33] and of service times with a tail that is heavier than $\mathrm{e}^{-\sqrt{u}}$; see [24] and finally [2]. An explanation why it is impossible to get (53) when $\overline{G}(u)$ is lighter than $\mathrm{e}^{-\sqrt{u}}$ as $u \to \infty$ can be found in [1, Section 6.2].

More or less equivalently, very similar conditions (basically either intermediate regular variation or a tail heavier than $\mathrm{e}^{-\sqrt{u}}$, and some moment conditions) have been provided [10, 9] under which

$$
\mathbb{P}\{N > u\} \sim \frac{1}{1-\rho}\overline{G}\left[\frac{u(1-\rho)}{\lambda}\right] \quad \text{as} \quad u \to \infty. \tag{54}
$$

19

The precise conditions for both formulas to hold are rather delicate; for details we refer to Theorem 1.2 of [2] (for B) and Proposition 3.2 of [9] (for N). In the remainder of this section we assume that these conditions are satisfied, so that (53) and (54) hold. For clarity, the statements of the main results (Theorems 3 and 4) appear in Subsection 6.1 while the proofs are included in Subsection 6.2.

6.1 Results and some open problems

6.1.1 LCFS-PR

Assume that $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ is such that $v_i > 0$ for every $1 \leq i \leq n$. For the statement of the upcoming Theorem 3, we introduce the set $\mathcal{D} \equiv \{w_1, \ldots, w_n\} \cup \{w_1 + x, \ldots, w_n + x\}$ and its cardinality $d \equiv |\mathcal{D}|$. Also, let $\widetilde{w}_1 < \widetilde{w}_2 < \ldots < \widetilde{w}_d$ be the elements of $\mathcal{D}$ in an increasing order and for each $1 \leq j \leq d - 1$, let q_j be the number of k's ($1 \leq k \leq n$) such that $[\widetilde{w}_j, \widetilde{w}_{j+1}] \subseteq [w_k, w_k + x]$, i.e.,

$$q_j \equiv |\{1 \leq k \leq n; w_k \leq \widetilde{w}_j, \widetilde{w}_{j+1} \leq w_k + x\}| . \tag{55}$$

Respectively, denote the set $\mathcal{J} \equiv \{1 \leq j \leq d - 1; q_j > 0\}$ which is not empty. To see that, if $1 \leq j \leq d - 1$ and $\widetilde{w}_j = w_k$ for some $1 \leq k \leq n$, then $\widetilde{w}_{j+1} \leq w_k + x$ and hence $[\widetilde{w}_j, \widetilde{w}_{j+1}] \subseteq [w_k, w_k + x]$ which implies that $q_j \geq 1$.

Theorem 3 *Assume that (53) holds. Then:*

1. *For each $1 \leq k \leq n$, $E_k(x, \boldsymbol{v})$ has a subexponential distribution such that*

$$\mathbb{P}\{E_k(x, \boldsymbol{v}) > u\} \sim \frac{\lambda x}{1 - \rho} \overline{G}[u(1 - \rho)] \quad as \quad u \to \infty. \tag{56}$$

2. *$\mathcal{E}_L(x, \boldsymbol{v})$ has a subexponential distribution such that*

$$\mathbb{P}\{\mathcal{E}_L(x, \boldsymbol{v}) > u\} \sim \frac{\lambda}{1 - \rho} \sum_{j \in \mathcal{J}} (\widetilde{w}_{j+1} - \widetilde{w}_j) \overline{G}\left[\frac{u(1 - \rho)}{q_j}\right] \quad as \quad u \to \infty. \tag{57}$$

Remark 13 Since $\overline{G}(\cdot)$ is nonincreasing, (57) implies that

$$\mathbb{P}\{\mathcal{E}_L(x, \boldsymbol{v}) > u\} = \Theta\left\{\overline{G}\left[\frac{u(1 - \rho)}{\max\{q_j; j \in \mathcal{J}\}}\right]\right\} \quad as \quad u \to \infty. \tag{58}$$

This means that as $u \to \infty$, $\mathbb{P}\{\mathcal{E}_L(x, \boldsymbol{v}) > u\}$ vanishes at the same rate (up to a multiplicative constant) that $\overline{G}\left[\frac{u(1-\rho)}{\max\{q_j;j\in\mathcal{J}\}}\right]$ tends to zero.

Remark 14 Observe that $\mathcal{E}_L(x, \boldsymbol{v}) = \sum_{k=1}^{n} E_k(x, \boldsymbol{v})$ is a sum of dependent random variables. There is an extensive literature about the tail asymptotics of sums of dependent random variables, e.g., [14, 18, 31, 32] and the references therein. Having said that, we are not aware of any existing result which might be useful in order to prove Theorem 3.

