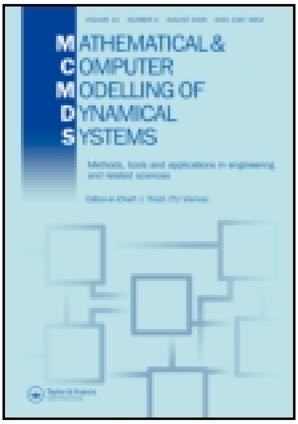

Mathematical and Computer Modelling of Dynamical Systems: Methods, Tools and Applications in Engineering and Related Sciences
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nmcm20

Optimal control with monotonicity constraints for a parallel-server loss channel serving multi-class jobs

Wanyang Dai[a]

[a] Department of Mathematics and State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210093, China
Published online: 25 Aug 2013.

Mathematical and Computer Modelling of Dynamical Systems, 2014
Vol. 20, No. 3, 284–315, http://dx.doi.org/10.1080/13873954.2013.831359

Optimal control with monotonicity constraints for a parallel-server loss channel serving multi-class jobs

Wanyang Dai*

Department of Mathematics and State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210093, China

(*Received 1 October 2012; final version received 31 July 2013*)

We study a parallel-server loss channel serving multi-class jobs, which appears in many real-world systems, e.g., cloud computing, multi-input multi-output (MIMO) orthogonal frequency division multiplexing (OFDM), and call centre. An α-discounted optimal control with monotonicity constraints (OCMC) model over infinite time horizon is established by using the physical queueing model with linear revenue function. Existence of a solution to the OCMC model is proved, whose optimal value provides an upper bound of the corresponding values of the physical queueing model under Markovian decision rules. Algorithms with lower complexity in solving the OCMC model are proposed, which are further used to design an admission control policy for the loss channel. Furthermore, a simulation algorithm is proposed to implement the designed policy. Performance comparisons through numerical examples are conducted among our newly designed policy, the first-in first-out (FIFO) policy, an arbitrarily selected (AS) policy, and the Markov decision process (MDP) based threshold policy. Advantages and disadvantages of these policies are identified under different channel parameters and channel (e.g., Markovian and non-Markovian) conditions. Particularly, we find out that our designed policy outperforms the other three policies when the traffic intensity is relatively large, and the differences of the revenues per unit of time and the penalty costs among different classes of jobs are large.

Keywords: optimal control with monotonicity constraints; parallel-server; loss channel; multi-class jobs; Pontryagin's maximum principle; linear programming; Markov decision process; first-in first out

1. Introduction

Parallel-server systems appear in many real-world systems, e.g., cloud computing, multi-input multi-output (MIMO) orthogonal frequency division multiplexing (OFDM), and call centre (see, e.g., Bodas *et al.* [1], Dai [2], Dai and Feng [3], Gans *et al.* [4], Harrison [5], Ormeci *et al.* [6], Ormeci and Wal [7], Zhang *et al.* [8]). Especially, in the next generation broadband wireless systems (e.g., WiMax [9]), the MIMO OFDM channel is divided into a number of parallel sub-channel (sub-band) consisting of a fixed number of sub-carriers. Each sub-channel is a server that can be assigned to one and at most one user during each time slot.

Existing parallel-server systems can be classified into three categories: the zero buffer-regime such as in Dai [2], Dai and Feng [3], Ormeci *et al.* [6], Ormeci and Wal [7],

*Email: nan5lu8@netra.nju.edu.cn

the small-buffer regime such as in Bodas *et al.* [1], the large-buffer regime such as in Bassamboo *et al.* [10,11] and Harrison [5].

The discussions in [3] and [6,7] are both focused on systems with two-classes of jobs and aim to design optimal threshold policies to conduct Markov decision process (MDP) based admission controls (see, e.g., Altman [12] and Lippman [13]). One of the main differences between the discussions in [3] and [6,7] is the way of receiving rewards. The policy found in [3] is based on random payments depending on service durations while the payments in the studies of [6,7] are deterministic constants. Therefore, the study in [3] is different from the ones in [6,7]. But all the techniques used in [3] and [6,7] are certain structured value function based MDPs, which implies that certain conditions should be imposed on related value functions. Although decision-making using a structured value function based MDP is a popular method in many fields recently, the size of related decision space is limited owing to these imposed conditions. Furthermore, when the parallel-server system is used to serve general number (≥ 3) of job classes, the computation of the corresponding MDP algorithm becomes heavy. Hence, in the conference report of Dai [2], we turn to derive a constrained optimal control (COC) model over infinite time horizon and with linear revenue function for the general parallel-server channel in which we allow the service rates to be class-dependent and suppose that each service rate corresponding to an individual job class is a constant. In a wireless system, this condition is often related to the well-known quasi-static channel, i.e., the channel is assumed to be fixed for all transmissions over the period of interest.

In the current paper, we will establish an optimal control with monotonicity constraints (OCMC) model by imposing equality constraints based on non-smooth monotone control functions to replace the inequality constraints as used in Dai [2]. This newly established OCMC model is more naturally related to the physical queueing system. Furthermore, we will provide more detailed analysis and design resolution algorithms for the OCMC model and the loss channel.

We set up and thoroughly justify an exact connection between the physical queueing related optimal control model and an OCMC model over infinite time horizon when the revenue function is linear. This connection transforms the original stochastic optimal control problem to a deterministic OCMC one whose optimal value provides an upper bound of the corresponding values of physical queueing model under Markovian decision rules. Furthermore, the OCMC model is solvable by developing Pontryagin's maximum principle and linear programming (LP) combined algorithms that are of tolerable computational complexity. In addition, these algorithms are further used to design an admission control policy and to develop a simulation algorithm for the loss channel. Performance comparisons through numerical examples are conducted among our newly designed policy, the first-in first-out (FIFO) policy, an arbitrarily selected (AS) policy, and the MDP based threshold policy. Advantages and disadvantages of these policies are identified under different channel parameters and channel (e.g., Markovian and non-Markovian) conditions. Particularly, we find out that our designed policy outperforms the other three policies when the traffic intensity is relatively large, and the differences of the revenues per unit of time and the penalty costs among different classes of jobs are large.

One way to interpret the phenomenon is as follows. Although our derived OCMC model is approximate for a non-Markovian system, it can still be viewed as an approximating fluid model for both Markovian and non-Markovian parallel-server systems under suitable *time/space re-scaling* owing to the functional strong law of large numbers (see, e.g., the related discussions in Bassamboo *et al.* [10,11], Bäuerle [14], Dai [15], Dai [16], etc.).

The fluid approximation is a very simple first-order approximation of the original queueing system and it has been used to design the optimal polices for various stochastic networks (see, e.g., Bassamboo *et al.* [10,11], Bäuerle [14], and the references therein). However, our derived α-discounted OCMC model is subject to the monotonicity and mixed control-state inequality constraints over infinite time horizon. Hence, it is different from the so-called *separated continuous linear program* (SCLP) (see, e.g., Bäuerle [14], Luo and Bertsimas [17]). Furthermore, our algorithms are iterative ones with respect to time evolving in an online fashion, which employ the LP as a subroutine at each time instant by applying Pontryagin's maximum principle so the algorithms are of lower computational complexity at each time instant. In addition, our derived OCMC model is also different from the so-called *pointwise stationary fluid models* as studied in Bassamboo *et al.* [10,11], where the instantaneous flow-balance conditions are imposed and a stationary way of time/space re-scaling to the objective functionals is used in designing asymptotically optimal policies over a fixed and finite planning horizon. Nevertheless, our OCMC model can capture the non-stationary characteristic of the original α-discounted optimal control problem, e.g., the feature of the initial *warm-up* period of the system.

Finally, we provide a proof concerning the existence of solution to our OCMC model, which provides a theoretical basis for our study. Moreover, the proof generalizes the existing discussions such as Fleming and Rishel [18], Hellwig [19], Kugelmann and Pesch [20,21], Portal [22], Tu [23], etc., in several ways since our time horizon is infinite and the model is with the monotonicity and control-state constraints.

The remainder of the paper is organized as follows. In Section 2, we describe our physical system and present our main result to establish the OCMC model when the revenue function is linear. In Section 3, we design an admission control policy through numerical schemes in solving the OCMC model. In Section 4, we propose a simulation algorithm via the admission policy and demonstrate the effectiveness of the OCMC model via numerical examples. In Section 6, we prove the existence of solution to our OCMC model. Finally, in Section 5, we prove our main theorem.

2. The OCMC model

2.1. Physical description and stochastic dynamics of the channel

We study a loss channel consisting of n identical and independent servers for an incoming job being either served immediately or discarded. There are K classes of external arriving jobs to the channel. For each class $i \in \{1, ..., K\}$, the arrival stream follows a Poisson process with rate λ_i. An arrival job for each class $i \in \{1, \ldots, K\}$ to the channel is either rejected from service with penalty cost l_i or accepted into service experiencing exponentially distributed amount of service time with rate μ_i according to some admission control policy that will be addressed later. The arrival processes and service times are assumed to be mutually independent.

Now, let $X_i(t)$ for each $i \in \{1, \ldots, K\}$ be the number of class i jobs being served in the channel at time t, which takes values in the state space $S_i = \{0, 1, \ldots, n\}$, and let $X(t) = (X_1(t), \ldots, X_K(t))$ be the corresponding vector form taking values in the following K-dimensional set:

$$S \equiv \left\{(x_1, \ldots, x_K) : \sum_{i=1}^{K} x_i \leq n, x_i \in S_i, i = 1, \ldots, K\right\}. \tag{1}$$

Furthermore, let $h(X(t))$ be the induced revenue per unit of time and $h(x)$ is assumed to be a linear function with respect to $x \in S$, i.e., $h(x) = \sum_{i=1}^{K} r_i x_i$. The number r_i is the revenue per unit of time that is induced from class i jobs being served at the channel. In addition, there is a fixed cost c per unit of time associated with the channel. In the end, let $N_i(t)$ denote the total number of rejected class i customers at the channel by time t. Then we can write down the value function (expected infinite horizon discounted profit) for given discount factor α, initial state $x \in S$, and admission control policy π as follows:

$$v^{\pi}(x) \equiv E_x^{\pi}\left[\int_0^{\infty} e^{-\alpha t} h(X(t))dt - \int_0^{\infty} e^{-\alpha t} \sum_{i=1}^{K} l_i dN_i(t) - \int_0^{\infty} e^{-\alpha t} c dt\right]. \tag{2}$$

An admission control policy π specifies at each time instant whether an arrived job to the channel is accepted into service or is discarded. Furthermore, we confine our discussions to Markovian policies for which actions taken at a particular decision epoch depend only on the current state of the system (see, e.g., Sennott [24]). The aim of our study is to find out an optimal admission control policy π^* such that

$$v^{\pi^*}(x) = \max_{\pi \in \Pi} v^{\pi}(x), \tag{3}$$

where the set Π is the one of all Markov decision policies.

Next, define $X_i^{\pi^*}(t)$ to be the number of class i jobs being served in the channel at time t under the optimal policy π^*, $N_i^{\pi^*}(t)$ to be the total number of rejected class i jobs at the channel by time t, and $S_i^{\pi^*}(t)$ to be the total number of class i jobs that finish services by time t. Then, the stochastic dynamics of $X_i^{\pi^*}(t)$ can be written as follows:

$$\begin{cases} X_i^{\pi^*}(t) = R_i(t) - S_i^{\pi^*}(t) - N_i^{\pi^*}(t) + X_i^{\pi^*}(0), \\ X_i^{\pi^*}(0) = x_i, \end{cases} \tag{4}$$

where $R_i(t)$ is the total number of arrivals of class i jobs by time t, i.e.,

$$R_i(t) = \max\left\{m : \sum_{j=1}^{m} u_i^j \leq t\right\} \tag{5}$$

and u_i^j for $j \in \{1, 2, \ldots\}$ is the inter-arrival time sequence of class i jobs.

Finally, since an exponentially distributed random variable (r.v.) is memoryless and because both the inter-arrival times of Poisson processes and service times are exponentially distributed, the above channel system is memoryless (or called Markovian). In the sequel, such a system is called *a Markovian model*. Nevertheless, when the inter-arrival times of an input process and/or service times are not exponentially distributed, these r.v.s are not memoryless. Hence, even we still employ a Markovian admission control policy and keep the independence assumption on the arrival processes and service times unchanged, the corresponding system is not memoryless. Consequently, such a system is called *a non-Markovian model* in the following discussion.

2.2. OCMC model formulation

First, we present the value function defined in Equation (2) for an initial state $x \in S$ with policy π^* as follows:

$$v^{\pi^*}(x) \equiv E^{\pi^*}\left[\int_0^\infty e^{-\alpha t}\sum_{i=1}^K r_i X_i^{\pi^*}(t)dt - \int_0^\infty e^{-\alpha t}\sum_{i=1}^K l_i dN_i^{\pi^*}(t) - \frac{c}{\alpha}\right]. \tag{6}$$

Furthermore, let $\mathcal{U}_K[0,\infty)$ be the space of all K-vector functions denoted by $m(t) = (m_1(t), \ldots, m_K(t))$ that are continuous in t, uniformly bounded by n, and satisfy

$$|m_i(t+\Delta t) - m_i(t)| \leq C\Delta t, \quad i \in \{1, \ldots, K\} \tag{7}$$

for certain constant $C \geq 0$. Then, we have the following theorem.

