Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2026 (44) 3** : 1–8.
ISSN-0037-8712
doi:10.5269/bspm.79731

Asymptotic Analysis of the Processor Sharing Multi-Queue

Amal Ezzidani, Mohamed Ghazali and Abdelghani Ben Tahar

ABSTRACT: Queueing theory is a key tool for analyzing complex systems like cloud computing and networks. It helps understand how delays, congestion, and resource sharing behave under different regimes. This paper studies the asymptotic behavior of the fluid model solution associated with a network of processor sharing multi-queues. This model is particularly relevant to modern applications where multiple tasks share processing resources. The network consists of J queues, each with a single server, an infinite waiting room and arbitrary interarrival and service time distributions. Under the processor-sharing discipline, all customers present in a queue are served simultaneously. In this system, customers may arrive at a queue either from outside the system or from the previous queue. Upon completing service at one queue, customers proceed to the next. Our results show that, as time approaches infinity, the fluid model solution converges in the critical regime and grows asymptotically linearly with time in the supercritical regime.

Key Words: Processor Sharing, multi-queue, asymptotic behavior, critical regime, supercritical regime.

Contents

1. Introduction

Open queueing networks refer to mathematical models used to represent and analyze various computer systems, particularly those where multiple tasks or processes interact with each other in a dynamic manner. As a result, the processor sharing (PS) discipline is a specific way of managing and allocating resources (such as CPU time) in these queueing networks. In PS, all tasks or processes present in a queue are served simultaneously, with each receiving a fraction of the available resources proportional to its size. Moreover, the relevance of PS extends beyond queueing theory and finds application in various computer applications, as demonstrated in Kleinroc [1]. For real-world examples, the paper by Moscholios [2] proposes teletraffic models in which arriving calls follow a random process and compete for service in the cell under the bandwidth sharing policy. Other authors [3] address the challenge of task scheduling in cloud computing by employing an analytical approach based on queuing theory.

To study the evolution of such systems, a fluid model solution constitutes a mathematical or analytical representation aimed at simplifying the analysis. Moreover, several authors have delved into the examination of the asymptotic behavior of fluid model solutions. Chen et al. [4] initiated this exploration by studying a fluid approximation for a PS queue, which focussed on approximating the queue length process. Besides, Gromoll et al. [5] studied a heavily loaded PS queue. They studied the fluid (or law of large numbers) approximation and continued by establishing the asymptotic behavior of critical fluid model solutions in [6].On the other hand, Puha et al. [7] has established an analogous results for supercritical fluid model solutions. They shown that the crucial differentiation between critical and strictly supercritical fluid models lies in the behavior of the total mass. That is, for a solution that begins from zero, the total mass increases over time in the case of strictly supercritical models. However, for

2020 *Mathematics Subject Classification:* 68M20, 90B22.
Submitted October 26, 2025. Published January 22, 2026

critical models, the total mass remains consistently at zero. Consequently, these findings are particularly intriguing and motivate us to further extend the analysis by examining the asymptotic behavior of the fluid model solution for a finite sequence of queues. Specifically, we investigate the limiting behavior of the queue length, the total number of arrivals, and the number of departures in each queue as time approaches infinity, under both the critical regime (where the service and arrival rates are equal) and the supercritical regime (where the service rate is lower than the arrival rate).

In this paper, we consider a network composed of J queues, indexed by $j = 1, \dots, J$. Each queue (denoted by j) has a single server and an infinite waiting room. Following the PS rule, all customers present in each queue are served simultaneously. Customers arrive at queue j from an external source following a renewal process and have a general service time distribution. After receiving their service, customers exit queue j and proceed to queue $j+1$.

This paper establishes two main theorems. The first one (Theorem 3.1) demonstrates that, under mild assumptions, the critical fluid model solution converges, and the queue length associated with each queue becomes asymptotically constant. The second theorem (Theorem 3.2) characterizes the asymptotic behavior in the supercritical regime, showing that the fluid model solution grows linearly with time. These results extend the foundational analyses developed in [6] and [7].

The paper is organized in the following manner. Section 2 contains the model description and the definition of fluid model solution. Section 3 states the main results along with their proofs. Section 4 is devoted to examples. Finally, Section 5 provides concluding remarks.

