



HAL
open science

Stability analysis of two-class retrial systems with constant retrial rates and general service times

Konstantin Avrachenkov, Evsey Morozov, Ruslana Nekrasova

► **To cite this version:**

Konstantin Avrachenkov, Evsey Morozov, Ruslana Nekrasova. Stability analysis of two-class retrial systems with constant retrial rates and general service times. *Performance Evaluation*, 2023, 159, pp.102330. 10.1016/j.peva.2022.102330 . hal-04372687

HAL Id: hal-04372687

<https://inria.hal.science/hal-04372687v1>

Submitted on 4 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Stability analysis of two-class retrial systems with constant retrial rates and general service times[☆]

Konstantin Avrachenkov^{a,1,*}, Evsey Morozov^{b,c,d}, Ruslana Nekrasova^{b,c}

^a*Inria Sophia Antipolis, France*

^b*Institute of Applied Mathematical Research, KarRC of RAS, Petrozavodsk, Russia*

^c*Petrozavodsk State University, Petrozavodsk, Russia*

^d*Moscow Center for Fundamental and Applied Mathematics, Moscow State University, Moscow, Russia*

Abstract

We establish stability criterion for a two-class retrial system with Poisson inputs, general class-dependent service times and class-dependent constant retrial rates. We also characterise an interesting phenomenon of partial stability when one orbit is tight but the other orbit goes to infinity in probability. All theoretical results are illustrated by numerical experiments.

Keywords: Retrial queues, Constant retrial rate, Multi-class queues, Stability, Partial stability

2020 MSC: 60K25, 60J10, 60J27, 60J74, 60G50

1. Introduction

In this work, we consider a two-class retrial system with a single server and no waiting space associated with the server. If an incoming job finds the server busy, the job goes to the orbit associated with its class. The jobs blocked
5 on a class-dependent orbit attempt to access the server after class-dependent exponential retrial times in FIFO manner. This is different from the classical

[☆]The research of EM was published with the financial support of the Ministry of Education and Science of the Russian Federation as part of the program of the Moscow Center for Fundamental and Applied Mathematics under the agreement No. 075-15-2022-284 .

*Corresponding author

Email addresses: k.avrachenkov@inria.fr (Konstantin Avrachenkov),
emorozov@karelia.ru (Evsey Morozov), ruslana.nekrasova@mail.ru (Ruslana Nekrasova)

retrial policy when each job retries independently with respect to the other orbit jobs. The jobs initially arrive to the system according to Poisson processes. The service times are generally distributed. The arrival processes as well as service
10 times are class-dependent.

An interested reader can find the description of various types of retrial systems and their applications in the books and surveys [1, 2, 3, 4, 5]. Specifically, the multi-class retrial systems with constant retrial rate can be applied to computer networks [6, 7], wireless networks [8, 9, 10, 11, 12, 13] and call centers
15 [14, 15].

Let us outline our contributions and the structure of the paper. After providing a formal description of the system in Section 2, we first establish equivalence in terms of stability between the original continuous-time system and a discrete-time system embedded in the departure instants, see Section 3.1. Then,
20 in Section 3.2 we prove the stability criterion for our retrial system. We also give an extension of the stability criterion to the modified system with balking, which is useful for modelling two-way communication systems. In Section 4 we characterize a very interesting regime of partial stability when one orbit is tight and the other orbit goes to infinity in probability. In particular, we show that
25 in this regime, as time progresses, the original two-orbit system becomes equivalent to a single orbit system. Curiously enough, this new single orbit system gains in stability region due to the jobs lost at infinity. Namely, the stability of one orbit is attained in part due to a ‘displacement’ of the customers going from other (unstable) orbit, and it gives a new insight to the transience phenomena.

We mention the most related works in the next paragraph, leaving the detailed description of related works and the comparison of various stability conditions to Section 5. Then, in Section 6 we illustrate all theoretical results by simulations with exponential and Pareto distributions of the service times. We conclude in Section 7 with future research directions.
30

Stability conditions for the single-class retrial systems with constant retrial rates have been investigated in [16, 17, 6, 18, 19]. In [20] the authors have established the necessary stability conditions for the present system that coincide
35

with the sufficient conditions obtained here. In fact, the necessary conditions have been obtained for N -orbit systems, with $N \geq 2$. We would like to note that the proof of the necessary conditions turns out to be much less challenging than the proof of sufficiency of the same conditions. In [7] the necessary and sufficient conditions have been established by algebraic methods for the case of two classes in a completely Markovian setting with the same service rates. In [8] the author, also in the Markovian setting, has generalized the model of [7] to the case of coupled orbits and different service rates. Then, the author of [21] conjectured sufficient conditions for the two-class retrial system in the case of general service times. In [22], using an auxiliary majorizing system, the authors have obtained sufficient stability conditions for a very general multi-class retrial system with $N \geq 2$ classes. Their sufficient conditions coincide with the necessary conditions of the present model in the case of homogeneous classes. Recently, the authors of [15] have also obtained sufficient (but not generally necessary) conditions for the multi-class retrial system with balking. To the best of our knowledge, in this paper we for the first time establish stability criterion for the two-class retrial system with constant retrial rates and general service times. We credit this to the combination of the regenerative approach [23] with the Foster-Lyapunov approach for stability analysis of random walks [24]. We would like to emphasize that the verification of the general conditions from [24] is not at all straightforward. Furthermore, we refined the case of transience described in [24] with the novel notion of partial stability. The concept of partial (local) stability has been studied in [22] in the context of retrial systems and in [25] in a more general context of Markov chains. In the present work, we combine the two approaches to obtain a detailed characterisation of the phenomenon of partial stability in multi-class retrial systems.

2. System description

Consider a single-server two-class retrial queueing system with constant retrial rates. The system has two Poisson inputs with class-dependent rates λ_k

and generic service times $S^{(k)}$ with distribution functions F_k , $k = 1, 2$. There is no waiting space but two orbits. Define the basic three-dimensional process

$$\mathbf{X}(t) = (N(t), X^{(1)}(t), X^{(2)}(t)), \quad t \geq 0, \quad (1)$$

65 where $N(t) = 1$ if the server is busy at instant t^- ($N(t) = 0$, otherwise) and $X^{(k)}(t)$ is the state (size) of orbit k at instant t^- , $k = 1, 2$. If an incoming customer of class k comes to the system and sees that the server is busy, he/she goes to the k -th orbit. The class- k customers retries from orbit k in FIFO manner with exponential retrial times with rate α_k .

In general, the continuous-time process $\{\mathbf{X}(t), t \geq 0\}$ is not Markovian. Now let us construct a discrete-time process, embedded in the process $\{\mathbf{X}(t)\}$ at the departure instants, which turns out to constitute a Markov chain. Denote by $\{D_n, n \geq 1\}$ the sequence of the departure instants, and let $X_n^{(k)} = X^{(k)}(D_n)$ be the number of customers in orbit k just after the n -th departure, $k = 1, 2$. Construct the following two-dimensional discrete-time process

$$\mathbf{X}_n = (X_n^{(1)}, X_n^{(2)}), \quad n \geq 1. \quad (2)$$

It is easy to check that the process $\{\mathbf{X}_n\}$ is a homogeneous irreducible aperiodic Markov chain (MC). Let us define the increments

$$\Delta_{n+1}^{(k)} = X_{n+1}^{(k)} - X_n^{(k)}, \quad k = 1, 2, \quad (3)$$

and introduce the sequence of vectors

$$\mathbf{\Delta}_{n+1} = (\Delta_{n+1}^{(1)}, \Delta_{n+1}^{(2)}), \quad n \geq 1.$$

Then, the dynamics of MC $\{\mathbf{X}_n\}$ dynamics is described by

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{\Delta}_{n+1}, \quad (4)$$

70 where the distribution of $\mathbf{\Delta}_{n+1}$ depends on the value of \mathbf{X}_n only.

2.1. Transition probabilities of the embedded MC

Denote by \mathcal{I}_n and \mathcal{B}_n , the idle and busy periods of the server between the n -th and the $(n + 1)$ -st departures, respectively, $n \geq 1$. Thus, the $(n + 1)$ -st

departure instant can be recursively presented as

$$D_{n+1} = D_n + \mathcal{I}_n + \mathcal{B}_n.$$

Then, let \mathcal{I} and \mathcal{B} be the corresponding generic times. Next, define by $\mathcal{A}_n^{(k)}$ the event, when the $(n+1)$ -st customer in the server belongs to class k , $k = 1, 2$. Note, that on the event $\mathcal{A}_n^{(k)}$, \mathcal{B} is distributed as service time $S^{(k)}$. On the other
75 hand, the distribution of \mathcal{I} depends on the state of the orbits: idle/busy. Now we consider these cases separately.

1. *Both orbits are empty.* In this case $\mathbf{X}_n = (0, 0)$ and the server stays idle until the next arrival. Thus, the idle period \mathcal{I} is exponentially distributed with the rate $\lambda_1 + \lambda_2$ and the mean $\mathbf{E}\mathcal{I} = 1/(\lambda_1 + \lambda_2)$.

80 Denote by $p_l^{(k)}(i)$ the probability that i customers join the orbit k in the interval $[D_n, D_{n+1})$, provided that both orbits are empty at instant D_n and the $(n+1)$ -st customer arriving to the server is from class l . Thus, for $k = 1, 2$; $l = 1, 2$; $i \geq 0$ we can write

$$p_l^{(k)}(i) = \int_0^\infty e^{-\lambda_k t} \frac{(\lambda_k t)^i}{i!} dF_l(t). \quad (5)$$

In fact, with probability

$$\frac{\lambda_l}{\lambda_1 + \lambda_2} p_l^{(k)}(i), \quad (6)$$

the class- l customer occupies the server and i customers join the orbit $k = 1, 2$,
85 resulting in $X_{n+1}^{(k)} = i$.

2. *Only the first orbit is empty.* Note that in this case $\mathbf{E}\mathcal{I} = 1/(\lambda_1 + \lambda_2 + \alpha_2)$. Thus, with probability

$$\frac{\lambda_l}{\lambda_1 + \lambda_2 + \alpha_2} p_l^{(k)}(i), \quad (7)$$

the newly arriving class- l customer occupies the server and i customers join the k -th orbit, resulting in $X_{n+1}^{(1)} = i$ (resp., $X_{n+1}^{(2)} = X_n^{(2)} + i$). Moreover, with probability

$$\frac{\alpha_2}{\lambda_1 + \lambda_2 + \alpha_2} p_2^{(k)}(i), \quad (8)$$

an *orbital* class-2 customer occupies the server and this results in $X_{n+1}^{(1)} = i$ ($X_{n+1}^{(2)} = X_n^{(2)} - 1 + i$).

