Approximation of an $M/M/s$ queue by the $M/M/\infty$ one using the operator mathod

A Bareche[1], M Cherfaoui[2] and D Aïssani[1]

[1] Research Unit LaMOS (Modeling and Optimization of Systems), University of Bejaia, 06000 Bejaia, Algeria
[2] Department of Mathematics, University of Biskra, 07000 Biskra, Algeria

E-mail: `aicha_bareche@yahoo.fr`

Abstract. In this paper, we provide an approximate analysis of an $M/M/s$ queue using the operator method (strong stability method). Indeed, we use this approach to study the stability of the $M/M/\infty$ system (ideal system), when it is subject to a small perturbation in its structure ($M/M/s$ is the resulting perturbed system). In other words, we are interested in the approximation of the characteristics of an $M/M/s$ system by those of an $M/M/\infty$ one. For this purpose, we first determine the approximation conditions of the characteristics of the perturbed system, and under these conditions we obtain the stability inequalities for the stationary distribution of the queue size. To evaluate the performance of the proposed method, we develop an algorithm which allows us to compute the various obtained theoretical results and which is executed on the considered systems in order to compare its output results with those of simulation.

1. Introduction

Multi-servers queues are analytically tractable queues of practical importance. However, the analytical results of such systems are only available in terms of Laplace transforms or/and generating functions which are often cumbersome and not useful in practice. To avoid and circumvent this problem, many authors use numerical and approximation methods to analyze such type of systems (see [5, 6, 7, 8]). On the other hand, many situations are modeled by a multi-server queue, but for which we seek to have the average number of waiting customers to tend to zero. This situation coincides with an infinite-servers queue that has the non waiting property.

It is why we focus, in this work, on the strong stability method (see [1, 11]) which allows us to make both qualitative and quantitative analysis helpful in understanding complicated models by more simpler ones for which an evaluation can be made. This method, also called "method of operators" can be used to investigate the ergodicity and stability of the stationary and non-stationary characteristics of the imbedded Markov chains (see [11]). In contrast to other methods, it supposes that the perturbations of the transition kernel are small with respect to some norms in the operators space. This stringent condition gives better stability estimates and enables us to find precise asymptotic expansions of the characteristics of the perturbed system.

The applicability of the strong stability method is well proved and documented in various fields and for different purposes. In particular, it has been applied to several proposals (see for example [2, 3, 4, 12]).

This paper aims to study the strong stability of an $M/M/\infty$ system (ideal system) after a small perturbation of its structure (the $M/M/s$ is the resulting perturbed system). We first clarify the conditions for such an approximation, then we provide an upper bound for the norm of the difference between the two stationary distributions of the considered systems. Note that there is a first attempt to study this particular case of approximation of the $M/M/s$ queue by the $M/M/\infty$ one, using the theory of Markov chains and the special norm L_1 (see [9]). The authors of this latter work had produced two important results. The first one is summarized in the upper bound of the absolute difference (L_1 norm) between the stationary probabilities of the $M/M/s$ and $M/M/\infty$ systems, and in the second point the authors have proved that this difference tends to zero when the number of servers tends to infinity. Unlike [9], we use here the general weight norm ($\|.\|_v$) instead of the norm L_1 to approach the $M/M/s$ system by the $M/M/\infty$ one. Moreover, we provide the domain within which this approximation is valid.

2. Description of $M/M/s$ and $M/M/\infty$ models

Let us consider an $M/M/s$ system with s servers, where inter-arrival times are independently distributed with an exponential distribution $E_\lambda(t)$ and mean inter-arrival time $1/\lambda$, and service times are distributed with $E_\mu(t)$ (exponential with parameter μ).