2. Fluid model for the PS multi-queue

In this section we give the description of the fluid model of a network of the PS multi-queue consisting of J queues. The model, as presented in Figure 1, has three parameters, $\alpha \in \mathbb{R}_+^{*J}$, a vector of Borel

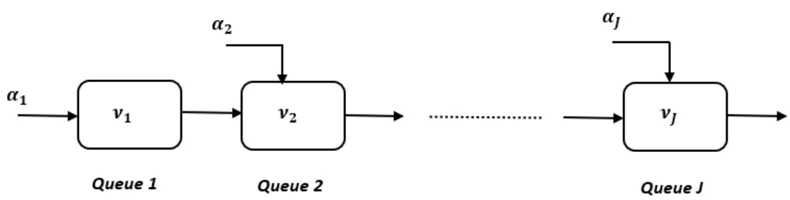

Figure 1: The processor sharing multi-queue

probability measure ν on $\mathbb{R}_+^{*J}$ where $[\nu_j(\{0\}) = 0]$ and the component ν_j which has a finite first moment $[\langle \chi, \nu_j \rangle < \infty]$ for each $j = 1, \dots, J$ as well as the nonnegative matrix P is formed from the components p_{ji} where

$$p_{ji} = \begin{cases} 1 & \text{if } i = j+1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

The network is considered to be open, meaning that the matrix $Q = I + P' + (P')^2 + \dots$ is finite where P' represents the transpose of the matrix P. This condition is equivalent to requiring that $(I - P')$ is invertible. These parameters refer to queueing system parameters. In particular, α corresponds to the exogenous arrival rate, the probability measure ν corresponds to the distribution of the i.i.d. service times and the matrix P represents the routing probability between queues. Thus, the load factor of each station j is given by:

$$\rho_j = \langle \chi, \nu_j \rangle \sum_{i=1}^{j} \alpha_i.$$

If $\rho_j = 1$ for all $j = 1, \dots, J$, the triple (α, ν, P) is called a critical data. Otherwise if there is j such that $\rho_j > 1$ then the data (α, ν, P) is called a supercritical data.

At each queue j, any customer presents at time zero is called an initial customer and has an i.i.d. service time distributed according to the probability measure ν_j^0. Moreover, we assume that all queues are initially non-empty.

Let $\mathcal{M}$ denote the space of finite, nonnegative Borel measures on $\mathbb{R}_+$ endowed with the topology of weak convergence and $\mathcal{M}^J$ is the Cartesian product of J with itself J times. We express $\langle g, \mu \rangle$ as the integral of a Borel measurable function g with respect to the measure μ, where $\mu \in \mathcal{M}$ and g is integrable with respect to μ.

For a given data (α, ν, P) and an initial state $\xi \in \mathcal{M}^J$, a fluid model solution is a triple $(A(t), D(t), \mu(t))$ of two real continuous and nondecreasing componentwise functions $A, D : \mathbb{R}_+ \to \mathbb{R}_+^J$, and one measure-valued vectors of continuous mappings $\mu : \mathbb{R}_+ \to \mathcal{M}^J$ such that, for every j, the following equations hold

$$A_1(t) = \alpha_1 t \quad \textit{and} \quad A_j(t) = \alpha_j t + D_{j-1}(t) \tag{2.2}$$

$$Z_j(t) = Z_j(0) + A_j(t) - D_j(t) \tag{2.3}$$

$$\langle 1_{[x,\infty)}, \mu_j(t) \rangle = \langle 1_{[x+S_j(t),\infty)}, \xi_j \rangle + \int_0^t \langle 1_{[x+S_j(t)-S_j(s),\infty)}, \nu_j \rangle \, dA_j(s) \tag{2.4}$$

for all $t, x \in \mathbb{R}_+$. Here:

$$S_j(t) = \int_0^t \frac{1}{Z_j(u)} du \tag{2.5}$$

such that for all $0 \leq s < t$, $\inf\limits_{s \leq u \leq t} \langle 1, Z_j(u) \rangle > 0$, and $S_j(\tau) - S_j(s) = \infty$ for all $s < \tau$ such that $\langle 1, Z_j(\tau) \rangle = 0$.