6.1.2 FCFS

The analogue of Theorem 3 for the externalities under the FCFS discipline is the following Theorem 4.

Theorem 4 *Assume that (54) holds. In addition, denote*

$$\omega_1 \equiv \int_0^1 \liminf_{u \to \infty} \frac{\overline{G}(u/y)}{\overline{G}(u)} \mathrm{d}y \quad , \quad \omega_2 \equiv \int_0^1 \limsup_{u \to \infty} \frac{\overline{G}(u/y)}{\overline{G}(u)} \mathrm{d}y. \qquad (59)$$

Then $\mathcal{E}_F(x, \boldsymbol{v})$ has a subexponential distribution such that

$$\frac{\lambda(v + x\omega_1)}{(1 - \rho)} \leq \liminf_{u \to \infty} \frac{\mathbb{P}\{\mathcal{E}_F(x, \boldsymbol{v}) > ux\}}{\overline{G}\left[\frac{ux(1-\rho)}{\lambda}\right]} \qquad (60)$$

$$\leq \limsup_{u \to \infty} \frac{\mathbb{P}\{\mathcal{E}_F(x, \boldsymbol{v}) > ux\}}{\overline{G}\left[\frac{ux(1-\rho)}{\lambda}\right]} \leq \frac{\lambda(v + x\omega_2)}{(1 - \rho)}.$$

Remark 15 There are cases in which $\omega_1 = \omega_2$. In such cases, Theorem 4 yields a first order approximation of the tail distribution of $\mathcal{E}_F(x, \boldsymbol{v})$. For example, when $G(\cdot)$ is regularly varying, i.e.,

$$\overline{G}(u) \sim L(u)u^{-\alpha} \quad \text{as} \quad u \to \infty \qquad (61)$$

for some $\alpha > 0$ and $L : (0, \infty) \to (0, \infty)$ such that $L(uy) \sim L(u)$ as $u \to \infty$ for every $y > 0$ (i.e., $L(\cdot)$ is slowly varying at infinity). In that case, standard calculations yield that $\omega_1 = \omega_2 = \frac{1}{1+\alpha}$.

6.1.3 Open problem: Is $\mathcal{E}(x, v)$ bivariate regularly varying?

Theorem 3 and Theorem 4 include certain conditions on heavy-tailed $G(\cdot)$ under which the (univariate) tail distributions of $\mathcal{E}_L(x, \boldsymbol{v})$ and $\mathcal{E}_F(x, \boldsymbol{v})$ are just as heavy as $\overline{G}(\cdot)$. It would be interesting to also study the *bivariate* tail asymptotics of $(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v}))$. For example, assume that $G(\cdot)$ is regularly varying; does it imply that the joint distribution of $(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v}))$

is also regularly varying? For details about regularly varying random vectors, see, e.g., [27, Section 2].

We believe that this query should start by figuring out whether a regularly varying $G(\cdot)$ implies a bivariate regularly varying distribution of (B, N). Then, once such a relation is established, it might be applied in order to derive conditions on the service distribution under which the distribution of $(\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v}))$ is bivariate regularly varying. To the best of our knowledge, there are no existing results in this direction.

6.1.4 Open problem: the stationary distribution as an initial condition

If the additional customer arrives when the M/G/1 LCFS-PR queue is in steady state, then she sees a geometric(ρ) distributed number of customers (cf. Subsection 4.4.1). If, moreover, the service time of the additional customer has distribution $G(\cdot)$, then all customers already present experience the same M/G/1 busy period as extra delay. One can now use the abovementioned known results about the tail behavior of such a busy period to conclude that the externalities distribution is just as heavy-tailed as $G(\cdot)$. For FCFS, however, the situation is more complicated. Now the remaining workload v of the customers already present plays a role in the externalities. Considering the tail behavior of (cf. (10)) $\int_v^{v+x} S_2(y)\mathrm{d}y$ when v, x are heavy-tailed random variables, the 'principle of the one big jump' (cf. [13]) suggests that the workload v will be dominant, and will determine the tail behavior of the externalities. For example, when $\overline{G}(\cdot)$ is a regularly varying function at infinity of index $-\nu \in (-2, -1)$, then the steady-state workload v is regularly varying of index $1 - \nu$ [6], and we conjecture that $\mathcal{E}_F$ is then also regularly varying at infinity of index $1 - \nu$, i.e., one degree heavier than the service time and the busy period.

6.2 Proofs

In general, the proof of Theorem 3 (resp. Theorem 4) is based on the decomposition which was stated in Theorem 1 (resp. [23, Theorem 2]) with an application of some known results from the theory about subexponential distributions. The next lemma is a summary of the existing results which are needed.

Lemma 1 *Assume that X is a nonnegative random variable with a subexponential distribution. In addition, let Y be a nonnegative random variable which is independent from X. Then:*

(i) For every $h > 0$, hX also has a subexponential distribution.