Let X_k be the number of customers in the system just prior to the arrival of the kth customer. Therefore, $X = (X_k; k = 0, 1, \ldots)$ is an homogeneous Markov chain with a state space $\mathbf{N} = \{0, 1, 2, \ldots\}$ and with a transition operator $P = (P_{ij})_{i,j\geq 0}$ where:

$$P_{ij} = \begin{pmatrix}
\frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\
\frac{\mu}{\lambda+\mu} & 0 & \frac{\lambda}{\lambda+\mu} & 0 & \cdots & 0 & 0 & 0 & \cdots \\
0 & \frac{2\mu}{\lambda+2\mu} & 0 & \frac{\lambda}{\lambda+2\mu} & \ddots & 0 & 0 & 0 & \ddots \\
0 & 0 & \frac{3\mu}{\lambda+3\mu} & 0 & \ddots & \ddots & 0 & 0 & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ddots & \frac{s\mu}{\lambda+s\mu} & 0 & \frac{\lambda}{\lambda+s\mu} & 0 & \ddots \\
0 & 0 & 0 & 0 & \ddots & \frac{s\mu}{\lambda+s\mu} & 0 & \frac{\lambda}{\lambda+s\mu} & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{pmatrix} \tag{1}$$

Consider also an $M/M/\infty$ system, with the same distributions of arrivals and service times as in the previous system, and with no waiting room.

Let $\tilde{X}_k$ be the number of customers in the system just prior to the arrival of the kth customer. Therefore, $\tilde{X} = (\tilde{X}_k; k = 0, 1, \ldots)$ is an homogeneous Markov chain with a state space $\mathbf{N} = \{0, 1, 2, \ldots\}$ and with a transition operator $\tilde{P} = (\tilde{P}_{ij})_{i,j\geq 0}$ where:

$$\tilde{P}_{ij} = \begin{pmatrix}
\frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\
\frac{\mu}{\lambda+\mu} & 0 & \frac{\lambda}{\lambda+\mu} & 0 & \cdots & 0 & 0 & 0 & \cdots \\
0 & \frac{2\mu}{\lambda+2\mu} & 0 & \frac{\lambda}{\lambda+2\mu} & \ddots & 0 & 0 & 0 & \ddots \\
0 & 0 & \frac{3\mu}{\lambda+3\mu} & 0 & \ddots & \ddots & 0 & 0 & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ddots & \frac{s\mu}{\lambda+s\mu} & 0 & \frac{\lambda}{\lambda+s\mu} & 0 & \ddots \\
0 & 0 & 0 & 0 & \ddots & \frac{(s+1)\mu}{\lambda+(s+1)\mu} & 0 & \frac{\lambda}{\lambda+(s+1)\mu} & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{pmatrix} \tag{2}$$

Let π and $\tilde{\pi}$ be, respectively, the stationary distributions of the $M/M/s$ and $M/M/\infty$ systems.

3. The strong stability concept

3.1. Preliminary notations

Let $\mathcal{M} = \{\nu_j\}$ be the space of finite measures on $\mathbf{N}$, and let $\mathcal{N} = \{f(j)\}$ be the space of bounded measurable functions on $\mathbf{N}$. We associate with each transition kernel P the linear mapping:

$$(\mu P)_k = \sum_{j \geq 0} \mu_j P_{ik}, \tag{3}$$

$$(Pf)(k) = \sum_{i \geq 0} f(i) P_{ki}. \tag{4}$$

Introduce on $\mathcal{M}$ the υ-norm of the form:

$$\|\nu\|_\upsilon = \sum_{j \geq 0} \upsilon(j) |\nu_j|, \tag{5}$$

where $\upsilon(k) = \beta^k$, for all $k \in \mathbf{N}$ and $\beta > 1$ is a real parameter. This norm induces in the space $\mathcal{N}$ the norm

$$\|f\|_\upsilon = \sup_{k \geq 0} \frac{|f(k)|}{\upsilon(k)}. \tag{6}$$

Moreover, for all $\nu \in \mathcal{M}$ and $f \in \mathcal{N}$, the symbols νf and $f \circ \nu$ denote respectively the summation and the kernel defined as below

$$\nu f = \sum_{k=0}^{+\infty} f(k) \nu_k, \tag{7}$$

$$(f \circ \mu)(k, j) = f(k) \mu_j, \text{ for all } (k, j) \in \mathbf{N} \times \mathbf{N}. \tag{8}$$

Let us consider $\mathcal{B}$, the space of linear operators, with the norm

$$\|Q\|_\upsilon = \sup_{k \geq 0} \frac{1}{\upsilon(k)} \sum_{j \geq 0} \upsilon(j) Q_{kj}. \tag{9}$$

3.2. Strong stability criteria

For more details on the strong stability method, see [1, 10, 11].

