

Asymptotically optimal parallel resource assignment with interference

I.M. Verloop^{1,2} , R. Núñez-Queija^{2,3}

¹BCAM – Basque Center for Applied Mathematics, Derio, Spain

²CWI, Amsterdam, The Netherlands

³University of Amsterdam, The Netherlands

Abstract

Motivated by scheduling in cellular wireless networks and resource allocation in computer systems, we study a service facility with two classes of users having heterogeneous service requirement distributions. The aggregate service capacity is assumed to be largest when both classes are served in parallel, but giving preferential treatment to one of the classes may be advantageous when aiming at minimization of the number of users, or when classes have different economic values, for example.

We set out to determine the allocation policies that minimize the total number of users in the system. For some particular cases we can determine the optimal policy exactly, but in general this is not analytically feasible. We then study the optimal policies in the fluid regime, which prove to be close to optimal in the original stochastic model. These policies can be characterized by either linear or exponential switching curves. We numerically compare our results with existing approximations based on optimization in the heavy-traffic regime. By simulations we show that, in general, our simple computable switching-curve strategies based on the fluid analysis perform well.

1 Introduction

In many practical applications where resources must be allocated to several contending users or tasks, the service capacity itself may be affected by the scheduling policy deployed. Our work is motivated by two specific application areas. In third generation wireless networks, neighboring base stations may interfere with each other when transmitting simultaneously. When one base station is not active, other base stations can work at higher rates, see for example [7, 8]. For data applications, base stations may coordinate transmissions (i.e., transmit simultaneously or alternatingly) so as to improve the use of the shared spectrum. A second motivating application is the scheduling of resources in computer systems (or Web servers) where jobs must be routed to one of several servers, see for example [27, 28]. There, the capacity depends on the allocation when servers are specialized for certain tasks.

Scheduling of resources with policy-dependent capacities has attracted much attention in recent years. Most of the results concern stochastic stability properties of such systems. Due to the dependence of capacity on the service policy, even this most basic performance measure is a non-trivial task to determine. In [11] bounds for stability in a general class of systems with policy-dependent capacity have been determined. In the specific context of wireless networking, stability of utility-based allocation strategies was shown to be intimately related with the shape of the feasible capacity region [10], i.e., the set of simultaneously achievable transmission rates for all users. With a convex capacity region, the system is stabilized by any such allocation

strategy, but this is not the case for non-convex capacity regions. These results were later generalized to non-convex and time-varying capacity regions in [22], showing the precise conditions for stability of utility-based strategies under quite general assumptions on the time-variations. Stability conditions for non utility-based strategies, for example threshold-based policies, were investigated in [28, 34].

As may be expected from the complexity of determining stability, results on the flow level *performance* in terms of system delay or system occupancy are scarce. In this paper, we focus on a particular model with simultaneous resource sharing that turns out to be equivalent to a parallel-server model where user classes can be served in parallel, all by a dedicated server, or where several servers can be simultaneously allocated to one class only. This type of models is known to be notoriously hard to analyze, as is illustrated by special cases (including the so-called coupled-processors model) requiring the solution of a Riemann-Hilbert boundary value problem [13, 16].

Most results on flow-level performance in parallel-server models concentrate on a specific class of scheduling policies. For example, besides determining the stability conditions, the authors in [28] investigate the performance of threshold-based policies. One main observation there is that finding reasonable values for the thresholds is not trivial since performance as well as stability can be quite sensitive to the threshold values. Approximations for mean response times are given in [27]. A general class of threshold-based priority policies for multi-class parallel-server networks is also proposed in [33]. For these strategies, the authors derive approximate formulas for the queue lengths and illustrate how these can be used to obtain reasonable threshold values. In [7, 8] a parallel two-server model is analyzed under the policy that always serves both classes in parallel whenever both are present, and a diffusion approximation for the queue lengths is found for a specific heavy-traffic setting.

Our goals here are to study the structural properties of optimal scheduling policies in a parallel-server model, and to determine computable approximations that are close to optimality. Our objective is to minimize (in some appropriate sense) the total number of users. A crucial observation when addressing optimality is that, in general, users will have class-specific sizes, so that few users of one class can typically add up to the same amount of work as many of another class. On one hand, it seems reasonable to maximize the departure rate of users, by serving the “small” users first. In the short run, this will keep the number of users in the system at a low level, thus shortening overall delays. On the other hand, it is also desirable to deploy the highest possible total service capacity. That will minimize the volume of back-logged work and drain the system at maximum rate, thus ensuring maximum stability. In general, finding the optimal trade-off between these two intrinsically different objectives is a challenging task.

Determining the exact optimal policy in a parallel-server model has so far proved analytically infeasible. Most research on this area has focused on heavily-loaded systems under a (complete) resource pooling condition for which asymptotically optimal policies in heavy traffic are determined [1, 5, 6, 19, 20, 24, 32]. In [1, 19, 20] several discrete-review policies are proposed (the system is reviewed at discrete points in time, and decisions are based on the queue lengths at the revision moment) and are proved to be asymptotically optimal in heavy traffic. In [24, 32] a generalized $c\mu$ -rule is proposed (including the Max-Weight policy as a special case) that myopically maximizes the rate of decrease of certain instantaneous holding cost. This policy is robust in the sense that it only depends on the departure rates and the cost function, and it is proved that this policy minimizes the cumulative cost over any finite interval in a heavily-loaded system. In [5, 6], the authors prove that threshold-based strategies minimize the scaled total number of users in a heavy-traffic setting. The order of magnitude of the optimal thresholds as functions of the traffic load can be determined, but this does not give a recipe to choose good threshold values in moderately-loaded regimes. In [33], the authors propose values for the threshold, which can be found by solving a minimization problem.