(ii) For every $h > 0$,

$$\mathbb{P}\{X > u\} \sim \mathbb{P}\{X > u + h\} \quad as \quad u \to \infty. \qquad (62)$$

(iii) If Y is bounded, then XY has a subexponential distribution.

(iv) Assume that Y is unbounded and there are $0 < z_1 < z_2 < \infty$ such that

$$z_1 \leq \frac{\mathbb{P}\{X > u\}}{\mathbb{P}\{Y > u\}} \leq z_2 \qquad (63)$$

for every $u > 0$. Then, Y has a subexponential distribution.

(v) Assume that Y is unbounded such that

$$\sup_{u > 0} \frac{\mathbb{P}\{Y > u\}}{\mathbb{P}\{X > u\}} < \infty, \qquad (64)$$

and for every $h > 0$,

$$\mathbb{P}\{Y > u\} \sim \mathbb{P}\{Y > u + h\} \quad as \quad u \to \infty. \qquad (65)$$

Then,

$$\mathbb{P}\{X + Y > u\} \sim \mathbb{P}\{X > u\} + \mathbb{P}\{Y > u\} \quad as \quad u \to \infty. \qquad (66)$$

(vi) Assume that Y is distributed according to a Poisson distribution with rate $\gamma > 0$. If $X, X_1, X_2, \ldots$ is an iid sequence of random variables, which are independent from Y, then $S_Y \equiv X_1 + X_2 + \ldots + X_Y$ has a subexponential distribution such that

$$\mathbb{P}\{S_Y > u\} \sim \gamma \mathbb{P}\{X > u\} \quad as \quad u \to \infty. \qquad (67)$$

Proof: [**Lemma 1**] (i) is an immediate consequence of the definition in (52). (ii) is [13, Lemma 3.2], (iii) is [4, Corollary 2.5], (iv) is due to [26, Theorem 2.1(a)], (v) is [12, Lemma 2] and (vi) is due to [11, Theorem 3]. ∎

Proof: [**Theorem 3**] We provide only a proof for (57) since the proof of (56) follows from similar arguments and it is more standard.

It is given that $G(\cdot)$ is subexponential. Thus, due to (53), Lemma 1(iv) implies that B has a subexponential distribution. Next, for every $1 \leq i \leq d - 1$, denote

$$\hat{S}_i \equiv S_1(\widetilde{w}_{i+1}) - S_1(\widetilde{w}_i). \qquad (68)$$

Therefore, Theorem 1 yields that

$$\mathbb{P}\left\{\mathcal{E}_L(x, \boldsymbol{v}) > u\right\} = \mathbb{P}\left\{\sum_{k=1}^{n}\left[S_1\left(x + w_k\right) - S_1\left(w_k\right)\right] > u - nx\right\}$$

$$= \mathbb{P}\left\{\sum_{j \in \mathcal{J}} q_j \hat{S}_j > u - nx\right\}. \tag{69}$$

Notably, since $S_1(\cdot)$ has stationary independent increments, then:

1. $\left\{q_j \hat{S}_j; j \in \mathcal{J}\right\}$ is a collection of independent random variables.

2. For each $j \in \mathcal{J}$, the random variable $q_j \hat{S}_j$ has a compound Poisson distribution with rate $\lambda\left(\widetilde{w}_{j+1} - \widetilde{w}_j\right)$ and jumps which have the distribution of $q_j B$.

Especially, observe that for each $j \in \mathcal{J}$, Lemma 1(vi) and the fact that $q_j B$ has a subexponential distribution (due to Lemma 1(i)) imply that $q_j \hat{S}_j$ has a subexponential distribution such that

$$\mathbb{P}\left\{q_j \hat{S}_j > u\right\} \sim \lambda\left(\widetilde{w}_{j+1} - \widetilde{w}_j\right) \mathbb{P}\left\{q_j B > u\right\}$$

$$\sim \frac{\lambda\left(\widetilde{w}_{j+1} - \widetilde{w}_j\right)}{1 - \rho} \overline{G}\left[\frac{u(1 - \rho)}{q_j}\right] \quad \text{as } u \to \infty. \tag{70}$$

Let $r \equiv |\mathcal{J}|$ and without loss of generality assume that $q_1 \leq q_2 \leq \ldots \leq q_r$. In particular, observe that for every $1 \leq i_1 < i_2 \leq r$,

$$\sup_{u>0} \frac{\overline{G}\left[\frac{u(1-\rho)}{q_{i_1}}\right]}{\overline{G}\left[\frac{u(1-\rho)}{q_{i_2}}\right]} \leq 1. \tag{71}$$