The symmetric case, when only the second orbit is empty, is treated similarly.

3. *Both orbits are busy.* In the case when class-1 retrial attempt is not successful, we have $X_{n+1}^{(1)} = X_n^{(1)} + i$, $i \geq 0$, with probability

$$\frac{\lambda_1}{(\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2)} p_1^{(1)}(i) + \frac{\lambda_2 + \alpha_2}{(\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2)} p_2^{(1)}(i). \quad (9)$$

(The probability of $X_{n+1}^{(2)} = X_n^{(2)} + i$ is symmetrically defined.) If class- k retrial attempt is successful, we have $X_{n+1}^{(k)} = X_n^{(k)} + i - 1$ with the corresponding probability

$$\frac{\alpha_k}{(\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2)} p_k^{(k)}(i). \quad (10)$$

3. Stability criterion

90 3.1. Stability of the embedded MC and the underlying continuous-time process

In this section we establish a connection between the notion of *stability* (*ergodicity*) of the embedded MC introduced in the previous section and the concept of *positive recurrence*, which is an analogous notion of *stability* for regenerative processes in continuous time. Although it seems quite intuitive that
95 stability of the embedded MC implies the positive recurrence of the underlying continuous-time process and vice versa, it is instructive to give a formal proof of this fact.

Recall the definition of the basic three-dimensional process

$$\mathbf{X}(t) = (N(t), X^{(1)}(t), X^{(2)}(t)), \quad t \geq 0,$$

where $N(t)$ is the indicator function of the server occupancy at time instant t^- and $X^{(k)}(t)$ is the size of orbit k . Denote by t_n the arrival instants of the
100 (superposed) Poisson process and let $\hat{\mathbf{X}}_n = \mathbf{X}(t_n^-)$, $n \geq 1$. We stress that the new hat-notation reflects the fact that the discrete-time process $\{\hat{\mathbf{X}}_n\}$ in general is *not a Markov chain* and evidently differs from the original Markov chain $\{\mathbf{X}_n\}$ obtained by embedding at the *departure instants*.

The process $\{\mathbf{X}(t)\}$ is *regenerative* with the *regeneration instants* $\{T_n\}$ defined recursively as

$$T_{n+1} = \min_i(t_i > T_n : \hat{\mathbf{X}}_i = \mathbf{0}), \quad T_0 = 0, \quad n \geq 0. \quad (11)$$

We note that the equality $\hat{\mathbf{X}}_i = \mathbf{0}$ is component-wise and that the regeneration epoch T_n represents the n th arrival instant when a customer meets the system totally idle, and the distribution of the process $\{\mathbf{X}(t)\}$ after the instant T_n is independent of n and the prehistory preceding instant T_n . (A detailed description of the regenerative method can be found in Chapter VI of the fundamental book [26].)

We assume that the first customer arrives in the empty system at instant $t = 0$. Such a setting is called *zero initial state* [27], and the corresponding regenerative process is called *zero-delayed* [26]. Denote by T generic *regeneration period* (which is distributed as any difference $T_{n+1} - T_n$). Then the regenerative process is called *positive recurrent* if $\mathbf{E}T < \infty$. Denote by τ the generic interarrival (exponential) time in the superposed Poisson input process, which has rate $\lambda = \sum \lambda_k$. Because the input process is Poisson, then the regeneration period T is *non-lattice*. (In this particular model the latter means that the regeneration period length has absolutely continuous distribution.) In this case the positive recurrence implies the existence of the stationary distribution of the process $\mathbf{X}(t)$ as $t \rightarrow \infty$ and hence the stability of the system [26]. To study stability, it is much more convenient to work with a one-dimensional process

$$Z(t) = N(t) + X_1(t) + X_2(t), \quad t \geq 0,$$

counting the total number of customers in the system, which is *regenerative with the same regeneration instants* (11). In the following lemma we establish the equivalence between the stability of the embedded MC and the stability of the original continuous-time process for the case of zero initial state.

Lemma 1. *The zero-initial state Markov chain \mathbf{X} is positive recurrent if and only if the process $\{\mathbf{X}(t)\}$ is positive recurrent, that is, if $\mathbf{E}T < \infty$.*

PROOF. If the process $\{\mathbf{X}(t)\}$ is positive recurrent, then it follows by a regenerative argument [26] that the *stationary probability* P_0 , the probability of the system being totally idle, exists and is equal to

$$P_0 = \lim_{t \rightarrow \infty} \mathbb{P}(Z(t) = 0) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{1}(Z(u) = 0) du = \frac{\mathbb{E}\tau}{\mathbb{E}T} > 0, \quad \text{w.p. 1,} \quad (12)$$

where τ denotes the generic inter-arrival time in the superposed Poisson input process. On the other hand, $\{\hat{\mathbf{X}}_n\}$ is the embedded discrete-time regenerative process with the regeneration instants

$$\theta_{n+1} = \min(i : i > \theta_n, \hat{\mathbf{X}}_i = \mathbf{0}), \quad \theta_0 = 0, \quad n \geq 0, \quad (13)$$

with θ denoting generic regeneration period of this discrete-time process. Namely, the generic regeneration cycle is given by $\theta =_{st} \theta_{n+1} - \theta_n$. It is well known that the discrete-time length θ of the regeneration cycle is connected with the continuous-time length T by the following stochastic equality [26] (Chapter X, Propositions 3.1 and 3.2):

$$T =_{st} \sum_{k=1}^{\theta} \tau_k \quad \left(\sum_{\emptyset} := 0 \right),$$

where $\tau_k = t_{k+1} - t_k$ is the k -th inter-arrival interval and the summation index θ is a (randomized) stopping time. It then immediately follows from the Wald's identity that

$$\mathbb{E}T = \mathbb{E}\tau \mathbb{E}\theta.$$

Note that, because $0 < \mathbb{E}\tau < \infty$, then $\mathbb{E}T = \infty$ implies $\mathbb{E}\theta = \infty$, and vice versa. Thus we obtain that $\mathbb{E}\theta < \infty$, and hence the positive recurrence of the basic process $\{\mathbf{X}(t)\}$ implies the positive recurrence of the process $\{\hat{\mathbf{X}}_n\}$ embedded (in the basic process) at the arrival instants. Conversely, $\mathbb{E}\theta < \infty$ implies $\mathbb{E}T < \infty$ as well.

It remains to connect the process $\{\hat{\mathbf{X}}_n\}$ with the embedded MC we studied above. Note that, as in (11), regenerations $\{\theta_n\}$, defined in (13), are generated by the arrivals meeting empty system. On the other hand, θ represents *both the number of arrivals and the number of departures* from the system within

a *continuous-time regeneration period* T . Thus, θ is also generic regeneration period of the embedded MC $\{\mathbf{X}_n\}$. It is worth mentioning that, at each instant of time, the index of a customer which sees the empty system (and generating a new regeneration period of the processes $\{\mathbf{X}(t)\}$ and $\{\hat{\mathbf{X}}_n\}$), differs from the index of a customer *leaving empty system* by no more than one. Thus we obtain that

$$\hat{\pi}_0 = \lim_{n \rightarrow \infty} \mathbb{P}(\hat{\mathbf{X}}_n = \mathbf{0}) = \lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{X}_n = \mathbf{0}) =: \pi_0 = \frac{1}{\mathbb{E}\theta} > 0.$$

130 Thus, $\pi_0 > 0$, and because the MC we consider is aperiodic and irreducible, it is also ergodic. Therefore we see that the two concepts of stability (in continuous and discrete times) agree and the lemma hereby is proven. \square

We note that, at the first sight, the equality $\pi_0 = \hat{\pi}_0$ seems rather surprising because π_0 relates to the MC while $\hat{\pi}_0$ relates to the regenerative process $\{\hat{\mathbf{X}}_n\}$ 135 which in general is not Markov.

3.2. Stability of the embedded Markov chain

The results of this section are based on the general stability conditions for two-dimensional MCs, obtained in [24].

Let us first introduce some additional notations for the embedded Markov 140 chain (2) needed for the application of the results from [24]. Specifically, let us derive the drifts of the embedded MC in various regions of the state space.

Denote by M_k^{01} the mean number of customers joining the class- k orbit in the time interval $[D_n, D_{n+1})$, provided $X_n^{(1)} = 0, X_n^{(2)} > 0$. Recall $\mu_k = 1/ES^{(k)}$ and denote

$$\rho_k = \frac{\lambda_k}{\mu_k}, \quad \hat{\rho}_k = \frac{\alpha_k}{\mu_k}, \quad k = 1, 2.$$

Then it follows from (7)–(8) after a simple algebra that

$$M_1^{01} = \sum_{i \geq 1} i \left[\frac{\lambda_1 p_1^{(1)}(i)}{\lambda_1 + \lambda_2 + \alpha_2} + \frac{(\lambda_2 + \alpha_2) p_2^{(1)}(i)}{\lambda_1 + \lambda_2 + \alpha_2} \right] = \frac{\lambda_1(\rho_1 + \rho_2 + \hat{\rho}_2)}{\lambda_1 + \lambda_2 + \alpha_2}, \quad (14)$$

$$M_2^{01} = \frac{\lambda_2(\rho_1 + \rho_2 + \hat{\rho}_2) - \alpha_2}{\lambda_1 + \lambda_2 + \alpha_2}. \quad (15)$$

Similarly, denote by M_k^{10} the mean number of customers joining the class- k orbit in the time interval $[D_n, D_{n+1})$, provided $X_n^{(1)} > 0$, $X_n^{(2)} = 0$, $k = 1, 2$. Then
 145 by analogy with (14) and (15), we obtain

$$M_1^{10} = \frac{\lambda_1(\rho_1 + \rho_2 + \hat{\rho}_1) - \alpha_1}{\lambda_1 + \lambda_2 + \alpha_1}, \quad (16)$$

$$M_2^{10} = \frac{\lambda_2(\rho_1 + \rho_2 + \hat{\rho}_1)}{\lambda_1 + \lambda_2 + \alpha_1}. \quad (17)$$

Continuing in the same way, we denote by M_k^{11} the mean number of customers joining the class- k orbit in the time interval $[D_n, D_{n+1})$, if $X_n^{(1)} > 0$, $X_n^{(2)} > 0$, and obtain

$$M_1^{11} = \frac{(\lambda_1 + \alpha_1)\rho_1 + \lambda_1(\rho_2 + \hat{\rho}_2) - \alpha_1}{\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2}, \quad (18)$$

$$M_2^{11} = \frac{(\lambda_2 + \alpha_2)\rho_2 + \lambda_2(\rho_1 + \hat{\rho}_1) - \alpha_2}{\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2}. \quad (19)$$

150 Our further analysis is based on Theorem A presented in the Appendix, a statement from [24]. Note, that in the general case, Theorem A is applicable under some additional technical conditions (see Appendix A), which hold automatically when the input is Poisson. Denote $\rho = \rho_1 + \rho_2$, $\hat{\rho} = \hat{\rho}_1 + \hat{\rho}_2$. Now we are in a position to state our central result.