In this paper, we consider a parallel two-server model with two traffic classes that can be served either in parallel or alternately. The highest service capacity is achieved when serving both classes in parallel, but with asymmetric service requirements, the user departure rate may be larger when serving one class only. For some special cases the optimal policy can be determined exactly, but this is not possible in general. In a similar setting, [4] states that switching-curve policies are optimal (a proof will be included in a forthcoming paper by the authors of [4]). Numerical experiments included for illustration in the present paper indeed support this optimality. In order to find computable approximations for the optimal policies we study the model in a fluid-limit regime for which we show that the optimal policy is characterized by a linear switching curve. The optimal switching curves in the *fluid regime* can be used to determine asymptotically fluid optimal policies for the *stochastic model*. These policies are characterized by either linear or exponential switching curves. Our analysis is inspired by that in [17, 18] where a multi-class tandem-network is studied. By simulations we compare these asymptotically fluid optimal switching-curve policies with threshold-based policies [5, 6] and Max-Weight policies [24, 32] which are optimal in heavy traffic. We show that the fluid-based and threshold-based policies give good performance in general, while significant improvements over Max-Weight policies can be achieved.

It is worth noting that the optimal policies studied in this paper rely on centralized control. In practice, centralized control may require a prohibitive amount of overhead. However, knowledge of the (centralized) optimum is extremely valuable to (numerically) estimate the scope for improvement of decentralized control policies. For example, in the application area of bandwidth-sharing networks, it was found numerically that certain distributed schemes may actually be close to the theoretical (centralized) optimum [35, 36].

The paper is organized as follows. In Section 2 we describe the model and state some preliminary results. Section 3 contains our optimality results for the stochastic model. The fluid analysis and the asymptotically fluid optimal policies are presented in Section 4. For comparison we briefly discuss optimal policies in heavy traffic using the results of [5, 6] and [24, 32] in Section 5. Numerical experiments and concluding remarks can be found in Sections 6 and 7.

2 Model description and preliminaries

We consider the following model. There are two classes of users. Class- i users, $i = 1, 2$, arrive according to independent Poisson processes with rate λ_i and have exponentially distributed service requirements with mean $1/\mu_i$, $i = 1, 2$. Without loss of generality, we assume throughout the paper that $\mu_1 \geq \mu_2$, that is the service requirements of class-1 users are relatively small. Define the traffic load of class i as $\rho_i := \frac{\lambda_i}{\mu_i}$. At any time, either one class can be served individually with capacity 1, or both classes 1 and 2 can be served in parallel with capacities c_1 and c_2 respectively, $c_i \leq 1$, or the system is idling (not serving any class), or any convex combination of these four.

For a given policy π , denote by $s_i^\pi(t)$ the service capacity devoted to class i at time t . We assume that $s_i^\pi(t) = 0$ when $N_i(t) = 0$. In addition, we assume the process $s_i^\pi(t)$ to be right continuous with left limits. The vector $s^\pi(t) = (s_1^\pi(t), s_2^\pi(t))$ lies in the capacity region S , which is defined as the convex hull of the

set $\{(0, 0), (1, 0), (0, 1), (c_1, c_2)\}$ (see Figure 1 in the case $c_1 + c_2 > 1$). Note that the total (service) capacity $s_1^\pi(t) + s_2^\pi(t)$, that is, the speed at which the total amount of backlogged work in the system decreases, is not constant in time. Depending on the decision taken at time t , it may vary between 0 and $\max(1, c_1 + c_2)$. The rate at which users leave the system is $\mu_1 s_1^\pi(t) + \mu_2 s_2^\pi(t)$, which we refer to as the (user) departure rate.

Let $S_i^\pi(t) := \int_0^t s_i^\pi(u) du$ denote the cumulative amount of capacity devoted to class i in the time interval $(0, t)$ under policy π . Let $A_i(u, t)$ be the amount of class- i work that arrived during the

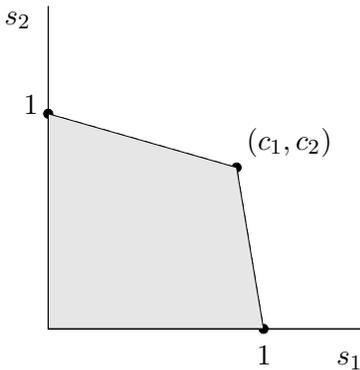


Figure 1: Capacity region S when $c_1 + c_2 > 1$.

time interval $(u, t]$. Then, the workload in class i at time t can be written as

$$W_i^\pi(t) := W_i(0) + A_i(0, t) - S_i^\pi(t). \quad (1)$$

Denote by $N_i^\pi(t)$ the number of class- i users at time t , and let $N^\pi(t) = (N_1^\pi(t), N_2^\pi(t))$. We further define N_i^π and N^π as random variables with the corresponding equilibrium distributions (when they exist).