Thus, Lemma 1(v) implies that

$$\mathbb{P}\left\{\sum_{i=1}^{2} q_i \hat{S}_i > u\right\} \sim \frac{\lambda}{1 - \rho} \sum_{j=1}^{2}(\widetilde{w}_{j+1} - \widetilde{w}_j)\overline{G}\left[\frac{u(1 - \rho)}{q_j}\right] \quad \text{as } u \to \infty. \tag{72}$$

In particular, (72) and the assumption that $G(\cdot)$ is subexponential imply that $q_1 \hat{S}_1 + q_2 \hat{S}_2$ is a nonnegative random variable which is independent from $q_3 \hat{S}_3$ and satisfies the condition (65). Therefore, Lemma 1(v) with (71) may be applied once again in order to deduce that

$$\mathbb{P}\left\{\sum_{i=1}^{3} q_i \hat{S}_i > u\right\} \sim \frac{\lambda}{1 - \rho} \sum_{j=1}^{3}(\widetilde{w}_{j+1} - \widetilde{w}_j)\overline{G}\left[\frac{u(1 - \rho)}{q_j}\right] \quad \text{as } u \to \infty \tag{73}$$

24

such that $q_1\hat{S}_1 + q_2\hat{S}_2 + q_3\hat{S}_3$ is a nonnegative random variable which is independent of $q_4\hat{S}_4$ and satisfies the condition (65). Then, this argument may be repeated until getting

$$\mathbb{P}\left\{\sum_{i=1}^{r} q_i\hat{S}_i > u\right\} \sim \frac{\lambda}{1-\rho}\sum_{j=1}^{r}(\widetilde{w}_{j+1} - \widetilde{w}_j)\overline{G}\left[\frac{u(1-\rho)}{q_j}\right] \quad \text{as} \quad u \to \infty \quad (74)$$

such that $\sum_{i=1}^{r} q_i\hat{S}_i$ satisfies the condition (65). Therefore, (57) follows via (69) with Lemma 1(ii). Finally, since $q_r\hat{S}_r$ has a subexponential distribution, (57) implies that $\mathcal{E}_L(x, \boldsymbol{v})$ has a subexponential distribution via Lemma 1(iv) (see also Remark 13). ∎

The proof of Theorem 4 is based on an application of [23, Theorem 2] which is now provided.

Theorem 5 [Jacobovic & Mandjes 2022]
Assume that M_1, M_2, $(\psi_{m,1})_{m=1}^{\infty}$, $(\psi_{m,2})_{m=1}^{\infty}$, $(\varphi_m)_{m=1}^{\infty}$ are independent random variables such that:

1. *$M_1 \sim Poi(\lambda v)$ and $M_2 \sim Poi(\lambda x)$.*

2. *$(\psi_{m,1})_{m=1}^{\infty}$ and $(\psi_{m,2})_{m=1}^{\infty}$ are two iid sequences of random variables which are distributed like N.*

3. *$(\varphi_m)_{m=1}^{\infty}$ is an iid sequence of random variables which are distributed uniformly on $[0,1]$.*

Then,

$$\mathcal{E}_F(x, \boldsymbol{v}) \overset{d}{=} x\left(\sum_{m=1}^{M_1} \psi_{m,1} + \sum_{m=1}^{M_2} \psi_{m,2}\varphi_m\right). \quad (75)$$

Proof: [Theorem 4] Assume that $v > 0$. Otherwise, when $v = 0$, the proof follows from the same arguments and becomes much simpler.

Consider the random variables which appear in the statement of Theorem 5. In addition, $G(\cdot)$ has a subexponential distribution such that (54) is valid and hence Lemma 1(iv) implies that N has a subexponential distribution. Therefore, Lemma 1(vi) implies that $\sum_{m=1}^{M_1} \psi_{m,1}$ has a subexponential distribution such that

$$\mathbb{P}\left\{\sum_{m=1}^{M_1} \psi_{m,1} > u\right\} \sim \frac{\lambda v}{1-\rho}\overline{G}\left[\frac{u(1-\rho)}{\lambda}\right] \quad \text{as} \quad u \to \infty. \quad (76)$$

25

In addition, as it was shown above, $\psi_{1,2}$ has a subexponential distribution and φ_1 is bounded from above. As a result, Lemma 1(iii) implies that $\psi_{1,2}\varphi_1$ has a subexponential distribution. Consequently, Lemma 1(vi) yields that $\sum_{m=1}^{M_2} \psi_{m,2}\varphi_m$ has a subexponential distribution such that

$$\mathbb{P}\left\{\sum_{m=1}^{M_2} \psi_{m,2}\varphi_m > u\right\} \sim \lambda x \mathbb{P}\left\{\psi_{1,2}\varphi_1 > u\right\} \quad \text{as} \quad u \to \infty. \qquad (77)$$