Theorem 1. *Two-class retrial system with constant retrial rates, Poisson inputs, general service times and exponential retrials is ergodic if and only if*

$$\rho < \min\left(\frac{\alpha_1}{\lambda_1 + \alpha_1}, \frac{\alpha_2}{\lambda_2 + \alpha_2}\right). \quad (20)$$

155 **PROOF.** Note that the defined above drifts M_k^{01} , M_k^{10} , M_k^{11} , $k = 1, 2$, correspond to the drifts used in the statement of Theorem A (see the appendix).

First, let us consider the conditions for the case (a) of Theorem A. Specifically, the condition $M_1^{11} < 0$ takes the following form

$$\rho_1\left(\lambda_1 + \alpha_1 + (\lambda_2 + \alpha_2)\frac{\mu_1}{\mu_2}\right) < \alpha_1, \quad (21)$$

while the condition $M_2^{11} < 0$ takes the form

$$\rho_2\left(\lambda_2 + \alpha_2 + (\lambda_1 + \alpha_1)\frac{\mu_2}{\mu_1}\right) < \alpha_2. \quad (22)$$

160 The inequalities (21) and (22) can be further rewritten as

$$\lambda_k(\rho + \hat{\rho}) < \alpha_k, \quad k = 1, 2.$$

Next, the first condition $M_1^{11}M_2^{10} - M_2^{11}M_1^{10} < 0$ in (67) becomes, after a tedious algebra (see Appendix B for details),

$$\rho < \frac{\alpha_1}{\lambda_1 + \alpha_1}, \quad (23)$$

while the second condition $M_2^{11}M_1^{01} - M_1^{11}M_2^{01} < 0$ in (67) can be transformed to

$$\rho < \frac{\alpha_2}{\lambda_2 + \alpha_2}. \quad (24)$$

Thus we can express the three ergodic cases (a.1), (b.1) and (c.1) of Theorem A (see inequalities (67), (69) and (71) in Appendix A) in terms of our system parameters:

Case (a.1)

$$\begin{cases} \lambda_k(\rho + \hat{\rho}) < \alpha_k, \quad k = 1, 2, \\ \rho < \frac{\alpha_k}{\lambda_k + \alpha_k}, \quad k = 1, 2; \end{cases} \quad (25)$$

Case (b.1)

$$\begin{cases} \lambda_1(\rho + \hat{\rho}) \geq \alpha_1, \\ \lambda_2(\rho + \hat{\rho}) < \alpha_2, \\ \rho < \frac{\alpha_1}{\lambda_1 + \alpha_1}; \end{cases} \quad (26)$$

Case (c.1)

$$\begin{cases} \lambda_1(\rho + \hat{\rho}) < \alpha_1, \\ \lambda_2(\rho + \hat{\rho}) \geq \alpha_2, \\ \rho < \frac{\alpha_2}{\lambda_2 + \alpha_2}. \end{cases} \quad (27)$$

Our next goal is to simplify the ergodicity conditions (25)–(27). Towards
 165 this end, we rewrite the system (25) in terms of functions of α_1 and α_2 , assuming
 other parameters fixed. The first pair of inequalities in (25) can be transformed

to

$$\alpha_2 < \frac{(1-\rho_1)\mu_2}{\lambda_1}\alpha_1 - \rho\mu_2 =: g_1(\alpha_1), \quad (28)$$

$$\alpha_2 > \frac{\lambda_2}{(1-\rho_2)\mu_1}\alpha_1 + \frac{\lambda_2}{1-\rho_2}\rho =: g_2(\alpha_1), \quad (29)$$

while inequalities $\rho < \alpha_k/(\lambda_k + \alpha_k)$, $k = 1, 2$, become

$$\alpha_k > \frac{\rho}{1-\rho}\lambda_k, \quad k = 1, 2. \quad (30)$$

For fixed values λ_k, μ_k , such that $\rho < 1$, the right hand sides of (28), (29) are
 170 the increasing linear functions of α_1 with a common point (α_1^*, α_2^*) , where

$$\alpha_k^* = \frac{\rho}{1-\rho}\lambda_k, \quad k = 1, 2. \quad (31)$$

The ergodic case (a.1), described by system (25), corresponds to the values of
 (α_1, α_2) such that

$$g_2(\alpha_1) < \alpha_2 < g_1(\alpha_1), \quad (32)$$

$$\alpha_1 > \alpha_1^*. \quad (33)$$

Next we show that the set of (α_1, α_2) corresponding to (32) and (33) is non-
 empty as long as $\rho < 1$. Assume on contrary that

$$g_2(\alpha_1) \geq g_1(\alpha_1), \quad (34)$$

thus (32) is violated. The inequality (34) for linear increasing functions g_1, g_2
 under conditions $\alpha_k > \alpha_k^*$, $k = 1, 2$ implies a similar relation for the coefficients
 in front of α_1 (see (28),(29)), that is

$$\frac{\lambda_2}{(1-\rho_2)\mu_1} \geq \frac{(1-\rho_1)\mu_2}{\lambda_1}. \quad (35)$$

Multiplying both parts of (35) by $(1-\rho_2)\lambda_1/\mu_2 \geq 0$, we obtain
 $\rho_1\rho_2 \geq (1-\rho_1)(1-\rho_2)$, which is equivalent to $\rho = \rho_1 + \rho_2 > 1$ and yields
 175 a contradiction.

Now, similarly, we describe the stability regions (b.1) and (c.1), presented
 in (26) and (27), in terms of the functions $g_1(\alpha_1)$ and $g_2(\alpha_1)$ as follows:

$$\text{Case (b.1):} \quad \alpha_2 \geq g_1(\alpha_1), \quad \alpha_1 > \alpha_1^*,$$

$$\text{Case (c.1):} \quad \alpha_2 \leq g_2(\alpha_1), \quad \alpha_2 > \alpha_2^*.$$

Next, by combining the three cases, we conclude that MC is ergodic, if $\alpha_1 > \alpha_1^*$ and $\alpha_2 > \alpha_2^*$, which is equivalent in fact to $\rho < \alpha_k / (\lambda_k + \alpha_k)$, $k = 1, 2$. Thus, 180 the conditions (25)–(27) can be written as (20).

So far we have shown that (20) is a sufficient condition for stability (ergodicity). To show that this condition is also necessary, we refer to the paper [20], where it is shown, in the adopted notation and with $\rho = \rho_1 + \dots + \rho_N$, that if N -class retrial system with Poisson inputs is ergodic then

$$\lambda_k \rho < \alpha_k (1 - \rho), \quad k = 1, \dots, N. \quad (36)$$

(Indeed, in paper [20], we have applied an equivalent notion of positive recurrence in the framework of the regenerative approach, see Lemma 1 above.) We can rewrite (36) (for $N=2$) as

$$\lambda_k \rho < \alpha_k (1 - \rho), \quad k = 1, 2,$$

implying

$$\rho < \min_{k=1,2} \frac{\alpha_k}{\lambda_k + \alpha_k}. \quad (37)$$

Thus (37) coincides with (20) and condition (20) is also the necessary stability condition. \square

Remark. *It follows from (37) that if the two-class retrial system under consideration is ergodic, then $\rho < 1$.*

185 3.3. Stability of a system with balking

We can assume an extra feature in the system under consideration as follows. If a primary class- k customer meets busy server, he joins the corresponding orbit with a given (*balking*) probability $b_k \geq 0$, $k = 1, 2$ and leaves the system with probability $1 - b_k$. In this case, the stability condition of Theorem 1 transforms to

$$b_1 \rho_1 + b_2 \rho_2 < \min \left(\frac{\alpha_1}{\alpha_1 + \lambda_1}, \frac{\alpha_2}{\alpha_2 + \lambda_2} \right). \quad (38)$$

This is an immediate extension. Namely, taking into account balking policy, we redefine the transition probabilities (5) and the statement (38) is then proved

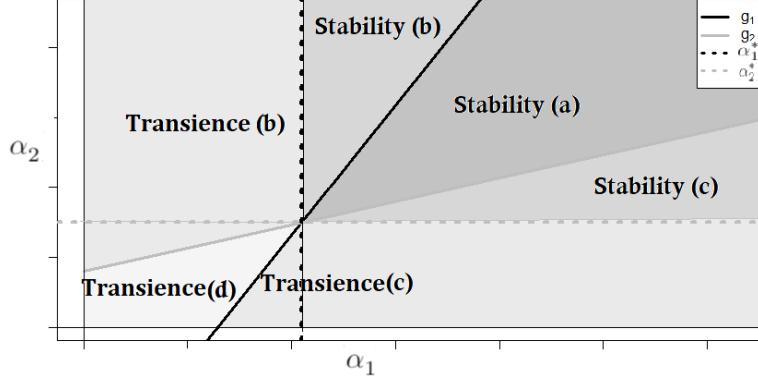


Figure 1: Stability/transience regions.

by the same arguments as Theorem 1. We note that this modification of the system is useful to model two-way communication systems, for more details see
 190 e.g., [13, 15].

4. Partial stability

Let us now discuss an effect of *partial stability* to the best of our knowledge first discovered in [22]. In the case of two classes of customers, the statement of Theorem 4 from [22] can be qualitatively formulated as follows: under some
 195 (given below) conditions, the class-1 orbit queue stays tight while the class-2 orbit queue goes to infinity in probability. (Of course, by symmetry, this can be formulated for the opposite case, when the second orbit is tight while the first orbit diverges.) By the evident reason, this statement can be regarded as the case of *partial stability*.

