

Article



Regenerative Analysis and Approximation of Queueing Systems with Superposed Input Processes

Irina Peshkova ^{1,†}, Evsey Morozov ^{1,2,3,†} and Michele Pagano ^{4,*,†}

- ¹ Department of Applied Mathematics and Cybernetics, Petrozavodsk State University, Lenin str. 33, 185910 Petrozavodsk, Russia; iaminova@petrsu.ru (I.P.); emorozov@karelia.ru (E.M.)
- ² Institute of Applied Mathematical Research, Karelian Research Centre of Russian Academy of Sciences, Pushkinskaja str. 11, 185910 Petrozavodsk, Russia
- ³ Moscow Center for Fundamental and Applied Mathematics, Lomonosov Moscow State University, GSP-1, Leninskie Gory, 119991 Moscow, Russia
- ⁴ Department of Information Engineering, University of Pisa, Via G. Caruso 16, 56122 Pisa, Italy
- * Correspondence: michele.pagano@unipi.it
- ⁺ These authors contributed equally to this work.

Abstract: A single-server queueing system with *n* classes of customers, stationary superposed input processes, and general class-dependent service times is considered. An exponential splitting is proposed to construct classical regeneration in this (originally non-regenerative) system, provided that the component processes have heavy-tailed interarrival times. In particular, we focus on input processes with Pareto interarrival times. Moreover, an approximating *GI/G/1*-type system is considered, in which the independent identically distributed interarrival times follow the stationary Palm distribution corresponding to the stationary superposed input process. Finally, Monte Carlo and regenerative simulation techniques are applied to estimate and compare the stationary waiting time of a customer in the original and in the approximating systems, as well as to derive additional information on the regeneration cycles' structure.

Keywords: superposition of arrival streams; regeneration; splitting; finite mixture

MSC: 60K25

1. Introduction

In this research, we consider a single-server FIFO (First In First Out) queueing system $\sum_{i=1}^{n} GI_i/G/1$, with a superposition of *n* independent renewal stationary inputs (components) corresponding to customers of different classes. The superposition may include component processes with general (component-dependent) interarrival times. The service time distributions are also assumed to be general and class-dependent. (In what follows, we use the terms 'distribution' and 'distribution function' as synonymous terms.)

The superposed input process system has been intensively studied since the 1970s (see, for example, [1–7], among others). It is well known that, unless the components are Poissonian, the superposed input process system is not (classically) regenerative. On the other hand, the regeneration property is quite important for both the stability and performance analysis of the queueing system. Indeed, it admits accurate estimation of stationary performance metrics on the basis of regenerative simulation [8–11], which allows one to use the independence of regeneration cycles (of a basic process) in the framework of the regenerative variant of the Central Limit Theorem (CLT) in order to estimate the required metrics with a given precision. (In Section 5, we discuss this topic in more detail.) What is especially important is that the values of a queueing process, within a regeneration cycle, are typically highly correlated, and a direct application of the classical variant of the CLT is not justified.



Citation: Peshkova, I.; Morozov, E.; Pagano, M. Regenerative Analysis and Approximation of Queueing Systems with Superposed Input Processes. *Mathematics* **2024**, *12*, 2202. https://doi.org/10.3390/ math12142202

Academic Editors: Anatoly Nazarov and Alexander Dudin

Received: 23 May 2024 Revised: 23 June 2024 Accepted: 25 June 2024 Published: 13 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

In general, the superposed process generated by independent renewal processes can be transformed, under mild assumptions, into a regenerative process with the so-called onedependent regeneration cycles (see [6,12]). This construction, which is based on splitting and coupling procedures, is rather complicated and hardly applied in practice to evaluate the required performance measures by simulation. (We will discuss it in Section 3.) Moreover, to apply this approach for accurate estimation, we must overcome an even bigger problem; we need to construct regeneration events of the queueing process (not only of the input process), which, in turn, requires the synchronization of a regeneration point of the input process with an empty state of the system. For this reason, we instead will base our analysis on the construction of the artificial regeneration obtained by *exponential splitting*. (For a deeper understanding of the splitting method, we refer readers to the fundamental monograph [13].) A key property opening the possibility of this construction in the initially non-regenerative process is that the interarrival distributions of the component processes are assumed to be *heavy-tailed*. Thus, in this research, we focused on the case where the superposition consists of n_1 Poisson inputs and n_2 inputs with Pareto interarrival times. The splitting approach, in keeping the distributional properties of the original process, allows one to use regenerative simulation for accurate estimations of the required stationary performance metrics. To the best of our knowledge, this approach is realized for the first time in systems with a superposed input process, and this finding is the main contribution of this research.

Our second main purpose is to study the basic system by means of another *classically* regenerative system with a single renewal input process. It is not a new approach as, for instance, it has been realized in a series of papers, such as [3,5,7]. In these works, the socalled stationary-interval method (based on the Palm approach) was used to approximate the distribution of the interarrival times in the superposition process by a stationary distribution with independent identically distributed (iid) intervals. We emphasize that such a replacement is indeed an approximation because the successive intervals in the superposed process are dependent in general. The latter issue has been analyzed in [1], in which the correlation structure of the successive interevent times was considered in detail. Another motivation of the approximation, as we mentioned, is that the target processes in the basic system (for example, workload or queue size) are not regenerative; however, a regenerative structure is highly desirable both to establish the stability of the system and to accurately estimate its stationary performance metrics. An issue in the study of queueing systems with superposed input (with independent renewal components) concerns the approximation of such a process by a Poisson process. It is a well-known problem going back to A. Khintchine [14]. More exactly, when the number of summands (components) increases, while the (properly defined) traffic intensity ρ remains bounded (and not close to 1), then the superposed process approaches a Poisson process, with the rate being the sum of the rates of the summands. On the other hand, it has been shown in [2] that, as the number *n* of the summands increases, this approximation works well only if the product $n(1-\rho)^2 \gg 1$. The latter result indicates that the Poissonian approximation is not suitable when the number of components is small and/or the traffic intensity ρ is close to 1, that is, the system is close to 'saturation'. The saturated system (in the so-called heavy-traffic regime) with the superposed input process is studied in the paper [15]. Notice also that the basic results for ordinary renewal process, obtained originally by W. Smith [16] and D. Cox [17], have been extended to the superposition of independent renewal processes in [4].

Thus, our second important contribution is the detailed analysis of the approximating GI/G/1-type system when the interarrival time distribution is a mixture containing Pareto distributions, and when the service time distribution is also a mixture of distributions that describe the service times of the customers in the component processes. Indeed, we show that, in the approximating system, the stationary workload distribution is analytically available when the service times of the different classes of customers have a common exponential distribution (Section 4.2); this is also a new contribution of this research. Moreover, we compare the accuracy of the approximation in terms of the Kolmogorov

distance between the (empirical) distributions of the stationary waiting times in the original and approximated systems (Section 6.1). In particular, in spite of the presence of heavytailed components in interarrival distribution, the simulation results show quite acceptable distance values (less than 5%) for most of the considered cases (when the traffic intensity $\rho < 0.6$), and it is consistent with the results of the abovementioned works [3,5,7]. Hence, the proposed approximation may be acceptable if the splitting construction turns out to be too complicated, e.g., if the number of component processes is very large. Another possible application of the approximating system is as a base for appropriate lower/upper bounds for the stationary workload process, exploiting the known monotonicity properties of the classical queueing system under some mild assumptions. (In this regard, see [18] and recent papers of one of the authors [19,20] on the monotonicity of systems with mixed service times.)

As mentioned above, the approximation of the superposed input process by a renewal process is indeed a mixed distribution, consisting of an appropriately weighted sum of the stationary distributions of the remaining renewal times (integrated-tail distributions) in the component stationary renewal processes. This distribution, based on the Palm probability concept, is well known (see, for instance, [7,17]) and is the exact one-dimensional distribution of the interarrival times, which, however, does not capture possible dependencies between the neighboring intervals. For this reason, reducing the superposed input process to an ordinary renewal process is indeed an approximation. We note that, in general, a dependence between the adjacent intervals cannot be ignored in the careful study of the superposed input, as shown in [5].

The rest of the paper is organized as follows. In Section 2, we give some important preliminary results from the theory of renewal and point processes with elements of the Palm theory. In Section 3, we outline the construction of *one-dependent regeneration* using the approach from [12] and highlight the difficulties arising in its practical realization in simulation. Nevertheless, we use some basic ideas of this approach to construct regeneration events by the exponential splitting of heavy-tailed interarrival times.

In Section 4, we examine an approximating GI/G/1 system with iid interarrival times, generated by the Palm distribution of the stationary intervals in the original superposed input process. We then focus on special cases where we find explicit expressions for the Laplace–Stieltjes transform (LST) of the stationary workload. In Section 5, we introduce exponential splitting and explain how it is used to obtain classical regenerations in the superposed process generated by Pareto components. Next, in Section 6, we analyze our theoretical findings through discrete-event simulations. Specifically, we first evaluate the accuracy of the input process approximation in terms of Kolmogorov distance for different system parameter settings. Then, we apply exponential splitting to construct a confidence estimate of the empirical mean of the stationary waiting time in the original system. Finally, Section 7 concludes the paper by providing an overview of its main contributions.