Theorem 2.1 *For each $t \in [0,\infty)$, let $m_i^{\pi^*}(t) = E[X_i^{\pi^*}(t)]$ be the expectation of $X_i^{\pi^*}(t)$ (or call $m_i^{\pi^*}(\cdot)$ the mean value function of $X_i^{\pi^*}(\cdot)$). Then, under the previous conditions, $m_i^{\pi^*} \in \mathcal{U}_K[0,\infty)$ and the value function in (6) has the following equivalent expression for each $x \in S$:*

$$v^{\pi^*}(x) = \sum_{i=1}^K \int_0^\infty e^{-\alpha t}\left[(r_i + \alpha l_i)m_i^{\pi^*}(t) + \alpha l_i \mu_i \int_0^t m_i^{\pi^*}(s)ds\right]dt \tag{8}$$
$$-\left(\sum_{i=1}^K \left(x_i l_i + \frac{\lambda_i l_i}{\alpha}\right) + \frac{c}{\alpha}\right).$$

Proof. The lengthy proof of Theorem 2.1 is presented in Section 5. □

As pointed in Section 1, the relationship in (8) may be not exactly true for a non-Markovian model defined at the end of Section 2.1. However, it can still be considered as a fluid approximating model to the non-Markovian system under suitable time/space re-scaling. The approximation can be justified by using the functional strong law of large numbers (see, e.g., the related discussions in Bassamboo *et al.* [10,11], Bäuerle [14], Dai [15], and Dai [16], nevertheless, how to exactly justify such an approximation theorem is not a concern of the current paper). Hence, based on Theorem 2.1, we establish our OCMC model as follows. First, define

$$J(m) \equiv \int_0^\infty f(y(t), m(t), t)dt, \tag{9}$$

$$f(y(t), m(t), t) \equiv \sum_{i=1}^K e^{-\alpha t}\left[(r_i + \alpha l_i)m_i(t) + \alpha l_i \mu_i y_i(t)\right], \tag{10}$$

$$y_i(t) = \int_0^t m_i(s)ds. \tag{11}$$

Second, let $\mathcal{N}_K[0,\infty)$ be the set of all the nonnegative non-decreasing continuous functions that satisfy (7). Third, by considering $m = (m_1, \ldots, m_K) \in \mathcal{U}_K[0,\infty)$, $z = (z_1, \ldots, z_K) \in$

$\mathcal{N}_K[0,\infty)$ as control functions, and $y = (y_1, \ldots, y_K)$ as state function, we can formulate the following deterministic linear optimal control problem by employing Theorem 2.1:

$$\sup_{(m,z)\in\mathcal{U}_K[0,\infty)\times\mathcal{N}_K[0,\infty)} J(m) \tag{12}$$

subject to the constraints for each $i \in \{1, \ldots, K\}$,

$$\dot{y}_i(t) = m_i(t),\ y_i(0) = 0,\ m_i(0) = x_i, \tag{13}$$

$$\begin{cases} g_i(y,z,m,t) & \equiv m_i(t) \geq 0, \\ g_{K+i}(y,z,m,t) & \equiv z_i(t) \geq 0, \\ g_{2K+i}(y,z,m,t) & \equiv (x_i + \lambda_i t) - (m_i(t) + \mu_i y_i(t) + z_i(t)) = 0, \\ g_0(y,z,m,t) & \equiv n - \sum_{i=1}^{K} m_i(t) \geq 0, \end{cases} \tag{14}$$

where the third and fourth constraints in (14) are obtained by taking expectations in both sides of (4) (e.g., $z_i(t) = E[N_i(t)]$), and the fact that the channel has n servers with no buffering capacity.

Furthermore, we note that the feasible set determined by the constraint in (14) is nonempty at any time $t \in [0,\infty)$. Hence, we can talk about the existence of solution to the OCMC problem in (12)–(14), which is addressed in Section 2.3. In addition, the established OCMC model in (12)–(14) requires that some control variables (e.g., $z_i(\cdot)$ with $i \in \{1, \ldots, K\}$) should satisfy monotonicity constraints, which attracts some recent research interests in both control theory and applications (see, e.g., Hellwig [19] and Ruiz del Portal [22]).

2.3. *Existence of optimal solution and upper bound on Markovian rules*

First, consider a time interval $[T_1, T_2]$ with $T_1 < T_2$ and $T_1, T_2 \in [0,\infty]$ (denote it as $[T_1, T_2]$ if $T_2 = \infty$). For $k \in \{0, 1, \ldots\}$ and $K \in \{1, 2, \ldots\}$, let $C_K^{(k)}[T_1, T_2]$ be the space of all K-vector functions defined on $[T_1, T_2]$ and having continuous derivatives up to the kth-order, which is endowed with the unified norm for both $T_2 < \infty$ and $T_2 = \infty$,

$$\begin{aligned} \|y\|_C \equiv & \sum_{n=1}^{N} 2^{-n} \left(1 \bigwedge \sup_{T_1 \leq t \leq n} \sum_{j=0}^{k} \left\| y^{(j)}(t) \right\| \right) \\ & + 2^{-(N+1)} \left(1 \bigwedge \sup_{T_1 \leq t < T_2 \wedge (N+1)} \sum_{j=0}^{k} \left\| y^{(j)}(t) \right\| \right), \end{aligned} \tag{15}$$

where $\|\cdot\|$ denotes the Euclidean norm in R^K and $N = \sup\{n : n \leq T_2\}$ with $n \in \{1, 2, \ldots,\}$, $y^{(j)}(t)$ denotes the jth-order derivative of $y(t)$ at time t and $y^{(0)}(t) = y(t)$ (here, we note that all vectors associated in the paper are explained as row-vectors).

Now, take $m = (m_1, \ldots, m_K) \in C_K^{(0)}[0,\infty)$ and $z \in C_K^{(0)}[0,\infty)$ as control functions, $y = (y_1, \ldots, y_K) \in C_K^{(1)}[0,\infty)$ as state function. Then, we can rewrite the OCMC model in (12)–(14) as follows:

$$\sup_{(m,z)\in C_K^{(0)}[0,\infty)\times C_K^{(0)}[0,\infty)} J(m) \tag{16}$$

subject to

$$\text{The constraints in (13)–(14) and } (m, z) \in \mathcal{U}_K[0,\infty) \times \mathcal{N}_K[0,\infty). \tag{17}$$

Then, we have the following theorem.

Theorem 2.2 *There is at least one optimal solution* $(y^*, m^*, z^*) \in C_K^{(1)}[0,\infty) \times C_K^{(0)}[0,\infty) \times C_K^{(0)}[0,\infty)$ *for the OCMC model defined in* (16)–(17). *Furthermore, the optimal value is finite and is an upper bound to the corresponding values of the physical queueing system under Markovian decision rules*.

The proof of Theorem 2.2 is provided in Section 6.

3. An admission policy by numerical schemes to the OCMC model

First, it follows from the definition of $f(y(t), m(t), t)$ in (10) that the following convergence is uniform in all y and m owing to the fact that $m(\cdot)$ is bounded over $[0,\infty)$:

$$\int_t^{\infty} f(y(s), m(s), s)ds \to 0 \quad \text{as} \quad t \to \infty. \tag{18}$$

Thus, we can take a sufficiently large time $T > 0$ to replace the infinity for our purpose. Furthermore, note that any function in $\mathcal{U}_K[0,T]$ can be uniformly approximated by a sequence of piecewise continuous functions bounded by n over $[0,T]$ owing to Ergoroff's Theorem (see, e.g., p. 73 of Royden [25]), and $z^{\pi^*}(\cdot) = (z_1^{\pi^*}(\cdot), \ldots, z_K^{\pi^*}(\cdot)) \in \mathcal{N}_K[0,\infty)$ can be approximated by a sequence of piecewise linear nondecreasing functions. Then, we can estimate $z(\cdot)$ in the following way:

$$\dot{z}_i(t) = v_i(t),\ \dot{z}_i(0) = 0, \quad v_i(0) = 0, \tag{19}$$

almost everywhere (a.e.) in $t \in [0,T]$, where $v(\cdot) = (v_1(\cdot), \ldots, v_K(\cdot))$ is a nonnegative piecewise continuous K-vector function. Hence, we turn to find a piecewise continuous function as our approximating control function for the OCMC problem. In doing so, we replace the constraints related to z_i by two different approximating forms that allow us to design implementable and effective numerical algorithms.

The first approximating form is actually taking derivative of $y(t)$ by using (11) to get corresponding control inequality constraints to replace the constraints in (14), which leads to our first numerical algorithm (Algorithm 3.1) to solve the OCMC problem in (12)–(14). The second approximating form is directly corresponding to the numerical integrations of $y(t)$ and $z(t)$ by applying (11) and (19), which leads to our second numerical algorithm (Algorithm 3.2) to solve the OCMC problem in (12)–(14). From Remark 4.1 presented in Section 4.2, one can see that Algorithm 3.1 and Algorithm 3.2 display different advantages related to different system parameters in real numerical implementations.

3.1. An algorithm based on control inequality constraints

First, it follows from the first and third constraints in (14), (11), and (19) that

$$\int_0^t ((x_i w(s) + \lambda_i) - \mu_i m_i(s) - v_i(s))\, ds = m_i(t) \geq 0, \tag{20}$$

where $w(t)$ is a fast decaying function in terms of $t \in [0, T]$ such that

$$\int_0^t x_i w(s) ds \approx x_i \quad \text{for all} \quad t \in (0, T].$$

Then, we can find our approximating control function for the OCMC problem with constraints related to z_i replaced by (19) and (21) for $i \in \{1, \ldots, K\}$ a.e. in $t \in [0, T]$,

$$\begin{cases} g_i(y, z, m, v, t) & \equiv v_i(t) \geq 0, \\ g_{K+i}(y, z, m, v, t) & \equiv (x_i + \lambda_i t) - (m_i(t) + \mu_i y_i(t) + z_i(t)) = 0, \\ g_{2K+i}(y, z, m, v, t) & \equiv (x_i w(t) + \lambda_i - \mu_i m_i(t)) \vee 0 - v_i(t) \geq 0, \\ g_0(y, z, m, v, t) & \equiv nw(t) - \sum_{i=1}^K (x_i w(t) + \lambda_i - \mu_i m_i(t)) \vee 0 \\ & \quad + \sum_{i=1}^K v_i(t) \geq 0, \end{cases} \tag{21}$$

where the third constraint in (21) is naturally obtained from the constraint in (20) such that $v(t) \geq 0$ for each $t \in [0, T]$ is not violated, the fourth constraint in (21) is obtained by the fourth constraint in (14) and the third constraint in (21).

In this case, the Pontryagin's maximum principle with monotonicity and equality and inequality constraints (see, e.g., Hellwig [19], Theorems 6.1–6.2 in pp. 153–161 of Tu [23]) can be applied. More precisely, the corresponding Hamiltonian is given by

$$\begin{aligned} H(y, m, z, v, p, q, t) = p_0 f(y(t), m(t), t) + \sum_{i=1}^K p_i(t) m_i(t) \\ + \sum_{i=K+1}^{2K} p_i(t) v_i(t) + \sum_{j=0}^{3K} q_j(t) g_j(y, m, z, v, t). \end{aligned} \tag{22}$$

In the Equation (22), p_0 is a constant (and without loss of generality, we suppose it to be the unity), $p_i(t)$ with $i \in \{1, \ldots, 2K\}$ and $q_j(t)$ with $j \in \{0, 1, \ldots, 3K\}$ are continuous functions over $[0, T]$, satisfying

$$\frac{dp_i(t)}{dt} = -\frac{\partial H}{\partial y_i} = -\alpha l_i \mu_i e^{-\alpha t} + q_{K+i}(t) \mu_i, \tag{23}$$

$$\frac{dp_{K+i}(t)}{dt} = -\frac{\partial H}{\partial z_i} = q_{K+i}(t), \tag{24}$$

$$0 = \frac{\partial H}{\partial m_i} = (r_i + \alpha l_i) e^{-\alpha t} + p_i(t) + q_0(t) \mu_i - q_{K+i}(t) - q_{2K+i}(t), \tag{25}$$

$$0 = \frac{\partial H}{\partial v_i} = p_{K+i}(t) + q_i(t) + q_0(t) - q_{2k+i}(t), \tag{26}$$

$$0 = q_j(t) g_j(y, m, z, v, t), \quad q_j(t) \geq 0, \quad g_j(y, m, z, v, t) \geq 0. \tag{27}$$

Hence, the optimal control function at each time $t \in [0, T]$ can be obtained by the LP problem subject to the constraints (13), (19), and (21),

$$\max_{(m,v)} H(y,m,z,v,p,q,t), \tag{28}$$

where $y = y^*$ and $z = z^*$ correspond to the optimal path by time t, and

$$H(y,m,z,v,p,q,t) = \sum_{i=1}^{K} \left(e^{-\alpha t}\alpha l_i \mu_i y_i(t) + \eta_i(t) m_i(t) + \zeta_i(t) v_i(t)\right), \tag{29}$$

$$\eta_i(t) = e^{-\alpha t}(r_i + \alpha l_i) + p_i(t), \tag{30}$$

$$\zeta_i(t) = p_{K+i}(t) \tag{31}$$

for $i \in \{1, \ldots, K\}$. Note that the optimal solutions to the LP problem in (28) at each time t are located on the boundary of the constraint set (see, e.g., Corollary 32.3.4 in p. 345 of Rockafeller [26] and Figure 1 in the current paper). Hence, the extreme points of the feasible sets are the possible optimal solutions. Therefore, it follows from Theorem 6.2 in p. 161 of Tu [23], we can take $q_0(t) = q_{2K+i}(t) = 0$ for each $i \in \{1, \ldots, K\}$ in order to get an optimal solution since they are either zero or the free variables (whose values can be arbitrarily taken) owing to (27). Thus, it follows from (23) to (26) that

$$\frac{dp_i(t)}{dt} = -\alpha l_i \mu_i e^{-\alpha t} + q_{K+i}(t)\mu_i, \tag{32}$$

$$\frac{dp_{K+i}(t)}{dt} = q_{K+i}(t), \tag{33}$$

$$0 = (r_i + \alpha l_i)e^{-\alpha t} + p_i(t) - q_{K+i}(t), \tag{34}$$

$$0 = p_{K+i}(t) + q_i(t) \tag{35}$$

with $p_i(T) = p_{K+i}(T)$ owing to the transversality conditions in Theorem 6.2 of Tu [23] (or Theorem 6.2 in p. 21 of Fleming and Soner [27]). Hence, for each $t \geq 0$, we have

$$p_i(t) = \frac{r_i \mu_i}{\mu_i + \alpha}\left(e^{\mu_i(t-T)-\alpha T} - e^{-\alpha t}\right), \tag{36}$$

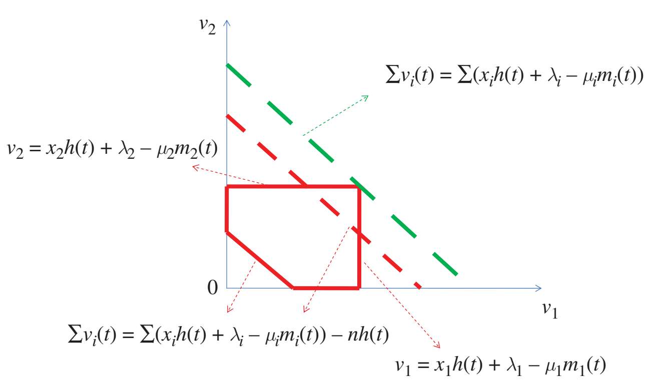

Figure 1. A typical feasible convex set for the loss control function at time t.

$$p_{K+i}(t) = -\frac{r_i}{\mu_i+\alpha}\left(e^{-\alpha T} - e^{\mu_i(t-T)-\alpha T}\right) + \frac{1}{\alpha}\left((r_i+\alpha l_i) - \frac{r_i\mu_i}{\mu_i+\alpha}\right)\left(e^{-\alpha T} - e^{-\alpha t}\right). \tag{37}$$

In addition, owing to the linearity of f in (10), the Erdmann–Weierstrass conditions required in Theorem 6.2 of Tu [23] are satisfied by the obtained $p(t)$ and $q(t)$. Therefore, we have the following Bang–Bang control based algorithm.