200 The purpose of this section is firstly to show that, in terms of the embedded MC $\{\mathbf{X}_n\}$, the partial stability corresponds to the transient case (c.2) of Theorem A, i.e., $M_1^{11} < 0$, $M_2^{11} \geq 0$ and condition $M_2^{11}M_1^{01} - M_1^{11}M_2^{01} > 0$.

Secondly, by establishing a relation with a single-orbit system, we shall show how to describe the long run behaviour of the stable orbit.

205 Note that the stability conditions which correspond to transience case (c.2) can be defined in terms of the load and rate coefficients as follows:

$$\alpha_1 > \lambda_1(\rho + \hat{\rho}), \quad (39)$$

$$\alpha_2 \leq \lambda_2(\rho + \hat{\rho}), \quad (40)$$

$$\rho > \frac{\alpha_2}{\lambda_2 + \alpha_2}. \quad (41)$$

It is important to note that (39) can be written as

$$\rho_1 < \frac{\hat{\rho}_1}{\rho + \hat{\rho}_1 + \hat{\rho}_2} < 1. \quad (42)$$

Now we show that, provided conditions (39) and (41) hold, then they imply condition (40), which turns out to be *redundant*.

210 Next consider in detail inequalities (39)–(41) in all three possible sub-cases when $\rho_1 < 1$.

Sub-case 1: $\rho_1 < 1, \rho < 1$. In this case, it is convenient to rewrite conditions (39)–(41) as follows:

$$\alpha_1 > \frac{\lambda_1}{(\mu_1 - \lambda_1)} \cdot \frac{\mu_1}{\mu_2} \alpha_2 + \frac{\lambda_1}{\mu_1 - \lambda_1} \mu_1 \rho, \quad (43)$$

$$\alpha_1 \geq \frac{(\mu_2 - \lambda_2)}{\lambda_2} \cdot \frac{\mu_1}{\mu_2} \alpha_2 - \mu_1 \rho, \quad (44)$$

$$\alpha_2 < \frac{\rho}{1 - \rho} \lambda_2 =: \alpha_2^*, \quad (45)$$

respectively. Next assume that the following relation holds between the r.h.s. of conditions (43) and (44)

$$\frac{\lambda_1}{(\mu_1 - \lambda_1)} \cdot \frac{\mu_1}{\mu_2} \alpha_2 + \frac{\lambda_1}{\mu_1 - \lambda_1} \mu_1 \rho \leq \frac{(\mu_2 - \lambda_2)}{\lambda_2} \cdot \frac{\mu_1}{\mu_2} \alpha_2 - \mu_1 \rho. \quad (46)$$

215 After some algebra, (46) transforms to the inequality $\alpha_2 \geq \lambda_2 \rho / (1 - \rho)$, which contradicts (45). Thus, we have

$$\alpha_1 > \frac{\lambda_1}{(\mu_1 - \lambda_1)} \cdot \frac{\mu_1}{\mu_2} \alpha_2 + \frac{\lambda_1}{\mu_1 - \lambda_1} \mu_1 \rho > \frac{(\mu_2 - \lambda_2)}{\lambda_2} \cdot \frac{\mu_1}{\mu_2} \alpha_2 - \mu_1 \rho, \quad (47)$$

and inequality (43) implies inequality (44); or equivalently, (39) implies (40). Therefore, the latter condition is redundant.

220 *Sub-case 2:* $\rho_1 < 1, \rho_2 > 1$. In this case, from the condition (40) we have $\lambda_2(\rho + \hat{\rho}_1) > \alpha_2(1 - \rho_2)$. Note

$$\lambda_2(\rho + \hat{\rho}_1) > 0 > \alpha_2(1 - \rho_2). \quad (48)$$

Thus in this sub-case (40) always holds, and as a consequence is redundant.

Sub-case 3: $\rho_1 < 1, \rho_2 < 1, \rho > 1$. In this case, conditions (43), (44) remain unchanged, while condition (45) becomes

$$\alpha_2 > \frac{\rho}{1 - \rho} \lambda_2. \quad (49)$$

225 As in Sub-case 1, it is easy to check that, provided inequality (49) holds, the condition (47) holds as well, and thus condition (44), or equivalently, condition (40) is redundant in this sub-case as well.

Consequently a pair of conditions $M_1^{11} < 0$ and $M_2^{11}M_1^{01} - M_1^{11}M_2^{01} > 0$ (or its analogues (39) and (41)) define the transience case (c.2) in our model.

Before we formulate next results, let us recall the definition of the failure rate

$$r(x) := \frac{f(x)}{1 - F(x)},$$

230 of a non-negative absolutely continuous distribution F with density f , defined for all x such that $1 - F(x) > 0$. We say that a distribution belongs to class \mathcal{D} if its failure rate satisfies $\inf_{x \geq 0} r(x) > 0$. (Some, fairly common, distributions satisfying this requirement can be found in [22].)

Theorem 2. *If, in the initially empty system, conditions*

$$\alpha_1 > \lambda_1(\rho + \hat{\rho}), \quad \rho > \alpha_2/(\lambda_2 + \alpha_2), \quad (50)$$

235 *hold and the distribution F_k of service times of class- k customers belongs to class \mathcal{D} , $k = 1, 2$, then the 1-st orbit is tight and the 2-nd orbit increases in probability, that is $X^{(2)}(t) \Rightarrow \infty$.*

PROOF. Recall notation

$$\rho = \rho_1 + \rho_2, \quad \hat{\rho} = \frac{\alpha_1}{\mu_1} + \frac{\alpha_2}{\mu_2}.$$

and also denote

$$P_L = \frac{\rho + \hat{\rho}}{1 + \rho + \hat{\rho}}. \quad (51)$$

Following [22], we consider an auxiliary two-class system, denoted by $\tilde{\Sigma}$, with the same service times as in the original system and with two ‘infinitely loaded’ orbits. Thus, as the original system, the system $\tilde{\Sigma}$ has two Poisson inputs from ‘outside’ (‘exogenous’ customers) with the rates λ_k and, in addition, two Poisson inputs with the rates α_k which we call the retrial inputs, $k = 1, 2$. We ‘color’ each new (exogenous) class- k job meeting server busy, and then it joins the queue in orbit k . At that orbit the colored jobs have a priority: a non-colored job from orbit k cannot attempt to occupy the server while colored jobs present in orbit k . In other words, colored jobs stand ahead in the (infinitely long) orbit queue. It is intuitive (and proved in [22]) that the colored class- k orbit size (in the system $\tilde{\Sigma}$) stochastically dominates class- k orbit size in the original system, provided $F_k \in \mathcal{D}$, $k = 1, 2$. (We stress that, in the system $\tilde{\Sigma}$ the class- k retrial jobs continue to arrive to the server even if class- k colored jobs are exhausted.) Then it follows that P_L defined in (51) is the stationary probability that a customer meets server busy in the system $\tilde{\Sigma}$. Moreover, it is shown in [22] (Theorem 4 there) that, if the system is initially empty, and the following conditions hold:

$$P_L < \frac{\alpha_1}{\lambda_1 + \alpha_1}, \quad (52)$$

$$\rho \geq \frac{\alpha_2}{\lambda_2 + \alpha_2}, \quad (53)$$

then the first orbit is tight while $X^{(2)}(t) \Rightarrow \infty$.

Remark. The proof of tightness in [22] is based on the monotonicity property mentioned above which in turn has been proven for $F_k \in \mathcal{D}$ only. We believe that the latter requirement is only technical and indeed is not needed for stability. This conjecture, in particular, is supported in Section 6.2 by a numerical example with Pareto service time distribution which does not belong to class \mathcal{D} .

Thus, our goal is to show that conditions (50) of Theorem 2 coincide with conditions (52), (53), and, for this purpose, we write conditions (50) separately

265 as

$$\alpha_1 > \lambda_1(\rho + \hat{\rho}), \quad (54)$$

$$\rho > \frac{\alpha_2}{\lambda_2 + \alpha_2}. \quad (55)$$

Because

$$P_L = \frac{\rho + \hat{\rho}}{1 + \rho + \hat{\rho}},$$

it is straightforward to show that (52) coincides with (54). It remains to note that (55) is a particular case of condition (53). Thus, conditions (50) define transience case (c.2) and simultaneously are the assumptions of Theorem 2. Hence, this means that the first orbit is tight and $X^{(2)}(t) \Rightarrow \infty$. \square

270 Theorem 2 (as well as Theorem 4 in [22]) shows that, unlike conventional retrial systems, in the constant retrial rate system, stability/instability may happen *locally*. Denote

$$P_B^{(2)} = \frac{\rho + \hat{\rho}_2}{\rho_2 + \hat{\rho}_2 + 1}. \quad (56)$$

It is shown in [22] that, under conditions (52) and (53), $P_B^{(2)}$ is the limiting fraction of the mean busy time of the server, that is, in evident notation,

$$\lim_{t \rightarrow \infty} \frac{EB(t)}{t} = P_B^{(2)}.$$

Next, we note that the condition (53) can be written as

$$\frac{\rho + \hat{\rho}_2}{\rho_2 + \hat{\rho}_2 + 1} = P_B^{(2)} \leq \rho. \quad (57)$$

275 In particular, if the inequality (53) is strict, then $P_B^{(2)} < \rho$. On the other hand, if (53) holds as equality, then the limiting fraction of busy time $P_B^{(2)}$ becomes equal to ρ . This surprising result has the following intuitive explanation (first remarked in [20]). Assume the inequality (53) is strict and rewrite it as

$$\lambda_2 \rho > \alpha_2(1 - \rho). \quad (58)$$

Then it follows that the rate of the input of the (blocked) customers to orbit 2 is bigger than the (potential) rate of the output from this orbit. In this case one expects that orbit 2 size increases in such a way that a *non-negligible* fraction of the blocked class-2 customers joining the ‘end’ of the infinitely increasing orbit 2, that is, those customers ‘go to infinity’ and never reach the server in the future. In other words, this fraction of customers can be treated as ‘lost’ or ‘disappeared’ work. As a result, the limiting fraction of the ‘processed’ work becomes less than ρ .