2. Preliminary Results from the Theory of Renewal and Point Processes

In this section, we present theoretical results from the theory of renewal and point processes [21,22]. These results are then utilized in the analysis of queueing systems with the superposed input process generated by independent renewal processes. Among the most important sources on renewal theory, besides the pioneering work [23], we also mention the recognized monographs [24,25].

To start, let us construct a stationary renewal process. To do this, consider iid nonnegative random variables (r.v.s) $\{X_k, k \ge 2\}$ and an independent r.v. $X_1 \ge 0$. We denote by *F* the distribution function of an r.v. *X*, which is a stochastic copy of any X_k , $k \ge 2$. It is assumed that the renewal rate $\lambda = 1/EX < \infty$. Let

$$G(x) = \lambda \int_0^x (1 - F(u)) du, \ x \ge 0,$$
(1)

be the distribution of the first r.v. X_1 . Also, denote by $T_k = \sum_{1 \le i \le k} X_i$ the renewal instants, where $k \ge 1$. Hence, $N_s(t) := \max\{k > 0 : T_k \le t\}$ represents the number of renewal events that occur in the time interval [0, t], with $N_s(0) = 0$. The process $\{N_s(t), t \ge 0\}$ is a time-stationary renewal process [22,26]. Specifically, the stationary renewal process satisfies $EN_s(t) = \lambda t$, $t \ge 0$, and the distribution of the increments $N_s(t+u) - N_s(u)$ does not depend on $u \ge 0$ for each fixed t. There are other equivalent statements to identify a stationary renewal process (see, for example, Chapter 2 in [25]). It is also known that (1) represents the distribution of the stationary remaining (and attained) renewal time, and it is referred to as the integrated-tail distribution [12].

In the following analysis, we will consider n independent stationary renewal processes that model the arrivals of different classes of customers. Then, the renewal points $\{T_k^{(i)}, k \ge 1\}$ correspond to the arrival instants of class-i customers from process i, and let $X^{(i)}$ denote the generic interarrival time (with a rate of $\lambda_i = 1/\mathsf{E}X^{(i)}$) in process i, where $i = 1, \ldots, n$. These processes will compose a merged superposed input process. To provide a more detailed description, let $\mathcal{T} = \{T_k^{(i)}, k \ge 1; i = 1, \ldots, n\}$ be the overall set of renewal points for all n processes, and let $\{T_k\}$ be the ordered sequence of these renewal instants. In other words, T_1 is the first smallest element of \mathcal{T} , T_2 is the second smallest element, and so on.

Assuming, for simplicity, that there are no batch arrivals, it follows that the sequence $\{T_k\}$ satisfies

$$0 < T_1 < T_2 < \cdots \tag{2}$$

It is clear that, unlike a renewal process, the distances between events (arrivals) in the superposed process are, in general, dependent.

Since the component renewal processes are assumed to be time-stationary, then the superposed process, defined (in evident notation) as

$$N_s(t) := N_s^{(1)}(t) + \dots + N_s^{(n)}(t), \ t \ge 0,$$
(3)

is also time-stationary (counting), but not renewal in general. (A detailed proof can be found, for example, in [21], Chapter 1.)

To explain this process in more detail, we recall some findings from the theory of stationary point processes, while maintaining the previously established notation. Following [21,25], we consider a set of points (occurrence times) $\{T_k\}$ on the positive real line that satisfy (2), with the condition that $T_0 = 0$ is an occurrence time. This process is assumed to be interval-stationary and is referred to as the stationary Palm process. We also consider a counting process $\{N(t), t \ge 0\}$, which records the number of points in the interval [0, t], formally defined as

$$N(t) = \max\{k \ge 0 : T_k \le t\}, \ t \ge 0,$$
(4)

with the assumption that N(0) = 1. Let $X_k := T_k - T_{k-1}$ denote the interval between the *k*th and (k - 1)st points, where $k \ge 1$. These intervals are identically distributed with distribution function *F*. This new process is also referred to as the stationary Palm process [7,21,22].

In the following discussion, we will distinguish between the Palm process (4) and the related time-stationary (also counting) process $\{N_s(t), t \ge 0\}$. Note that an interval sequence $\{X_i\}$ (forming the Palm process) is stationary if the joint distribution of the *k*-tuple $(X_{i_1+h}, \ldots, X_{i_k+h})$ is independent of the positive integer *h* for all integers k > 0and all non-negative (i_1, \ldots, i_k) . It is worth pointing out that the Palm process $\{N(t)\}$ and the associated sequence $\{X_k\}$ are usually not both stationary at the same time [7]. However, there is a one-to-one correspondence between the stationary counting process $\{N(t)\}$ and the stationary sequence $\{X_k\}$ based on the Palm theory [21]. To establish this correspondence, let us introduce the *Palm function* $\varphi_k(t) = P(N(t) = k)$, which represents the probability of having k points ($k \ge 0$) in a time interval of length t provided that the beginning of the interval is an occurrence point. Hence, Palm functions describe the dynamics of the Palm process (4). The distribution of the process { $N_s(t)$ } associated with the Palm process (4) can be expressed via Palm functions by the following relations [7,14]:

$$P(N_{s}(t) = k) = \lambda \int_{0}^{t} (\varphi_{k-1}(u) - \varphi_{k}(u)) du, \ k \ge 1,$$

$$P(N_{s}(t) = 0) = 1 - \lambda \int_{0}^{t} \varphi_{0}(u) du =: 1 - G(t),$$
(5)

$$G(t) = \lambda \int_{0}^{t} (1 - F(x)) dx, \ t \ge 0,$$
(6)

implying, in particular, the following key connection:

$$\varphi_0(t) = 1 - F(t) =: \overline{F}(t), \ t \ge 0,$$
(7)

where, as already stated, *F* is the distribution function of the intervals in the stationary Palm process. (In what follows, for any distribution function *F*, we denote by \overline{F} its tail.) Note that, by construction, the stationary process $\{N_s(t)\}$ satisfies both required conditions (see [7]):

$$\mathsf{E}N_s(t)<\infty,\quad \lim_{t o\infty}N_s(t)/t=:\lambda<\infty.$$

Recall that $\{N_s^{(i)}(t)\}$ represents the *i*th component (stationary renewal process) and we will supply with index *i* the corresponding quantities in process *i*. Note that, by (1), the first interval in the process *i* has distribution

$$G_i(t) = \lambda_i \int_0^t (1 - F_i(u)) du, \tag{8}$$

and rate $\lambda_i = 1/\mathsf{E}X^{(i)}$, i = 1, ..., n. Let $\lambda = \lambda_1 + \cdots + \lambda_n$. It is known [7,14,21] that the Palm functions $\varphi_k(t)$ of the superposed stationary Palm process $\{N(t)\}$ are expressed via the original time-stationary processes $\{N_s^{(i)}(t)\}$ and their stationary Palm versions $\{N^{(i)}(t)\}$ as follows:

$$\varphi_{k}(t) = \sum_{i=1}^{n} \frac{\lambda_{i}}{\lambda} \mathsf{P}\left(\sum_{j=1}^{i-1} N_{s}^{(j)}(t) + N^{(i)}(t) + \sum_{j=i+1}^{n} N_{s}^{(j)}(t) = k\right)$$

$$= \sum_{i=1}^{n} \frac{\lambda_{i}}{\lambda} \mathsf{P}(N^{(i)}(t) = k_{i}) \prod_{j,j \neq i}^{n} \mathsf{P}(N_{s}^{(j)}(t) = k_{j}),$$
(9)

where the summation is taken over all different partitions of $k \ge 0$ such that $k = k_1 + \cdots + k_n$. In particular, by (8), the tail of the distribution *F* satisfies [7]:

$$\overline{F}(t) = \varphi_0(t) = \mathsf{P}(N(t) = 0) = \sum_{i=1}^n \frac{\lambda_i}{\lambda} (1 - F_i(t)) \prod_{j,j \neq i} G_j(t)$$
$$= \sum_{i=1}^n \frac{\lambda_i}{\lambda} (1 - F_i(t)) \prod_{j,j \neq i} \lambda_j \int_t^\infty (1 - F_j(u)) du, \ t \ge 0.$$
(10)

It is easy to verify that $\overline{F}(0) = \varphi(0) = 1$, as required. We emphasize that the Palm process $\{N(t)\}$ is interval-stationary, while the process $\{N_s(t)\}$ is time-stationary. Fur-

thermore, since, for a single component process with interevent distribution *F* and rate λ , we have

$$\lambda \int_0^\infty \overline{F}(t)dt = 1,$$

Equation (5) implies that $P(N_s(t) = 0) = G(t)$, satisfying (6). Additionally, expression (10) has a straightforward intuitive interpretation.