Remark 2.1 *With resource allocation in computer systems in mind, it is more natural to view the model as an equivalent parallel-server model with two servers and two classes, as depicted in Figure 2. Server 1 can either serve class 1 with capacity c_1 , or class 2 with capacity $1 - c_1$. Similarly, server 2 can either serve class 2 with capacity c_2 , or class 1 with capacity $1 - c_2$. Hence, when the two servers are dedicated to their own classes, classes 1 and 2 are served in parallel with capacities c_1 and c_2 , respectively. When instead both servers are allocated to the same class, this class is served with capacity 1. (In our setting, both servers can work together on one single user, thus achieving a service capacity of 1 even when there is only one user in the system.) Note that, although uncommon in this setting, it is no restriction to require that the service capacity obtained by combining the two servers equals 1 irrespective of the queue being served. In fact, this can be achieved for any parallel-server model by normalizing the service requirements.¹*

For any point in time, one needs to decide how the service capacity should be divided between the two classes. The objective of the paper is to identify scheduling policies that in some appropriate sense minimize the total number of users in the system. We focus on policies that only use knowledge of the past evolution of the number of users. Since the service requirements and inter-arrival times are exponentially distributed, the Markov property implies that we only need to consider policies that base decisions on the number of users present in the various classes. In particular, we exclude anticipating policies, i.e., policies that have knowledge available of the remaining service requirements. The set of these Markovian non-anticipating policies is denoted by Π . We call a policy $\tilde{\pi} \in \Pi$ average optimal when $\tilde{\pi} = \operatorname{argmin}_{\pi \in \Pi} \limsup_{m \rightarrow \infty} \frac{1}{m} \mathbb{E}(\int_0^m (N_1^\pi(t) + N_2^\pi(t)) dt)$. A policy $\tilde{\pi} \in \Pi$ is stochastically optimal when $N_1^{\tilde{\pi}}(t) + N_2^{\tilde{\pi}}(t) \leq_{st} N_1^\pi(t) + N_2^\pi(t)$, for

¹One may think of μ_i to be the user departure rate of class i when served exclusively (with normalized service capacity 1). Then c_1 and c_2 may be adjusted so that $\mu_i c_i$ equals the user departure rate of class i when the two classes are served simultaneously. To be specific, consider a parallel two-server model where \tilde{C}_i is the service capacity in queue i when allocated both servers, \tilde{c}_1 and \tilde{c}_2 are the service capacities in both queues under parallel service, and $1/\tilde{\mu}_1$ and $1/\tilde{\mu}_2$ are the mean service requirements. The queue length process is then equivalent with our normalized system when setting $\mu_i = \tilde{\mu}_i \tilde{C}_i$ and $c_i = \tilde{c}_i / \tilde{C}_i$.

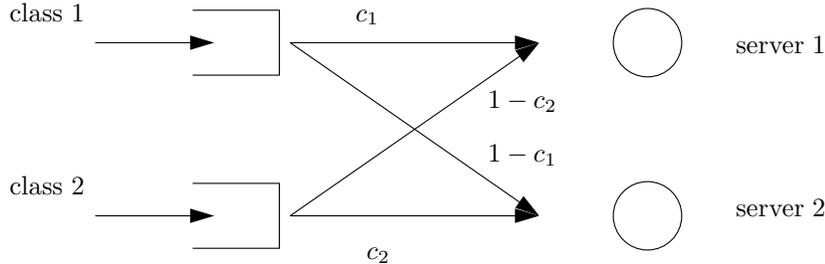


Figure 2: Parallel two-server model.

all $t \geq 0$, $\pi \in \Pi$, whenever $N^{\tilde{\pi}}(0) = N^{\pi}(0)$. By definition, for two positive random variables X and Y , we use $X \leq_{st} Y$ to denote that $\mathbb{P}(X > s) \leq \mathbb{P}(Y > s)$ for all $s \geq 0$. A stochastically optimal policy, if it exists, is automatically average optimal as well.

In the paper we assume $c_1 + c_2 > 1$. However, before proceeding let us briefly consider the situation $c_1 + c_2 \leq 1$. In the latter case, the policy that gives preemptive priority to class 1 (the class with the highest departure rate) is stochastically optimal. (In fact, this result holds for any shape of the capacity region where the points $(1, 0)$ and $(0, 1)$ are not dominated by any other element in the capacity region.) Intuitively, this can be understood by noting that if $c_1 + c_2 \leq 1$, then serving class 1 exclusively will maximize the rate at which the total workload in the system decreases. At the same time, since $\mu_1 \geq c_1\mu_1 + c_2\mu_2$, serving class 1 myopically maximizes the departure rate. A formal proof can be obtained along the lines of Proposition 3.3 below using dynamic programming. Average optimality is actually rather easy to deduce and we give its proof here: Denote by $\pi^{(1)}$ the policy that gives preemptive priority to class 1. Then for any policy $\pi \in \Pi$, if at time $t = 0$ the workloads satisfy

$$W_1^{\pi^{(1)}}(t) \leq W_1^{\pi}(t), \quad (2)$$

$$W_1^{\pi^{(1)}}(t) + W_2^{\pi^{(1)}}(t) \leq W_1^{\pi}(t) + W_2^{\pi}(t), \quad (3)$$