If, instead of the inequality (53) (or (58)) we assume the equality, that is

$$\lambda_2\rho = \alpha_2(1 - \rho), \quad (59)$$

then the orbit-2 size increases in general *in probability* only (i.e., ‘slower’ than in the previous case). In this case the ‘disappearing’ fraction of class-2 customers joining the infinitely increasing orbit 2 turns out to be *negligible*, and it follows that the limiting fraction of the mean busy time (‘busy probability’) $\mathbf{P}_B^{(2)} = \rho$. It is because ‘almost all’ incoming work is processed and as a result the stationary busy probability in this case coincides with the stationary busy probability in the positive recurrent case.

Again assume that the conditions (52) and (53) hold, and that $\mathbf{P}_B^{(2)} \geq \mathbf{P}_L$, that is

$$(\rho + \hat{\rho}_2)(1 + \hat{\rho} + \rho) \geq (\rho + \hat{\rho})(\rho_2 + \hat{\rho}_2 + 1).$$

After a simple algebra, we obtain the inequality $\lambda_1(\rho + \hat{\rho}) \geq \alpha_1$, contradicting (52). Thus $\mathbf{P}_B^{(2)} < \mathbf{P}_L$. This result has the following intuitive explanation in the bufferless setting. Note that the loss probability \mathbf{P}_L in the auxiliary system, by PASTA, is also the limiting fraction of the server busy time. Because the first orbit in the original system is *tight* then it follows that the fraction of the server idle time in the original system is non-negligible, and as a result, this (limiting) fraction $\mathbf{P}_B^{(2)}$ is strictly dominated by the probability \mathbf{P}_L .

Note that the simulation results in Section 6 illustrate the related phenomena of partial stability in regions 6 and 8 (see Figure 2 below).

4.1. Relation to a single-orbit system

In this section, we establish an intuitively expected result that the stability conditions from Theorem 2 coincide with stability conditions of the following associated single-server two-class system: while class-1 customers meeting server busy join the orbit, class-2 customers arrive as if the 2nd orbit would be permanently busy. In other words, the 2-nd orbit becomes the source of the Poisson process with rate α_2 . Intuitively, the associated system can be considered as a ‘limit’ of the original system under conditions of Theorem 2. Next, we formally justify this intuition. In the limit system, class-2 customers arrive to the server with the input rate

$$\tilde{\lambda}_2 = \lambda_2 + \alpha_2,$$

and leave the system if they find the server busy, while the external class-1 customers arrive to the server with the input rate λ_1 . Thus, the limit system can also be viewed as a single-orbit retrial system with an additional exogenous Poisson input with rate $\tilde{\lambda}_2$. In this single-orbit system, we denote by $\{X(t)\}$ the evolution of the orbit size, and let X_n be the orbit size just after the n -th departure, $n \geq 1$. Recall the probabilities $p_l^{(1)}(i)$, $l = 1, 2$ from (5) and note that, if $X_n = 0$, then $X_{n+1} = i$ with the probability

$$\frac{\lambda_1}{\lambda_1 + \tilde{\lambda}_2} p_1^{(1)}(i) + \frac{\tilde{\lambda}_2}{\lambda_1 + \tilde{\lambda}_2} p_2^{(1)}(i), \quad i \geq 0. \quad (60)$$

If $X_n > 0$, then $X_{n+1} = X_n + i$ with probability

$$\frac{\lambda_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} p_1^{(1)}(i) + \frac{\tilde{\lambda}_2}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} p_2^{(1)}(i),$$

and $X_{n+1} = X_n + i - 1$ with probability

$$\frac{\alpha_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2}.$$

This gives the (conditional) mean orbit size:

$$\begin{aligned}
\mathbb{E}[X_{n+1}|X_n = i] &= \sum_{j \geq 0} \frac{\alpha_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} p_1^{(1)}(i)(i - 1 + j) \\
&+ \sum_{j \geq 0} \left(\frac{\lambda_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} p_1^{(1)}(i) + \frac{\tilde{\lambda}_2}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} p_2^{(1)}(i) \right) \\
&= i + \frac{\lambda_1 + \alpha_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} \rho_1 + \frac{\tilde{\lambda}_2}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} \frac{\lambda_1}{\tilde{\lambda}_2} \rho_2 - \frac{\alpha_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2}.
\end{aligned}$$

Thus, the negative drift takes place (i.e., $\mathbb{E}[X_{n+1}|X_n = i] < i$) if

$$\frac{\lambda_1 + \alpha_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} \rho_1 + \frac{\tilde{\lambda}_2}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2} \frac{\lambda_1}{\tilde{\lambda}_2} \rho_2 < \frac{\alpha_1}{\lambda_1 + \alpha_1 + \tilde{\lambda}_2}.$$

The latter inequality, after a simple algebra, becomes

$$\rho_1(\rho + \rho_1) < \hat{\rho}_1, \tag{61}$$

305 and coincides with condition (54) implying tightness of the 1-st orbit in partially stable scenario for the two-orbit system.

Next we establish a stronger result: the *weak convergence* of the original two-orbit system to the associated single-orbit system. Towards this end, we first prove a monotonicity property of the two-dimensional embedded MC $\{\mathbf{X}_n\}$. Let $\mathbf{x} = (x_1, x_2)$ denote a point of \mathbb{Z}_+^2 . Then, for all $\mathbf{x} > (0, 0)$ and $i, j \geq -1$ we define the probabilities

$$\mathbb{P}(i, j) := \mathbb{P}(X_1^{(1)} = x_1 + i, X_1^{(2)} = x_2 + j | X_0^{(1)} = x_1, X_0^{(2)} = x_2),$$

where, by definition, $\mathbb{P}(-1, -1) = 0$. Define the total input rate to the server by $\Lambda = \lambda_1 + \lambda_2 + \alpha_1 + \alpha_2$, if the both orbits are non-empty. We have the following four alternative cases implying the event

$$\mathcal{A} = \{X_1^{(1)} = x_1 + i, X_1^{(2)} = x_2 + j\},$$

1. with probability $\lambda_1/\Lambda p_1^{(1)}(i)p_1^{(2)}(j)$, when a class-1 new arrival occupies the server, where we use the independence of inputs;
2. with probability $\lambda_2/\Lambda p_2^{(1)}(i)p_2^{(2)}(j)$, when a class-2 new arrival occupies the server;

310

3. with probability $\alpha_1/\Lambda p_1^{(1)}(i+1)p_1^{(2)}(j)$, when a class-1 orbit customer occupies the server;
4. and finally, with probability $\alpha_1/\Lambda p_2^{(1)}(i)p_2^{(2)}(j+1)$, when a class-2 orbit customer occupies the server.

315 We remind that the customers of different classes arrive independently.

Collecting together all possible cases, we obtain

$$\begin{aligned} P(i, j) &= \frac{1}{\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2} \left[\lambda_1 p_1^{(1)}(i) p_1^{(2)}(j) + \lambda_2 p_2^{(1)}(i) p_2^{(2)}(j) \right. \\ &\quad \left. + \alpha_1 p_1^{(1)}(i+1) p_1^{(2)}(j) + \alpha_2 p_2^{(1)}(i) p_2^{(2)}(j+1) \right], \end{aligned} \quad (62)$$

where $p_1^{(k)}(-1) = 0$, $p_2^{(k)}(-1) = 0$, $k = 1, 2$, by definition. Moreover, we have

$$\begin{aligned} P(i, -1) &= \frac{\alpha_2}{\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2} p_2^{(1)}(i) p_2^{(2)}(0), \quad i \geq 0; \\ P(-1, j) &= \frac{\alpha_1}{\lambda_1 + \lambda_2 + \alpha_1 + \alpha_2} p_2^{(1)}(0) p_2^{(2)}(j), \quad j \geq 0. \end{aligned}$$

Now, for arbitrary $\mathbf{y} \in \mathbb{Z}_+^2$, we define the set

$$C_{\mathbf{y}} = \{\mathbf{z} \in \mathbb{Z}_+^2 : \mathbf{z} \geq \mathbf{y}\}.$$

Then, following [25], we must establish the following monotonicity property of MC $\{\mathbf{X}_n\}$:

$$P(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \mathbf{x}) \geq P(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \hat{\mathbf{x}}), \quad \mathbf{x} \geq \hat{\mathbf{x}}, \quad \mathbf{y} \in \mathbb{Z}_+^2. \quad (63)$$

Next, fix arbitrary \mathbf{x} , $\hat{\mathbf{x}} > (0, 0)$ and calculate the left hand side of (63):

$$\begin{aligned} P(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \mathbf{x}) &= \sum_{i \geq y_1 - x_1} \sum_{j \geq y_2 - x_2} P(X_1^{(1)} = x_1 + i, X_2^{(1)} = x_2 + j | \mathbf{X}_0 = \mathbf{x}) \\ &= \sum_{i \geq y_1 - x_1} \sum_{j \geq y_2 - x_2} P(i, j). \end{aligned} \quad (64)$$

Taking into account that $x_1 \geq \hat{x}_1$, $x_2 \geq \hat{x}_2$ and (62), the expression (64) can be

320 represented as

$$\begin{aligned}
\mathbb{P}(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \mathbf{x}) &= \sum_{i \geq y_1 - x_1} \left(\sum_{j=y_2-x_2}^{y_2-\hat{x}_2-1} \mathbb{P}(i, j) + \sum_{j \geq y_2 - \hat{x}_2} \mathbb{P}(i, j) \right) \\
&= \sum_{i \geq y_1 - x_1} \sum_{j=y_2-x_2}^{y_2-\hat{x}_2-1} \mathbb{P}(i, j) + \sum_{i \geq y_1 - x_1} \sum_{j \geq y_2 - \hat{x}_2} \mathbb{P}(i, j) \\
&= \sum_{i \geq y_1 - x_1} \sum_{j=y_2-x_2}^{y_2-\hat{x}_2-1} \mathbb{P}(i, j) + \sum_{j \geq y_2 - \hat{x}_2} \sum_{i=y_1-x_1}^{y_1-\hat{x}_1-1} \mathbb{P}(i, j) \\
&+ \mathbb{P}(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \hat{\mathbf{x}}),
\end{aligned}$$

which implies the monotonicity property (63).

As the embedded two-dimensional MC $\{\mathbf{X}_n\}$ satisfies the monotonicity property and the second orbit grows to infinity in probability by Theorem 2, we can apply Theorem B from the Appendix, to state the following result.