The above equations can be interpreted as follows: $A_j(t)$ represents the number of arrivals at queue j by time t, $D_j(t)$ signifies the number of departures from queue j by time t, and $Z_j(t)$ denotes the number of customers present in queue j at time t. In addition, $S_j(t)$ is the accumulated service quantity received by a customer from the beginning of the observation up to time t in queue j.

Since $\langle 1_{[x,\infty)}, \xi_j \rangle = \mathbb{P}(\nu_j^0 > x) Z_j(0)$ for all $x \geq 0$, then, for $x = 0$, Eq. (2.4) gives the evolution for $Z_j(t) := \langle 1, \mu_j(t) \rangle$:

$$Z_j(t) = \mathbb{P}(\nu_j^0 > S_j(t)) Z_j(0) + \int_0^t \mathbb{P}(\nu_j > S_j(t) - S_j(s)) \, dA_j(s). \tag{2.6}$$

3. Main results

In this section we state and prove the limit of the triple $(A_j(t), D_j(t), Z_j(t))$ as $t \to \infty$ for each $j = 1, \ldots, J$ for both critical and supercritical data.

Define $\mathcal{M}^{c,J} = \{\xi \in \mathcal{M}^J : \xi_j(\{x\}) = 0 \text{ for all } x \in \mathbb{R}_+ \text{ and } j = 1, \ldots, J\}$.

Theorem 3.1 *Let (α, ν, P) be a critical data and $\xi \in \mathcal{M}^{c,J}$. For each $j = 1, \ldots, J$, let $(A_j(t), D_j(t), Z_j(t))$ the triple defined by (2.2), (2.3) and (2.6). Assume that $\langle \chi, \nu_j^0 \rangle < \infty$, $\langle \chi, \nu_j \rangle < \infty$ and $\langle \chi^2, \nu_j \rangle < \infty$ for all $j = 1, \ldots, J$. Then*

$$Z_j(t) \to Z_j(\infty) \quad \textit{as} \quad t \to \infty \tag{3.1}$$

where

$$Z_1(\infty) = \frac{2 \langle \chi, \nu_1^0 \rangle Z_1(0)}{\langle \chi^2, \nu_1 \rangle \alpha_1} \quad \textit{and} \tag{3.2}$$

$$Z_j(\infty) = \frac{2 \left(\langle \chi, \nu_j^0 \rangle Z_j(0) + \langle \chi, \nu_j \rangle \sum_{i=1}^{j-1} (Z_i(0) - Z_i(\infty)) \right)}{\langle \chi^2, \nu_j \rangle \sum_{i=1}^{j} \alpha_i} \tag{3.3}$$

for $j \geq 2$. Moreover,

$$\lim_{t \to \infty} \frac{A_j(t)}{t} = \lim_{t \to \infty} \frac{D_j(t)}{t} = \sum_{i=1}^{j} \alpha_i.$$

Proof: We use the induction on j, in fact, for $j=1$ this convergence is proved in Theorem 4.1 of [6]. Fix $j=2,\dots,J$ and assume that (3.1) holds for $i=1,\dots,j-1$. Since the function $S_j(\cdot)$ defined by (2.5) is continuous and strictly increasing, we have the same for the function $T_j(\cdot)$ defined by $T_j(t)=S_j^{-1}(t)$ for all $t\geq 0$. Moreover, $\lim_{t\to\infty} T_j(t)=\infty$. By changing of variable $t\to T_j(t)$ in Eq. (2.6), then we obtain:

$$Z_j(T_j(t)) = \mathbb{P}(\nu_j^0>t)Z_j(0) + \int_0^t \mathbb{P}(\nu_j>t-s)\ A_j'(T_j(s))\ T_j'(s)\ ds. \tag{3.4}$$