Let us give the definition of the strong stability (qualitative aspect) for an homogeneous Markov chain in the phase state $(\mathbf{N}, \mathcal{B}(\mathbf{N}))$ with respect to the υ-norm. Here $\mathcal{B}(\mathbf{N})$ is the σ-algebra generated by the singletons $\{j\}$.

Definition 1 *(see [1, 11]) A Markov chain X with a transition kernel P and an invariant measure π is said to be strongly υ-stable with respect to the norm $\|.\|_\upsilon$ if $\|P\|_\upsilon < \infty$ and each stochastic kernel Q on the space $(\mathbf{N}, \mathcal{B}(\mathbf{N}))$ in some neighborhood $\{Q : \|Q - P\| < \epsilon\}$ has a unique invariant measure $\mu = \mu(Q)$ and $\|\pi - \mu\|_\upsilon \to 0$ as $\|Q - P\|_\upsilon \to 0$*

4. Strong stability in the $M/M/\infty$ system

4.1. Strong stability conditions

Proofs of the results provided in this section, are based on theoretical results (theorems and corollaries) given in [1, 10].

Theorem 1 *Suppose that in the $M/M/\infty$ system, the condition $\lambda/\mu < 1$ is fulfilled. Then, there exists $\beta \in \,]1, \mu/\lambda[$ such that*

$$\rho = \left(\frac{1}{\beta} \left(\frac{\mu + \lambda \beta^2}{\lambda + \mu} \right) \right) < 1. \tag{10}$$

In addition, for all β such that $1 < \beta < \mu/\lambda$ the embedded Markov chain $\tilde{X}$ is $\upsilon-$strongly stable for the test function $\upsilon(k) = \beta^k$.

4.2. Strong stability estimates
Theorem 2 *Suppose that the conditions of the preceding theorem hold. Then, for all $1 < \beta < \mu/\lambda$, we have:*

$$\|P - \widetilde{P}\|_v = \frac{\alpha(1 + \beta^2)}{\beta(s + \alpha)}, \tag{11}$$

and under the condition:

$$s \geq \alpha\left(\left(\frac{1 + \beta^2}{\beta - 1}\right)\left(\frac{1 + \alpha}{1 - \alpha\beta}\right)\left(1 + e^{\alpha(\beta-1)}\right) - 1\right), \tag{12}$$

we have

$$\|\pi - \tilde{\pi}\|_v \leq \frac{\alpha(1 + \alpha)(1 + \beta^2)\left(1 + e^{\alpha(\beta-1)}\right)\left(e^{\alpha(\beta-1)}\right)}{(s + \alpha)(\beta - 1)(1 - \alpha\beta) - \alpha(1 + \alpha)(1 + \beta^2)\left(1 + e^{\alpha(\beta-1)}\right)} = B_\beta. \tag{13}$$

Where $\alpha = \lambda/\mu$.

5. Numerical application
To be able to put into practice the previous theoretical results concerning the approximation of the $M/M/s$ system by the $M/M/\infty$ one, we have developed the following algorithm:

5.1. Algorithm
Step 1. Introduce input parameters: inter-arrivals mean rate λ, number of servers s and service mean rate μ.

Step 2. Verify the existence of β_0
    if $\mu > \lambda$ **then**
    the system is stable and **goto** step 3
    else   *disp 'the system is not stable'* and **goto** step 7.

Step 3. Determine the constant $\beta_0 := \arg\max_\beta \rho < 1$.

Step 4. Determine the constant $\beta_{min} := \arg\min_\beta \{1 \leq \beta \leq \beta_0 /$ the lowest value of β satisfying the condition **(12)**$\}$.