325 **Theorem 3.** *Let $\pi = \{\pi_n, n \geq 0\}$ be the stationary distribution of the orbit size in the limiting single-orbit system at the departure instants, and let the assumptions of Theorem 2 hold. Then, $X_n^{(1)} \rightarrow X^{(1)}$ in the total variation norm, where $X^{(1)}$ has the distribution π .*

As a final remark of this section, let us explain why the probability $\mathbb{P}_B^{(2)}$ satisfies expression (57). When the auxiliary single-orbit system is stable, $\mathbb{P}_B^{(2)}$ first must include fraction ρ_1 of time when the server is occupied by class-1 customers. Next, the other fraction of time, $1 - \rho_1$, is devoted to serving class-2 customers. When the server is working as the auxiliary system with input rate $\lambda_2 + \alpha_2$ and service rate μ_2 , the loss probability equals

$$\frac{(\lambda_2 + \alpha_2)/\mu_2}{1 + (\lambda_2 + \alpha_2)/\mu_2} = \frac{\rho_2 + \hat{\rho}_2}{1 + \rho_2 + \hat{\rho}_2}.$$

Now collecting together both these fractions, we easily obtain that

$$\rho_1 + (1 - \rho_1) \frac{\rho_2 + \hat{\rho}_2}{1 + \rho_2 + \hat{\rho}_2} = \frac{\rho + \hat{\rho}_2}{\rho_2 + \hat{\rho}_2 + 1} = \mathbb{P}_B^{(2)},$$

as intuitively expected. Note that in the analysis above we implicitly used the
330 PASTA property allowing in our case to equate fraction of class-2 arrivals which

meet server busy by other class-2 customers and the fraction of time when server is occupied by class-2 customers.

5. Comparison with known stability results

In this section, we compare the obtained stability criterion (20) with earlier obtained stability conditions mentioned in the introduction. In the papers [15, 20] the following *necessary* stability condition

$$\rho := \sum_{k=1}^K \rho_k < \min_{1 \leq k \leq K} \frac{\alpha_k}{\lambda_k + \alpha_k},$$

has been obtained for a bufferless K -class retrial system, in which class- k customers follow Poisson input with rate λ_k , have i.i.d. general service times with the mean μ_k and retrial rate α_k . As we see, for $K = 2$ classes, this necessary condition coincides with stability criterion (20). On the other hand, the *sufficient* stability condition from [15] has the form

$$\rho < \min_k \left(\frac{\alpha_k}{\alpha_k + \lambda} \right),$$

where $\lambda = \sum_{k=1}^K \lambda_k$, and definitely less tight than condition (20) (for $K = 2$).

335 We note that for a single-class system, condition (20) becomes (in an evident notation)

$$\rho < \frac{\alpha}{\lambda + \alpha}, \tag{65}$$

and coincides with stability condition obtained in a few previous papers [16, 17, 19, 20]. Note that in [17] a renewal input is allowed while service time is exponential. On the other hand, the work [19] allows both general service time and a general renewal input. We note that condition (65), written as

$$\lambda \rho < \alpha(1 - \rho),$$

has a very clear intuitive interpretation: input rate to the orbit (generated by customers meeting the busy server) must be less than the output rate from the orbit (the rate of successful attempts). Of course a similar interpretation

340 holds for stability conditions $\rho < \alpha_k/(\lambda_k + \alpha_k)$, for each orbit in the multi-class system. One more interesting interpretation of condition (65) is the following: when the input is Poisson, by PASTA property, ρ is the probability that an arriving customer meets the busy server and joins the orbit, while the r.h.s of (65) equals

$$\frac{\alpha}{\lambda + \alpha} = \text{P}(\xi < \tau),$$

345 that is the probability that the retrial time ξ is less than the (remaining) inter-arrival time τ and thus the *orbital customer* occupies the server. As a result, the orbit size decreases, and this negative drift implies stability of the system. We also note that other related stability results can be found in the reference lists of [22, 19].

350 Finally, we would like to mention a series of recent works devoted to regenerative stability analysis of the multiclass retrial systems with *coupled orbits* (or *state-dependent retrial rates*), being a far-reaching generalization of the constant retrial rate systems, in which the retrial rate of each orbit depends on the binary state (busy or idle) of all other orbits, see [28, 29, 30, 31, 32, 33]. In particular, the analysis there is based on PASTA and a coupling procedure connecting the real processes of the retrials with the independent Poisson processes corresponding to various ‘configurations’ of the (binary) states of the orbits.

6. Simulations

6.1. Exponential service times

First we consider the case of exponential service times with corresponding service rates μ_1, μ_2 and consider a particular case when $\lambda_1 = 2, \lambda_2 = 0.5, \mu_1 = 4, \mu_2 = 2$. Thus, in this case we have

$$\rho_1 = 0.5, \quad \rho_2 = 0.25, \quad \rho = 0.75.$$

360 All the experiments were based on 100 000 arrivals.

The regions in Figure 2 illustrate various cases of Theorems 1 and A, corresponding to the values of the orbit rates presented in Table 1.

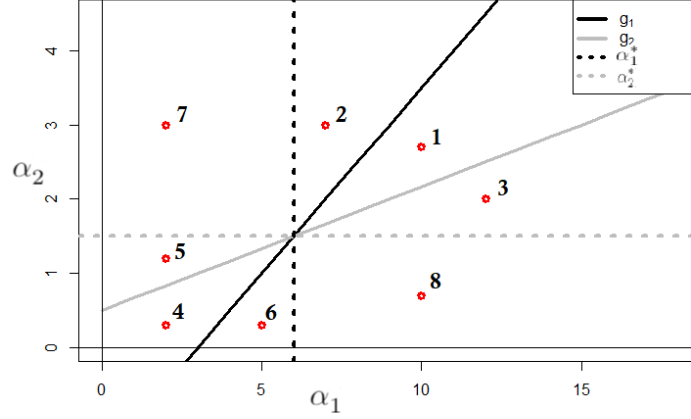


Figure 2: Stability regions for $\rho = 0.75$, exponential service times

Point	1.	2.	3.	4.	5.	6.	7.	8.
α_1	10.0	7.0	12.0	2.0	2.0	5.0	2.0	10.0
α_2	2.7	3.0	2.0	0.3	1.2	0.3	3.0	0.7
Case	(a.1)	(b.1)	(c.1)	(d.1)	(b.2)	(c.2)	(b.2)	(c.2)
	stable	stable	stable	trans.	trans.	trans.	trans.	trans.

Table 1: Retrial rate points in the system with exponential service times, $\rho = 0.75$.

Note that (c.2)–transience region (see Theorem A in the appendix) is further divided into two cases: Case 6 when

$$\alpha_1 \leq \alpha_1^* = \frac{\lambda_1 \rho}{1 - \rho},$$

or equivalently

$$\rho \geq \frac{\alpha_1}{\lambda_1 + \alpha_1},$$

and case 8 when

$$\rho < \frac{\alpha_1}{\lambda_1 + \alpha_1}.$$

Both these examples correspond to a partially stable system.

Thus we obtain a *very interesting result*: in Case 6 both stability conditions
 365 $\rho < \alpha_k/(\lambda_k + \alpha_k)$, $k = 1, 2$, are violated, while the 1-st class orbit remains tight.

The same phenomenon arises in (b.2)–transience region, which is divided by
 the horizontal line (see (45))

$$\alpha_2 = \alpha_2^* = \frac{\lambda_2 \rho}{1 - \rho},$$

into two sub-regions as well. (Also see Figure 2 and equation (59).)

These results have the following intuitive explanation. For instance, in zone
 6 customers from (more 'aggressive') orbit 1 overtake class-2 customers resulting
 in a 'disappearance' of a part of the work generated by class-2 customers, see
 370 the discussion around formulae (58) and (59). As a result, the freed server
 capacity turns out to be enough to process all incoming class-1 customers and
 keeping orbit 1 tight (even if $\rho \geq \alpha_1/(\lambda_1 + \alpha_1)$, i.e., under violation of the
 class-based drift condition for orbit 1). Symmetrically, the stability (tightness)
 of class-2 orbit in zone 5 is provided by pushing out class-1 customers by class-2
 375 customers.

Next, we consider a scenario with non-zero initial conditions for both orbits:

$$X^{(1)}(t_1^-) = X^{(2)}(t_1^-) = 1000,$$

and explore the orbit behavior, setting values (α_1, α_2) from Table 1. The sim-
 ulation results are presented as a 2D plot in Figure 3. Note that cases 2, 5 and
 7 are symmetric to 3, 6 and 8, respectively.

Figure 3.(a) depicts the empirical means obtained by averaging over 100
 380 independent trajectories, while Figure 3.(b) demonstrates the corresponding
 results based on only one realisation. We observe that the orbit dynamics in
 both figures are in agreement.

Both Cases 6 and 8 correspond to the phenomenon of partial stability (only
 the second orbit grows, while the first unloads). Note that the bold black
 385 line, corresponding to Case 6, majorizes the bold grey line, which describes the
 configuration 8. Such results are explained by the fact that unlike case 8, in
 case 6 the condition $\rho < \alpha_1/(\lambda_1 + \alpha_1)$ is violated.