Since a renewal process is a particular case of the point process, it is worth mentioning that *F* represents the exact marginal (one-dimensional) distribution of the interarrival time in the superposed Palm process $\{N(t)\}$. However, this distribution does not capture the dependence between adjacent intervals. As a result, expression (10) only provides an approximation of the original superimposed process, embedded at the instances of the points' occurrences.

Therefore, it is crucial to examine the accuracy of this approximation using various metrics. Specifically, as detailed in Section 6, the approximation (10) of the original $\sum_{i=1}^{n} G_i/G_i/1$ system yields an acceptable Kolmogorov distance between the empirical distributions of the stationary waiting times when $\rho < 0.6$.

Remark 1. Representations (10), (9) hold for any superposition in which the components $\{N_s^{(i)}(t), t \ge 0\}$ are time-stationary processes, not necessary renewal. In this general case, the intensities are defined as $\lambda_i := \mathsf{E}N_s^{(i)}(1), i = 1, ..., n$ [7].

3. Splitting of a Superposed Process

In general, for superposed processes, it is only possible to construct the so-called one-dependent regenerations. In this section, we will briefly describe this construction following [12]. Roughly speaking, in a one-dependent regenerative process, two consecutive regeneration cycles are dependent, but the cycles separated by at least one cycle are independent. However, as we will discuss below, this approach is challenging to apply in simulations. Therefore, in Section 5, we will introduce exponential splitting, which can be effectively applied when the renewal intervals follow heavy-tailed distributions [22].

Let us describe the construction of one-dependent regenerations for the superposition of two renewal processes, with inter-renewal time distributions F_i , i = 1, 2. Firstly, we assume that both these processes are positive recurrent, i.e., the mean inter-renewal time is finite for both processes. Furthermore, let us assume that F_1 satisfies the following condition:

$$F_1(dx) \ge cdx \text{ for } x \in (a, a+2b) \tag{11}$$

for some positive constants a, b, and c. This means that F_1 is a particular case of a spread-out distribution F. (In the general case, some convolution of F with itself must satisfy (11); see [22].)

Denote by $R_i(t)$ the (right-continuous) remaining renewal time at instant t in the *i*th renewal process, i = 1, 2; $t \ge 0$. Then, under assumption (11), the distribution of $R_1(t)$ has the following smoothness property:

$$\mathsf{P}(R_1(t) \in B) \ge \delta \nu(B), \ t \ge C,\tag{12}$$

for each Borel set *B* on (0, b), where ν is the uniform measure on (0, b), $\delta > 0$, and *C* is some constant. (Condition (12) is sometimes called minorization.) That is, for large enough *t*, the distribution of $R_1(t)$ has a uniform component. Let $R_t = (R_1(t), R_2(t))$ and denote by $(t_n^{(i)}, n \ge 1)$ the renewal points of the process i = 1, 2. Let

$$t_0 = 0, \ L_0 = R_1(t_0), \ t_1 = \min(t_n^{(2)} : t_n^{(2)} \ge t_0 + L_0 + C),$$
 (13)

where *C* satisfies (12). Also, denote by $\{U_i, i \ge 1\}$ the iid r.v.s with uniform distribution ν (in what follows, we will denote it as $U \sim \nu$.) Also, let $\{V_i, i \ge 1\}$ be iid 0 - 1 Bernoulli r.v.s such that $P(V = 1) = \delta$. Take

$$R_1(t_1) = U_1 V_1 + (1 - V_1) Z_1,$$

where the r.v. Z_1 has the rest distribution (see (12))

$$\frac{\mathsf{P}(R_1(t)\in \cdot) - \delta\nu(\cdot)}{1-\delta}.$$
(14)

Continuing as in (13), define recursively

$$t_{k} = \min(t_{i}^{(2)} : t_{i}^{(2)} \ge t_{k-1} + L_{k-1} + C),$$

$$L_{k} = R_{1}(t_{k}) = U_{k}V_{k} + (1 - V_{k})R_{k}, \ k \ge 1,$$
(15)

where U_k , V_k , R_k are independent of all U_l , V_l , R_l , l < k. Eventually, let

$$\widehat{T}_1 = \min(t_i : V_i = 1), \ \widehat{T}_{k+1} = \min(t_i > \widehat{T}_k : V_i = 1), \ k \ge 1.$$
 (16)

Then, the sequence $\{\hat{T}_k\} \subset \{t_k^{(2)}\}$ and moreover, at each instant \hat{T}_k , the remaining times vector has distribution

$$R_{\widehat{T}_{k}} = (R_{1}(\widehat{T}_{k}), R_{2}(\widehat{T}_{k})) \sim \nu \otimes F_{2}, \ k \ge 1,$$

$$(17)$$

which is independent of the pre-history prior to instant \hat{T}_{k-1} . (Here, $\nu \otimes F_2$ denotes the distribution generated by the marginal distributions ν and F_2 .) Thus, $\{\hat{T}_k\}$ are one-dependent regeneration points for the superposed remaining time process $\{R_t\}$, and, according to (14)–(16), they constitute a subsequence of the renewal points of input 2. The extension of this construction to arbitrary n renewal processes is straightforward [12].

Remark 2. It is worth pointing out that, under classical regeneration, the distribution of $R_{\widehat{T}_k}$ in (17) would be independent of the pre-history prior to instant \widehat{T}_k .

The primary challenge in practically implementing regenerations in the superposed process is that the constants in Equation (12) are not explicitly defined. An even more difficult problem is the need to synchronize these regeneration points with empty states of the queueing system to obtain regenerations of the entire system. However, in some cases, particularly when the inter-renewal intervals follow heavy-tailed distributions, this difficulty can be overcome. In this case, as we will show in Section 5, by utilizing the aforementioned construction of one-dependent regenerations, the splitting procedure can be effectively applied to construct the classical regenerations of the superposed process.

4. GI/G/1 Approximation

In this section, we assume that in the basic system $\sum_{i=1}^{n} G_i / G_i / 1$ (further denoted by Σ) with *n* independent renewal processes, the *i*th component has iid interarrival times $\{\tau_k^{(i)}, k \ge 2\}$ with distribution A_i , and the distribution of the first interarrival time (up to the first arrival) satisfies (cf. (1), (8))

$$\mathsf{P}(\tau_1^{(i)} \le x) = \frac{1}{\mathsf{E}\tau_2^{(i)}} \int_0^x (1 - A_i(u)) du, \ i = 1, \dots, n.$$

Customers from input *i* (called class-*i* customers) have iid (class-dependent) service times $\{S_k^{(i)}, k \ge 1\}$ with a general distribution B_i , i = 1, ..., n. In the following analysis, we will focus on exponential and Pareto distributions for A_i . Then, for specific cases, we

will construct the Palm distribution (10) to use it as the distribution of the iid interarrival times in an approximating GI/G/1 queueing system (denoted by $\hat{\Sigma}$).

4.1. Exponential–Pareto Input Process

Assume that out of the total *n* inputs, the first n_1 components have exponential interarrival times, $\tau_k^{(i)} \sim Exp(\lambda_i)$, $k \ge 1$; $i = 1, ..., n_1$,

$$A_i(x) = 1 - e^{-\lambda_i x}, \ x \ge 0, \ \lambda_i > 0, \ i = 1, \dots, n_1,$$
 (18)

while the remaining inputs have interarrival times following a type II Pareto distribution, $Pareto(x_0, \alpha_j), j = 1, ..., n_2$, i.e.,

$$\mathsf{P}(\tau_k^{(j)} \le x) =: A_j(x) = 1 - \left(\frac{x_0}{x_0 + x}\right)^{\alpha_j}, \ x \ge 0, \ k \ge 1, \ x_0 > 0, \ \alpha_j > 1, \ j = 1, \dots, n_2.$$
(19)

Let $\mu_i := 1/ES_k^{(i)}$ denote the mean service rate of the *i*-th class customers, where i = 1, ..., n. In the following, we use index *i* to represent the quantities associated with the Poisson inputs and index *j* to represent the quantities related to Pareto input processes. By substituting the distributions (18) and (19) into Formula (10), we can obtain the distribution of the stationary renewal input process in the following form:

$$A(x) = 1 - pe^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha - n_2} - (1 - p)e^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha - n_2 + 1},$$
(20)

where

$$\lambda = \lambda_1 + \cdots + \lambda_{n_1}, \ \alpha = \alpha_1 + \cdots + \alpha_{n_2},$$

and

$$p = \frac{\lambda x_0}{\lambda x_0 + \alpha - n_2}.$$
(21)