Step 5. Determine the constant $\beta_{max} := \arg\max_\beta \{1 \leq \beta \leq \beta_0 /$ the highest value of β satisfying the condition **(12)**$\}$.

Step 6. Determine the constant $\beta_{opt} :=$ the value of β minimizing the bound **(13)** where $\beta_{min} \leq \beta \leq \beta_{max}$.

Step 7. end.

For numerical applications, we consider the two following situations:

5.2. Situation 1: Variation of the number of servers s of the $M/M/s$ system
We fix the service rate $\mu = 10$ and some values for λ (for example $\lambda = 1$ (ie. the load $\alpha = \lambda/\mu = 0.1$) and $\lambda = 3$ (ie. $\alpha = 0.3$)), and for each fixed value of α we vary the values of the number of servers s. Then, for each value of s we compute the value of β minimizing the bound (13). For this value of β noted $\beta_{optimal}$, we compute the minimal bound $B_{\beta_{optimal}}$ (algorithmic error). The obtained results are listed in table 1 (case $\alpha = 0.1$) and table 2 (case $\alpha = 0.3$). The variations of the bound $B_{\beta_{Optimal}}$ in function of s for the case $\alpha = 0.1$ (respectively the case $\alpha = 0.3$) are represented in figure 1 (respectively figure 3). Figure 2 (respectively figure 4) compares between the algorithmic and simulation errors in function of s for the case $\alpha = 0.1$ (respectively the case $\alpha = 0.3$).

5.3. Situation 2: Variation of the load $\alpha = \lambda/\mu$ of the $M/M/s$ system
We fix the service rate $\mu = 20$, the number of servers $s = 20$, and we vary λ from 1 to 10 with step 0.5. Then, for each fixed value of $\alpha = \lambda/\mu$ we compute the value of β minimizing the bound (13). For this value of β noted $\beta_{optimal}$, we compute the minimal bound $B_{\beta_{optimal}}$ (algorithmic error). The obtained results are listed in table 3. The variations of the bound $B_{\beta_{Optimal}}$ in function of α are represented in figure 5. Figure 6 compares between the algorithmic and simulation errors in function of α.

Table 1. Values of $\beta_{optimal}$ and $B_{\beta_{optimal}}$ (case $\alpha = 0.1$).

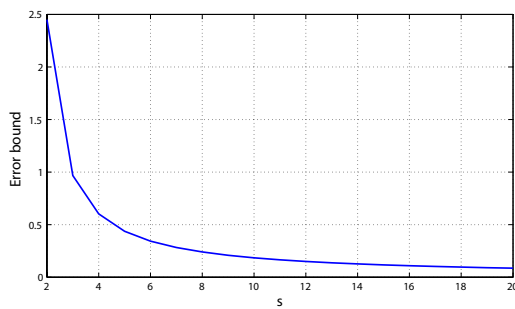| s	$\beta_{optimal}$	Error $(B_{\beta_{optimal}})$
1	$\times$	$\times$
2	1.9890	2.4503
3	1.9620	0.9668
4	1.9490	0.6020
5	1.9420	0.4370
6	1.9360	0.3430
7	1.9330	0.2823
8	1.9300	0.2398
9	1.9280	0.2084
10	1.9260	0.1843
11	1.9250	0.1652
12	1.9240	0.1497
13	1.9230	0.1369
14	1.9220	0.1260
15	1.9210	0.1168
16	1.9210	0.1088
17	1.9200	0.1019
18	1.9200	0.0958
19	1.9190	0.0903
20	1.9190	0.0855

Figure 1. Variations of the bound $B_{\beta_{Optimal}}$ in function of s (case $\alpha = 0.1$).
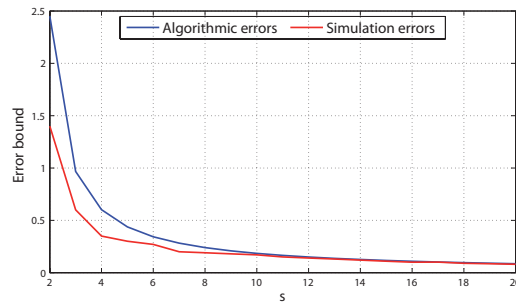

Figure 2. Comparison of algorithmic and simulation errors in function of s (case $\alpha = 0.1$).