To evaluate the probability in the RHS, Fatou's lemma and (54) yield

$$\liminf_{u\to\infty} \frac{\mathbb{P}\left\{\psi_{2,1}\varphi_1 > u\right\}}{\mathbb{P}\left\{\psi_{2,1} > u\right\}} = \liminf_{u\to\infty} \int_0^1 \frac{\mathbb{P}\left\{N > \frac{u}{y}\right\}}{\mathbb{P}\left\{N > u\right\}} \mathrm{d}y \geq \omega_1. \qquad (78)$$

In a similar fashion, we may apply the reverse Fatou's lemma with (54) in order to get

$$\limsup_{u\to\infty} \frac{\mathbb{P}\left\{\psi_{2,1}\varphi_1 > u\right\}}{\mathbb{P}\left\{\psi_{2,1} > u\right\}} = \limsup_{u\to\infty} \int_0^1 \frac{\mathbb{P}\left\{N > \frac{u}{y}\right\}}{\mathbb{P}\left\{N > u\right\}} \mathrm{d}y \leq \omega_2. \qquad (79)$$

Now, for every $u > 0$

$$\frac{\mathbb{P}\left\{\sum_{m=1}^{M_2} \psi_{2,m}\varphi_m > u\right\}}{\mathbb{P}\left\{\sum_{m=1}^{M_1} \psi_{1,m} > u\right\}} \leq \frac{\mathbb{P}\left\{\sum_{m=1}^{M_2} \psi_{2,m} > u\right\}}{\mathbb{P}\left\{\sum_{m=1}^{M_1} \psi_{1,m} > u\right\}}, \qquad (80)$$

and observe, using Lemma 1(vi), that the upper bound in (80) converges to $\frac{x}{v}$ as $u \to \infty$. Therefore, deduce that

$$\sup_{u>0} \frac{\mathbb{P}\left\{\sum_{m=1}^{M_2} \psi_{2,m}\varphi_m > u\right\}}{\mathbb{P}\left\{\sum_{m=1}^{M_1} \psi_{1,m} > u\right\}} < \infty \qquad (81)$$

and hence Lemma 1(v) implies that

$$\mathbb{P}\left\{\mathcal{E}_F(x, \boldsymbol{v}) > ux\right\} \sim \frac{\lambda v}{1-\rho} \overline{G}\left[\frac{ux(1-\rho)}{\lambda}\right] + \lambda x \mathbb{P}\left\{\psi_{1,2}\varphi_1 > ux\right\} \quad \text{as} \quad u \to \infty \qquad (82)$$

from which (60) follows.

Due to Lemma 1(iv), this yields that $\mathcal{E}_F(x, \boldsymbol{v})/x$ has a subexponential distribution, and hence $\mathcal{E}_F(x, \boldsymbol{v})$ also has a subexponential distribution via Lemma 1(i). ∎

7 Functional CLT

The model which was described in Section 2 is characterized by the triplet $(\lambda, G(\cdot), \boldsymbol{v})$. Fix $\boldsymbol{v} = (v_1, \ldots, v_n)$ with positive coordinates, and consider a sequence (in m) of models

$$(\lambda_m, G_m(\cdot), \boldsymbol{v}), \quad m \geq 1,$$

such that the m-th model is associated with an arrival rate λ_m and service distribution $G_m(\cdot)$. Respectively, for each $m \geq 1$ denote:

$$\mu_{k,m} \equiv \int_0^\infty t^k \mathrm{d}G_m(t), \quad k \geq 1,$$

$$\rho_m \equiv \lambda_m \mu_{1,m}.$$

In addition, for each $m \geq 1$, denote the externalities processes under LCFS-PR in the m-th model by $\mathcal{E}_L^{(m)}(\cdot, \boldsymbol{v}, n)$ and let B_m be a random variable which is distributed like the length of the busy period in the m-th model. In particular, the fixed-point relation (28) yields that for each $m \geq 1$ (see also [5, p. 251, (II.4.66)]

$$\sigma_m \equiv \mathbb{E}B_m^2 = \frac{\mu_{2,m}^3}{(1 - \rho_m)^3}, \tag{83}$$

$$\nu_m \equiv \mathbb{E}B_m^3 = \frac{1}{(1 - \rho_m)^5} \left[3\lambda_m \mu_{2,m}^2 + \mu_{3,m}(1 - \rho_m) \right], \tag{84}$$

and hence

$$\frac{\nu_m}{\sqrt{\lambda_m \sigma_m^3}} = 3\sqrt{\frac{\lambda_m}{\mu_{2,m}^5 (1 - \rho_m)}} + \frac{\mu_{3,m}}{\mu_{2,m}^{4.5}} \sqrt{\frac{1 - \rho_m}{\lambda_m}}, \quad m \geq 1. \tag{85}$$