Note that the tail distribution of (20) is the following mixture:

$$\overline{A}(x) = p\overline{F}_1(x) + (1-p)\overline{F}_2(x),$$
(22)

of the two-tail distributions

$$\overline{F}_{1}(x) = e^{-\lambda x} \left(\frac{x_{0}}{x_{0}+x}\right)^{\alpha-n_{2}},$$

$$\overline{F}_{2}(x) = e^{-\lambda x} \left(\frac{x_{0}}{x_{0}+x}\right)^{\alpha-n_{2}+1},$$
(23)

with the mixing proportion p defined by (21). Denote by τ the (generic) interarrival time following distribution (20) in the renewal process defined by distribution (22). Let distribution F_1 define an r.v. Y, and distribution F_2 define an r.v. Z. Then, the r.v. τ can be expressed as a two-component mixture:

$$\tau = IY + (1 - I)Z,\tag{24}$$

where *I* is the indicator function with P(I = 1) = p and

$$Y = \min(Y_1, Y_2) \text{ with } Y_1 \sim Exp(\lambda), \quad Y_2 \sim Pareto(x_0, \alpha - n_2),$$

$$Z = \min(Y_1, Z_1) \text{ with } Z_1 \sim Pareto(x_0, \alpha - n_2 + 1).$$

Moreover, the tail distribution of the service time is also determined by the following mixture of the predefined class-dependent distributions $\{B_k\}$, namely:

$$\overline{B}(x) = \sum_{i=1}^{n_1} p_i \overline{B}_i(x) + \sum_{j=1}^{n_2} p_j \overline{B}_j(x),$$
(25)

where

$$p_i = \frac{\lambda_i x_0}{\lambda x_0 + \alpha - n_2}, \ i = 1, \dots, n_1; \ p_j = \frac{\alpha_j - 1}{\lambda x_0 + \alpha - n_2}, \ j = 1, \dots, n_2$$

Thus, instead of the original system Σ , we consider a classical system $\hat{\Sigma}$, in which the renewal input process has interarrival time distribution (20) and the service times have a tail distribution defined by the mixture (25). Now, we give the stability condition of the new system $\hat{\Sigma}$. This system is regenerative, and regenerations (of all processes describing its dynamics) occur when the arrivals meet the system idle. Denote by T the generic regeneration period, i.e., the distance between two neighbor regeneration points. It is well known that the stability of a regenerative system is equivalent to positive recurrence, in which case ET < ∞ [22]. Denote by *S* the service time with tail distribution (25); recalling (24), the stability criterion of the system $\hat{\Sigma}$ is (see Section 2.2, Chapter 2 in [26]):

$$\rho := \frac{\mathsf{E}S}{\mathsf{E}\tau} < 1,\tag{26}$$

where

$$\mathsf{E}S = \sum_{i=1}^{n_1} p_i \mathsf{E}S^{(i)} + \sum_{j=1}^{n_2} p_j \mathsf{E}S^{(j)} = \sum_{i=1}^{n_1} \frac{p_i}{\mu_i} + \sum_{j=1}^{n_2} \frac{p_j}{\mu_j}.$$
 (27)

The first step in calculating $E\tau$ is to determine EY:

$$\begin{aligned} \mathsf{E}Y &= \mathsf{E}\min(Y_{1},Y_{2}) = \int_{0}^{\infty} x f_{Y_{1}}(x) \left(\int_{x}^{\infty} f_{Y_{2}}(y) dy \right) dx + \int_{0}^{\infty} y f_{Y_{2}}(y) \left(\int_{y}^{\infty} f_{Y_{1}}(y) dx \right) dy \\ &= \int_{0}^{\infty} x \lambda e^{-\lambda x} x_{0}^{\alpha - n_{2}}(x_{0} + x)^{-\alpha + n_{2}} dx + \int_{0}^{\infty} x e^{-\lambda x} (\alpha - n_{2}) x_{0}^{\alpha - n_{2}}(x_{0} + x)^{-\alpha + n_{2} - 1} dx \\ &= \frac{1}{\lambda} \Big(1 - (\alpha - n_{2}) e^{\lambda x_{0}} (\lambda x_{0})^{\alpha - n_{2}} \Gamma(-\alpha + n_{2}, \lambda x_{0}) \Big). \end{aligned}$$

Similarly, we find

$$\mathsf{E} Z = x_0 e^{\lambda x_0} (\lambda x_0)^{\alpha - n_2} \Gamma(-\alpha + n_2, \lambda x_0),$$

resulting in

$$\mathsf{E}\tau = p\mathsf{E}Y + (1-p)\mathsf{E}Z = \frac{p}{\lambda} = \frac{x_0}{\lambda x_0 + \alpha - n_2}.$$
(28)

Let

$$\rho_i := \frac{\mathsf{E}S^{(i)}}{\mathsf{E}\tau^{(i)}} = \frac{\lambda_i}{\mu_i}, \quad i = 1, \dots, n_1, \quad \rho_j := \frac{\mathsf{E}S^{(j)}}{\mathsf{E}\tau^{(j)}} = \frac{\alpha_j - 1}{x_0\mu_j}, \quad j = 1, \dots, n_2;$$

then, it is easy to check, using (27) and (28), that

$$\rho = \sum_{i=1}^{n_1} \rho_i + \sum_{j=1}^{n_2} \rho_j$$

Remark 3. Condition (26) is also the stability criterion of the original system with the stationary input process, not necessary renewal. In this case, the proof is based on the so-called Loynes' construction; see, e.g., Chapter 2 in [21].

4.2. Case 1: Common Exponential Service Time

An important special case of the system is the model $\sum_{i=1}^{n} G_i/M/1$ in which all arrivals have the same (iid) exponential service times, i.e., $\mu_1 = \cdots = \mu_n = \mu$. Such a setting has been motivated in the paper [5] where it is compared with an M/M/1 queue for n > 10. In this case, the approximating system is a stable classical GI/M/1 queue ($\rho < 1$), for which the analytical form of the stationary workload distribution is known [27]. Additionally, the LST obtained for the input stream of the form (20) allows us to compare simulation results for the waiting time distribution of the queueing system fed by the actual superposed input and its renewal approximation with the corresponding analytical solution (see Section 6.1).

Denote by $\psi_X(z)$ the LST of an r.v. *X* with distribution *F*:

$$\psi_F(z) = \mathsf{E} e^{-zX} = \int_0^\infty e^{-zx} dF(x), \ z \ge 0.$$

Then, the distribution of the stationary waiting time *W* can be expressed as

$$F_{W}(x) = 1 - \sigma e^{-\mu(1-\sigma)x}, \ x > 0,$$
(29)

where σ is the root inside the unit circle of the equation

$$z - \psi_A(\mu(1-z)) = 0.$$
 (30)

After some simple algebra, we can obtain the LST $\psi_A(z)$ of the distribution (20) as follows:

$$\begin{split} \psi_A(z) &= -\int_0^\infty e^{-zx} d\overline{A}(x) = 1 - z \int_0^\infty e^{-zx} \overline{A}(x) dx \\ &= 1 - pz \int_0^\infty e^{-zx} e^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha - n_2} dx - (1 - p)z \int_0^\infty e^{-zx} e^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha - n_2 + 1} dx \\ &= 1 - pz e^{x_0(z + \lambda)} x_0^{\alpha - n_2} (z + \lambda)^{\alpha - n_2 - 1} \Gamma(-\alpha + n_2 + 1, x_0(z + \lambda)) - \\ &- (1 - p)z e^{x_0(z + \lambda)} x_0^{\alpha - n_2 + 1} (z + \lambda)^{\alpha - n_2} \Gamma(-\alpha + n_2, x_0(z + \lambda)), \end{split}$$

where

$$\Gamma(\xi, x) = \int_{x}^{\infty} e^{-t} t^{\xi - 1} dt$$

is the upper incomplete Gamma function. Finally, applying the known property of the Gamma function

$$\Gamma(\xi+1,x)=\xi\Gamma(\xi,x)+x^{\zeta}e^{-x},$$

we obtain

$$\psi_{A}(z) = 1 - \frac{pz}{z+\lambda} + \frac{z}{z+\lambda} e^{x_{0}(z+\lambda)} x_{0}^{\alpha-n_{2}}(z+\lambda)^{\alpha-n_{2}} \times \Gamma(-\alpha+n_{2}, x_{0}(z+\lambda))(p(\alpha-n_{2}) - (1-p)x_{0}(z+\lambda)).$$
(31)

The explicit expression (31) of the LST allows us to find the moments of the r.v. τ . For example, the mean is given by

$$\mathsf{E}\tau = -\psi_A'(0) = p/\lambda = \frac{x_0}{\lambda x_0 + \alpha - n_2},$$

and this matches the expression in (28). Then, by utilizing the LST (31), we can numerically solve (30) and use the solution σ to obtain the target distribution function in (29).

4.3. Case 2: Class-Independent Inputs

Assume that

$$\lambda_1 = \cdots = \lambda_{n_1} =: \lambda$$
 and $\alpha_1 = \cdots = \alpha_{n_2} =: \alpha$.