Algorithm 3.1 *The algorithm consists of the following three parts*:

- *Partition* $[0,T]$ *into* $\{[t_l, t_{l+1}), l = 0,1,2,\ldots,L\}$ *with equal gauge* $\Delta t = t_{l+1} - t_l$ *and* $t_0 = 0$, *and for each* $l \geq 1$, *compute* $w(t_l)$, *e.g.*, $w(t_l)$ *can be taken in the following form for good performance*:

$$w(t_l) = \begin{cases} \dfrac{0.2}{(1-e^{-0.2\times 5})e^{0.2t_l}} & \text{if} \quad t_l \leq 8, \\ \dfrac{1}{t_l^{1.01}} & \text{if} \quad t_l > 8, \end{cases}$$

and calculate

$$\phi_i(t_l) = (x_i w(t_{l-1}) + \lambda_i - \mu_i m_i(t_{l-1})) \vee 0 \text{ for } \quad i \in \{1,\ldots,K\}.$$

- *Let* $v_i(t_0) = 0$ *for each* $i \in \{1,\ldots,K\}$, *and for each* t_l *with* $l \in \{1,\ldots,L\}$, *let* $\zeta_i(t_l)$ *be given in* (31) *and* (37). *Then*
  (1) *Take* $v_i(t_l) = \phi_i(t_l)$ *for each* $i \in \{1,\ldots,K\}$ *satisfying* $\zeta_i(t_l) \geq 0$;
  (2) *Let* $i_s \in \{1,\ldots,K\}$ *with* $s = 1,\ldots,K_1$ *for some* $K_1 \in \{0,1,\ldots,K\}$ *denote the index such that* $\zeta_{i_s}(t_l) < 0$, *take* $(v_{i_1}(t_l),\ldots,v_{i_{K_1}}(t_l))$ *to be the solution to the following LP problem*:

$$\max \sum_{s=1}^{K_1} \zeta_{i_s}(t_l) v_{i_s}(t_l),$$

$$s.t. \quad \begin{cases} \sum_{s=1}^{K_1} v_{i_s}(t_l) = \left(\sum_{s=1}^{K_1} \phi_{i_s}(t_l) - nw(t_l)\right) \vee 0, \\ 0 \leq v_{i_s}(t_l) \leq \phi_{i_s}(t_l); \end{cases}$$

- *Let* $m_i(t_0) = x_i$ *and* $y_i(t_0) = 0$ *with each* $i \in \{1,\ldots,K\}$, *and for each* $l \geq 1$, *take*

$$m_i(t_l) = \frac{1}{1+\mu_i \Delta t}\left(m_i(t_{l-1}) + \Delta t(\lambda_i - v_i(t_l))\right),$$

$$y_i(t_l) = y_i(t_{l-1}) + m_i(t_l)\Delta t.$$

Note that, in the third part of Algorithm 3.1, we have used numerical integrations of $y(t)$ and $z(t)$ to compute $m_i(t_l)$ by using (11), (19), and the third constraint in (14) as follows:

$$m_i(t_l) = x_i + \lambda_i t_l - \mu_i \sum_{j=0}^{l} m_i(t_j)\Delta t - \sum_{j=0}^{l} v_i(t_j),$$

where Δt should be sufficiently small owing to the convergence consideration of numerical integrations.

Remark 3.1 *Note that Algorithm* 3.1 *is an iterative one in terms of time parameter t_l. $\zeta_i(t_l)$ for each $l \in \{0, 1, \ldots, \}$ and each $i \in \{1, \ldots, K\}$ can be precalculated by using* (31) *and* (37) *to save computational time at each t_l. Hence, for an online threshold policy based on the values of $v_i(t_l)$, we only care about the computational complexity of the algorithm at each instant t_l, which is mainly dominated by the time in solving the LP problem appeared in the second part of the algorithm. If the Karmarkar algorithm* (*see, e.g., Karmarkar* [28] *and Megiddo* [29]) *is employed, the complexity of Algorithm* 3.1 *at each time t_l is bounded by $O(K^{3.5}B_1) + O(K)$, where B_1 is the number of bits in the input of the LP appeared in Algorithm* 3.1 *at each time instant t_l. Note that B_1 depends on the number of servers* (*i.e., $B_1 = O(n)$*) *owing to the constraints in the second step of Algorithm* 3.1. *Nevertheless, if we employ the constrained MDP to design a threshold policy, the complexity for the related LP in computing the value function for a given initial state at each time t_l will be $O((nK)^{3.5}B_2)$* (*see, e.g., pp.* 12–16 *of Altman* [12], *Dai and Feng* [3]), *where B_2 is the number of bits in the input of the LP and is independent of n and K. Therefore, if the number of job classes is large and when the arrival and service rates are high, our algorithm has less complexity than the constrained MDP based algorithm. Furthermore, by sufficient conditions of optimality similarly as discussed in Theorem* 2.5 *of Yong and Zhou* [30], *one can see that the control function constructed by using Algorithm* 3.1 *is indeed an approximated solution to the OCMC model in* (12), *which subjects to the constraints* (13) *and* (19)–(21).

3.2. An alternative algorithm based on control-state inequality constraints

An alternative to Algorithm 3.1 can be proposed via replacing the third and fourth pure control oriented inequalities in (21) by the following mixed control and state oriented inequalities due to the third part in Algorithm 3.1, i.e., for the partition $\{[t_l, t_{l+1}), l = 0, 1, 2, \ldots, L\}$ with $t_0 = 0$ and each $l \geq 1$,

$$\begin{cases} g_{2K+i}(y, z, m, v, t_l) & \equiv \phi_i(t_l) - v_i(t_l)\Delta t/(1 + \mu_i \Delta t) \geq 0, \\ g_0(y, z, m, v, t_l) & \equiv n - \sum_{i=1}^{K} \phi_i(t_l) + \sum_{i=1}^{K} v_i(t_l)\Delta t/(1 + \mu_i \Delta t) \geq 0, \end{cases} \tag{38}$$

where, for $i \in \{1, \ldots, K\}$,

$$\phi_i(t_l) = \frac{1}{1 + \mu_i \Delta t}\left(x_i + \lambda_i t_l - \mu_i y_i(t_{l-1}) - z(t_{l-1})\right) \vee 0.$$

Then, we have the following algorithm.

Algorithm 3.2 *The algorithm consists of the following two parts*:

- *Let $v_i(t_0) = 0$ for each $i \in \{1, \ldots, K\}$, and for each t_l with $l \in \{1, \ldots, L\}$, let $\zeta_i(t_l)$ be given in* (31) *and* (37). *Then*
  (1) *Take $v_i(t_l) = \phi_i(t_l)(1 + \mu_i \Delta t)/\Delta t$ for each $i \in \{1, \ldots, K\}$ satisfying $\zeta_i(t_l) \geq 0$;*
  (2) *Let $i_s \in \{1, \ldots, K\}$ with $s = 1, \ldots, K_1$ and some $K_1 \in \{0, 1, \ldots, K\}$ denote the index such that $\zeta_{i_s}(t_l) < 0$, take $(v_{i_1}(t_l), \ldots, v_{i_{K_1}}(t_l))$ to be the solution to the following LP problem*:

$$\max \sum_{s=1}^{K_1} \zeta_{i_s}(t_l) v_{i_s}(t_l)$$

$$s.t. \begin{cases} \sum_{s=1}^{K_1}(1/(1+\mu_{i_s}\Delta t))v_{i_s}(t_l) = \left(\sum_{s=1}^{K_1}(\phi_{i_s}(t_l)-n)/\Delta t\right) \vee 0, \\ 0 \le v_{i_s}(t_l) \le \phi_{i_s}(t_l)(1+\mu_{i_s}\Delta t)/\Delta t; \end{cases}$$

- *Let* $m_i(t_0) = x_i$ *and* $y_i(t_0) = 0$ *with each* $i \in \{1, \ldots, K\}$, *and for each* $l \ge 1$, *take* $m_i(t_l)$ *and* $y_i(t_l)$ *as given in Algorithm* 3.1.

3.3. An online admission control policy

Based on the loss rate functions determined in the previous subsections, we can design an online admission control policy for the parallel-server loss channel as follows.

Admission Policy 3.1 *The policy consists of the following five parts*:

(1) *At each time* t_l *with* $l \in \{0, 1, \ldots, L\}$, *compute* (*or input*) $v(t_l)$ *by using Algorithm* 3.1 (*or Algorithm* 3.2);
(2) *Define two sequences of random times* $\{\alpha_l, l \in \{1, 2, \ldots\}\}$ *and* $\{\beta_l, l \in \{0, 1, 2, \ldots\}\}$ *with* $\beta_0 = 0$ *as*

$$\alpha_l = \inf\left\{ t > \beta_{l-1}, \sum_{i=1}^{K} X_i(t) = K \right\}, \ l \in \{1, 2, \ldots\},$$

$$\beta_l = \inf\left\{ t > \alpha_l, \sum_{i=1}^{K} X_i(t) < K \right\}, \quad l \in \{1, 2, \ldots\};$$

(3) *If* $t \in [\beta_{l-1}, \alpha_l]$ (*the system is not full in this time interval*) *with* $l \in \{1, 2, \ldots\}$ *and* $v_i(t) > 0$ (*the loss ratio is positive at time t*) *for an* $i \in \{1, \ldots, K\}$, *the system rejects an arrived class i job at time t*;
(4) *If* $t \in [\beta_{l-1}, \alpha_l]$ *with* $l \in \{1, 2, \ldots\}$ *and* $v_i(t) = 0$ *for an* $i \in \{1, \ldots, K\}$, *the system accepts an arrived class i job at time t*;
(5) *If* $t \in [\alpha_l, \beta_l]$ (*the system is full in this time interval*) *with* $l \in \{1, 2, \ldots\}$, *the system rejects an arrived job at time t.*

4. Numerical simulations and performance comparisons

4.1. A simulation algorithm

To implement Policy 3.1, we consider a 2-class and n-server parallel system and propose a simulation algorithm by partitioning $[0, T]$ as in Algorithm 3.1 and Algorithm 3.2.

Algorithm 4.1 *The simulation algorithm consists of the following five steps*:

(1) *Counting* $\xi(k)$ *for each* $k \in \{1, 2, \ldots\}$ *and* $\xi_i(k)$ *for each* $i \in \{1, 2\}$, *which denote the arrival time of the kth job* (*either a class*-1 *job or a class*-2 *job*) *and the arrival time of the kth class i job, respectively*;
(2) *For each* $i \in \{1, 2\}$, $r \in \{1, \ldots, n\}$, *and* $k \in \{1, 2, \ldots\}$, *define*

$$\tau_i^r = \xi(k-1) + \gamma_i^r,$$

where γ_i^r is the remaining service time required for server r to finish serving a class i job, which starts at time $\xi(k-1)$;

(3) *Let J_i denote the profit associated with class* i *jobs and take $J_i = 0$ initially. Then, for each $k \in \{1, 2, \ldots\}$ such that $\xi(k) \leq T$, if $\tau_i^r \geq \xi(k)$, take*

$$J_i = J_i + r_i \left(e^{-\alpha\xi(k-1)} - e^{-\alpha\xi(k)}\right),$$

otherwise, take

$$J_i = J_i + r_i \left(e^{-\alpha\xi(k-1)} - e^{-\alpha\tau_i^r}\right);$$

(4) *Let $q(t)$ be the number of servers occupied by a job at time $t \in [t_{l-1}, t_l)$. If $t_{l-1} \leq \xi(k) < t_l$ and $\xi(k) = \xi_i(k_1)$ for some $i \in \{1, 2\}$, $k_1 \in \{1, \ldots, k\}$, then*
   (a) *if $v_i(t) = 0$ and $q(t) < n$, the system accepts the arrived class i job at time $\xi(k)$ and assigns an available server $r \in \{1, \ldots, n\}$ to serve the job while updating $\gamma_i^r = u_i^j$ for some $j \in \{1, 2, \ldots\}$*
   (b) *otherwise, the system rejects the arrived class i job at time $\xi(k)$ while taking $L_i = 0$ initially and for each $k \geq 1$ such that $\xi(k) \leq T$, taking*

$$L_i = L_i + l_i e^{-\alpha\xi(k)},$$

where L_i is the lost value associated with class i jobs;

(5) *Finally, compute the optimal total profit by*

$$L = \sum_{i=1}^{2}(J_i - L_i) - \frac{c}{\alpha}.$$

4.2. *Comments to Algorithm 4.1 and its implementations*

In the following three subsections, we implement Policy 3.1 by using the proposed simulation algorithm for a 2-class and 7-sever parallel system when the system is either of or not of the Markovian property. In these simulations, we find out that Policy 3.1 implemented by applying Algorithm 4.1 outperforms several ones when the traffic intensity

$$\rho = \lambda_1/\mu_1 + \lambda_2/\mu_2 \tag{39}$$

is relatively large, and the differences of the revenues per unit of time and the penalty costs among different classes of jobs are large. Here, for each $x \in S$, the comparison criterion is the relative error between $v^{\pi^*}(x)$ (the numerically computed (or approximated) optimal value obtained under Policy 3.1) and $v(x)$ obtained by one of the following three policies:

(1) The FIFO policy, i.e., the policy by taking $v_1(t) = v_2(t) = 0$ for all $t \in [0, T]$ and accepting an arrived job at time t if there is a server available at time t;
(2) The AS policy: the system accepts an arrived job at time t in all odd indexed interval $[t_l, t_{l+1})$ for $l \in \{0, 1, \ldots, L-1\}$ and rejects otherwise;
(3) The MDP based threshold policy in Dai and Feng [3].