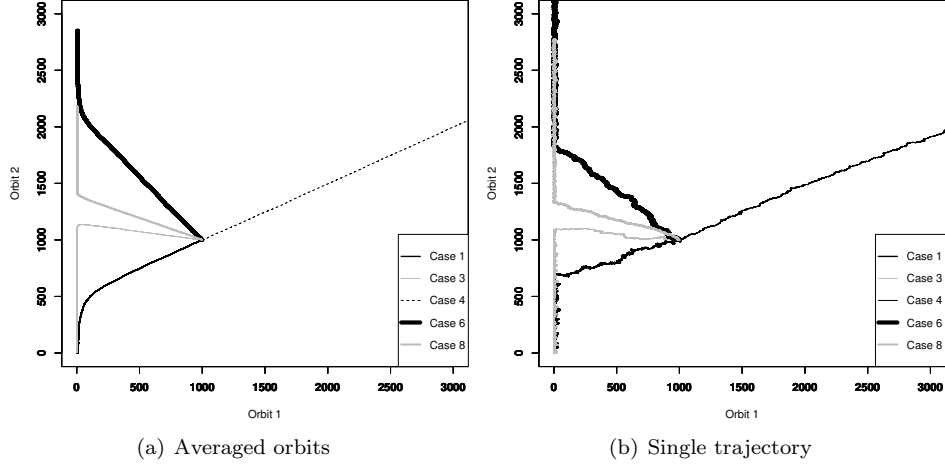


Figure 3: Orbits' dynamics, non-zero initial conditions and exponential service times

6.2. Pareto service times

Now we consider Pareto class-dependent distributions of service times with shape parameter $x_k > 0$ and degree parameter $\beta_k > 0$, $k = 1, 2$. Namely,

$$P(S^{(k)} \leq x) = 1 - (x_k/x)^{\beta_k}, \quad x > x_k,$$

and, consequently,

$$ES^{(k)} = \frac{x_k \beta_k}{\beta_k - 1}.$$

Then, we set again $\lambda_1 = 2$, $\lambda_2 = 0.5$ and set $x_1 = 0.125$, $x_2 = 0.4$, $\beta_1 = 2$, $\beta_2 = 5$, which yields $ES^{(1)} = 0.25$, $ES^{(2)} = 0.5$. Thus, similarly to the case of exponential service times, we obtain

$$\rho_1 = 0.5, \quad \rho_2 = 0.25, \quad \rho = 0.75.$$

Next, we consider the values of retrial rates (α_1, α_2) , corresponding to all cases presented in Table 1.

Simulations for the model with Pareto service time distributions and non-zero initial conditions: $X^{(1)}(t_1^-) = X^{(2)}(t_1^-) = 1000$ are presented in Figure 4

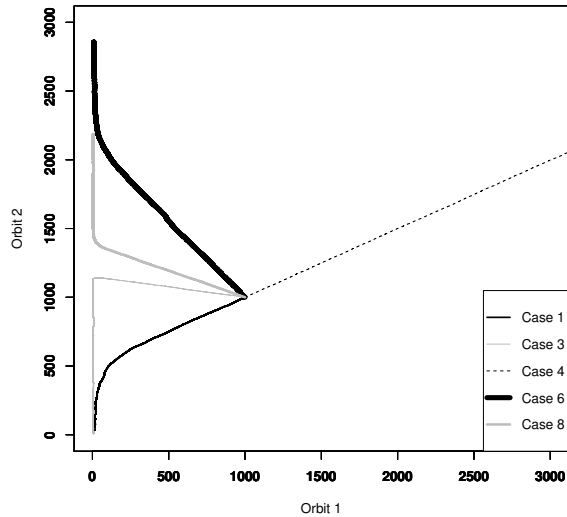


Figure 4: Averaged orbits for non-zero initial conditions and Pareto service times

and follow closely the results for exponential service times. It is worth mention-
 ing that simulation results illustrate the same phenomenon of partial stability in
 395 cases 6 and 8 as we detected for exponential service times although the Pareto
 distribution *does not belong to class \mathcal{D}* . It shows that the latter requirement is
 rather technical and the statement of Theorem 2 should hold for a wider class
 of distribution functions.

7. Conclusion and future research

400 We have established the ergodicity criterion for two-class single-server retrial
 systems with constant retrial rates. The proof of the criterion is based on a
 combination of the regenerative approach for continuous-time systems [26, 23]
 and the ergodicity criterion for discrete-time random walks in a two-dimensional
 quadrant [24]. Essentially, the latter result is only applicable to the case of two
 405 classes. Therefore, a very interesting and at the same time quite challenging
 open research direction is to extend the obtained ergodicity criterion to the case

of more than two classes. We would like to note that the adaptation of [24] to our system has not been straightforward and required a non-trivial algebraic transformations.

410 In addition to the ergodicity criterion, we have also investigated deeper a novel phenomenon of partial (or local) stability which was detected in [22] for the first time. Namely, under certain conditions, one class can be stable (tight), while the orbit size of the other class can grow to infinity in probability. This also implies asymptotic equivalence of the original two-class retrial system and
415 the limiting single-class retrial system with an additional exogenous Poisson input. The convergence has been rigorously justified and some interesting insights are drawn. One surprising insight is that in some situations class-based drift conditions are violated and at the time the system remains partially stable. We have provided an explanation for this interesting phenomenon. In fact, this
420 phenomenon indicates that the classification of transience cases in [24] can be further refined, which is another very interesting future research direction.

Appendix A

In this appendix we present known results which are used in the main part of the paper.

425 First we consider two-dimensional MC $\{\mathbf{X}_n\}$ and mention the conditions for applicability of a basic theorem from [24].

Condition A. (*lower boundedness condition*)

$$\begin{cases} \mathbb{P}\left(X_{n+1}^{(1)} = i, X_{n+1}^{(2)} = j \mid X_n^{(1)} > 0, X_n^{(2)} > 0\right) = 0, & \text{if } i < -1, \quad \text{or } j < -1, \\ \mathbb{P}\left(X_{n+1}^{(1)} = i, X_{n+1}^{(2)} = j \mid X_n^{(1)} > 0, X_n^{(2)} = 0\right) = 0, & \text{if } i < -1, \quad \text{or } j < 0, \\ \mathbb{P}\left(X_{n+1}^{(1)} = i, X_{n+1}^{(2)} = j \mid X_n^{(1)} = 0, X_n^{(2)} > 0\right) = 0, & \text{if } i < 0, \quad \text{or } j < -1. \end{cases}$$

Note that the definitions of the transition probabilities in Section 2.1 automatically imply the validity of Condition A for our embedded MC.

Recall that $\Delta_{n+1} = \mathbf{X}_{n+1} - \mathbf{X}_n$.

Condition B. (*first moment condition*)

$$\mathbb{E}\left[\|\Delta_{n+1}\| \mid \mathbf{X}_n = (k, l)\right] \leq C < \infty, \quad \forall (k, l) \in \mathbb{Z}_+^2, \quad (66)$$

430 where $\|\cdot\|$ denotes the Euclidean norm and $C > 0$ is a constant.

Since

$$\begin{aligned} \mathbb{E}\left[\|\Delta_{n+1}\|\right] &\leq \mathbb{E}[\Delta_{n+1}^{(1)} + \Delta_{n+1}^{(2)}] \\ &\leq 2 \max(\lambda_1, \lambda_2, \alpha_1, \alpha_2) \max(\mathbb{E}S^{(1)}, \mathbb{E}S^{(2)}) < \infty, \end{aligned}$$

the condition (66) holds for all sets of (finite) parameters for our MC.

Now recall that $M_k^{01}, M_k^{10}, M_k^{11}$ denote the mean increments (drifts) of the k -th component of the MC $\{\mathbf{X}_n\}$ between the n -th and $(n+1)$ -st departure instants, given the respective conditions $(X_n^{(1)} = 0, X_n^{(2)} > 0)$, $(X_n^{(1)} > 0, X_n^{(2)} = 0)$ or $(X_n^{(1)} > 0, X_n^{(2)} > 0)$.

Theorem A. (See Theorem 3.3.1 in [24]) Let a MC $\{\mathbf{X}_n\}$ be aperiodic and irreducible and let Conditions A and B hold.

(a) If $M_1^{11} < 0, M_2^{11} < 0$, then MC $\{\mathbf{X}_n\}$ is

1. ergodic (positive recurrent) if

$$\begin{cases} M_1^{11}M_2^{10} - M_2^{11}M_1^{10} < 0, \\ M_2^{11}M_1^{01} - M_1^{11}M_2^{01} < 0, \end{cases} \quad (67)$$

2. non-ergodic if either

$$M_1^{11}M_2^{10} - M_2^{11}M_1^{10} \geq 0 \quad \text{or} \quad M_2^{11}M_1^{01} - M_1^{11}M_2^{01} \geq 0. \quad (68)$$

440 (b) If $M_1^{11} \geq 0$, $M_2^{11} < 0$, then MC $\{\mathbf{X}_n\}$ is

1. ergodic if

$$M_1^{11}M_2^{10} - M_2^{11}M_1^{10} < 0, \quad (69)$$

2. transient if

$$M_1^{11}M_2^{10} - M_2^{11}M_1^{10} > 0. \quad (70)$$

(c) If $M_1^{11} < 0$, $M_2^{11} \geq 0$, then MC $\{\mathbf{X}_n\}$ is

1. ergodic if

$$M_2^{11}M_1^{01} - M_1^{11}M_2^{01} < 0, \quad (71)$$

2. transient if

$$M_2^{11}M_1^{01} - M_1^{11}M_2^{01} > 0. \quad (72)$$

(d) If $M_1^{11} \geq 0$, $M_2^{11} \geq 0$, $M_1^{11} + M_2^{11} > 0$, then MC $\{\mathbf{X}_n\}$ is transient.

Next we present partial stability results from [25] for a two-dimensional MC.

Theorem B. (See Proposition 2 in [25]) *For a MC \mathbf{X}_n , with a state space*

445 \mathbb{Z}_+^2 , *assume the following:*

1. $X_n^{(2)} \Rightarrow \infty$ in probability, as $n \rightarrow \infty$, given $X_0^{(2)} = 0$, $X_0^{(1)} = 0$;
2. $\mathbb{P}(X_{n+1}^{(1)} = j | X_n^{(1)} = i, X_n^{(2)} = l) = p_{ij}$ for all values of $l > 0$, where p_{ij} are transition probabilities of an ergodic MC with the unique stationary distribution $\pi = \{\pi_j\}$;
3. MC \mathbf{X}_n is monotone. Namely,

$$\mathbb{P}(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \mathbf{x}) \geq \mathbb{P}(\mathbf{X}_1 \in C_{\mathbf{y}} | \mathbf{X}_0 = \hat{\mathbf{x}}), \quad \mathbf{x} \geq \hat{\mathbf{x}}, \quad \mathbf{y} \in \mathbb{Z}_+^2.$$

Then, for all initial states i_0, j_0 :

$$\sup_j |\mathbf{P}(X_n^{(1)} = j | X_0^{(1)} = i_0, X_0^{(2)} = j_0) - \pi_j| \rightarrow 0, \quad n \rightarrow \infty, \quad (73)$$

450 *i.e.*, $X_n^{(1)}$ converges in distribution to π in the total variation norm.