Then, the tail distribution of the interarrival time becomes (cf. (23)):

$$\overline{A}(x) = p e^{-\lambda n_1 x} \left(\frac{x_0}{x_0 + x}\right)^{n_2(\alpha - 1)} + (1 - p) e^{-\lambda n_1 x} \left(\frac{x_0}{x_0 + x}\right)^{n_2(\alpha - 1) + 1},$$
(32)

where

$$p = \frac{n_1 x_0 \lambda}{n_1 x_0 \lambda + n_2 (\alpha - 1)}.$$

Since there are only two customer classes, the service time distribution becomes a two-component mixture:

$$\overline{B}(x) = p\overline{B}_1(x) + (1-p)\overline{B}_2(x), \ x \ge 0.$$

4.4. Case 3: Two Pareto Inputs

Consider the superposition of two (stationary) renewal processes with interarrival times $\tau^{(i)} \sim Pareto(x_0, \alpha_i)$, $\alpha_i > 1$, i = 1, 2. By using (10), it is easy to show that the stationary sequence of interarrival times in the new system $\hat{\Sigma}$ follows a Pareto distribution, i.e.,

$$\overline{A}(x) = \left(\frac{x_0}{x_0 + x}\right)^{\alpha_1 + \alpha_2 - 1}, \quad x \ge 0.$$
(33)

In this two-class scenario, the service time distribution can be expressed as a twocomponent mixture with the mixing proportion

$$p = \frac{\alpha_1 - 1}{\alpha_1 + \alpha_2 - 2}$$

and the stability condition (26) becomes

$$ho = rac{lpha_1 - 1}{x_0 \mu_1} + rac{lpha_2 - 1}{x_0 \mu_2} < 1.$$

4.5. *Case* 4: $n_1 = 1, n_2 = 1$

Finally, we consider the superposition of two inputs with exponential and Pareto interarrival times, i.e.,

$$A_1(x) = 1 - e^{-\lambda x}$$
 and $A_2(x) = 1 - \left(\frac{x_0}{x_0 + x}\right)^{\alpha}$.

The interarrival time Palm distribution (10) becomes

$$\overline{A}(x) = pe^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha - 1} + (1 - p)e^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha},\tag{34}$$

where

$$p = \frac{\lambda x_0}{\alpha - 1 + \lambda x_0}.$$
(35)

We define $\overline{F}_1(x)$ and $\overline{F}_2(x)$ as:

$$\overline{F}_1(x) = e^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha - 1} \text{ and } \overline{F}_2(x) = e^{-\lambda x} \left(\frac{x_0}{x_0 + x}\right)^{\alpha}.$$

Then, the r.v. τ with distribution satisfying (34) can be expressed as a two-component mixture (24) of an r.v. Y with distribution F_1 and an r.v. Z with distribution F_2 (with the mixing proportion p satisfying (35)), where

$$Y = \min(Y_1, Y_2)$$
 with $Y_1 \sim Exp(\lambda)$, $Y_2 \sim Pareto(x_0, \alpha - 1)$

and

$$Z = \min(Y_1, Z_1)$$
 with $Z_1 \sim Pareto(x_0, \alpha)$.

In this case, the stability condition (26) becomes

$$\rho = \frac{\lambda}{\mu_1} + \frac{\alpha - 1}{x_0 \mu_2} < 1.$$

5. Exponential Splitting and Regenerative Estimation

In this section, we first present the construction of artificial regeneration points based on the exponential splitting of the interarrival time density. Then, we outline the regenerative approach to estimate stationary performance indicators of the queueing system with the superposed input process. (For details on regenerative estimation, see, for instance, Chapter IV in [11]).

5.1. Exponential Splitting

The splitting method, in its general form, is described in [13], while the specific variant used in this paper relies on an idea presented in [28]. The method involves replacing the original process with a stochastically equivalent process (defined on an enlarged probability space) with some desired property. In our case, it is the memoryless property that leads to a regenerative structure. This allows us to apply regenerative simulation for constructing confidence estimates and accurately analyzing the stationary performance metrics of the original process. Next, we will explain the splitting procedure, which is the core of the method.

Assume that the density *g* of an absolutely continuous r.v. *T* satisfies the inequality

$$g \ge \delta g_0,$$
 (36)

where $\delta \in (0, 1)$ is a constant and g_0 is some density. Now, let us define a new Bernoulli r.v. I_T (referred to as the splitting indicator) such that $P(I_T = 1) = \delta$. We can construct a new r.v. T' as follows:

$$\Gamma' = I_T T_0 + (1 - I_T) T_1, \tag{37}$$

where the r.v. T_0 has density g_0 , and the r.v. T_1 has density (c.f. (14))

$$g_1 = \frac{f - \delta g_0}{1 - \delta}.\tag{38}$$

As we mentioned earlier, the purpose of such a transformation is to let the r.v. T' have a desired property of the distribution g_0 with a probability of at least δ . In our case, it is the memoryless property that allows us to construct regeneration points. To achieve this, we

focus on exponential splitting; a positive r.v. *T* is exponentially split if there exist constants $\eta > 0$ and $\delta \in (0, 1)$ such that

$$g(x) \ge \delta \eta e^{-\eta x}, \quad x \ge 0. \tag{39}$$

That is, the splitting representation T' of the r.v. T consists of an exponential r.v. T_0 with rate η (and density $g_0(x) = \eta e^{-\eta x}$, $x \ge 0$), and an r.v. T_1 which, according to (38), has density

$$g_1(x) = \frac{g(x) - \delta \eta e^{-\eta x}}{1 - \delta}, \quad x \ge 0.$$
 (40)

The applicability of exponential splitting to the superposition of two Pareto inputs (see Section 6.2) is based on the following lemma.

Lemma 1. Assume that an r.v. T follows a Pareto (x_0, α) distribution, with the density function given by

$$g(x) = \alpha x_0^{\alpha} (x_0 + x)^{-(\alpha + 1)}, \ \alpha > 0, \ x_0 > 0, \ x \ge 0.$$
(41)

Then, under the condition

$$\alpha + 1 \le \eta x_0 \le \alpha / \delta, \tag{42}$$

where η and δ are defined in (39), the r.v. T' has distribution (37), where T_0 is an exponential r.v. with parameter η and inequality (39) holds. Additionally, T_1 has the following distribution:

$$G_1(x) = 1 - \frac{1}{1 - \delta} \left(\frac{x_0}{x_0 + x} \right)^{\alpha} + \frac{\delta}{1 - \delta} e^{-\eta x}, \quad x \ge 0,$$
(43)

with density g_1 satisfying (40).

Proof. Recall that the failure rate of an r.v. $X \ge 0$ with distribution *F* and density *f* is defined as:

$$r_X(x) = \frac{f(x)}{\overline{F}(x)}, \ x \ge 0.$$

Consider now two r.v.s $X \sim Exp(\eta)$ and $Y \sim Pareto(x_0, \alpha + 1)$. For any $x \ge 0$, the failure rate order, denoted by $X \le_r Y$, implies the following inequality between the failure rates:

$$r_X(x) = \eta \ge \frac{\alpha + 1}{x_0} \ge \frac{\alpha + 1}{x_0 + x} = r_Y(x), \ x \ge 0.$$

This implies the stochastic order, $X \le Y$ [18]. In turn, it means that the corresponding tail distributions satisfy the following inequality:

$$e^{-\eta x} \le \left(\frac{x_0}{x_0 + x}\right)^{\alpha + 1}, \ x \ge 0.$$

If $x_0 \delta \eta \leq \alpha$, then the inequality

$$\frac{x_0\delta\eta}{\alpha}e^{-\eta x} \le \left(\frac{x_0}{x_0+x}\right)^{\alpha+1}$$

holds as well. Therefore,

$$\delta \eta e^{-\eta x} \leq \frac{\alpha}{x_0} \left(\frac{x_0}{x_0+x}\right)^{\alpha+1} = g(x),$$

and inequality (39) is satisfied. \Box

Figure 1 illustrates the role of conditions (42) showing the graphs of the Pareto density (41) (with $x_0 = 1$ and $\alpha = 3$) and exponential functions $\delta \eta e^{-\eta x}$ for different values of η and δ .



Figure 1. Graphical interpretation of inequalities (39).

The plots highlight that the exponential function is upper-bounded by the given Pareto distribution only when both inequalities (42) are met (blue curve with $\delta = 0.75$, $\eta = 4$). Instead, the green and orange curves, violating the l.h.s. and r.h.s of (42), respectively, do not satisfy the splitting condition (39).

5.2. Regenerative Simulation and Estimation

From the simulation viewpoint, splitting means that, instead of generating an r.v. T with density g, a triple (T_0 , T_1 , I) is generated and then the r.v. T' is constructed as in (37). The property (36), together with the specially constructed measure (38), ensures that the simulated value T' is stochastically equivalent to T.