Mathematically, for each $x \in S$, the relative error is expressed as

$$error(x) = \left(v^{\pi^*}(x) - v(x)\right)/v^{\pi^*}(x). \tag{40}$$

Remark 4.1 *From our simulation experiences, we see that $v_2(t) = 0$ for all $t \in [0, T]$ when λ_2 is relatively small (e.g., approximately below 5) and $v_2(t) > 0$ for a small number of $t \in [0, T]$ when λ_2 is relatively large (e.g., approximately above 5). However, when we implement Policy 3.1 by taking $v_2(t) = 0$ for all $t \in [0, T]$, good performance is reached for a system with λ_2 relatively large. This is because when λ_2 increases, the loss rate of class-2 jobs increases solely owing to system blocking. Thus, the system accepts the small number of class-2 jobs (or a portion of the small number) corresponding to $v_2(t) > 0$ is a reasonable strategy. In the following simulation examples, when $v_2(t)$ is taken to be 0 for all $t \in [0, T]$, $v_1(t)$ will converge to a constant (e.g., 0 or 1) in all cases. The phenomenon indicates that the dynamical feasible control region as shown in Figure 1 will converge to a fixed one as $t \to \infty$. In addition, the limit of dynamical optimal control pair $(v_1(t), v_2(t))$ is located in one of the extreme points of the limit feasible region. Essentially, the phenomenon roughly reflects the principle of LP or Bang–Bang control for linear optimal control system as discussed in Section 3.*

Furthermore, when Algorithm 3.1 is employed in the simulations, good performance is reached for all cases in terms of the traffic intensity ρ and the computed loss rate functions. However, when the system is overloaded (i.e., $\rho > 1$), the summation of computed mean functions of two classes of jobs at each time t may exceed the upper bound $n = 7$. When Algorithm 3.2 is employed in the simulations, good performances are reached if the system is critically loaded and overloaded (i.e., $\rho \geq 1$), and Policy 3.1 reduces to FIFO if $\rho < 1$. Nevertheless, the computed variables satisfy all the constraints as stated in the OCMC model.

Finally, the numerical comparisons provided in the following three subsections are conducted when the system is overloaded, critically loaded, or under-loaded in terms of the traffic intensity ρ in (39). In all these simulations, the sever number is fixed to be $n = 7$, the number of job classes is fixed to be $K = 2$, the time horizon is 400 (i.e., $T = 400$), the time gauge is 400/10,000, the discounted rate $\alpha = 0.73$, the fixed cost is $c = 1000$, the revenue per unit of time is 1000 for class-1 jobs and is 6000 for class-2 jobs, the penalty cost is 200 for each class-1 job and is 5000 for each class-2 job (except the cost in Example 4.10), and the number of simulation run times is 1000. Furthermore, to be simple for notations, we use 'exp' to denote 'exponential' in all the related figures.

Here, we note that $\alpha = 0.73$ is an arbitrary choice. If one chooses a larger α, the convergence of the numerical integration related to (18) will be more faster in suitable accuracy. If one chooses a smaller α, T may be chosen larger than the current value 400 to guarantee the convergence of the associated numerical integration related to (18) in suitable accuracy. When the numbers of n and K increase, the corresponding computation run-time will increase. For the current purpose of illustration, we choose $n = 7$ and $K = 2$.

4.3. Algorithm 3.1 based simulation examples: Markovian cases and comparing with FIFO and AS policies

In this subsection, we implement the simulation algorithm based on Algorithm 3.1 and provide numerical comparisons when the job arrival processes are Poisson ones and the service times are exponentially distributed.

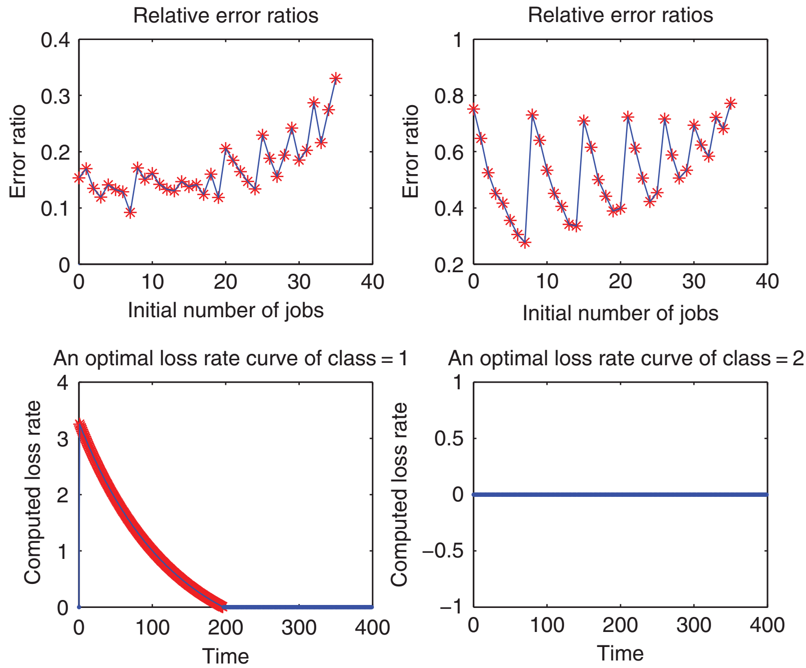

Figure 2. Alg 3.1, exp, $\rho = 1.1626$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 5.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

Example 4.1 (With FIFO, AS, Markovian, and Overloaded) *The traffic intensity is taken to be* $\rho = 1.1626$ *with* $\lambda_1 = 5.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$. *The related error ratios are displayed in Figure 2. In the figure, the error ratios showed in the left-upper graph are compared with the FIFO policy, the error ratios showed in the right-upper graph are compared with the AS policy. Furthermore, in the upper two graphs, each starred point corresponds to an initial state* (i,j) *that is arranged in the order of* $j = 0, 1, \ldots, n - i + 1$ *when* i *increases from* 0 *to* n. *In the lower two graphs, the curves represent loss rate functions, respectively, corresponding to class-*1 *jobs and class-*2 *jobs with initial state* (2,1)*, which are generated from Policy* 3.1 *and are drawn from their first* 400 *computed points. If* $v_i(t_l) > 0$, *a red 'star' point is drawn and otherwise, a blue 'dot' point is drawn.*

Example 4.2 (With FIFO, AS, Markovian, and Critically Loaded) *Besides* $\rho = 1$ *and* $\lambda_1 = 4.0622$, *the other data are the same as in Example* 4.1. *The related error ratios are displayed in Figure* 3 *with the same explanation as in Example* 4.1.

Example 4.3 (With FIFO, AS, Markovian, and Heavy Traffic) *Besides* $\rho = 0.9675$ *and* $\lambda_1 = 3.8622$, *the other data are the same as in Example* 4.1. *The related error ratios are displayed in Figure* 4 *with the same explanation as in Example* 4.1.

4.4. Algorithm 3.1 based simulation examples: non-Markovian cases and comparing with FIFO and AS policies

As pointed out at the end of Subsection 2.1, when the Poisson arrival processes and/or the exponentially distributed service times imposed in the previous discussions are

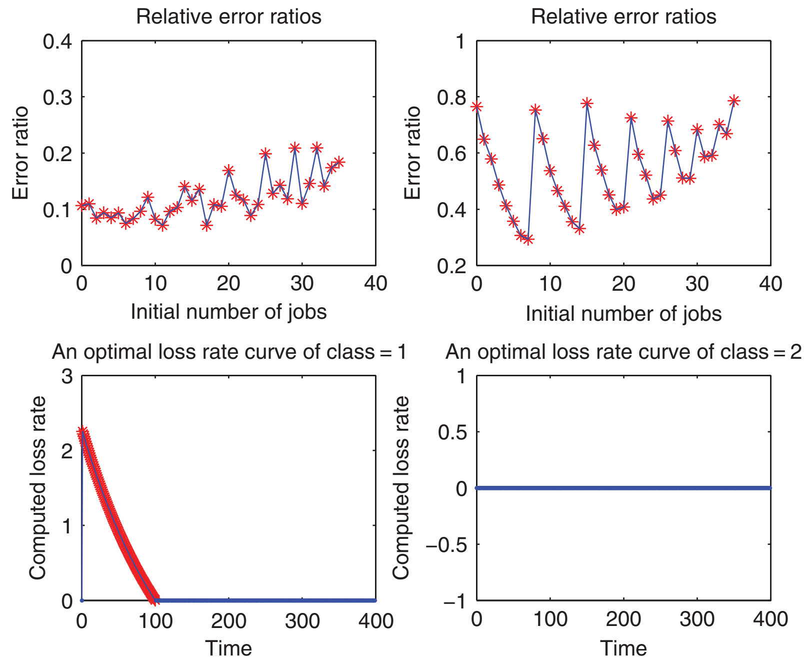

Figure 3. Alg 3.1, exp, $\rho = 1.0000$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 4.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

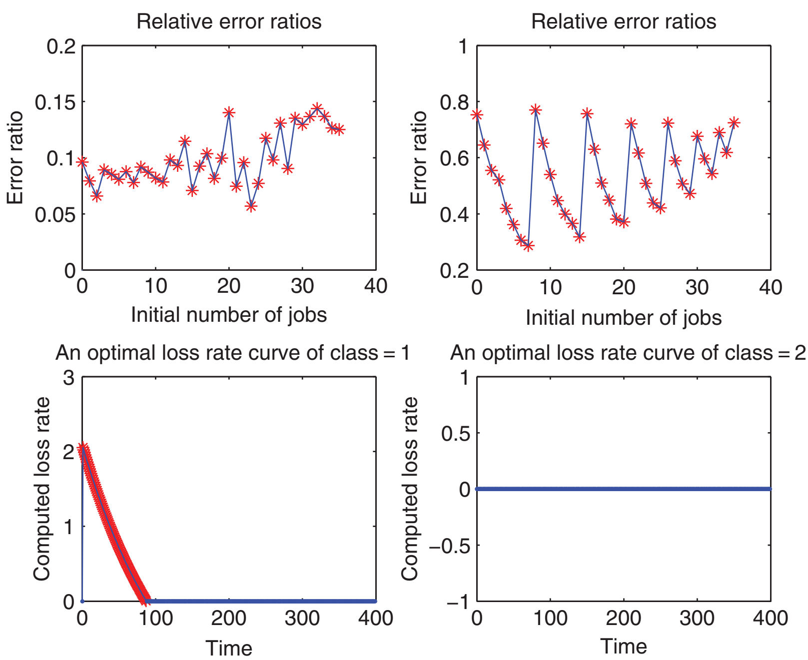

Figure 4. Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

replaced by renewal processes and/or generally distributed service times, the formerly discussed system does not satisfy Markov (memoryless) property. The conditions used in Theorem 2.1 are not satisfied by such a generalized queueing system. However, the following Examples 4.4–4.6 indicate that our OCMC based Policy 3.1 still performs well for the previously studied 2-class and 7-server parallel system with renewal arrival processes that are of uniformly distributed job interarrival times. One way to interpret this phenomenon is as given in the introduction of the paper. Here, we use the following numerical Examples 4.4–4.6 to illustrate the effectiveness of Policy 3.1 for a non-Markovian model with uniformly distributed job interarrival times.

Example 4.4 (With FIFO, AS, Non-Markovian, and Overloaded) *Besides the uniformly distributed interarrival times, the other data are the same as in Example* 4.1. *The related error ratios are displayed in Figure* 5.

Example 4.5 (With FIFO, AS, Non-Markovian, and Critically Loaded) *Besides the uniformly distributed interarrival times, the other data are the same as in Example* 4.2. *The related error ratios are displayed in Figure* 6.

Example 4.6 (With FIFO, AS, Non-Markovian, and Heavy-Traffic) *Besides the uniformly distributed interarrival times, the other data are the same as in Example* 4.3. *The related error ratios are displayed in Figure* 7.