Eq. (2.2) can be rewritten as

$$A_j(t) = \sum_{i=1}^{j}\alpha_i\ t + \sum_{i=1}^{j-1}(Z_i(0)-Z_i(t)).$$

Then Eq. (3.4) becomes

$$Z_j(T_j(t)) = \mathbb{P}(\nu_j^0>t)Z_j(0) - \int_0^t \mathbb{P}(\nu_j>t-s)\sum_{i=1}^{j-1}Z_i(T_j(s))'\ ds + \sum_{i=1}^{j}\alpha_i\int_0^t \mathbb{P}(\nu_j>t-s)\ T_j'(s)\ ds.$$

As a result:

$$Z_j(T_j(t)) = \mathbb{P}(\nu_j^0>t)Z_j(0) - \int_0^t \mathbb{P}(\nu_j>t-s)\sum_{i=1}^{j-1}Z_i(T_j(s))'ds + \rho_j\int_0^t \frac{\mathbb{P}(\nu_j>t-s)}{\langle\chi,\nu_j\rangle}\ T_j'(s)\ ds.$$

On the other hand, since $Z_j(T_j(t))=\langle 1,\mu_j(T_j(t))\rangle = T_j'(t)$ and $\rho_j=1$, then we have

$$T_j'(t) = h_j(t) + \int_0^t K_j(t-s)\ T_j'(s)\ ds, \tag{3.5}$$

where $h_j(t)=\mathbb{P}(\nu_j^0>t)Z_j(0) - \int_0^t \mathbb{P}(\nu_j>t-s)\ \sum_{i=1}^{j-1}Z_i(T_j(s))'ds$ and $K_j(t)=\frac{\mathbb{P}(\nu_j>t)}{\langle\chi,\nu_j\rangle}$. Applying the key renewal Theorem to Eq. (3.5), we have

$$\begin{aligned}
\lim_{t\to\infty} T_j'(t) &= \frac{\int_0^\infty h_j(t)dt}{\int_0^\infty tK_j(t)dt} = \frac{\langle\chi,\nu_j^0\rangle Z_j(0)}{\frac{1}{2\langle\chi,\nu_j\rangle}\langle\chi^2,\nu_j\rangle} \\
&\quad - \frac{\int_0^\infty\int_0^t \mathbb{P}(\nu_j>t-s)\left(\sum_{i=1}^{j-1}\left(Z_i(T_j(s))\right)'\right)dsdt}{\frac{1}{2\langle\chi,\nu_j\rangle}\langle\chi^2,\nu_j\rangle} \\
&= \frac{\langle\chi,\nu_j^0\rangle Z_j(0) - \langle\chi,\nu_j\rangle(\sum_{i=1}^{j-1}(Z_i(\infty)-Z_i(0)))}{\frac{1}{2\langle\chi,\nu_j\rangle}\langle\chi^2,\nu_j\rangle} \\
&= \frac{\langle\chi,\nu_j^0\rangle Z_j(0) - \langle\chi,\nu_j\rangle(\sum_{i=1}^{j-1}(Z_i(\infty)-Z_i(0)))}{\frac{1}{2}\langle\chi^2,\nu_j\rangle\ \sum_{i=1}^{j}\alpha_i}.
\end{aligned}$$

On the other hand, by induction and using (2.2) and (2.3), we obtain $\lim_{t\to\infty}\frac{A_j(t)}{t} = \lim_{t\to\infty}\frac{D_j(t)}{t} = \sum_{i=1}^{j}\alpha_i$. □

To introduce Theorem 3.2, we first review some background. Let (α,ν,P) be a data. For each $j=1,\dots,J$, and by [8], there exists a unique positive real number m_j solution to the equation:

$$m_j = \left(1-\int_0^\infty e^{-m_js}\nu_j(ds)\right)a_j, \tag{3.6}$$

such that $\rho_1 > 1$ *and* $\rho_j - \langle \chi, \nu_j \rangle \sum_{i=1}^{j-1} m_i > 1$, where

$$a_1 = \alpha_1 \quad \textit{and} \quad a_j = \alpha_j + \sum_{i=1}^{j-1} (\lambda_i - m_i), \tag{3.7}$$

$$\lambda_1 = \alpha_1 \quad \textit{and} \quad \lambda_j = \alpha_j + \sum_{i=1}^{j-1} (\alpha_i - m_i), \tag{3.8}$$

for $j = 2, \dots J$. Moreover, define $p_j : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$p_j(x) = a_j \int_x^\infty 1_{[x,\infty)}(y)\; e^{m_j(x-y)}\; \nu_j(dy), \tag{3.9}$$

for all $x \in \mathbb{R}_+$. For each $j = 1, \dots, J$, let $\varsigma_j \in \mathcal{M}$ denotes the measure that is absolutely continuous with respect to Lebesgue's measure where its Radon–Nikodym derivative is $p_j(\cdot)$:

$$\varsigma_j(A) = \int_A p_j(x)\; dx \quad \textit{for all measurable} \;\; A \subseteq \mathbb{R}_+. \tag{3.10}$$