Denote by $\{t_k\}$ the arrival instances in the superposed input process. To construct regeneration points of the input process, denoted by $\{\gamma_k\}$, we first select and fix an arbitrary non-Poisson component process i_0 (with the arrival instances $\{t_k^{(i_0)} \equiv t_k^{(0)}, k \ge 1\}$). Now consider the sequence of events:

$$\mathcal{E}_{k}^{(0)} = \{\text{the } j\text{th component process has exponential phase at instant } t_{k}^{(0)}, \forall j \neq i_{0}\}, k \ge 1.$$
(44)

This event means that at the *k*th arrival instant of input i_0 , the interarrival times in all other components have an exponential 'phase', obtained by splitting (i.e., the corresponding splitting indicators are equal to 1 simultaneously).

For any integer $k \ge 1$, we introduce the index $n_0(k)$ as $t_k = t_{n_0(k)}^{(0)}$. Now, we define the regeneration instances $\{\gamma_k\}$ of the superposed input as follows: $\gamma_0 = 0$,

$$\gamma_{k+1} = \min\{i : 1(t_{n_0(i)}^{(0)} > \gamma_k) \cdot 1(\mathcal{E}_{n_0(i)}^{(0)}) = 1\}, \ k \ge 0.$$
(45)

It is easy to verify that the difference $\Delta_k := \gamma_{k+1} - \gamma_k$ is the number of arrivals in the superposed process between the *k*th and the (k + 1)st occurrence of the events $\mathcal{E}^{(0)}$. These events are generated by the arrivals of class- i_0 customers, which meet the interarrival times of all other components in the exponential phase. It is easy to understand that $\{\gamma_k\}$ are

indeed the regeneration points of the superposed input process, Δ_k is the (discrete) length of the *k*th regeneration cycle of the input process, and the increments { Δ_k } are iid.

If there are $n_2 \ge 2$ Pareto input components, then, according to the described construction, the density of each component $j \ne i_0$ must satisfy the inequalities

$$\eta_j x_0 \ge \alpha_j + 1, \ \delta_j \eta_j x_0 \le \alpha_j, \tag{46}$$

for some constants $\delta_j \in (0, 1)$, $\eta_j > 0$, $j = 1, \ldots, n_2$.

For each $j \neq i_0$, the interarrival times $\tau_k^{(j)}$ are drawn with probability δ_j from the distribution $Exp(\eta_j)$, and with probability $1 - \delta_j$ from the distribution (43) with parameters $\alpha_j, x_0, \delta_j, \eta_j$. Then, we construct the instances $\{\gamma_k\}$ following (45).

Now, we return to the system Σ described in Section 4. In this system, the classical regeneration points are arrival points satisfying the following two conditions: (i) they are regeneration instants of the input process (i.e., $\{\gamma_k\}$); (ii) at these arrival instants, the server is idle, i.e.,

$$\beta_0 = 0, \ \beta_{k+1} = \min\{\gamma_i > \beta_k : W_{\gamma_i} = 0\}, \ k \ge 0,$$
(47)

where W_i is the waiting time of the *i*th customer in the queue. Let $\widehat{\Delta}_k := \beta_{k+1} - \beta_k$ be the length of the *k*th regeneration cycle of type (47); then, { $\widehat{\Delta}_k$ } are iid.

To estimate the workload process by the regenerative method [11], we define the iid sequence $\{Y_i\}$ of the accumulated workload within regeneration cycles,

$$Y_{j} = \sum_{i=\beta_{j}}^{\beta_{j+1}-1} W_{i},$$
(48)

where W_i is the waiting time of the *i*th customer within the regeneration cycle j, $j \ge 1$. It is worth mentioning that we use a *regenerative version* of the CLT (called RCLT) because correlations between the values of W_i belonging to the same regeneration cycle are not captured by the classical CLT. To obtain a regenerative estimate of the mean stationary waiting time, the iid observations $(Y_i, \hat{\Delta}_i), j \ge 1$ are used to estimate the ratio [11]

$$r = \frac{\mathsf{E}Y_1}{\mathsf{E}\widehat{\Delta}_1}$$

Then, applying the RCLT to the iid zero mean variables $\{Y_i - r\hat{\Delta}_i, i \ge 1\}$ and assuming that $E(Y_1 - r\hat{\Delta}_1)^2 < \infty$, we obtain the $(1 - 2\vartheta)$ % confidence interval (CI) for the parameter r in the following form [11]:

$$\tilde{r}_N \pm \frac{h_{\theta} \sqrt{\tilde{S}_N^2}}{\tilde{\Delta}_N \sqrt{N}} \bigg], \tag{49}$$

where $\tilde{r}_N := \tilde{Y}_N / \tilde{\Delta}_N$, $\tilde{\Delta}_N$ and \tilde{Y}_N are the sample means of $\{\tilde{\Delta}_i\}$ and $\{Y_i\}$, respectively, *N* is the number of regeneration cycles, \tilde{S}_N^2 is the estimate of the second moment

$$\mathsf{E}(Y_1 - r\widehat{\Delta}_1)^2 = \operatorname{Var}(Y_1) - 2r\operatorname{cov}(Y_1, \widehat{\Delta}_1) + r^2 \operatorname{Var}(\widehat{\Delta}_1), \tag{50}$$

and

$$h_{\vartheta} = \Phi^{-1}((1-\vartheta)/2),$$

where $\Phi(x)$ is the Laplace function.

The finiteness of $E(Y_1 - r\hat{\Delta}_1)^2$, containing the 'cycle' variables Y_1 and $\hat{\Delta}_1$, can be easily expressed in terms of the moments of the queueing parameters (service and interarrival times). Indeed, it follows from [29] that (50) is finite if both the service time and the interarrival time have finite second moments. In the simulation settings reported in Section 6.2, this condition holds since we consider exponential service times and the

superposition of two Pareto inputs, which is characterized by a Pareto distribution of the interarrival time with shape parameter $\alpha_{eq} = \alpha_1 + \alpha_2 - 1 \ge 3$ in all our experiments.

Finally, the finiteness of the term $cov(Y_1, \widehat{\Delta}_1)$ follows by the finiteness of $E(Y_1^2)$, $E(\widehat{\Delta}_1^2)$ and Hölder's inequality.

Remark 4. It is useful to note that the width of the constructed CI, and hence the accuracy of the estimation, depends critically on the number of observed regeneration cycles N rather than on the number of observations, as in classical iid estimation based on the CLT. In general, this makes estimation through regenerative simulation more time-consuming, which is expected when dealing with correlated data. However, in all our experiments, a large number of regeneration cycles was obtained in a very limited time.

Remark 5. In general, the ratio estimate r, while strongly consistent, is biased. This bias arises from estimating a nonlinear function (the ratio r) based on the 'scale' of regeneration cycles. The detailed proof of the interval estimate (49) can be found in [11] (Propositions 4.1 and 4.2).

6. Simulation Results

The aim of this section is two-fold: to highlight the accuracy of the approximations presented in Section 4, focusing on the waiting time distribution for various types of arrival processes, and to verify the applicability of exponential splitting to superposed processes, as described in Section 5.

6.1. Waiting Time Analysis in Presence of Superposed Processes

Initially, we consider queueing systems fed by two independent input processes with Pareto interarrival times. In all the simulations, 10¹⁰ arrivals are generated and exponential service times are assumed.

Using the notation from Section 4.4, we compare the actual arrival process with a renewal process having the same interarrival time distribution, specifically $Pareto(x_0, \alpha_1 + \alpha_2 - 1)$. The first set of simulations, carried out with $\alpha_1 = 2$ and $x_0 = 1$, evaluates the accuracy of the approximation by comparing the Kolmogorov distance between the corresponding waiting time distributions for different values of ρ and α_2 . As highlighted in Figure 2, the approximation works better for low utilization values and high α_2 values, i.e., when more moments of the underlying Pareto distribution are finite.



Figure 2. Kolmogorov distance between the waiting time distributions of the superposed process and its approximation as a function of ρ and α_2 .

The assumption of common exponential service times allows us to compare the sim-
ulation estimates with the exact expression for the distribution function of the stationary
waiting time, given by Formula (29). The results are summarized in Table 1, where W ,
W_A , and W_T denote the waiting time distributions when the queue is fed by the actual
superposed process, its renewal approximation, and the corresponding theoretical approxi-
mation (using distribution (29) with the appropriate value of the parameter σ), respectively.
For completeness, the values of μ and σ are also reported. The results in the table confirm
the validity of the simulation approach (implemented using ad hoc software written in C)
and the trends (in terms of α_2 and ρ) highlighted in Figure 2.

α2	ρ	μ	σ	$d(W, W_A)$	$d(W, W_T)$	$d(W_T, W_A)$
8	0.8	10	0.8196	0.02369	0.02368	1.307×10^{-5}
3	0.75	4	0.8078	0.07127	0.07126	$1.306 imes10^{-5}$
2	0.67	3	0.7641	0.08289	0.08288	$1.030 imes10^{-5}$
3	0.5	6	0.5766	0.02633	0.02632	$6.125 imes 10^{-6}$
4	0.5	8	0.5587	0.01757	0.01756	$5.919 imes10^{-6}$
5	0.5	10	0.5475	0.01253	0.01252	$6.097 imes10^{-6}$
10	0.5	20	0.5244	0.00401	0.00400	$5.044 imes10^{-6}$
3	0.3	10	0.3655	0.00949	0.00949	$2.680 imes 10^{-6}$

Table 1. Kolmogorov distance for different simulation settings (in all cases, $\alpha_1 = 2$ and $x_0 = 1$).