Appendix B

Below we present some technical details of the proof of Theorem 1. Note that our goal is to apply the results of Theorem A to the system under consideration. First, consider case (a) from Theorem A and recall expressions (16)–(19) 455 for $M_1^{10}, M_2^{10}, M_1^{11}, M_2^{11}$, respectively. Conditions $M_1^{11} < 0, M_2^{11} < 0$ can be relatively easy rewritten as (21) and (22) in terms of the load coefficients. Next let us also represent the system (67) in terms of the load coefficients. Define the following auxiliary parameters

$$\begin{aligned} A_1 &= (\lambda_1 + \alpha_1)\lambda_1\mu_1 + \lambda_2\lambda_1/\mu_2, \\ A_2 &= (\lambda_1 + \alpha_1)\lambda_2/\mu_1 + \lambda_2\lambda_2/\mu_2, \\ A_3 &= \lambda_1 + \lambda_2 + \alpha_1, \\ A_4 &= \lambda_1 + \lambda_2 + \alpha_1 + \alpha_2. \end{aligned}$$

Thus, from (16)–(19) we obtain

$$\begin{aligned} M_1^{10} &= \frac{A_1}{A_3} - \frac{\alpha_1}{A_3}, \\ M_2^{10} &= \frac{A_2}{A_3}, \\ M_1^{11} &= \frac{A_1}{A_4} + \frac{\alpha_2\lambda_1/\mu_2 - \alpha_1}{A_4}, \\ M_2^{11} &= \frac{A_2}{A_4} + \frac{\alpha_2\lambda_2/\mu_2 - \alpha_2}{A_4}. \end{aligned}$$

460 Next, denote

$$\begin{aligned} B_1 &= \frac{\alpha_2\lambda_1/\mu_2 - \alpha_1}{A_4}, \\ B_2 &= \frac{\alpha_2\lambda_2/\mu_2 - \alpha_2}{A_4}, \\ B_3 &= \frac{\alpha_1}{A_3}. \end{aligned}$$

Thus, the condition $M_1^{11}M_2^{10} - M_2^{11}M_1^{10} < 0$ transforms to

$$\left(\frac{A_1}{A_4} + B_1\right)\frac{A_2}{A_3} - \left(\frac{A_2}{A_4} + B_2\right)\left(\frac{A_1}{A_3} - B_3\right) < 0.$$

After opening brackets, we obtain

$$B_1\frac{A_2}{A_3} + B_3\frac{A_2}{A_4} - B_2\frac{A_1}{A_3} + B_2B_3 < 0.$$

Then we substitute back the expressions for B_1, B_2, B_3 to get

$$\frac{(\alpha_2\lambda_1/\mu_2 - \alpha_1)A_2}{A_3A_4} + \frac{\alpha_1A_2}{A_3A_4} - \frac{(\alpha_2\lambda_2/\mu_2 - \alpha_2)A_1}{A_3A_4} + \frac{(\alpha_2\lambda_2/\mu_2 - \alpha_2)\alpha_1}{A_3A_4} < 0. \quad (74)$$

Multiplying both sides of (74) by $A_3A_4 > 0$, yields

$$A_2\alpha_2\lambda_1/\mu_2 - A_1\alpha_2(\lambda_2/\mu_2 - 1) + \alpha_1\alpha_2(\lambda_2/\mu_2 - 1) < 0.$$

Then, after substituting the expressions for A_1 and A_2 , we obtain the following inequality in terms of the original parameters

$$(\lambda_1 + \alpha_1)\lambda_1/\mu_1 + \lambda_2\lambda_1/\mu_2 + \alpha_1(\lambda_2/\mu_2 - 1) < 0,$$

which easily transforms to

$$(\lambda_1 + \alpha_1)(\lambda_1/\mu_1 + \lambda_2/\mu_2) < \alpha_1.$$

Thus, we obtain the condition

$$\rho < \frac{\alpha_1}{\lambda_1 + \alpha_1}.$$

After similar derivations, the condition $M_2^{11}M_1^{01} - M_1^{11}M_2^{01} < 0$ is transformed to

$$\rho < \frac{\alpha_2}{\lambda_2 + \alpha_2}.$$

Thus, the system (67) is equivalent to $\rho < \alpha_k/(\lambda_k + \alpha_k)$, $k = 1, 2$.

References

- [1] G. Falin, J. G. Templeton, Retrial queues, Vol. 75, CRC Press, 1997.

- 465 [2] G. Falin, A survey of retrial queues, *Queueing systems* 7 (2) (1990) 127–167.
- [3] J. R. Artalejo, Accessible bibliography on retrial queues, *Mathematical and computer modelling* 30 (3-4) (1999) 1–6.
- [4] J. R. Artalejo, Accessible bibliography on retrial queues: progress in 2000–2009, *Mathematical and computer modelling* 51 (9-10) (2010) 1071–1081.
- 470 [5] J. Artalejo, A. Gómez-Corral, *Retrial Queueing Systems: A Computational Approach*, Springer, 2008.
- [6] K. Avrachenkov, U. Yechiali, Retrial networks with finite buffers and their application to internet data traffic, *Probability in the Engineering and Informational Sciences* 22 (4) (2008) 519–536.
- 475 [7] K. Avrachenkov, P. Nain, U. Yechiali, A retrial system with two input streams and two orbit queues, *Queueing Systems* 77 (1) (2014) 1–31.
- [8] I. Dimitriou, A two-class retrial system with coupled orbit queues, *Probability in the Engineering and Informational Sciences* 31 (2) (2017) 139–179.
- [9] I. Dimitriou, A queueing system for modeling cooperative wireless networks with coupled relay nodes and synchronized packet arrivals, *Performance Evaluation* 114 (2017) 16–31.
- 480 [10] I. Dimitriou, Modeling and analysis of a relay-assisted cooperative cognitive network, *Proceedings Analytical and Stochastic Modelling Techniques and Applications (ASMTA)* (2017) 47–62.
- 485 [11] I. Dimitriou, On the power series approximations of a structured batch arrival two-class retrial system with weighted fair orbit queues, *Performance Evaluation* 132 (2019) 38–56.
- [12] I. Dimitriou, T. Phung-Duc, Analysis of cognitive radio networks with cooperative communication, in: *Proceedings of the 13th EAI International Conference on Performance Evaluation Methodologies and Tools (Value-Tools)*, 2020, pp. 192–195.
- 490

- [13] E. Morozov, T. Phung-Duc, Regenerative analysis of two-way communication orbit-queue with general service time, *Proceedings International Conference Queueing Theory and Network Applications 10932* (2018) 22–32.
- 495 [14] T. Phung-Duc, W. Rogiest, Y. Takahashi, H. Bruneel, Retrial queues with balanced call blending: analysis of single-server and multiserver model, *Annals of Operations Research* 239 (2) (2016) 429–449.
- [15] E. Morozov, A. Rumyantsev, S. Dey, T. Deepak, Performance analysis and stability of multiclass orbit queue with constant retrial rates and balking,
500 *Performance Evaluation* 134 (2019) 102005.
- [16] G. Fayolle, A simple telephone exchange with delayed feedbacks, in: *Proc. of the international seminar on Teletraffic analysis and computer performance evaluation*, 1986, pp. 245–253.
- [17] R. Lillo, A G/M/1-queue with exponential retrial, *Top* 4 (1996) 99–120.
- 505 [18] K. Avrachenkov, U. Yechiali, On tandem blocking queues with a common retrial queue, *Computers & Operations Research* 37 (7) (2010) 1174–1180.
- [19] K. Avrachenkov, E. Morozov, Stability analysis of GI/GI/c/K retrial queue with constant retrial rate, *Mathematical Methods of Operations Research* 79 (3) (2014) 273–291.
- 510 [20] K. Avrachenkov, E. Morozov, R. Nekrasova, B. Steyaert, Stability analysis and simulation of N-class retrial system with constant retrial rates and Poisson inputs, *Asia-Pacific Journal of Operational Research* 31 (02) (2014) 1440002.
- [21] I. Dimitriou, A two-class queueing system with constant retrial policy and
515 general class dependent service times, *European Journal of Operational Research* 270 (3) (2018) 1063–1073.
- [22] K. Avrachenkov, E. Morozov, B. Steyaert, Sufficient stability conditions for multi-class constant retrial rate systems, *Queueing Systems* 82 (1-2) (2016) 149–171.

- 520 [23] E. Morozov, B. Steyaert, *Stability Analysis of Regenerative Queueing Models: Mathematical Methods and Applications*, Springer, 2021.
- [24] G. Fayolle, V. A. Malyshev, M. V. Menshikov, *Topics in the Constructive Theory of Countable Markov Chains*. 1st edn., Cambridge University Press, 1995.
- 525 [25] I. Adan, S. Foss, S. Shneer, G. Weiss, Local stability in a transient Markov chain, *Statistics and Probability Letters* 165 (108855) (2020) 1–6.
- [26] S. Asmussen, *Applied Probability and Queues*. 2nd edn., Springer, New York, 2003.
- [27] E. Morozov, R. Delgado, Stability analysis of regenerative queues, *Automation and remote control* (2009) 1977–1991.
- 530 [28] E. Morozov, I. Dimitriou, Stability analysis of a multiclass retrial system with coupled orbit queues, *Proceedings of 14th European Workshop, EPEW 10497* (2017) 85–98.
- [29] E. Morozov, T. Morozova, Analysis of a generalized retrial system with coupled orbits, *Proceeding 23rd Conference of Open Innovations Association (FRUCT)* (2018) 253–260.
- 535 [30] E. Morozov, T. Morozova, I. Dimitriou, A multiclass retrial system with coupled orbits and service interruptions: Verification of stability conditions., *Proceedings of the 24th Conference of Open Innovations Association* 24 (2019) 75–81.
- 540 [31] E. Morozov, T. Morozova, A coupling-based analysis of a multiclass retrial system with state-dependent retrial rates, *Proceedings 14th International Conference on Queueing Theory and Network Applications* 11688 (2019) 34–50.
- 545 [32] E. Morozov, T. Morozova, The remaining busy time in a retrial system with unreliable servers, *Proceedings International Conference on Distributed Computer and Communication Networks* 12563 (2020) 555–566.

- [33] E. Morozov, T. Morozova, I. Dimitriou, Simulation of multiclass retrial system with coupled orbits, Proceedings of the First International Workshop on Stochastic Modeling and Applied Research of Technology Petrozavodsk (2018) 6–16.