Table 2. Values of $\beta_{optimal}$ and $B_{\beta_{optimal}}$ (case $\alpha = 0.3$).

s	$\beta_{optimal}$	Error $(B_{\beta_{optimal}})$
≤ 9	$\times$	$\times$
10	1.6040	22.3974
11	1.5960	7.6876
12	1.5880	4.6363
13	1.5830	3.3175
14	1.5780	2.5822
15	1.5730	2.1133
16	1.5700	1.7883
17	1.5660	1.5498
18	1.5630	1.3674
19	1.5610	1.2233
20	1.5590	1.1067

Remark 1 *The symbol ($\times$) in the the different tables indicates that the system is unstable for the fixed parameters and for the test function $v(k) = \beta^k$.*
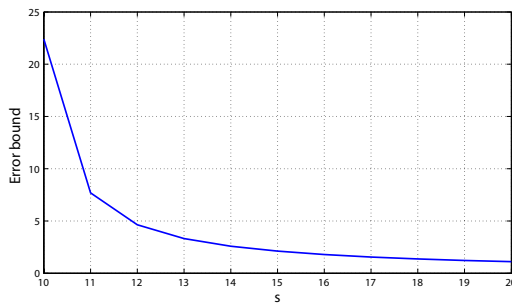
Figure 3. Variations of the bound $B_{\beta_{Optimal}}$ in function of s (case $\alpha = 0.3$).
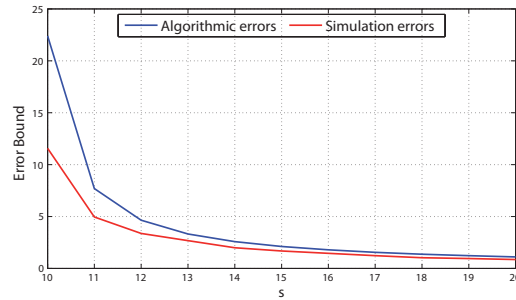

Figure 4. Comparison of algorithmic and simulation errors in function of s (case $\alpha = 0.3$).

Table 3. Values of $\beta_{optimal}$ and $B_{\beta_{optimal}}$ (case $\mu = 20$ and $s = 20$).

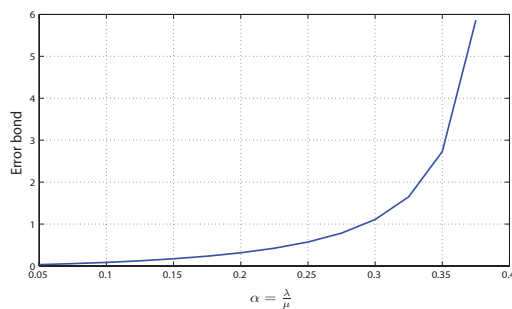| α	$\beta_{optimal}$	Error ($B_{\beta_{optimal}}$)
0.0500	2.1050	0.0322
0.0750	2.0020	0.0556
0.1000	1.9190	0.0855
0.1250	1.8490	0.1234
0.1500	1.7900	0.1718
0.1750	1.7380	0.2340
0.2000	1.6940	0.3152
0.2250	1.6540	0.4234
0.2500	1.6190	0.5716
0.2750	1.5870	0.7837
0.3000	1.5590	1.1067
0.3250	1.5330	1.6488
0.3500	1.5100	2.7282
0.3750	1.4890	5.8599
0.4000	1.4710	107.3026
$\alpha \geq 0.4250$	$\times$	$\times$

Figure 5. Variations of the bound $B_{\beta_{Optimal}}$ in function of α.
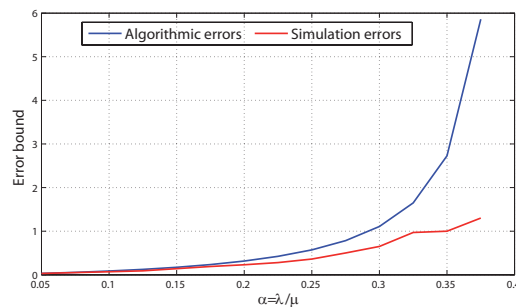

Figure 6. Comparison of algorithmic and simulation errors in function of α.