Theorem 6 *For $m \geq 1$, define a stochastic process (in x):*

$$\hat{\mathcal{E}}_{\boldsymbol{v}}^{(m)}(x) \equiv \frac{\mathcal{E}_L^{(m)}(x, \boldsymbol{v}) - \mathbb{E}\mathcal{E}_L^{(m)}(x, \boldsymbol{v})}{\sqrt{\lambda_m \sigma_m}}, \quad x \geq 0.$$

In addition, let $W(\cdot)$ be a standard Wiener process on $[0, \infty)$ and define a stochastic process (in x):

$$H_{\boldsymbol{v}} \equiv \sum_{k=1}^n [W(x + w_k) - W(w_k)], \quad x \geq 0.$$

Assume that all of the following conditions are satisfied:

(I) $\lambda_m \to \infty$ *as* $m \to \infty$,

(II) *There is* $m' \geq 1$ *such that* $\rho_m < 1$ *for every* $m \geq m'$,

(III) $\dfrac{\nu_m}{\sqrt{\lambda_m \sigma_m^3}} \to 0$ *as* $m \to \infty$.

Then,

$$\hat{\mathcal{E}}_{\boldsymbol{v}}^{(m)}(\cdot) \Rightarrow H_{\boldsymbol{v}}(\cdot) \quad as \quad m \to \infty,$$

where $\Rightarrow$ *denotes weak convergence on* $\mathcal{D}[0, \infty)$ *equipped with the uniform metric (on compacta).*

Proof: For each $m \geq 1$, let $S^{(m)}(\cdot)$ be a compensated compound Poisson process with rate λ_m and jumps which are distributed like B_m.

Let $K > 0$ and according to [23, Theorem 5], existence of Conditions (I)-(III) all together implies that

$$\frac{S^{(m)}(\cdot)}{\sqrt{\lambda_m \sigma_m}} \Rightarrow W(\cdot) \quad as \quad m \to \infty,$$

where $\Rightarrow$ denotes weak convergence in $\mathcal{D}[0, \infty)$ equipped with the uniform metric on compacta. Since the limit process is concentrated on $C[0, x + v]$, the representation theorem [30, Theorem IV.13] yields that there is a probability space with stochastic processes $\widetilde{W}(\cdot)$ and $\left\{ \widetilde{S}_m(\cdot); m \geq 1 \right\}$ such that:

1. $\widetilde{W}(\cdot) \stackrel{d}{=} W(\cdot)$.

2. $\widetilde{S}_m(\cdot) \stackrel{d}{=} S^{(m)}(\cdot)$ for every $m \geq 1$.

3.
$$\lim_{m \to \infty} \sup_{0 \leq y \leq v+x} \left| \frac{\widetilde{S}_m(y)}{\sqrt{\lambda_m \sigma_m}} - \widetilde{W}(y) \right| \to 0 \quad , \quad \text{w.p. 1.} \tag{86}$$

Then, notice that

$$\sup_{0 \leq x \leq K} \left| \frac{1}{\sqrt{\lambda_m \sigma_m}} \sum_{i=1}^{n} \left[\widetilde{S}_1(x + w_i) - \widetilde{S}_1(w_i) \right] - \sum_{i=1}^{n} \left[W(x + w_i) - W(w_i) \right] \right|$$

$$\leq 2n \sup_{0 \leq y \leq v+x} \left| \frac{\widetilde{S}_m(y)}{\sqrt{\lambda_m \sigma_m}} - \widetilde{W}(y) \right|, \tag{87}$$

28

and the RHS tends to zero with probability one. As a result, due to Theorem 1, deduce that the sequence of processes $\hat{\mathcal{E}}_{\boldsymbol{v}}^{(m)}(\cdot)$ admits weak convergence to $H_{\boldsymbol{v}}(\cdot)$ on $\mathcal{D}[0, K]$ equipped with the uniform metric. Since this convergence holds for every positive K and the limiting process $H_{\boldsymbol{v}}(\cdot)$ is concentrated in $C[0, \infty)$, this convergence can be extended to $\mathcal{D}[0, \infty)$ via [30, Theorem V.23] and the result follows. ∎

7.1 The conditions

The pre-conditions of Theorem 6 are not straightforward. Namely, Condition (I) states that $\lambda_m \to \infty$ as $m \to \infty$. Thus, Condition (II) implies that $\mu_{1,m}$ vanishes as $m \to \infty$. In addition, since $\lambda_m \to \infty$ as $m \to \infty$, it must be that $\mu_{2,m} \to \infty$ as $m \to \infty$ in order to make the first term of (85) vanish as $m \to \infty$.

Below, there is an example in which the pre-conditions of Theorem 6 are all satisfied. This illustrates that Theorem 6 is informative. Having said that, it might be interesting to look for more general conditions under which the asymptotic distribution of the externalities process can be identified.