It's obvious that $\int_0^\infty p_j(x)\; dx = m_j = \langle 1, \varsigma_j \rangle$.

Theorem 3.2 *Let (α, ν, P) be a supercritical data and $\xi \in \mathcal{M}^{c,J}$ be an initial state. For each $j = 1, \dots, J$, let $(A_j(t), D_j(t), Z_j(t))$ the triple defined by (2.2), (2.3) and (2.6). Then for each $j = 1, \dots, J$ we have:*

$$Z_j(t)/t \to m_j \qquad as \qquad t \to \infty, \tag{3.11}$$

where, m_j defined by (3.6). As a consequence,

$$\lim_{t\to\infty} A_j(t)/t = \lambda_j \quad \textit{and} \quad \lim_{t\to\infty} D_j(t)/t = \lambda_j - m_j. \tag{3.12}$$

Proof: For $j = 1$, Theorem 3.2 follows from Theorem 3.5 of [7]. For $j \geq 2$, we use a change of variable $T_j(t) \to t$, one has

$$\lim_{t\to\infty} Z_j(t)/t = \lim_{t\to\infty} Z_j(T_j(t))/T_j(t) = T_j'(t)/T_j(t). \tag{3.13}$$

Based on the arguments in Lemmas C.2 and C.3 of [9], and using Eq. (3.5), we obtain the following as $m \to m_j^+$:

$$\mathcal{L}(T_j)(m) \sim \frac{\mathcal{L}(h_j)(m_j)}{\left(\rho_j - \langle \chi, \nu_j \rangle \sum_{i=1}^{j-1} m_i\right) \mathcal{L}'(\nu_j)(m_j)} \frac{1}{m - m_j},$$

with $\mathcal{L}(f)(m_j) = \int_0^\infty e^{-m_j t} df(t)$ is the Laplace-Stieltjes transform of any function f and $\mathcal{L}'(\nu_j)(t) = \frac{d\mathcal{L}(\nu_j)(t)}{dt}$. Then, the Tauberian Theorem gives us $T_j(t) \sim d_j \exp(m_j t)$ where

$$d_1 = \frac{m_1^{-1} \left(1 - \mathcal{L}(\nu_1^0)(m_1)\right) Z_1(0)}{-1 - \mathcal{L}(\nu_1)(m_1)\alpha_1}, \tag{3.14}$$

and for $j \geq 2$,

$$d_j = \frac{m_j^{-1}}{-1 - \mathcal{L}(\nu_j)(m_j)\; \lambda_j} \Bigg[\left(1 - \mathcal{L}(\nu_j^0)(m_j)\right)\; Z_j(0) + (1 - \mathcal{L}(\nu_j)(m_j)) \sum_{i=1}^{j-1} \left(Z_i(0) + m_i \mathcal{L}(T_i)(m_j) - \mathcal{L}(Z_i(T_j))(m_j)\right) \Bigg]. \tag{3.15}$$

Thus, Eq. (3.13) becomes

$$\lim_{t\to\infty} Z_j(t)/t = \lim_{t\to\infty} \frac{m_j d_j e^{m_j t}}{d_j e^{m_j t}} = m_j. \tag{3.16}$$