then the same is true for all $t \geq 0$. These inequalities hold sample-path wise (for all t), and they imply stochastic inequalities for the *workload processes*. Multiplying (2) by $\mu_1 - \mu_2 \geq 0$ and (3) by μ_2 and adding the two inequalities gives that $\mu_1 W_1^{\pi^{(1)}}(t) + \mu_2 W_2^{\pi^{(1)}}(t) \leq \mu_1 W_1^{\pi}(t) + \mu_2 W_2^{\pi}(t)$. Since we have exponentially distributed service requirements and we consider only non-anticipating policies, we obtain $\mathbb{E}(W_i^{\pi}(t)) = \frac{1}{\mu_i} \mathbb{E}(N_i^{\pi}(t))$, so that $\mathbb{E}(N_1^{\pi^{(1)}}(t)) + \mathbb{E}(N_2^{\pi^{(1)}}(t)) \leq \mathbb{E}(N_1^{\pi}(t)) + \mathbb{E}(N_2^{\pi}(t))$, for all $t \geq 0$ and for all policies $\pi \in \Pi$. In particular, policy $\pi^{(1)}$ is average optimal.

As mentioned before, in the remainder of the paper we will focus on the unsolved case $c_1 + c_2 > 1$. In this case, the total service capacity is largest when both classes are served in parallel. For application in wireless networks, this represents the joint capacity when both base stations transmit in parallel, and in computer scheduling it corresponds to dedicated specialized servers.

2.1 Stability

For a given policy π , the system is called stable when the process $N^{\pi}(t)$ is positive recurrent. Since $c_1 + c_2 > 1$, the policy that serves classes 1 and 2 in parallel, whenever possible, minimizes the total workload in the system at every moment in time. Hence, this policy will keep the system stable whenever possible. Under this policy, the model becomes a coupled-processors

model for which the stability conditions are

$$\min\left(\frac{\rho_1}{c_1}, \frac{\rho_2}{c_2}\right) < 1 \quad \text{and} \quad (4)$$

$$\text{if } \frac{\rho_i}{c_i} < 1 \quad \text{then } \rho_j + \frac{\rho_i}{c_i}(1 - c_j) < 1, \quad i \neq j, \quad (5)$$

as proved in [13, 16]. Conditions (4) and (5) are therefore necessary conditions for the system to be stable. However, they do not guarantee stability for an arbitrary policy, and the exact (sufficient and necessary) stability conditions depend strongly on the scheduling policy used. Note that the load vectors (ρ_1, ρ_2) that satisfy the necessary stability conditions (4) and (5), are exactly those vectors that lie in the interior of the capacity region S depicted in Figure 1.

3 Optimality results

For a standard multi-class single-server queue it is well known that if class- i users have exponentially distributed service requirements with mean $1/\mu_i$, for all classes i , then the policy that gives preemptive priority to the class with the highest departure rate μ_i (the so-called μ -rule), is stochastically optimal [30]. The rationale behind this rule is that it maximizes the departure rate at all times. One might expect that such a rule is optimal in our model as well. The μ -rule would amount to choosing the allocation $s(t)$ that maximizes the user departure rate, $\mu_1 s_1(t) + \mu_2 s_2(t)$, at any time t . Unfortunately, the total service capacity, $s_1(t) + s_2(t)$, depends on the chosen allocation as well. For example, serving class i only decreases the total amount of work at rate 1, while serving both classes in parallel gives a decrease of the workload at rate $c_1 + c_2 > 1$. Therefore, the objective to maximize the user departure rate may be conflicting with that of maximizing the total service capacity used. The latter will minimize the total time needed to empty the system, which is advantageous in the long run, while the former is better in the short run.

Recall that we chose $\mu_1 \geq \mu_2$. If, in addition, $\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$, then there is no trade-off and it is intuitively clear that the policy that always serves classes 1 and 2 in parallel (whenever both are backlogged) is optimal, since this maximizes both the workload depletion rate and the departure rate. In Section 3.1 we show that the above described policy is in fact stochastically optimal.

When $\mu_1 \geq \mu_1 c_1 + \mu_2 c_2$, the highest departure rate is obtained when serving class 1 individually. It may therefore be better to sometimes serve class 1 individually, even if that does not maximize the rate at which the total work in the system decreases. Hence as the number of users varies, the system should dynamically switch between different allocations. This setting is included in Section 3.2.

3.1 Stochastic optimality when $\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$

In this section we show that when $(\mu_2 \leq) \mu_1 \leq \mu_1 c_1 + \mu_2 c_2$, the policy that serves both classes in parallel (whenever possible) is stochastically optimal. Although it seems natural to prove this using stochastic coupling techniques, we have not been able to find such a coupling. For that reason we resort to dynamic programming techniques. We choose a framework which is somewhat broader than strictly needed to prove the required stochastic optimality of the number of users (we only need a particular choice of the function $C(\cdot)$ below). Doing so, we emphasize the essential properties needed to prove stochastic optimality.