7.1.1 Example

For each $m \geq 1$, assume that $G_m(\cdot)$ is a distribution function of a random variable τ_m such that

$$\tau_m \equiv \begin{cases} 0 & \text{w.p} \quad 1 - m^{-p} \\ m^k & \text{w.p} \quad m^{-p} \end{cases} \tag{88}$$

for some $p, k > 0$. Then, for each $m \geq 1$, observe that

$$\mu_{1,m} = m^{k-p} \quad , \quad \mu_{2,m} = m^{2k-p} \quad , \quad \mu_{3,m} = m^{3k-p}. \tag{89}$$

Also, let $\lambda_m \equiv q m^{p-k}$ for some $q \in (0,1)$. Thus, in order to ensure that Condition (I) is satisfied, require that $p > k$ and observe that $\rho_m = q$ which means that Condition (II) is also valid. It is left to check Condition (III). To this end, note that

$$\frac{\lambda_m}{\mu_{2,m}^5 (1 - \rho_m)} = \frac{q}{1-q} m^{6p-11k} \tag{90}$$

and hence we require $6p - 11k < 0$. Similarly, note that

$$\frac{\mu_{3,m}}{\mu_{2,m}^{4.5}} \sqrt{\frac{1 - \rho_m}{\lambda_m}} = m^{-5.5k+3p} \sqrt{\frac{1-q}{q}} \tag{91}$$

and hence we require $-5.5k + 3p < 0$. Thus, for any $p, k > 0$ such that $\frac{6p}{11} < k < p$, all the pre-conditions of Theorem 6 are satisfied. Indeed, one may verify that for such p and k, we get that $\mu_{1,m} \to 0$ as $m \to \infty$ and $\mu_{2,m} \to \infty$ as $m \to \infty$.

7.2 Open problem: Is $\mathcal{E}(x, \boldsymbol{v})$ asymptotically Gaussian?

Theorem 6 is a functional CLT for the process $x \mapsto \mathcal{E}_L(x, \boldsymbol{v})$. In previous work [23, Theorem 4], there is an analogous result for $x \mapsto \mathcal{E}_F(x, \boldsymbol{v})$. Both of these univariate results are based on the bivariate decomposition which is given in Theorem 1 with an application of [23, Theorem 5]. Since the decomposition in Theorem 1 is bivariate, it looks promising to reach a unified version of both of these results, i.e., a bivariate functional CLT of $x \mapsto (\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v}))$. The proof of such a result might be based on a complicated bivariate version of [23, Theorem 5] which should be developed. Since we prefer not to shift the balance of the current work by developing this generalization of [23, Theorem 5], here we just mention the possibility of having a bivariate functional CLT of $x \mapsto (\mathcal{E}_L(x, \boldsymbol{v}), \mathcal{E}_F(x, \boldsymbol{v}))$ in passing.

8 Discussion

This work concerns the externalities in an M/G/1 queue under the LCFS-PR and FCFS disciplines. The cornerstone of the current analysis is Theorem 1: The externalities have a joint decomposition in terms of a bivariate compound Poisson process with an arrival rate λ and jumps which are distributed like (B, N). We demonstrated that this decomposition is useful via several applications, including: moments computations, externalities approximations as $x \to \infty$, tail asymptotics of the externalities and a functional CLT. The rest of this section is dedicated to a discussion about the special meaning of the externalities analysis under FCFS and LCFS-PR.

Observe that when FCFS is implemented, then the additional customer affects only customers who arrive *after* him. On the contrary, when LCFS-PR is implemented, the additional customer would affect only the customers who arrived *before* him. In that sense, FCFS and LCFS-PR are edge cases. As demonstrated in Subsection 3.2, there are service disciplines under which the additional customer affects both some customers who were in the queue at time $t = 0$ as well as some other customers who arrive after this time. The externalities in these intermediate cases have not been analysed. It might be interesting to examine in what sense their properties mitigate the properties of the externalities under FCFS and LCFS-PR.

Acknowledgments

References

[1] Asmussen, S., Klüppelberg, C., & Sigman, K. (1999). Sampling at subexponential times, with queueing applications. *Stochastic Processes and their Applications*, **79**, 265-286.

[2] Baltrunas, A., Daley, D.J., & Klüppelberg, C. (2004). Tail behaviour of the busy period of a GI/GI/1 queue with subexponential service times. *Stochastic Processes and their Applications*, **111**, 237-258.

[3] Chan, C. W., Huang, M., & Sarhangian, V. (2021). Dynamic server assignment in multiclass queues with shifts, with applications to nurse staffing in emergency departments. *Operations Research*, **69**, 1936-1959.