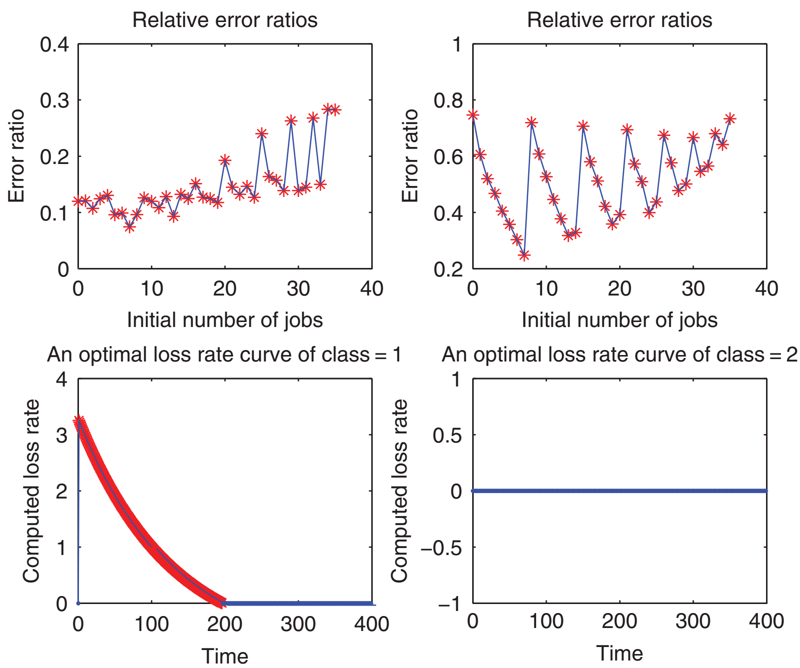

Figure 5. Alg 3.1, uniform, $\rho = 1.1626$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 5.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

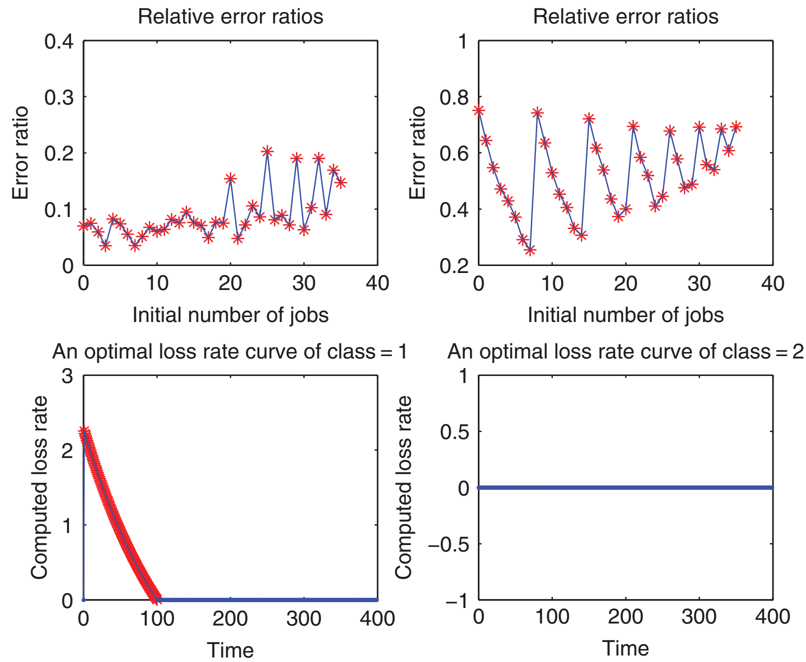

Figure 6. Alg 3.1, uniform, $\rho = 1$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 4.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

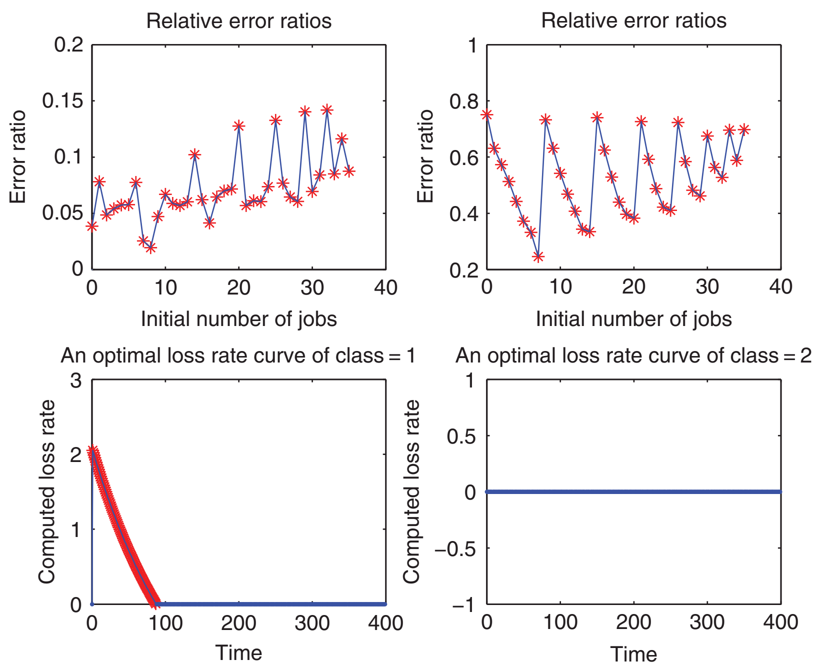

Figure 7. Alg 3.1, uniform, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

4.5. *Algorithm 3.2 based simulation examples: non-Markovian cases and comparing with FIFO and AS policies*

In this subsection, we implement the simulation algorithm based on Algorithm 3.2 instead of Algorithm 3.1 as used in the previous subsections when the system is non-Markovian. The different advantages of numerical performance between Algorithm 3.1 and Algorithm 3.2 are as explained in Remark 4.1.

Example 4.7 (With FIFO, AS, Non-Markovian, and Overloaded) *The system data are the same as in Example* 4.4 *and the related error ratios are displayed in Figure* 8.

Example 4.8 (With FIFO, AS, Non-Markovian, and Critically Loaded) *The system data are the same as in Example* 4.5 *and the related error ratios are displayed in Figure* 9, *where, in the left-lower graph, the curve is drawn from their* 10,000 *computed points.*

4.6. *Comparing with FIFO and MDP based threshold policies*

In this subsection, we employ Algorithm 3.1 to conduct the performance comparisons among our Policy 3.1, the FIFO policy, and the MDP based threshold policy proposed in Dai and Feng [3]. From the numerical Examples 4.9–4.12, we see that our Policy 3.1 together with Remark 4.1 outperforms both the FIFO policy and the MDP based policy in [3] when the difference of revenue rates between two classes of jobs is large. However, if the difference is small, the performance of the MDP based policy is fairly good. Furthermore, this phenomenon reveals an interesting issue in our study, i.e., the conditions (e.g., submodularity) imposed in [3] indeed have some impact on the design and

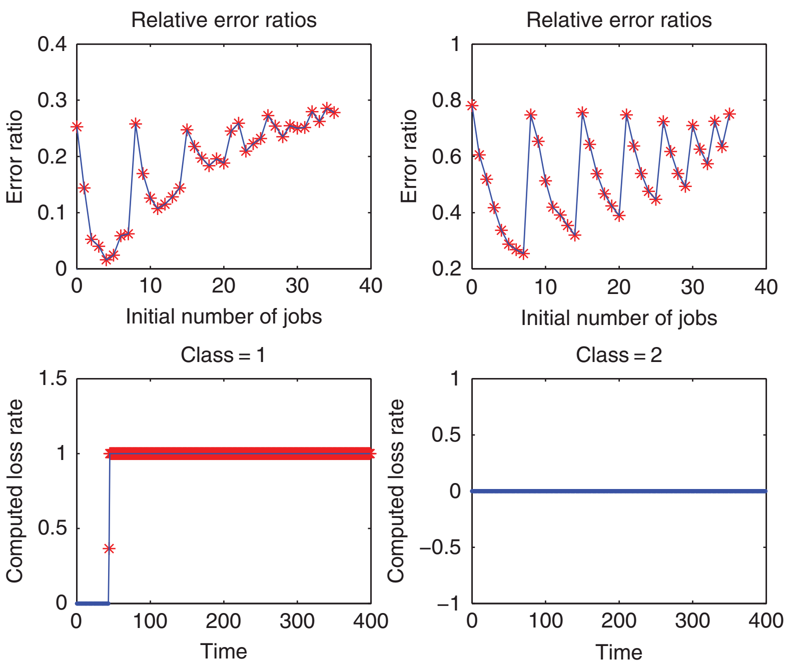

Figure 8. Alg 3.2, uniform, $\rho = 1.1626$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 5.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

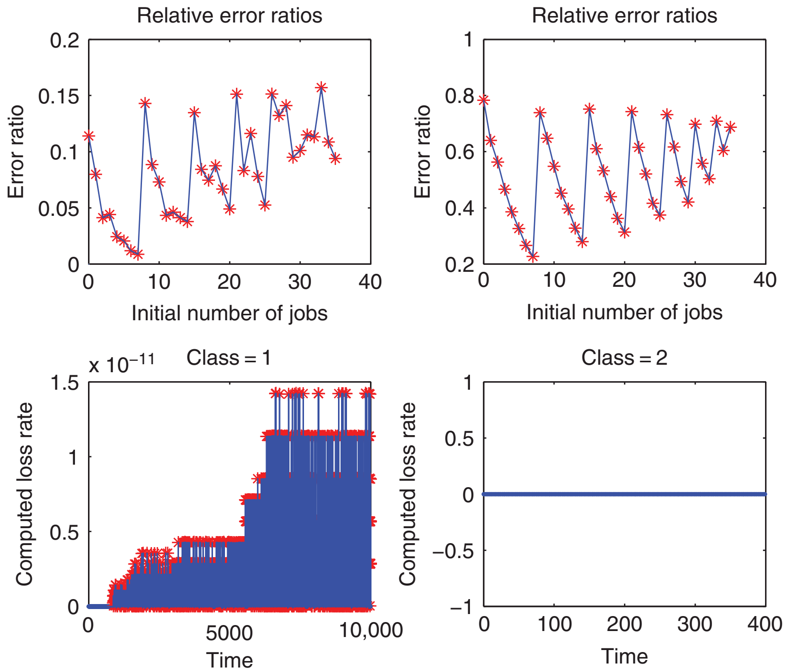

Figure 9. Alg 3.2, uniform, $\rho = 1$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 4.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

performance of the MDP based optimal policy. This is because the value function used in our OCMC model is not required to be submodular. Thus, we can identify an optimal value function and optimal admission control policy on an enlarged class of value functions. Actually, owing to Theorem 2.2, the optimal value to our OCMC model provide an upper bound to the corresponding values of our physical queueing system under Markovian decision rules. Therefore, it is possible that the approximate approach proposed in our current paper is better than the optimal policy resulting from a structured value function based MDP such as in the reference [3] in certain cases. Since using the class of structured value functions to design optimal policy for continuous time Markov decision problem is a popular method in many areas recently, our finding here should be valuable in this field.

Example 4.9 (With FIFO, MDP, and Critically Loaded) *The system data for the lower two graphs in Figure* 10 *are the same as those in Example* 4.2. *The system data for the upper two graphs in Figure* 10 *are the same as those in Example* 4.2 *except the revenue rate of the class-*2 *jobs, which is taken to be* 1300 *here. The left two graphs display the comparisons between our newly designed policy and the FIFO policy. The curves are drawn from their* 10,000 *computed points. Note that the slight difference between the left-lower graph in Figure* 10 *and the left-upper one in Figure* 3 *corresponding to Example* 4.2 *is mainly owing to the error ratio* $1/\sqrt{1000}$ *of random number generator used in Matlab Software Package since our number of simulation run times is* 1000 *and two different* 1000 *average sample paths are used. Furthermore, the error ratio measures the difference between a random variable and its counterpart corresponding to pseudo random numbers generated by a random number generator owing to central limit theorem, etc. However, the accuracy*

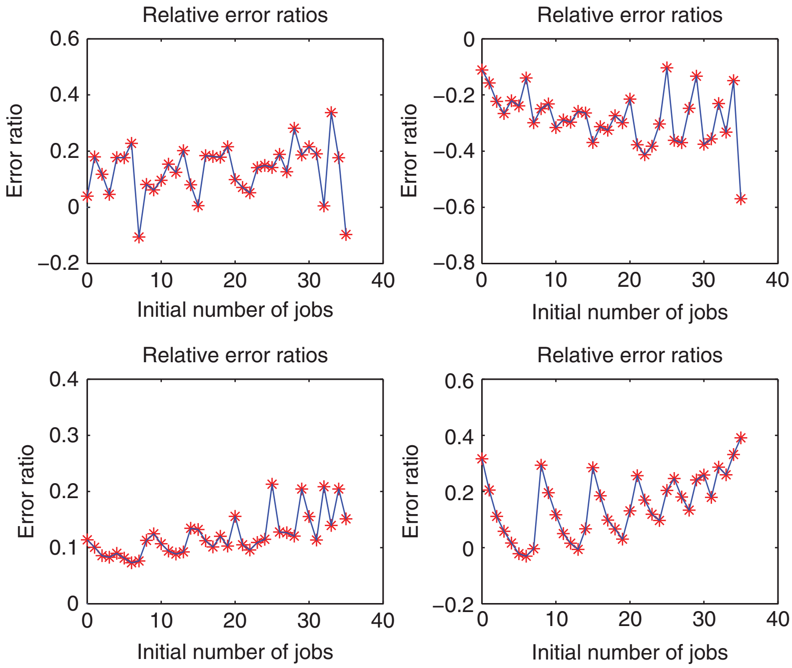

Figure 10. Upper two: Alg 3.1, exp, $\rho = 1.0000$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 4.0622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 1300$. Lower two: Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

is reasonably good for our current illustration purpose. The right two graphs display the comparisons between our newly designed policy and the MDP based threshold policy in [3]. *The number of iterative times for the MDP based algorithm is* 500.

Example 4.10 (With FIFO, MDP, and Critically Loaded) *The system data except the ones list below are the same as those in Example* 4.2. *The related error ratios are displayed in Figure* 11. *The revenue rates of the class-*2 *jobs for the upper two graphs are both* 2000 *and those for the lower two graphs are both* 7500. *Other interpretations are the same as in Example* 4.9.