We consider the uniformized Markov chain, which is equivalent to the original process, see [29, Section 11.5]. In the uniformized chain, the transition epochs (including 'dummy' transitions that do not alter the system state) are generated by a Poisson process of constant rate $\nu =$

$\lambda_1 + \lambda_2 + \mu_1(1 + c_1) + \mu_2(1 + c_2)$. Since ν is finite, we may assume $\nu = 1$ without loss of generality. We then focus on the discrete-time Markov chain embedded at transition epochs and, for transparency of notation, again denote the number of class- i users after k steps by $N_i(k)$, $i = 1, 2$. Let $x = (x_1, x_2) \in \mathbb{Z}_+^2$. We define the functions $V_k(x)$, $k = 0, 1, \dots$, as follows:

$$\begin{aligned}
V_0(x) &= C(x) \\
V_{k+1}(x) &= \lambda_1 V_k(x + e_1) + \lambda_2 V_k(x + e_2) \\
&\quad + \min_{s \in S} \left\{ \sum_{i=1,2} \mathbf{1}_{(x_i > 0)} \mu_i s_i V_k(x - e_i) + (1 - \lambda_1 - \lambda_2 - \sum_{i=1,2} \mathbf{1}_{(x_i > 0)} \mu_i s_i) V_k(x) \right\} \\
&= \lambda_1 V_k(x + e_1) + \lambda_2 V_k(x + e_2) + (\mu_1(1 + c_1) + \mu_2(1 + c_2)) V_k(x) \\
&\quad + \min_{s \in S} \left\{ \sum_{i=1,2} \mathbf{1}_{(x_i > 0)} \mu_i s_i (V_k(x - e_i) - V_k(x)) \right\}, \tag{6}
\end{aligned}$$

for $x_1, x_2 \geq 0$, $k = 0, 1, \dots$, with $C(\cdot) : \mathbb{Z}_+^2 \rightarrow \mathbb{R}$ a terminal cost function, S the capacity region, and e_i the i -th unit vector. The term $V_{k+1}(x)$ represents the minimum achievable expected terminal cost, when the system starts in state x at $k + 1$ steps from the horizon. In addition, a minimizing action in (6) is an optimal action at $k + 1$ steps from the horizon. Setting the cost function in this framework equal to $C(x) = \mathbf{1}_{(x_1 + x_2 > y)}$, we obtain that $V_{k+1}(x)$ represents the minimum achievable value for

$$\mathbb{P}(N_1(k + 1) + N_2(k + 1) > y | N(0) = x).$$

If we then show that for all $y \geq 0$ and all $k \in \{0, 1, \dots\}$, we can choose the same minimizing action in (6) (the optimal action may depend on the state x), then the corresponding policy is stochastically optimal at every instant in time. In the next two lemmas we establish convenient properties of $V_k(\cdot)$, under certain conditions on the function $C(\cdot)$.

Lemma 3.1 *If $C(x)$ is non-decreasing in x_1 and x_2 , then $V_k(x)$ is non-decreasing in x_1 and x_2 for all k .*

Proof: The statement follows directly from the definition of $V_k(\cdot)$. \square

The set S is convex, hence the minimizing action in (6) will be one of the extreme points of S . From Lemma 3.1 it can be concluded that the minimizer will not be $(0, 0) \in S$, since $\sum_{i=1,2} \mathbf{1}_{(x_i > 0)} \mu_i s_i (V_k(x - e_i) - V_k(x)) \leq 0$, for all $s \in S$. Hence, we can rewrite the function $V_{k+1}(\cdot)$ as follows:

$$\begin{aligned}
V_{k+1}(x) &= \lambda_1 V_k(x + e_1) + \lambda_2 V_k(x + e_2) \\
&\quad + \min \left(\mu_1 V_k((x_1 - 1)^+, x_2) + (\mu_2 + \mu_1 c_1 + \mu_2 c_2) V_k(x), \right. \\
&\quad \quad \mu_2 V_k(x_1, (x_2 - 1)^+) + (\mu_1 + \mu_1 c_1 + \mu_2 c_2) V_k(x), \\
&\quad \quad \left. \mu_1 c_1 V_k((x_1 - 1)^+, x_2) + \mu_2 c_2 V_k(x_1, (x_2 - 1)^+) + (\mu_1 + \mu_2) V_k(x) \right). \tag{7}
\end{aligned}$$

In the next lemma we will show that under certain conditions on $C(x)$, the minimizing action in (7) will be to always serve classes 1 and 2 in parallel, whenever possible. The proof uses Lemma 3.1 and may be found in Appendix A.

Lemma 3.2 *If $c_1 + c_2 \geq 1$ and $Z(x) = C(x)$ is non-decreasing in x_1 and x_2 and satisfies*

$$\begin{aligned}
&(\mu_1 + \mu_2) Z(x) + \mu_1 c_1 Z(x - e_1) + \mu_2 c_2 Z(x - e_2) \\
&\leq \min(\mu_1 Z(x - e_1) + (\mu_2 + \mu_1 c_1 + \mu_2 c_2) Z(x), \\
&\quad \mu_2 Z(x - e_1) + (\mu_1 + \mu_1 c_1 + \mu_2 c_2) Z(x)), \tag{8}
\end{aligned}$$

for $x_1, x_2 > 0$, then $Z(x) = V_k(x)$ satisfies (8) as well, for any $k \geq 0$.

We can now find a stochastically optimal policy when $(\mu_2 \leq)\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$.