[4] Cline, D. B., & Samorodnitsky, G. (1994). Subexponentiality of the product of independent random variables. *Stochastic Processes and their Applications*, **49**, 75-98.

[5] Cohen, J. W. (1969). *The Single Server Queue.* North-Holland Publishing Company.

[6] Cohen, J.W. (1973). Some results on regular variation for distributions in queueing and fluctuation theory. *Journal of Applied Probability*, **10**, 343-353.

[7] Dębicki, K., & Mandjes, M. (2015). *Queues and Lévy Fluctuation Theory.* Springer.

[8] De Meyer, A., & Teugels, J.L. (1980). On the asymptotic behaviour of the distributions of the busy period and service time in M/G/1. *Journal of Applied Probability*, **17**, 802-813.

[9] Denisov, D.E., & Shneer, V. (2010). Global and local asymptotics for the busy period of an M/G/1 queue. *Queueing Systems*, **64**, 383-393.

[10] Doney, R.A. (1989). On the asymptotic behaviour of first passage times for transient random walks. *Probability Theory and Related Fields*, **81**, 239-246.

[11] Embrechts, P., Goldie, C.M., & Veraverbeke, N. (1979). Subexponentiality and infinite divisibility. *Probability Theory and Related Fields*, **49**, 335-347.

[12] Embrechts, P., & Goldie, C. M. (1980). On closure and factorization properties of subexponential and related distributions. *Journal of the Australian Mathematical Society*, **29**, 243-256.

[13] Foss, S., Korshunov, D., & Zachary, S. (2011). *An Introduction to Heavy-tailed and Subexponential Distributions*. Springer.

[14] Geluk, J., & Tang, Q. (2009). Asymptotic tail probabilities of sums of dependent subexponential random variables. *Journal of Theoretical Probability*, **22**, 871-882.

[15] Goldie, C. M. (1978). Subexponential distributions and dominated-variation tails. *Journal of Applied Probability*, **15**, 440-442.

[16] Grandell, J. (1997). *Mixed Poisson Processes (Vol. 77)*. CRC Press.

[17] Haviv, M., & Ritov, Y. A. (1998). Externalities, tangible externalities and queue disciplines. *Management Science*, **44**, 850-858.

[18] Hazra, R. S., & Maulik, K. (2012). Tail behavior of randomly weighted sums. *Advances in Applied Probability*, **44**, 794-814.

[19] Hu, Y., Chan, C. W., & Dong, J. (2022). Optimal scheduling of proactive service with customer deterioration and improvement. *Management Science*, **68**, 2533-2578.

[20] Huang, J., Carmeli, B., & Mandelbaum, A. (2015). Control of patient flow in emergency departments, or multiclass queues with deadlines and feedback. *Operations Research*, **63**, 892-908.

[21] Jacobovic, R. (2022). Internalization of externalities in queues with discretionary services. *Queueing Systems*, **100**, 453-455.

[22] Jacobovic, R. (2022). Regulation of a single-server queue with customers who dynamically choose their service durations. *Queueing Systems*, **101**, 245–290.

[23] Jacobovic, R., & Mandjes, M. (2022). Externalities in queues as stochastic processes: The case of FCFS M/G/1. Pre-print, arXiv:2207.02599 [math.PR].

[24] Jelenkovic, P. R., & Momcilovic, P. (2004). Large deviations of square root insensitive random sums. *Mathematics of Operations Research*, **29**, 398-406.

[25] Kelly, F. P. (1976). The departure process from a queueing system. *Mathematical Proceedings of the Cambridge Philosophical Society*, **80**, 283-285.

[26] Klüppelberg, C. (1988). Subexponential distributions and integrated tails. *Journal of Applied Probability*, **25**, 132-141.

[27] Kulik, R., & Soulier, P. (2020). *Heavy-tailed Time Series.* Springer.

[28] Liu, Y., Sun, X., & Hovey, K. (2022). Scheduling to differentiate service in a multiclass service system. *Operations Research*, **70**, 527-544.

[29] Markowitz, H. M., & Todd, G. P. (2000). *Mean-variance analysis in portfolio choice and capital markets (Vol. 66)*. John Wiley & Sons.

[30] Pollard, D. (2012). *Convergence of Stochastic Processes.* Springer.

[31] Tang, Q. (2008). Insensitivity to negative dependence of asymptotic tail probabilities of sums and maxima of sums. *Stochastic Analysis and Applications*, **26**, 435-450.

[32] Zhang, Y., Shen, X., & Weng, C. (2009). Approximation of the tail probability of randomly weighted sums and applications. *Stochastic Processes and their Applications*, **119**, 655-675.

[33] Zwart, A. P. (2001). Tail asymptotics for the busy period in the GI/G/1 queue. *Mathematics of Operations Research*, **26**, 485-493.