Example 4.11 (With FIFO, MDP, and Heavy Traffic) *The system data for the lower two graphs in Figure* 12 *are the same as those in Example* 4.3. *The system data for the upper two graphs in Figure* 12 *are the same as those in Example* 4.3 *except the revenue rate of the class-*2 *jobs, which is taken to be* 1300 *here. The left two graphs display the comparisons between our newly designed policy and the FIFO policy. The curves are drawn from their* 10,000 *computed points. Note that the slight difference between the left-lower graph in Figure* 12 *and the left-upper one in Figure* 4 *corresponding to Example* 4.3 *is mainly owing to the error ratio* $1/\sqrt{1000}$ *of random number generator used in Matlab Software Package since our number of simulation run times is* 1000 *and two different average sample paths are used. However, the accuracy is reasonably good for our current illustration purpose. The right two graphs display the comparisons between our newly designed policy and the MDP based threshold policy in* [3]. *The number of iterative times for the MDP based algorithm is* 500.

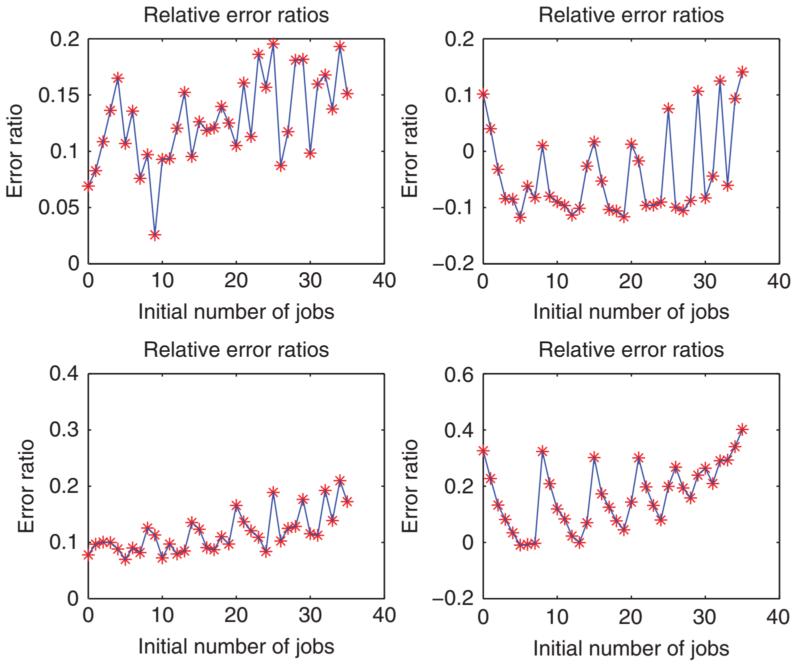

Figure 11. Upper two: Alg 3.1, exp, $\rho = 1.0000, n = 7, \alpha = 0.73, \lambda_1 = 4.0622, \lambda_2 = 2.0878, \mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 2000$. Lower two: Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 7500$.

Example 4.12 (With FIFO, MDP, and Heavy Traffic) *The system data except the ones list below are the same as those in Example* 4.3. *The related error ratios are displayed in Figure* 13. *The revenue rates of the class-*2 *jobs for the upper two graphs are both* 2000 *and those for the lower two graphs are both* 7500. *Other interpretations are the same as in Example* 4.11.

5. Proof of Theorem 2.1

5.1. Preliminary lemmas

We first note that all of the stochastic processes concerned in the following discussions are measurable functions from some probability space (Ω, F, P) into Skorohod topological space. Their paths are right-continuous with left-limits on $[0, \infty)$ and the space is endowed with the Skorohod topology (see, e.g., Ethier and Kurtz [31]).

Lemma 5.1 *For each fixed* $i \in \{1, \ldots, K\}$, *we have*

$$E \int_0^\infty e^{-\alpha t}(R_i(t) - \lambda_i t)dt = 0. \tag{41}$$

Proof. For each $i \in \{1, \ldots, K\}$, since $R_i(\cdot)$ is a Poisson process with right-continuous sample paths, then $R_i(\cdot)$ is progressively measurable in terms of the natural filtration $\mathcal{R} = \{\mathcal{R}_t = \sigma(R_i(s), 0 \leq s \leq t), t \geq 0\}$ (see, e.g., Theorem 1 in p. 38 of Chung [32]). Thus,

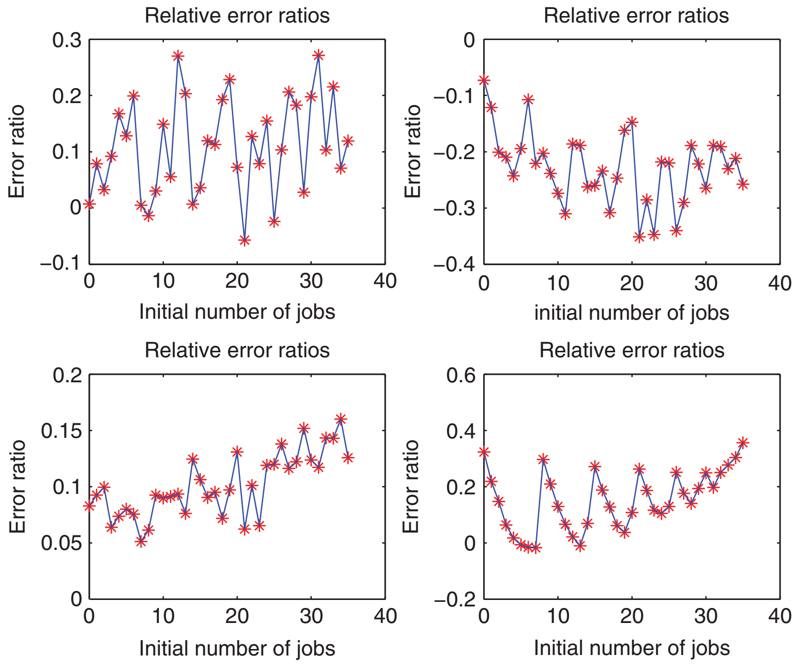

Figure 12. Upper two: Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 1300$. Lower two: Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 6000$.

$R_i(t, \omega)$ is nonnegative measurable on the product space $[0, \infty) \times \Omega$. Furthermore, note that $(\Omega, \mathcal{F}, P)$ and $([0, \infty), \mathcal{B}[0, \infty), dt)$ are two σ-finite measure spaces, where dt denotes the Lebesgue measure on $[0, \infty)$. Then, it follows from Tonelli's Theorem in p. 309 of Royden [25] that (41) is true. Hence, we finish the proof of Lemma 5.1. $\square$

Lemma 5.2 *For each fixed $i \in \{1, \ldots, K\}$, we have*

$$E\left(S_i^{\pi^*}(t) - \mu_i \int_0^t m_i^{\pi^*}(s)ds\right) = 0. \tag{42}$$

Proof. Let T_j be the departure time of the jth class i job that finishes service at the channel. Let $T_0 = 0$ and define $\eta_j \equiv T_j - T_{j-1}$ for $j = 1, 2, \ldots$. Then, we have

$$S_i^{\pi^*}(t) \equiv \max\{m : T_m \leq t\}, \quad T_m = \sum_{j=1}^{m} \eta_j. \tag{43}$$

Define

$$\xi_j \equiv \int_{T_{j-1}}^{T_j} \mu_i X_i^{\pi^*}(s)ds, \tag{44}$$

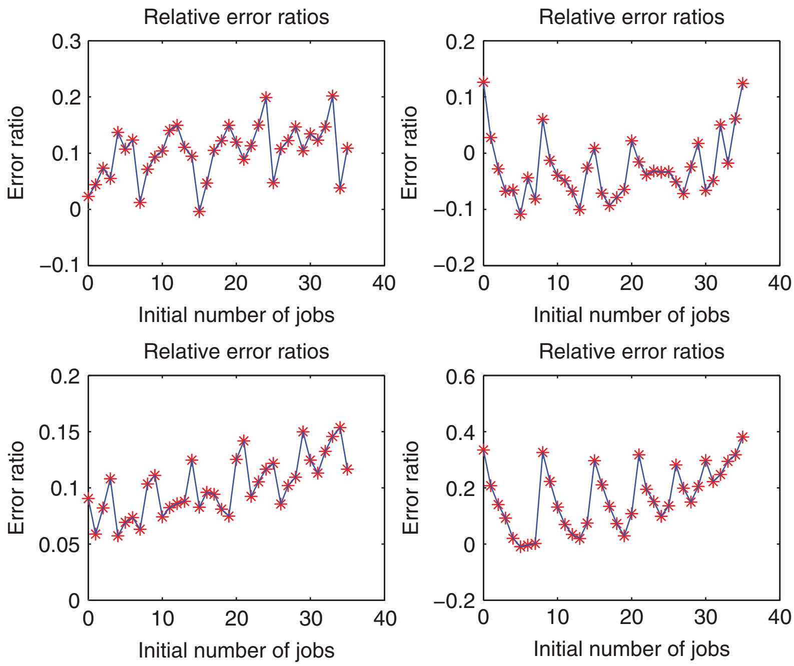

Figure 13. Upper two: Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 2000$. Lower two: Alg 3.1, exp, $\rho = 0.9675$, $n = 7$, $\alpha = 0.73$, $\lambda_1 = 3.8622$, $\lambda_2 = 2.0878$, $\mu_1 = \mu_2 = 6.15/n$, $c = 1000$, $l_1 = 200$, $l_2 = 5000$, $r_1 = 1000$, $r_2 = 7500$.

and let

$$\Xi_i^{\pi^*}(t) \equiv \max\{m : V_m \leq t\}, \quad V_m = \sum_{j=1}^{m} \xi_j. \tag{45}$$

Furthermore, let O_j denote the arrival time of the jth class i jobs that actually get into service under the policy π^* after arriving at the channel. Let $O_0 = 0$ and define $Z_j \equiv O_j - O_{j-1}$ for $j = 1, 2, \ldots$, which is the summation of a number of external interarrival times. Let $M_i^{\pi^*}(t)$ denote the total number of class i jobs that get into service under the policy π^* at the channel by time t, that is,

$$M_i^{\pi^*}(t) \equiv \max\{m : O_m \leq t\}, \quad O_m = \sum_{j=1}^{m} Z_j. \tag{46}$$

Then, we can express the left-side of (42) as

$$\begin{aligned} E&\left(S_i^{\pi^*}(t) - \mu_i \int_0^t m_i^{\pi^*}(s)ds\right) \\ &= E\left(\left(S_i^{\pi^*}(t) + 1\right) - V_{S_i^{\pi^*}(t)+1}\right) + E\int_t^{T_{S_i^{\pi^*}(t)+1}} \mu_i X_i^{\pi^*}(s)ds - 1 \\ &\equiv I + II - 1. \end{aligned} \tag{47}$$

In deriving the first equality of (47), we have used the fact that $X_i^{\pi^*}(\cdot)$ has right-continuous sample paths and is bounded by n. The remaining proof of the lemma is to show $I = 0$ and $II = 1$ for I and II that are defined in the right-side of (47).

First, we prove the claim that $I = 0$. By the similar way as developed in Dai [33] and Dai and Dai [34], let

$$\mathcal{G} = \left\{\mathcal{G}_t = \sigma\left(M_i^{\pi^*}(s), S_i^{\pi^*}(s), \Xi_i^{\pi^*}(s), X_i^{\pi^*}(0), 0 \leq s \leq t\right), t \geq 0\right\}. \tag{48}$$

Then, for each $(m_1, m_2, m_3) \in \{0, 1, \ldots\} \times \{0, 1, \ldots\} \times \{0, 1, \ldots,\}$, $(T_{m_1}, V_{m_2}, O_{m_3})$ is a multi-parameter $\mathcal{G}$-stopping time since, for example, $\{T_{m_1} \leq t\} = \{S_i^{\pi^*}(t) \geq m_1\} \in \mathcal{G}_t$ with $t \geq 0$. Since $ES_i^{\pi^*}(t)$ is bounded by $ER_i(t) = \lambda_i t < \infty$ for each fixed $t \geq 0$, we have that $S_i^{\pi^*}(t) < \infty$ almost surely (a.s.). Furthermore, the measure of the set $\{S_i^{\pi^*}(t) + 1 = \infty\}$ is zero. Thus, we conclude that $V_{S_i^{\pi^*}(t)+1}$ is $\mathcal{G}_t$-measurable since $S_i^{\pi^*}(t) + 1$ is also a $\mathcal{G}$-stopping time and the following expression holds:

$$V_{S_i^{\pi^*}(t)+1} = \sum_{m=1}^{\infty} V_m I_{\{S_i^{\pi^*}(t)+1=m\}} + V_\infty I_{\{S_i^{\pi^*}(t)+1=\infty\}}. \tag{49}$$

Next, define the filtration $\{\mathcal{G}_{T_m}, m \in \{1, 2, \ldots,\}\}$ associated with stopping times T_m for $m \in \{1, 2, \ldots\}$ by the σ-algebras,

$$\mathcal{G}_{T_m} \equiv \{B \in \mathcal{F} : B \cap \{T_m \leq s\} \in \mathcal{G}_s \quad \text{for all} \quad s \geq 0\}. \tag{50}$$

Let the filtration $\{\mathcal{G}_{(T_m,t)\}}, m \in \{1, 2, \ldots, t \geq 0\}$ be defined by the σ-algebra for each $t \geq 0$ and T_m,

$$\mathcal{G}_{(T_m,t)} \equiv \sigma\left(\mathcal{G}_{T_m} \cup \mathcal{G}_t\right). \tag{51}$$

Note that $M_i^{\pi^*}(\cdot)$ is progressively measurable in terms of the filtration $\mathcal{G}$ since it has right-continuous sample paths. Then, by Proposition 2.8.5 in p. 87 of Ethier and Kurtz [31], we know that $M_i^{\pi^*}(T_m)$ is $\mathcal{G}_{T_m}$-measurable (hence $\mathcal{G}_{(T_m,t)}$-measurable) and $\mathcal{G}_{T_{m-1}} \subset \mathcal{G}_{T_m} \subset \mathcal{G}_{(T_m,t)}$ since $T_{m-1} < T_m$. Furthermore, suppose that there are Y_m class i customers who get into service during $[T_{m-1}, T_m]$ at the channel, then we have

$$M_i^{\pi^*}(T_m) = M_i^{\pi^*}(T_{m-1}) + Y_m \equiv \max\left\{r : \sum_{j=1}^{r} Z_j \leq T_m\right\}. \tag{52}$$

Hence, we can conclude that Y_m is $\mathcal{G}_{(T_m,t)}$-measurable. Note that $S_i^{\pi^*}(t) + 1$ and $V_{S_i^{\pi^*}(t)+1}$ are also $\mathcal{G}_{(T_m,t)}$-measurable. Then, by using the following alternative expression of (4) derived from (43) and (46):

$$\begin{cases} X_i^{\pi^*}(t) = M_i^{\pi^*}(t) - S_i^{\pi^*}(t) + X_i^{\pi^*}(0), \\ X_i^{\pi^*}(0) = x_i, \end{cases} \tag{53}$$

we know that $X_i^{\pi^*}(T_{m-1})$ is $\mathcal{G}_{(T_m,t)}$-measurable.