Proposition 3.3 *Assume $c_1 + c_2 \geq 1$. If $(\mu_2 \leq)\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$, then it is stochastically optimal to serve both classes in parallel whenever possible.*

Proof: If $(\mu_2 \leq)\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$, then the cost function $C(x_1, x_2) = \mathbf{1}_{(x_1+x_2>y)}$ satisfies the conditions as given in Lemma 3.2, for all $y \geq 0$. From Lemma 3.2 we obtain that serving both classes in parallel (whenever possible) is always the minimizing action in (7) and hence the corresponding stationary policy is stochastically optimal. \square

3.2 General characterization of the average-optimal policy

Section 3.1 treats the case $\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$, for which a stochastically optimal policies exist. Since this may in general not be the case, we now discuss the general structure of an average-optimal policy.

When $\mu_1 > \mu_2$, maximizing the user departure rate would imply that an optimal policy will never serve class 2 individually when class 1 is also present. At the same time, serving class 2 individually does not give the highest possible total service capacity either, since $c_1 + c_2 > 1$. Therefore, it is natural that an optimal policy should never serve class 2 individually when there is also work of class 1 present. This fact is proved in Proposition 3.5. First we state a lemma that in fact holds for generally distributed inter-arrival times and service requirements, and in particular, holds irrespective of the values for μ_1 and μ_2 . The proof may be found in Appendix B.

Lemma 3.4 *(This lemma holds for generally distributed inter-arrival times and service requirements.) Assume $c_1 + c_2 > 1$. Let $\tilde{\pi}$ be a policy that sometimes does serve class 2 individually while there are class-1 users present. Define policy π to be the policy that uses the same allocation as $\tilde{\pi}$ when possible, except when policy $\tilde{\pi}$ serves class 2 individually. In that case policy π serves classes 1 and 2 in parallel (if possible).*

Consider the same realizations of the arrival processes and service requirements. Then the following sample-path inequalities hold:

$$S_1^\pi(t) \geq S_1^{\tilde{\pi}}(t) \tag{9}$$

$$S_1^\pi(t) + S_2^\pi(t) \geq S_1^{\tilde{\pi}}(t) + S_2^{\tilde{\pi}}(t) \tag{10}$$

$$(1 - c_2)S_1^\pi(t) + c_1 S_2^\pi(t) \geq (1 - c_2)S_1^{\tilde{\pi}}(t) + c_1 S_2^{\tilde{\pi}}(t), \tag{11}$$

for all $t \geq 0$.

Proposition 3.5 *Assume $\mu_1 \geq \mu_2$ and $c_1 + c_2 > 1$. For any policy $\tilde{\pi}$ that serves class 2 individually when there is work of class 1 present, there exists a modified policy π that never serves class 2 individually when class 1 is present and that does not do worse than $\tilde{\pi}$, i.e.,*

$$\mathbb{E}(N_1^\pi(t) + N_2^\pi(t)) \leq \mathbb{E}(N_1^{\tilde{\pi}}(t) + N_2^{\tilde{\pi}}(t)), \text{ for all } t \geq 0.$$

Proof: Let $\tilde{\pi}$ be a policy that sometimes does serve class 2 individually while there are class-1 users present. Define policy π as in Lemma 3.4 and hence the sample-path inequalities (9) and (10) hold. Multiplying (9) by $\mu_1 - \mu_2 \geq 0$ and (10) by μ_2 and adding the two inequalities gives that $\mu_1 S_1^\pi(t) + \mu_2 S_2^\pi(t) \geq \mu_1 S_1^{\tilde{\pi}}(t) + \mu_2 S_2^{\tilde{\pi}}(t)$ and hence by (1) we obtain

$$\mu_1 W_1^\pi(t) + \mu_2 W_2^\pi(t) \leq \mu_1 W_1^{\tilde{\pi}}(t) + \mu_2 W_2^{\tilde{\pi}}(t), \tag{12}$$

for all $t \geq 0$. Since we assumed exponentially distributed service requirements and we consider only non-anticipating policies, we have $\mathbb{E}(W_i^\pi(t)) = \frac{1}{\mu_i} \mathbb{E}(N_i^\pi(t))$. By taking expectations on

both sides in (12), we obtain $\mathbb{E}(N_1^\pi(t) + N_2^\pi(t)) \leq \mathbb{E}(N_1^{\tilde{\pi}}(t) + N_2^{\tilde{\pi}}(t))$. Hence policy π is not worse than $\tilde{\pi}$ and policy π never serves class 2 individually when there is work of class 1 present. \square

In Section 3.1 we explicitly found a stochastically optimal policy when $\mu_1 \leq \mu_1 c_1 + \mu_2 c_2$. Hence, the remaining interesting case is when $\mu_1 > \mu_1 c_1 + \mu_2 c_2$. Then, a stochastically optimal policy may not exist, due to the fact that there is a tradeoff when users of both classes are present: On one hand serving class 1 individually maximizes the user departure rate since $\mu_1 > \mu_1 c_1 + \mu_2 c_2$, which gives stochastic optimality in the short run. On the other hand, serving classes 1 and 2 simultaneously maximizes the speed at which the total workload in the system decreases. The latter policy would empty the system sooner and, hence, achieve a smaller number of users at the moment that it empties the system, compared to the policy that myopically maximizes the departure rate (which needs more time to completely drain the system). A stochastically optimal policy should achieve the lowest number of users at all times (in the sense of stochastic ordering), which is obviously not the case for the two described strategies.