5.4. Discussion of results
- The error (B_β) decreases exponentially with the increase of the number of servers s (see figures 1 and 3). Moreover, this error converges to zero when s becomes large. This constatation confirms the tendency of the $M/M/s$ system to behave as an $M/M/\infty$ one, when s is rather large. Note also that the system becomes stable starting from the value $s = 2$ for the case $\alpha = 0.1$ (see table 1), whereas it becomes stable starting from the value $s = 10$ for the case $\alpha = 0.3$ (see table 2). This highlights the role of the load α of the system

in its stability.

- The conditions and the stability bound of the system depend exponentially on the parameters of the system. Indeed, the error (B_β) increases exponentially with the increase of the load of the system α (see figure 5), where it passes from the stability state to the instability state starting from the value $\alpha = 0.4250$ (see table 3). This means that the stability of the system is strictly related not only to the number of its servers but also to its load.

- According to figures 2, 4 and 6, we notice that the simulation results are always lower than the algorithmic ones. This confirms that the bound (B_β) is an upper bound for the deviation $\|\pi - \tilde{\pi}\|_v$.

6. Conclusion

In this work, we applied for the first time the strong stability method on the $M/M/\infty$ queue which is subject to a small perturbation of its structure. The application of the method in question, allows us to determine the stability conditions of the $M/M/\infty$ system and to obtain stability bounds for the stationary characteristics of the $M/M/s$ system by those of the $M/M/\infty$ one. To validate and to illustrate the manner in which the theoretical results can be exploited in practice, we have presented numerical examples based on simulation studies. We may extend this analysis to the approximation case of the $GI/M/s$ system by the $GI/M/\infty$ one.

References

[1] Aïssani D and Kartashov N V 1983 Ergodicity and stability of Markov chains with respect to operator topology in the space of transition kernels *Compte Rendu Academy of Sciences U.S.S.R* A **11** 3–5.
[2] Bareche A and Aïssani D 2008 Kernel density in the study of the strong stability of the $M/M/1$ queueing system *Operations Research Letters* **36** 535–538.
[3] Benouaret Z and Aïssani D 2010 Strong stability in a two-dimensional classical risk model with independent claims *Scandinavian Actuarial Journal* **2** 83–92.
[4] Bouallouche-Medjkoune L and Aïssani D 2006 Measurement and performance of the strong stability method *Theor. Probab. Math. Statist.* **72** 1–9.
[5] Ferreira M A M 2013 *Infinite servers queue systems computational simulation* Proceedings: 12th Conference on Applied Mathematics APLIMAT 2013, At Bratislava, vol. 1, 1–16.
[6] Gross D, Shortie J F, Thompson J M and Harris C M 2013 *Bounds and Approximations. Fundamentals of Queueing Theory* John Wiley and Sons, Fourth edition.
[7] Guillemin F M, Mazumdar R R and Simonian A D 1996 On heavy traffic approximations for transient characteristics of $M/M/\infty$ queues *ournal of Applied Probability* **33(2)** 490–506.
[8] Halfin S and Whitt W 1981 Heavy traffic limits for queues with many exponenetial servers *Operation Research* **29(3)** 567–588.
[9] Helary J.M and Pedrono R 1983 *Recherche opérationnelle, Travaux Dirigés* Hermann.
[10] Kartashov N V 1986 Strong stability of Markov chains *Journal of Soviet Mathematics* **34** 1493–1498.
[11] Kartashov N V 1996 *Strong Stable Markov Chains* TbiMC Scientific Publishers, Utrecht, Kiev.
[12] Rabta B and Aïssani D 2005 Strong stability in an (R, s, S) inventory model *Int. J. Prod. Econ.* **4** 159171.