We now introduce a reference model for the purpose to prove the integrability of some related processes. In the model, the inter-arrival time sequence of class i jobs for each $i \in \{1, \ldots, K\}$ is given by $\{u_i^j, j = 1, 2, \ldots\}$ as in (5) and every arrived job is supposed to be permitted into service immediately. In other words, the reference model is the same as the original model except taking the number of servers to be infinity, i.e., $n = \infty$. For convenience, we denote the associated service time sequences by $\{v_i^j, j = 1, 2, \ldots\}$ for $i \in \{1, \ldots, K\}$ and define a multi-parameter filtration by $\mathcal{H} = \{\mathcal{H}_m = \sigma(u_i^1, \ldots, u_i^m; v_i^1, \ldots, v_i^m; i = 1, \ldots, K), m \geq 1\}$. Then, for each $t \geq 0$, we have

$$\begin{aligned} E\left|(S_i^{\pi^*}(t) + 1) - V_{S_i^{\pi^*}(t)+1}\right| &\leq x_i + E(R_i(t) + 1) + n\mu_i ET_{S_i^{\pi^*}(t)+1} \\ &\leq x_i + E(R_i(t) + 1) + n\mu_i \sum_{l=1}^{K} E\left(\sum_{j=1}^{R_i(t)+1} (u_l^j + v_l^j)\right) \\ &= x_i + E(R_i(t) + 1) + n\mu_i \sum_{l=1}^{K} E(u_l^1 + v_l^1) E(R_i(t) + 1) \\ &= x_i + (\lambda_i t + 1) + n\mu_i \sum_{l=1}^{K} (1/\lambda_l + 1/\mu_l)(\lambda_l t + 1) \\ &< \infty, \end{aligned} \tag{54}$$

where the first inequality in (54) follows from the fact that the total number of class i jobs that finish service by time t is less than the summation of the initial number of class i jobs and the total number of arrival class i jobs by time t, (54)–(55), and the fact that $X_i^{\pi^*}(t)$ is bounded by n; The second inequality in (54) follows from the fact that the departure time of the $(S_i^{\pi^*}(t) + 1)$th class i job that finishes service is less than the summation of arrival times of all the $(R_i(t) + 1)$th class i jobs and the total amount of time required to finish serving the number $\sum_{i=1}^{K} (R_i(t) + 1)$ of arrival jobs in the reference model; The first equality in (54) follows from the Wald's identity (see, e.g., Chow and Teicher [35]) since $R_i(t) + 1$ is a $\mathcal{H}$-stopping time with $E(R_i(t) + 1) = \lambda_i t + 1 < \infty$ and $\{u_l^j + v_l^j, j \in \{1, 2, \ldots, \}\}$ for each $l \in \{1, \ldots, K\}$ is an i.i.d. sequence of random variables owing to the independent assumptions among inter-arrival times and service times.

Therefore, we know that $(S_i^{\pi^*}(t) + 1) - V_{S_i^{\pi^*}(t)+1}$ and $V_{S_i^{\pi^*}(t)+1}$ are $L^1(\Omega, \mathcal{F}, P)$-integrable. Hence, for each $m \in \{1, 2, \ldots\}$, $\xi_m I_{\{S_i^{\pi^*}(t)+1 \geq m\}}$ is also $L^1(\Omega, \mathcal{F}, P)$-integrable since

$$E\left(\xi_m I_{\{S_i^{\pi^*}(t)+1 \geq m\}}\right) \leq E\left(\sum_{j=1}^{\infty} \xi_j I_{\{S_i^{\pi^*}(t)+1 \geq j\}}\right) = EV_{S_i^{\pi^*}(t)+1} < \infty.$$

Furthermore, since $(X_i^{\pi^*}(T_{m-1}), Y_m)$ is $\mathcal{G}_{(T_m,t)}$-measurable, we know that the following conditional expectations are well defined for $m \in \{1, 2, \ldots\}$ and $k, l \in \{0, 1, \ldots\}$ (see, e.g., p. 8 in Yong and Zhou [30])

$$E\left[\xi_m I_{\{S_i^{\pi^*}(t)+1 \geq m\}} \middle| X_i^{\pi^*}(T_{m-1}) = k, Y_m = l\right]. \tag{55}$$

Next, by Lemma 5.1 of Dai and Feng [3], we further divide the interval $[T_{m-1}, T_m)$ into sub-intervals $\{[T_m^p, T_m^{p+1}), p = 0, 1, \ldots, Y_m\}$ with $T_m^0 = T_{m-1}$ and $T_m^{Y_m+1} = T_m$ and at each time T_m^p for $p = 1, \ldots, Y_m$, there is a class i job that gets into service. Thus, we have

$$\eta_m = T_m - T_{m-1} = \sum_{p=0}^{Y_m} \left(T_m^{p+1} - T_m^p\right). \tag{56}$$

For convenience of notation, we use I_m to denote $I_{\{S_i^{\pi^*}(t)+1\geq m\}}$. Then, we have

$$\begin{aligned}
E[\xi_m I_m] &= E\left[I_m \sum_{p=0}^{Y_m} \int_{T_m^p}^{T_m^{p+1}} \mu_i X_i^{\pi^*}(s) ds \right] \\
&= E\left[I_m \sum_{p=0}^{Y_m} \mu_i \left(X_i^{\pi^*}(T_{m-1}) + p\right) \left(T_m^{p+1} - T_m^p\right) \right] \\
&= \sum_{l=0}^{\infty} \sum_{k=1}^{n} \mu_i E\left[I_m \left(k\left(T_m - T_{m-1}\right) + \sum_{p=1}^{l} \left(T_m - T_m^p\right) \right) \middle| X_i^{\pi^*}(T_{m-1}) = k, Y_m = l \right] \\
&\quad P\{X_i^{\pi^*}(T_{m-1}) = k, Y_m = l\}.
\end{aligned}$$

Remember that there is only one job that finishes service at time T_m (see Lemma 5.1 and its associated Remark 5.1 in Dai and Feng [3]). Thus $T_m - T_{m-1}$ is the time duration during which a job that is in the channel at time T_{m-1} is served at rate μ_i. Let W_m^j for $j \in \{1, \ldots, k\}$ denote the associated remaining service time after T_m, that is, the job that is in the channel at time T_{m-1} will finish service at time $T_m + W_m^j$. Note that both W_m^j and $T_m - T_{m-1} + W_m^j$ are exponentially distributed with rate μ_i. Similarly, $T_m - T_m^p$ for each $p \in \{1, \ldots, l\}$ is the time duration during which the job arrived at T_m^p is served at the rate μ_i. Let W_m^j for $j = k + p$ denote the associated remaining service time, i.e., the job arrived at T_m^p will finish service at $T_m + W_m^p$. Both W_m^j and $T_m - T_m^p + W_m^j$ are exponentially distributed with rate μ_i. Then, for $k \in \{1, \ldots, n\}$ and $l \in \{0, 1, \ldots\}$, we have

$$\begin{aligned}
&E\left[I_m \left(k\left(T_m - T_{m-1}\right) + \sum_{p=1}^{l} \left(T_m - T_m^p\right) \right) \middle| X_i^{\pi^*}(T_{m-1}) = k, Y_m = l \right] \\
&= \left(\sum_{j=1}^{k+l} E\left(T_m - T_m^l + W_m^j - W_m^j\right) \right) E\left(I_m | X_i^{\pi^*}(T_{m-1}) = k, Y_m = l\right) \\
&= \left(\sum_{j=1}^{k+l} E\left(I_{\{W_m^j=0\}} \left(T_m - T_m^l\right)\right) \right) E\left(I_m | X_i^{\pi^*}(T_{m-1}) = k, Y_m = l\right) \\
&= \frac{1}{\mu_i} E\left(I_m | X_i^{\pi^*}(T_{m-1}) = k, Y_m = l\right),
\end{aligned} \tag{57}$$

where in the first equality of (57), we have used the fact that $\{S_i^{\pi^*}(t) + 1 \geq m\} = \{T_{m-1} \leq t\}$ which is only dependent on $T_0, T_1, \ldots, T_{m-1}$ and independent of

$$\sum_{j=1}^{k} \left(T_m - T_{m-1} + W_m^j - W_m^j\right) + \sum_{p=1}^{l} \left(T_m - T_m^p + W_m^{k+p} - W_m^{k+p}\right),$$

and the fact that $T_m - T_{m-1} + W_m^j - W_m^j$ for each $j \in \{1, \ldots, k\}$ and $\{T_m - T_m^p + W_m^{k+p} - W_m^{k+p}\}$ for each $p \in \{1, \ldots, m\}$ are independent identically distributed (i.i.d.); in the second equality of (57), we have used the following fact for the equality:

$$E\left(I_{\{W_m^j \neq 0\}}\left(T_m - T_m^l + W_m^j - W_m^j\right)\right) = 0.$$

In the third equality, we have used the i.i.d property as explained for the first equality. Thus, we have

$$E[\xi_m I_m] = EI_m. \tag{58}$$

Furthermore, by (58), we can calculate I defined in (47) by

$$\begin{aligned} I &= E\left[\left(S_i^{\pi^*}(t) + 1\right) - V_{S_i^{\pi^*}(t)+1}\right] \\ &= \sum_{m=1}^{\infty} E[I_m] - \sum_{m=1}^{\infty} E\left[\xi_m I_m\right] \\ &= 0. \end{aligned}$$

Finally, we show that $II = 1$ claimed in (47). In fact, for each $m \geq 1$, we divide the interval $[t, T_m)$ into sub-intervals $\{[T_m^p, T_m^{p+1}), p = 0, 1, \ldots, Y_m\}$ with $T_m^0 = t$ and $T_m^{Y_m+1} = T_m$, which indicates that at each time point T_m^p for $p = 1, \ldots, Y_m$, there is a class i job that gets into service. Furthermore, define $\bar{I}_m = I_{\{S_i^{\pi^*}(t)+1=m\}}$ for each $m \in \{1, 2, \ldots\}$. Note that $\{S_i^{\pi^*}(t) + 1 = m\} = \{S_i^{\pi^*}(t) + 1 \geq m\} - \{S_i^{\pi^*}(t) + 1 \geq m + 1\} \in \mathcal{G}_t$ is independent of the exponentially distributed service times jobs experienced after time t. Then, by the similar discussion as before, we have

$$\begin{aligned} E\left(\int_t^{T_{S_i^{\pi^*}(t)+1}} \mu_i X_i^{\pi^*}(s)ds\right) &= \sum_{m=1}^{\infty} E\left(\bar{I}_m \sum_{p=0}^{Y_m} \mu_i \left(X_i^{\pi^*}(t) + p\right)\left(T_m^{p+1} - T_m^p\right)\right) \\ &= \sum_{m=1}^{\infty} E\bar{I}_m \\ &= 1. \end{aligned}$$

By the facts as shown above that $I = 0$ and $II = 1$, we have proved the claim given in (42) true. Hence, we complete the proof of Lemma 5.2. □

Lemma 5.3 *For each fixed $i \in \{1, \ldots, K\}$, we have*

$$E\int_0^{\infty} e^{-\alpha t}\left(S_i^{\pi^*}(t) - \mu_i \int_0^t m_i^{\pi^*}(s)ds\right) dt = 0. \tag{59}$$

Proof. Note that $S_i^{\pi^*}(\cdot)$ is measurable on $[0, \infty) \times \Omega$ since it has right-continuous sample paths. Then, by the fact that $S_i^{\pi^*}(t) \leq R_i(t)$ for all $t \geq 0$, we have

$$E\int_0^{\infty} e^{-\alpha t} S_i^{\pi^*}(t)dt \leq E\int_0^{\infty} e^{-\alpha t} R_i(t)dt < \infty,$$

that is, $e^{-\alpha t}S_i^{\pi^*}(t)$ is integrable on $[0,\infty)\times\Omega$. Furthermore, since $m_i^{\pi^*}(s)\leq n$, it follows from Tonelli Theorem that (59) is true. Thus, we complete the proof of Lemma 5.3. □

5.2. *Proof of Theorem 2.1*

Proof. First, since $N_i^{\pi^*}(t)$ for each $i\in\{1,\ldots,K\}$ is nondecreasing in t along each sample path and the total number of lost jobs cannot exceed the total number of arrived jobs over any time interval $[s,t]$ with $t\geq s\geq 0$, we have

$$z_i^{\pi^*}(t)-z_i^{\pi^*}(s)\leq\lambda_i(t-s), \tag{60}$$

where $z_i^*(t)=E\left[N_i^*(t)\right]$. Thus, by (4), (7), Lemma 5.2, we have, $m^{\pi^*}(\cdot)\in\mathcal{U}_K[0,\infty)$ and $z^{\pi^*}(\cdot)\in\mathcal{N}_K[0,\infty)$.