When seeking an average-optimal policy, by Proposition 3.5 we only need to consider policies that never serve class-2 users individually when there are also class-1 users present. The decision between whether to serve class 1 individually or classes 1 and 2 jointly is determined by the number of class-1 and class-2 users present in the system. Intuitively, one may expect that the optimal policy can be characterized by a switching curve, i.e., there exists a non-decreasing function h such that if $N_2 \geq h(N_1)$, then it is optimal to serve classes 1 and 2 in parallel, and otherwise it is optimal to serve class 1 individually. The authors in [4] state that for a model with slightly different behavior near the boundaries, the existence of such a switching curve can be proved using dynamic programming techniques. We expect that for our model, the existence of a switching curve can be proved using the same technique (see also [35] where this was done for a different model). However, dynamic programming techniques will not provide us with any information concerning the shape of the curve. Therefore, in the remainder of the paper we seek policies that are close to optimal by investigating two limiting regimes. In Section 4 this is done for a fluid scaled system and asymptotically fluid optimal switching curve policies are derived. Optimality results for the heavy-traffic regime are reviewed in Section 5.

4 Fluid analysis and asymptotic fluid optimality

In this section we consider the stochastic queue length processes under a fluid scaling and investigate close to optimal policies for the unsolved case $\mu_1 > c_1 \mu_1 + c_2 \mu_2$. In order to do so, it will be convenient to first study the related deterministic fluid control model. This will be done in Section 4.1. For this relatively simple model we derive an optimal control (which is characterized by a switching curve) and the corresponding optimal trajectory. In Section 4.2 we show that under certain switching curve policies in the stochastic process, the fluid scaled stochastic processes converge to the optimal trajectory as found for the deterministic fluid control model. In addition, we show that these switching curve policies are asymptotically fluid optimal (see Definition 4.9) in the stochastic model.

4.1 Optimal policies for the fluid control model

In this section we focus on the deterministic fluid control model, which arises from the original stochastic model by only taking into account the mean drifts. A fluid process is a solution $n(t) = (n_1(t), n_2(t))$ of the following equations:

$$n_i(t) = n_i + \lambda_i t - U_i(t) \mu_i - U_c(t) \mu_i c_i, \quad i = 1, 2, \quad (13)$$

$$n_i(t) \geq 0, \quad i = 1, 2. \quad (14)$$

Here $n = (n_1, n_2) \in \mathbb{R}_+^2$ and $U_j(t) = \int_0^t u_j(v)dv$, $j = 1, 2, c$, such that for all $v \geq 0$,

$$u_1(v) + u_2(v) + u_c(v) \leq 1, \quad (15)$$

$$u_j(v) \geq 0, \quad j = 1, 2, c, \quad (16)$$

and the functions $u_j(v)$ are measurable, $j = 1, 2, c$. The subscript c refers to “combined service”, i.e., serving both classes in parallel. We refer to $n_i(t)$ as the amount of class- i fluid in the system at time t . Note that $U_j(t)$ is Lipschitz continuous with constant less than or equal to 1, hence is differentiable almost everywhere. Then, $n_i(t)$ is differentiable almost everywhere as well, and

$$\frac{dn_i(t)}{dt} = \lambda_i - u_i(t)\mu_i - u_c(t)\mu_i c_i, \quad i = 1, 2, \quad (17)$$

at regular points (a regular point is a value of t at which $n_i(t)$ is differentiable). Under the stability conditions, the fluid model can be drained in finite time, as is stated in the following lemma.

Lemma 4.1 *If (4) and (5) are satisfied, then the policy that serves classes 1 and 2 in parallel whenever possible, drains the fluid model in finite time and keeps the system empty from that moment on.*

Proof: We consider the workload fluid processes $w_i(t) := \frac{n_i(t)}{\mu_i}$, $i = 1, 2$. From (17) we have $\frac{dw_i(t)}{dt} = \rho_i - u_i(t) - u_c(t)c_i$, $i = 1, 2$, at regular points. Focus on the policy that serves classes 1 and 2 in parallel whenever possible. So, when both $w_1(t) > 0$ and $w_2(t) > 0$, we have $u_c(t) = 1$. By (4), there is a class i with $\frac{\rho_i}{c_i} < 1$. Hence, $\frac{dw_i(t)}{dt} = \rho_i - c_i < 0$ and class i will eventually be drained to zero. When at that time the workload in class j ($j \neq i$) is strictly positive (while $w_i(t) = 0$), we have $u_c(t) = \frac{\rho_i}{c_i}$ and $u_j(t) = 1 - \frac{\rho_i}{c_i}$. From (5) this gives $\frac{dw_j(t)}{dt} = 0$ and $\frac{dw_j(t)}{dt} = \rho_j - 1 + \frac{\rho_i}{c_i} - \frac{\rho_i}{c_i}c_j = \rho_j + \frac{\rho_i}{c_i}(1 - c_j) - 1 < 0$. Hence, class j must eventually become empty as well. \square