The distribution functions for the worst and best cases are shown in Figures 3 and 4, respectively. In both cases, W_T completely overlaps with W_A (refer to the values of $d(W_T, W_A)$ in Table 1). The figures highlight the effect of correlation in the input process on the probability of an empty system, $\pi_0 = 1 - \sigma = F_W(0)$, and, more generally, on the shape of the distribution of W.



Figure 3. Waiting time distributions of the superposed process *W* and its approximation W_A for $\alpha_1 = 2$, $\alpha_2 = 2$, $x_0 = 1$, and $\mu = 3$.

Apart from the superposition of two Pareto distributions, it is also worth considering the superposition of two processes with Pareto and exponential interarrival time distributions. In these cases, the situation is even better than for the superposition of two Pareto inputs, as highlighted by Table 2, which reports the simulation parameters and the Kolmogorov distance values (as before, 10^{10} arrivals have been generated). For the sake of brevity, the waiting time distribution is shown only for the worst case, corresponding to $\alpha = 2$, $x_0 = 1$, $\lambda = 1$, and $\mu = 3$ (see Figure 5).



Figure 4. Waiting time distributions of the superposed process *W* and its approximation W_A for $\alpha_1 = 2$, $\alpha_2 = 10$, $x_0 = 1$, and $\mu = 20$.

α	x_0	λ	μ	σ	$d(W, W_A)$	$d(W, W_T)$	$d(W_T, W_A)$
2	1	1	3	0.7067	0.064	0.064	$9.84 imes 10^{-6}$
2	1	1	4	0.5480	0.032	0.032	$5.97 imes10^{-6}$
3	1	0.5	4	0.6879	0.023	0.023	$7.56 imes10^{-6}$
2	4	1	3	0.4252	0.001	0.006	$2.73 imes 10^{-6}$

Table 2. Kolmogorov distance for different simulation settings.



Figure 5. Waiting time distributions of the superposed process *W* and its approximation W_A for $\alpha = 2$, $x_0 = 1$, $\lambda = 1$, and $\mu = 3$.

To complete the overview of the approximations described in Section 4.4, we also consider the case of class-dependent service times. Table 3 reports the Kolmogorov distance $d(W, W_A)$ for different values of the service rates μ_1 and μ_2 (and hence of the utilization ρ). Since the focus is on the impact of the service rate (which is stochastically equivalent to a hyperexponential distribution with parameters μ_1 , μ_2 , and p), we consider two Pareto distributions with the same parameters $\alpha_1 = 2$, $\alpha_2 = 3$, and $x_0 = 1$ in all cases. As expected, the values of $d(W, W_A)$ are in good agreement with those in Figure 2 for $\alpha_2 = 3$ and the corresponding values of ρ .

μ_1	μ_2	ρ	$d(W, W_A)$
3	4	0.833	0.1178
4	5	0.650	0.0540
4	6	0.583	0.0454
4	7	0.536	0.0403
4	8	0.500	0.0370
4	9	0.472	0.0350
6	7	0.452	0.0226
6	8	0.417	0.0201
6	9	0.389	0.0184

Table 3. Kolmogorov distance for different simulation settings with class-dependent service time.

Going back to the example at the beginning of this section (see Table 1 for the simulation settings), Table 4 compares the mean values of the actual waiting time W and the waiting time W_A of the approximating system. While the Kolmogorov distance, reported in the first column of the table for clarity, was generally less than 5% in most cases, the relative error (RE) can still be significant, consistent with findings in the literature [3,5,7]. This highlights the importance of using such an approximation carefully, as a rough/preliminary estimate when accurate estimation is challenging.

$d(W, W_A)$	mean(W)	$mean(W_A)$	RE (%)
0.02369	0.49660	0.45420	8.54
0.07127	1.35830	1.05069	22.65
0.08289	1.48457	1.07954	27.28
0.02633	0.25972	0.22693	12.63
0.01757	0.17373	0.15822	8.93
0.01253	0.12958	0.12101	6.62
0.00401	0.05642	0.05512	2.29
0.00949	0.06193	0.05759	6.99

Table 4. Comparison of the mean waiting times for the simulation settings in Table 1.

6.2. Exponential Splitting

We consider the original system Σ fed by two Pareto input processes (with $\alpha_1 = 2$, $\alpha_2 = 3$, and $x_0 = 1$) and a common exponential service rate μ .

We carry out several sets of simulation to evaluate the effect of both the splitting parameters (namely, the choice of the i_0 process, δ and η) and the traffic intensity ρ .

First, we select the first component of the superposed input (with $\alpha_1 = 2$) as i_0 -process, so the exponential splitting is applied to the other flow (with $\alpha_2 = 3$). In the following tables, we denote by RegIn the number of regenerations in the superposed input (i.e., the number of events (45)), and by RegW the number of regenerations (47) of the queueing system.

Tables 5 and 6 refer to $\mu = 4$ (i.e., $\rho = 0.75$) and $\mu = 6$ (i.e., $\rho = 0.5$), respectively. In both tables, the estimations of the mean waiting time are the same for all the choices of the splitting parameters and coincide with the results obtained through traditional simulation (1.358 and 0.2597 for $\mu = 4$ and $\mu = 6$, respectively), so they are not explicitly reported. To evaluate the goodness of the estimates, the last column reports the half-widths of the 99% CI for the mean waiting time, estimated over RegW intervals (corresponding to a fixed number of arrivals, namely 10^{10} as already stated).

δ	η	RegIn	RegW	99%CI
0.10	4	166.65×10^6	$13.5 imes 10^6$	$4.59 imes10^{-4}$
0.33	4	$555.61 imes 10^6$	$45.1 imes 10^6$	$4.55 imes10^{-4}$
0.50	4	833.40×10^{6}	$67.5 imes 10^6$	$4.58 imes10^{-4}$
0.75	4	1250.05×10^{6}	$101.3 imes 10^6$	$4.58 imes10^{-4}$
0.10	5	$133.29 imes 10^6$	$9.1 imes10^6$	$4.34 imes10^{-4}$
0.33	5	$444.41 imes 10^6$	$30.4 imes 10^6$	$4.56 imes10^{-4}$
0.50	5	666.67×10^{6}	$45.6 imes10^6$	$4.55 imes10^{-4}$
0.10	6	111.11×10^{6}	$6.5 imes10^6$	$4.46 imes10^{-4}$

Table 5. Regenerative simulation for $i_0 = 1$ and $\mu = 4$ ($\rho = 0.75$).

Table 6. Regenerative simulation for $i_0 = 1$ and $\mu = 6$ ($\rho = 0.5$).

δ	η	RegIn	RegW	99%CI
0.10	4	166.67×10^6	$42.7 imes 10^6$	$4.29 imes 10^{-5}$
0.33	4	$555.61 imes10^6$	142. 4×10^{6}	$4.59 imes 10^{-5}$
0.50	4	$833.32 imes 10^6$	213. 6×10^{6}	$4.20 imes 10^{-5}$
0.75	4	1250.06×10^{6}	320.5×10^6	$4.57 imes 10^{-5}$
0.10	5	$133.34 imes10^6$	$29.8 imes10^6$	$4.47 imes10^{-5}$
0.33	5	$444.41 imes 10^6$	$99.5 imes10^6$	$4.27 imes10^{-5}$
0.50	5	666.60×10^6	149.3×10^{6}	$4.55 imes 10^{-5}$
0.10	6	111.10×10^6	22.1×10^6	$4.58 imes 10^{-5}$

By comparing Tables 5 and 6, we see that for fixed values of δ and η , the number of input regeneration points RegIn is approximately the same. This fact is not surprising since the input process does not depend on the service rate μ . On the other hand, RegW increases with μ since a lower traffic intensity implies a higher probability of an idle server. This increment, together with the decrease in the mean waiting time, leads to a difference in the width of the CIs of around one order of magnitude.

Finally, the simulation results in the case of splitting of the first component (with $\alpha_1 = 2$) are reported in Tables 7 and 8, which confirm the previous observations about RegIn, RegW, and the width of the CIs.