Moreover, by induction and using (2.2) and (2.3), the limits (3.12) holds. □

4. Examples

Example 4.1 Consider the PS system with two queues and the following critical data. Let $\alpha_1 = 1$ and ν_1 have a rate $\mu_1 = 1$ exponential distribution. Let $\alpha_2 = 2$ and ν_2 have a rate $\mu_2 = 3$ exponential distribution. We compute $\rho_1 = \frac{\alpha_1}{\mu_1} = 1$ and $\rho_2 = \frac{\alpha_1+\alpha_2}{\mu_2} = 1$. Besides ν_1^0 and ν_2^0 follow exponential distributions with rates $\mu_1^0 = 2$ and $\mu_2^0 = 2$, respectively. Figure 2 suggests that the queue lengths of each queue have finite limits: $Z_1(\infty) = \frac{1}{2}$ *and* $Z_2(\infty) = 2$. This example is interesting because both

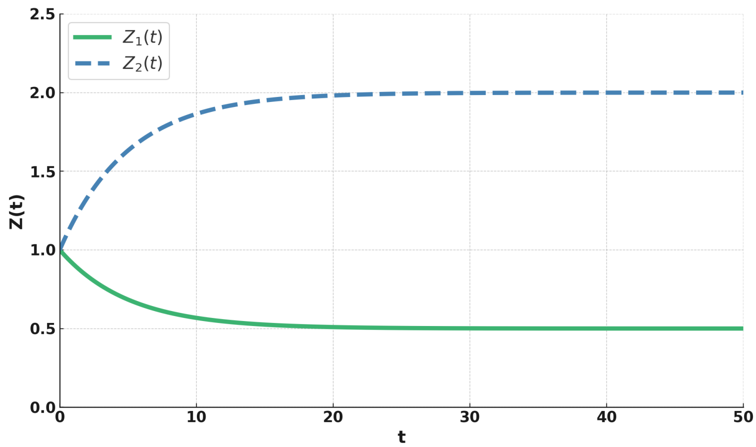

Figure 2: Total mass in both queues under critical condition and exponential service times

queues are critically loaded in the usual traffic-intensity sense: $\rho_1 = \rho_2 = 1$. Normally $\rho = 1$ signals a borderline (critical) case where queues can grow without bound. Yet the computed fluid limits show finite steady values, $Z_1(\infty) = \frac{1}{2}$ *and* $Z_2(\infty) = 2$. This outcome has the following important implications and interpretations:

1- *Stability in the fluid sense despite critical loads*: The finite limits mean that, under the model assumptions (PS discipline and the specified residual-service-rate parameters $\mu_1^0 = \mu_2^0 = 2$, the system admits nonzero but bounded equilibrium queue lengths. In other words, the system does not blow up even though the nominal load parameters ρ_i equal 1. This highlights that the usual $\rho < 1$ rule is a sufficient but not necessary condition for boundedness in more refined models.

2- *Why the two queues settle at different levels*:Queue 2 ends up larger (2 vs 1/2) because more work is fed into queue 2: $\alpha_2 = 2$ contributes to the cumulative arrival intensity seen by server 2, so even with the higher service rate $\mu_2 = 3$ the balance of arrivals and the sharing mechanism produces a larger equilibrium backlog. Intuitively, PS spreads capacity among all present jobs, but the net input to queue 2 (including feed-forward from queue 1 if present in the model) makes its long-run occupancy higher.

Example 4.2 Consider the PS system with two queues and the following supercritical parameters. Let $\alpha_1 = \frac{5}{4}$ and ν_1 be exponentially distributed with rate $\mu_1 = 1$. Let $\alpha_2 = 1$ and ν_2 be exponentially distributed with rate $\mu_2 = 1$. We compute the traffic intensity for the first queue as $\rho_1 = \frac{\alpha_1}{\mu_1} = \frac{5}{4} > 1$. Furthermore, let $m_1 = \frac{-1+\sqrt{6}}{2}$, which is the unique positive solution of the quadratic equation

$$x^2 + x - \frac{5}{4} = 0.$$

We also have

$$\rho_2 - \langle \chi, \nu_1 \rangle m_1 = \frac{11 - 2\sqrt{6}}{4} > 1.$$

Figure 3 illustrates that, as time $t \to \infty$, the queue lengths of both queues increase linearly on av-