Second, by the integration by parts theorem for the Riemann–Stieltjes integral, we a.s. have

$$\begin{aligned}\int_0^\infty e^{-\alpha t}dN_i^{\pi^*}(t)&=e^{-\alpha t}N_i^{\pi^*}(t)\Big|_0^\infty-\int_0^\infty N_i^{\pi^*}(t)(-\alpha)e^{-\alpha t}dt\\&=\int_0^\infty\alpha N_i^{\pi^*}(t)e^{-\alpha t}dt,\end{aligned} \tag{61}$$

where in the second equation, we have used the facts that $N_i^{\pi^*}(t)\leq R_i(t)$ and $R_i(t)$ has the same order of t (denoted by $O(t)$) a.s. for t large enough since $R_i(t)$ is a Poisson process (see, e.g., the functional law of the iterative logarithm presented in Theorem 5.13 of Chen and Yao [36]). Then, it follows from (6), (61), and Lemma 5.3 that the expression given in (8) is true. Hence, we complete the proof of Theorem 2.1. □

6. Proof of Theorem 2.2

First, let

$$J^*\equiv\sup_{(m,z)\in C_K^{(0)}[0,\infty)\times C_K^{(0)}[0,\infty)\ \text{satisfying (17)}}J(m). \tag{62}$$

Then, there exists a sequence of $\{(y_u,m_u,z_u),u\in\{1,2,\ldots\}\}\subset C_K^{(1)}[0,\infty)\times C_K^{(0)}[0,\infty)\times C_K^{(0)}[0,\infty)$ satisfying constraints (17) such that

$$J^*=\lim_{u\to\infty}J(y_u,m_u),$$

where we endow the $3K$-vector space $C_K^{(1)}[0,\infty)\times C_K^{(0)}[0,\infty)\times C_K^{(0)}[0,\infty)$ with the topology given by (15) (we here employ y_u,m_u,z_u for each $u\in\{1,2,\ldots\}$ to denote K-vector functions rather than components as used in other places of the paper). Thus, we claim that $\{(y_u,m_u,z_u),u=1,2,\ldots\}$ is relatively compact in $C_K^{(1)}[0,\infty)\times C_K^{(0)}[0,\infty)\times C_K^{(0)}[0,\infty)$, and owing to the constraints in (17), it only needs to be shown that $\{(y_u,z_u),u\in\{1,2,\ldots\}\}$ is relatively compact in $C_K^{(1)}[0,\infty)\times C_K^{(0)}[0,\infty)$.

Let $\beta_u(t) \equiv (y_u(t), \dot{y}_u(t), z_u(t))$ for all $u \in \{1, 2, ...\}$ and all $t \in [0, \infty)$. Owing to the constraints in (17), and particularly, the constraint in (7), $\{\beta_u, u \in \{1, 2, ...\}\}$ is equicontinuous over all finite interval $[0, N]$ with $N \in \{1, 2, ...\}$. Thus, for $N = 1$ and by Ascoli–Arzela Theorem (see, for example, Royden [25]), there exists a uniformly convergent subsequence β_{u_v} of β_u over $[0, 1]$, i.e., $\beta_{u_v}(t) \to \beta(t)$ uniformly for $t \in [0, 1]$ as $v \to \infty$, where $\beta(t)$ is a continuous function over $[0, 1]$. For convenience of notation, we still use β_u to denote this subsequence. Next, repeating the above procedure, we can obtain a further subsequence of β_{u_v} such that it converges uniformly to $\beta(t)$ over $[0, 2]$, ..., and continuing this procedure along $N \to \infty$, we can obtain a subsequence of the original $\beta_u(t)$ such that it converges uniformly on compact subset of $[0, T]$ to some $\beta(t)$. As before, we still use β_u to denote this subsequence and define $\beta \equiv (\xi_0, \xi_1, \xi_2)$. Hence, y_u is a Cauchy sequence in $C_K^{(1)}[0, \infty)$. Note that $C_K^{(1)}[0, \infty)$ is complete (see, e.g., the related discussion in p. 325 of Jacod and Shiryaev [37]), and $\beta_u \to \beta$ uniformly on compact set of $[0, \infty)$. Thus, we have that $y \in C_K^{(1)}[0, \infty)$ with $y = \xi_0, \dot{y} = \xi_1$. Furthermore, we have that (ξ_0, ξ_1, ξ_2) satisfies the constraints in (17).

Next, we show that $J(y, m, t) = J^*$ with $m = \dot{y}$. Since $y, y_u \in C_K^{(1)}[0, \infty)$ and satisfy the constraints in (17), we know that there is a positive constant c such that

$$\| \beta(t) \| \leq c(t + 1), \tag{63}$$

$$\| \beta_u(t) \| \leq c(t + 1) \quad \text{for all} \quad u \in \{1, 2, ...\}. \tag{64}$$

Furthermore, note that $\beta_u \to \beta$ uniformly on compact set of $[0, \infty)$. Then, it follows from the dominated convergence theorem,

$$\begin{aligned} |J^* - J(y, m, t)| &\leq \kappa \lim_{u \to \infty} \int_0^\infty e^{-\alpha t} (\| y_u(t) - y(t) \| + \| m_u(t) - m(t) \|) dt \\ &= 0, \end{aligned}$$

where κ is some nonnegative constant. Thus, $(y^*, m^*, z^*) \equiv (y, m, z) \in C_K^{(1)}[0, \infty) \times C_K^{(0)}[0, \infty) \times C_K^{(0)}[0, \infty)$ is a maximal point for the OCMC problem in (16)–(17). Therefore, it follows from (63) to (64) that $J^* < \infty$.

Finally, owing to Theorem 2.1, the pair $(m^{\pi^*}(\cdot), z^{\pi^*}(\cdot))$ corresponding to an optimal Markovian policy π^* satisfies the constraints in (13)–(14). Thus, $(m^{\pi^*}(\cdot), z^{\pi^*}(\cdot))$ is a feasible solution to the OCMC problem in (12)–(14) (and hence, the problem in (16)–(17)). Therefore, the optimal value to our OCMC model provides an upper bound to the corresponding values of the physical queueing system under Markovian decision rules. Concerning this point, readers are also referred to Section 4 in Bäuerle [14] for related discussion. Hence, we complete the proof of the theorem. □

7. Conclusion

In this research, a parallel-server loss channel serving multi-class jobs is studied. An OCMC model over infinite time horizon with non-smooth monotonicity and control-state inequality constraints is established and justified by using the physical queueing model with linear revenue function. Existence of a solution to the OCMC model is proved, whose optimal value provides an upper bound of the corresponding values of physical queueing model under Markovian decision rules. Algorithms with lower complexity in solving the OCMC model are proposed, which are further used to design an admission control policy for the loss channel. Furthermore, a simulation algorithm is proposed to implement the designed policy. Performance comparisons through numerical examples are conducted among our newly designed policy, the FIFO policy, the AS policy, and the MDP based threshold policy. Advantages and disadvantages of these policies are identified under different channel parameters and channel (e.g., Markovian and non-Markovian) conditions. In particular, we find out that our designed policy outperforms the other three policies when the traffic intensity is relatively large, and the differences of the revenues per unit of time and the penalty costs among different classes of jobs are large.

Finally, we note that the discussion in the paper is focused on the case that the revenue function is linear. The related discussion should be interesting and open for the case that the revenue function is nonlinear. Hence, at this moment, an MDP based method is still our major tool in solving such an admission control problem with nonlinear revenue function (see, e.g., the section of Concluding Remarks and Future Research in Dai and Feng [3]).

Acknowledgement

This project is supported by National Natural Science Foundation of China under grant No. 10971249.

References

[1] S. Bodas, S. Shakkottai, L. Ying, and R. Srikant, *Scheduling in Multi-channel Wireless Networks: Rate Function Optimality in the Small-buffer Regime*, Proceedings of ACM Sigmetrics, 2009, pp. 121–132.
[2] W. Dai, *Constrained Optimal Control for a Parallel-server Channel Serving Multi-class Jobs in Zero-buffer Regime*, Proceedings of 2011 World Congress of Engineering and Technology, Vol. 1, IEEE Computer Society Press, 2011, pp. 249–254.
[3] W. Dai and Y. Feng, *Stochastic optimal controls for parallel-server channels with zero waiting buffer capacity and multiclass customers*, Adv. Syst. Sci. Appl., 9 (4) (2009), pp. 627–646.
[4] N. Gans, G. Koole, and A. Mandelbaum, *Telephone call centers: Tutorial, review and research prospects*, Manufact. Service Oper. Manage. 5 (2) (2003), pp. 79–141.
[5] J.M. Harrison, *Heavy traffic analysis of a system with parallel servers: Asymptotic optimality of discrete-review policies*, Ann. Appl. Probability 8 (3) (1998), pp. 822–848.
[6] E.L. Ormeci, A. Burnetas, and J.V.D. Wal, *Admission policies for a two class loss system*, Stochastic Models 17 (4) (2001), pp. 513–539.
[7] E.L. Ormeci and J.V.D. Wal, *Admission policies for a two class loss system with general interaarival times*, Stochastic Models 22 (1) (2006), pp. 37–53.
[8] W. Zhang, X.G. Xia, and K.B. Letaief, *Space-time/frequency coding for MIMO-OFDM in next generation broadband wireless systems*, IEEE Wireless Commun. 14 (3) (2007), pp. 32–43.
[9] W. Forum. *Mobile WiMAX Part I: A Technical Overview and Performance Evaluation*, March 2006. White Paper.
[10] A. Bassamboo, J.M. Harrison, and A. Zeevi, *Dynamic routing and admission control in high-volume service systems: Asymptotic analysis via multi-scale fluid limits*, Queueing Syst. 52 (2006), pp. 249–285.
[11] A. Bassamboo, J.M. Harrison, and A. Zeevi, *Point stationary fluid models for stochastic processing networks*, Manufact. Service Oper. Manage. 11 (2009), pp. 70–81.
[12] E. Altman, *Constrained Markov Decision Processes*, Chapman & Hall/CRC, London, 1999.
[13] S. Lippman, *Applying a new device in the optimization of exponential queueing systems*, Oper. Res. 23 (1975), pp. 687–710.
[14] N.Bäuerle, *Optimal control of queueing networks: An approach via fluid models*, Adv. Appl. Prob. 34 (2002), pp. 313–328.
[15] J.G. Dai, *On positive Harris recurrence of multiclass queueing networks: A unified approach via fluid limit models*, Ann. Appl. Prob. 5 (1995), pp. 49–77.

[16] W. Dai, *Diffusion approximations for multiclass queueing networks under preemptive priority service discipline*, Appl. Math. Mech. 28 (10) (2007), pp. 1331–1342.
[17] X. Luo and D. Bertsimas, *A new algorithm for state-constrained separated continuous linear programming*, SIAM J. Control Optim. 37 (1998), pp. 177–210.
[18] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, Berlin, 1975.
[19] M. Hellwig, *A maximum principle for control problems with monotonicity constraints*, Preprints Max Planck Inst. Res. Collective Goods 4 (2008), pp 1–36.
[20] B. Kugelmann and H.J. Pesch, *A new general guidance method in constrained optimal control, Part 1: Numerical method*, J. Optim. Theory Appl. 67 (3) (1990), pp. 421–435.
[21] B. Kugelmann and H.J. Pesch, *New generation guidance method in constrained optimal control, Part 2: Application to space shuttle guidance*, J. Optim. Theory Appl. 67 (3) (1990), pp. 437–446.
[22] X. Ruiz del Portal, *Non-smooth monotonicity constraints in optimal control problems: Some economic applications*, Optim. Control Appl. Meth. 32 (2011), pp. 396–413.
[23] P.N.V. Tu, *Introductory Optimization Dynamics: Optimal Control with Economics and Management Science Applications*, Springer-Verlag, Berlin, 1984.
[24] L. Sennott, *Average reward optimization theory for denumerable state spaces*, in *Handbook of Markov Decision Processes Methods and Applications: International Series in Operations Research and Management Science*, E.A. Feinberg and A. Schwartz, eds., Kluwer Academic, Norwell, MA, 2002, pp. 153–172.
[25] H.L. Royden, *Real Analysis*, 3rd ed., Macmillan Publishing Company, New York, 1988.
[26] R.T. Rockafeller, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[27] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., Springer, New York, 2006.
[28] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica 4 (4) (1984) pp. 373–395.
[29] N. Megiddo, On the complexity of linear programming, in *Advances in Economic Theory*: *5th World Congress*, T. Bewley, ed., Cambridge University Press, 1987, pp. 225–268.
[30] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.
[31] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986.
[32] K.L. Chung, *Lectures from Markov Processes to Brownian Motion*, Springer-Verlag, New York, 1982.
[33] W. Dai, *Brownian approximations for queueing networks with finite buffers: Modeling, heavy traffic analysis and numerical implementations*, Ph.D. thesis, Georgia Institute of Technology, 1996. [Also published in UMI Dissertation Services, A Bell & Howell Company, 300 N. Zeeb Road, Ann Arbor, Michican 48106, U.S.A., 1997.]
[34] J.G. Dai and W. Dai, *A heavy traffic limit theorem for a class of open queueing networks with finite buffers*, Queueing Syst. 32 (1999), pp. 5–40.
[35] Y.S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability and Martingales*, 2nd ed., Springer-Verlag, New York, 1988.
[36] H. Chen and D.D. Yao, *Fundamentals of Queueing Networks*, Springer-Verlag, New York, 2001.
[37] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed., Springer-Verlag, Berlin, 2002.