δ	η	RegIn	RegW	99%CI
0.50	3	1111.11×10^{6}	$88.7 imes 10^6$	$4.56 imes10^{-4}$
0.10	4	$166.65 imes 10^6$	11.1×10^{6}	$4.51 imes10^{-4}$
0.33	4	$555.57 imes 10^6$	$37.1 imes 10^6$	$4.61 imes10^{-4}$
0.50	4	$833.35 imes 10^6$	$55.6 imes 10^6$	$4.55 imes10^{-4}$
0.10	6	111.11×10^6	$5.6 imes10^6$	$4.54 imes10^{-4}$
0.20	6	$222.18 imes 10^6$	$11.3 imes 10^6$	$4.60 imes10^{-4}$
0.33	6	370.41×10^6	$18.8 imes 10^6$	$4.56 imes 10^{-4}$

Table 7. Regenerative simulation for $i_0 = 2$ and $\mu = 4$ ($\rho = 0.75$).

Table 8. Regenerative simulation for $i_0 = 2$ and $\mu = 6$ ($\rho = 0.5$).

δ	η	RegIn	RegW	99%CI
0.50	3	1111.14×10^6	282.5×10^6	$4.52 imes 10^{-5}$
0.10	4	166.65×10^{6}	$36.7 imes 10^6$	$4.41 imes 10^{-5}$
0.33	4	$555.53 imes 10^6$	122.4×10^6	$4.23 imes10^{-5}$
0.50	4	833.36×10^{6}	$183.6 imes 10^6$	$4.18 imes10^{-5}$
0.10	6	$111.09 imes 10^6$	$19.5 imes 10^6$	$3.84 imes10^{-5}$
0.20	6	222.21×10^{6}	$39.0 imes 10^6$	$4.41 imes 10^{-5}$
0.33	6	370.35×10^6	$65.0 imes 10^6$	$4.43 imes 10^{-5}$

Focusing on the choice of the splitting parameters, the following remarks can be drawn from the simulation results. On the one hand, the larger the δ parameter (for a fixed value of η), the more frequent the exponential phase. As a result, RegIn also increases. On the other hand, the higher the η parameter (for a fixed δ), the lower the number of regenerations RegIn. Moreover, simulation results show that the highest number of regenerations RegW of the queueing system is achieved when equality is met in both inequalities (42) for all inputs $i \neq i_0$. It is mentioned in [11] that more frequent regenerations are intuitive and may be valuable when, for various reasons, the simulation time is limited and the number of regenerations becomes critical for constructing the CI.

7. Conclusions

This paper deals with queueing systems fed by superposed input processes generated by independent stationary renewal sources. It is well known that, in general, such systems are not classically regenerative; however, several analysis methods have been proposed in the literature and we discuss the two most promising ones. First, we describe a general approach that allows for the construction of the so-called one-dependent regeneration. Then, we focus on the superposed process containing heavy-tailed Pareto components and apply exponential splitting to construct classical regenerations of the system. These regenerations are used for confidence estimation of the mean waiting time by the regenerative version of the CLT. This is the main theoretical contribution of the research. Moreover, we identify the conditions under which exponential splitting results in the most frequent regenerations.

We also study an approximating GI/G/1 system with a single renewal process. This construction is based on Palm theory, and the corresponding stationary interarrival distribution takes the form of a mixture of the interarrival time distributions of the component processes. As a side contribution, in the special case of exponential, class-independent service times, the stationary waiting time distribution in the approximating system is explicitly derived.

Finally, an important element of this research is that the theoretical findings are supported by comprehensive numerical experiments, including both traditional Monte Carlo and regenerative simulation techniques.

Author Contributions: Conceptualization, writing—original draft preparation, E.M. and I.P.; writing—review and supervision, project administration, E.M.; editing, simulation and visualization M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Italian Ministry of Education and Research (MIUR) in the framework of the FoReLab project (Departments of Excellence) and by the University of Pisa in the framework of the PRA_2022_64 "hOlistic Sustainable Management of distributed softWARE systems (OSMWARE)" project.

Data Availability Statement: Data (simulation results) are contained within the article.

Acknowledgments: We would like to thank anonymous referees for their helpful comments. We acknowledge the use of Grammarly (https://app.grammarly.com/) and ChatGPT (https://chat. openai.com/) for English correction and rephrasing in some sections of the manuscript. These tools were not used to generate new content or ideas; they were employed solely to enhance the smoothness and correctness of the existing text.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- FIFO First In First Out
- iid independent and identically distributed
- r.v. random variable

- LST Laplace–Stiltjes Transform
- CLT Central Limit Theorem
- RE relative error
- CI confidence interval

References

- 1. Lawrence, A.J. Dependency of Intervals Between Events in Superposition Processes. J. R. Stat. Soc. Ser. B (Methodol.) 1973, 35, 306–315. [CrossRef]
- 2. Newell, G.F. Approximations for Superposition Arrival Processes in Queues. Manag. Sci. 1984, 30, 623–632. [CrossRef]
- 3. Whitt, W. Characterizing superposition arrival processes in packet multiplexers for voice and data. *J. Sel. Areas Commun.* **1986**, *4*, 833–846.
- 4. Lam, T.; Lehoczky, J.P. Superposition of Renewal Processes. Adv. Appl. Probab. 1991, 23, 64–85.
- 5. Albin, S.L. On Poisson Approximations for Superposition Arrival Processes in Queues. Manag. Sci. 1982, 28, 126–137. [CrossRef]
- 6. Asmussen, S.; Schmidli, H.; Schmidt, V. Tail probabilities for non-standard risk and queueing processes with subexponential jumps. *Adv. Appl. Probab.* **1999**, *31*, 422–447. [CrossRef]
- 7. Whitt, W. Approximating a point process by a renewal process, I: Two basic methods. Operat. Res. 1982, 30, 125–147. [CrossRef]
- 8. Shedler, G.S. *Regenerative Stochastic Simulation*; Academic Press Inc.: Cambridge, MA, USA, 1993.
- 9. Glynn, P.W. Some topics in regenerative steady-state simulation. Acta Appl. Math. 1994, 34, 225–236. [CrossRef]
- 10. Glynn, P.W.; Iglehart, D.L. Conditions for the applicability of the regenerative method. *Manag. Sci.* **1993**, *39*, 1108–1111. [CrossRef]
- 11. Asmussen, S.; Glynn, P.W. Stochastic Simulation: Algorithms and Analysis; Springer Nature: Cham, Switzerland, 2007.
- 12. Sigman, K. One-dependent regenerative processes and queues in continuous time. Math. Oper. Res. 1990, 15, 175–189. [CrossRef]
- 13. Thorisson, H. Coupling, Stationarity, and Regeneration, Probability and Its Applications; Springer: New York, NY, USA, 2000.
- 14. Khintchine, A.Y. Mathematical Methods in Queueing Theory; Griffin: London, UK, 1960.
- 15. Whitt, W. Queues with superposition arrival processes in heavy traffic. Stoch. Process. Their Appl. 1985, 21, 81–91. [CrossRef]
- 16. Smith, W.L. Renewal theory and its ramifications. J. R. Stat. Soc. (Ser. B) 1958, 20, 243–302. [CrossRef]
- 17. Cox, D.R. Renewal Theory; Methuen: London, UK, 1962.
- 18. Whitt, W. Comparing counting processes and queues. Adv. Appl. Prob. 1981, 13, 207–220. [CrossRef]
- Peshkova, I.; Morozov, E.; Maltseva, M. On Comparison of Multiserver Systems with Exponential-Pareto Mixture Distribution. In *Computer Networks. CN 2020. Communications in Computer and Information Science*; Gaj, P., Gumiński, W., Kwiecień, A., Eds.; Springer: Cham, Switzerland, 2020; Volume 1231.
- 20. Peshkova, I.V.; Morozov, E.V. On Comparison of Multiserver Systems with Multicomponent Mixture Distributions. *J. Math. Sci.* **2022**, 267, 260–272. [CrossRef]
- 21. Baccelli, F.; Bremaud, P. Elements of Queueing Theory. Palm Martingale Calculus and Stochastic Recurrences, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2003.
- 22. Asmussen, S. Applied Probability and Queues, 2nd ed.; Springer: New York, NY, USA, 2003.
- 23. Smith, W.L. Regenerative stochastic processes. Proc. R. Soc. (Ser. A) 1955, 232, 6–31.
- 24. Gut, A. Stopped Random Walks Limit Theorems and Applications, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.
- 25. Serfozo, R.F. Basics of Applied Stochastic Processes; Springer: Berlin/Heidelberg, Germany, 2009.
- 26. Morozov, E.; Steyaert, B. Stability Analysis of Regenerative Queueing Models; Springer: Cham, Switzerland, 2021.
- 27. Cohen, J.W. The Single-Server Queue, 2nd ed.; North-Holland: Amsterdam, The Netherlands, 1982.
- 28. Andronov, A. Artificial regeneration points for stochastic simulation of complex systems. In *Simulation Technology: Science and Art, 10th European Simulation Symposium*; Society for Computer Simulation International: Delft, The Netherlands, 1998; pp. 34–40.
- 29. Thorisson, H. The queue *GI/GI/1*: Finite moments of the cycle variables and uniform rates of convergence. *Stoch. Process. Their Appl.* **1985**, *19*, 85–99. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.