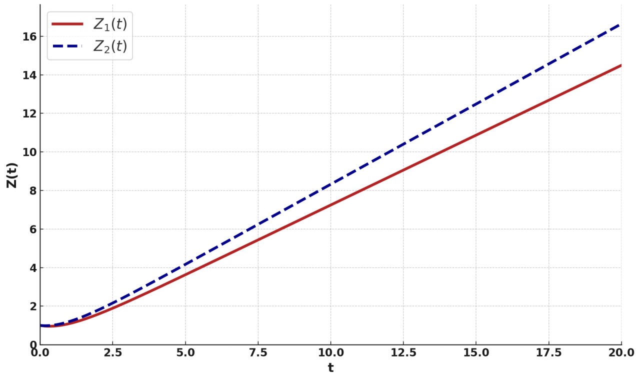

Figure 3: Total mass in both queues under supercritical critical data

erage. The corresponding asymptotic growth rates are $m_1 = \frac{11-2\sqrt{6}}{4} > 1$ for the first queue and $m_2 = \frac{-1+\sqrt{12-2\sqrt{6}}}{2}$ for the second queue. This behavior has two important implications:

1- *System instability*: Both queues accumulate customers indefinitely, which is characteristic of a supercritical system. Unlike the critical case where finite equilibria may exist, here the imbalance between arrival rates and service capacities forces persistent growth.

2- *Linear law of growth*: The fact that the growth is asymptotically linear rather than exponential highlights that, under the PS discipline, overload translates into a deterministic "drift" in the fluid limit. The constants m_1 and m_2 quantify the precise speed of divergence for each queue.

5. Conclusion

In this paper, we have analyzed the evolution of a PS multi-queue system using fluid model asymptotics, focusing on two operational regimes: critical and supercritical. In the critical regime, we demonstrated that each queue's state converges to a constant over time. Conversely, in the supercritical regime, we proved that each queue becomes transient, with queue lengths diverging over time. Comparatively, the critical regime allows for system stabilization and predictable long-term behavior, while the supercritical regime leads to instability and unbounded growth. Understanding these regimes provides valuable insights for designing and managing PS systems to ensure desired performance and stability.

References

1. L. Kleinrock, *Queueing systems, computer applications*, New York: Wiley, (1976).
2. I.D. Moscholios, V.G. Vassilakis, M.D. Logothetis, A.C. Boucouvalas, *State-dependent bandwidth sharing policies for wireless multirate loss networks*. IEEE Trans Wireless Commun., 16, 5481-5497, (2017).
3. M. Ghazali, A. Ben Tahar, *A Queuing Theory Approach to Task Scheduling in Cloud Computing with Generalized Processor Sharing Queue Model and Heavy Traffic Approximation*. IAENG International Journal of Computer Science, 51.10, (2024).
4. H. Chen, O. Kella and G. Weiss, *Fluid approximations for a processor-sharing queue*. Queueing Systems Theory Appl. 27, 99–125, (1997).
5. H. C. Gromoll, A. L. Puha, R. J. Williams, *The fluid limit of a heavily loaded processor sharing queue*. Ann. Appl. Probab., 12, 797–859, (2002).
6. A.L. Puha, R.J. Williams, *Invariant states and rates of convergence for the fluid limit of a heavily loaded processor sharing queue*. Queueing Systems Theory Appl., 14, 517-554, (2004).
7. A.L. Puha, A.L. Stolyar, R.J. Williams, *The fluid limit of an overloaded processor sharing queue*. Math. Oper. Res., 31(2), 316-350, (2006).
8. A. Ezzidani, A. Ben Tahar, M. Hanini. *Fluid model solution of feed-forward network of overloaded multiclass processor sharing queues*, J. Appl. Math. & Informatics, 42(2), pp.291-303, (2024).
9. A. Ben Tahar, A. Jean-Marie, *The fluid limit of the multiclass processor sharing queue*. Queueing Systems Theory Appl., 71(4), 347-404, (2012).

Amal Ezzidani,
Hassan First University of Settat, Faculty of Science and Technology
Morocco.
E-mail address: a.ezzidani@uhp.ac.ma

and

Mohamed Ghazali,
Hassan First University of Settat, Faculty of Science and Technology
Morocco.

and

Abdelghani Ben Tahar,
Hassan First University of Settat, Faculty of Science and Technology
Morocco.