A policy π for the fluid control model is described by the control functions $u_1^\pi(t)$, $u_2^\pi(t)$ and $u_c^\pi(t)$ (we also write $U_j^\pi(t) = \int_0^t u_j^\pi(v)dv$). A corresponding trajectory is denoted by $n^\pi(t)$. We are interested in finding an optimal fluid control that minimizes

$$\int_0^\infty (n_1^\pi(t) + n_2^\pi(t))dt, \quad \text{with } (n^\pi(t), u^\pi(t)) \text{ satisfying (13)–(16)}. \quad (18)$$

We denote such an optimal control by $u_j^*(t)$, $j = 1, 2, c$, and a corresponding optimal trajectory, by $n_i^*(t)$, $i = 1, 2$. Note that if (4) and (5) are satisfied, then $\min_\pi \int_0^\infty (n_1^\pi(t) + n_2^\pi(t))dt$ is finite due to Lemma 4.1. Before proceeding to find $n^*(t)$ and $u^*(t)$, we first prove in the next lemma that an optimal pair $(n^*(t), u^*(t))$ exists. In addition, the lemma states that if $n^*(t)$ is an optimal trajectory for the infinite horizon problem, then it is also optimal for the finite horizon problem whenever the horizon is large enough. This property will be useful to prove convergence of the stochastic model in Section 4.2.

Lemma 4.2 *If (4) and (5) are satisfied, then there exists a control $u^*(t)$ and a corresponding trajectory $n^*(t)$ that solves the minimization problem (18).*

In addition, there exists a function $H : \mathbb{R} \rightarrow \mathbb{R}$ such that,

$$\min_{n(t) \text{ s.t. (13)–(16)}} \int_0^D (n_1(t) + n_2(t))dt = \int_0^D (n_1^*(t) + n_2^*(t))dt = \int_0^\infty (n_1^*(t) + n_2^*(t))dt,$$

for all $D \geq H(n_1 + n_2)$.

Proof: By the Filippov-Cesari theorem [31, Chapter 2.8], there exists an optimal control and a corresponding optimal trajectory $n^{*D}(t)$ for the problem $\min_{n(t)} \text{ s.t. (13)-(16)} \int_0^D (n_1(t) + n_2(t)) dt$. For the moment, assume that there exists a function $H(\cdot)$ such that

$$n_1^{*D}(t) + n_2^{*D}(t) = 0, \quad \text{for all } H(n_1 + n_2) \leq t, \text{ with } n = (n_1, n_2) \text{ denoting the initial state.} \quad (19)$$

The proof of (19) will be given later on. From (19) we obtain

$$\begin{aligned} \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^\infty (n_1(t) + n_2(t)) dt &\geq \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^D (n_1(t) + n_2(t)) dt \\ &= \int_0^D (n_1^{*D}(t) + n_2^{*D}(t)) dt = \int_0^\infty (n_1^{*D}(t) + n_2^{*D}(t)) dt, \\ &\geq \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^\infty (n_1(t) + n_2(t)) dt, \end{aligned} \quad (20)$$

for all $D \geq H(n_1 + n_2)$. Hence, $n^{*D}(t)$ is an optimal solution of (18). In particular, this implies the existence result for the minimization problem (18). In addition, from (20) we obtain that for any optimal trajectory $n^*(t)$ of (20), it holds that

$$\begin{aligned} \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^\infty (n_1(t) + n_2(t)) dt &= \int_0^\infty (n_1^*(t) + n_2^*(t)) dt \geq \int_0^D (n_1^*(t) + n_2^*(t)) dt \\ &\geq \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^D (n_1(t) + n_2(t)) dt = \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^\infty (n_1(t) + n_2(t)) dt, \end{aligned}$$

for all $D \geq H(n_1 + n_2)$. This proves the lemma under the condition that there indeed exists a function $H(\cdot)$ satisfying (19). The latter will be shown in the remainder of the proof. We use similar arguments as in [23, Proposition 6.1].

Denote by π^p the policy that always serves classes 1 and 2 in parallel whenever possible. Let $n^p(t)$ be the trajectory that corresponds to policy π^p . Under the stability conditions we know that $n^p(t)$ hits zero after a finite time and then remains empty, see Lemma 4.1. Denote by $T^p(\tilde{n}, n')$ the time it takes for policy π^p to move from \tilde{n} to n' . Then, the depletion time, $T^p(\tilde{n}, 0)$, can be written as follows

$$T^p(\tilde{n}, 0) = T^p(\tilde{n}, \text{axes}) + \frac{y_1(\tilde{n})}{\mu_1(1 - \frac{\rho_2}{c_2}(1 - c_1) - \rho_1)} + \frac{y_2(\tilde{n})}{\mu_2(1 - \frac{\rho_1}{c_1}(1 - c_2) - \rho_2)}, \quad (21)$$

where $T^p(\tilde{n}, \text{axes}) = \min\left(\frac{\tilde{n}_1}{(\mu_1 c_1 - \lambda_1)^+}, \frac{\tilde{n}_2}{(\mu_2 c_2 - \lambda_2)^+}\right)$ is the time until the trajectory hits either one of the axes, and $y(\tilde{n})$ represents the point where the trajectory hits the axis when started in \tilde{n} . Note that $y_1(\tilde{n}) = \tilde{n}_1 - T^p(\tilde{n}, \text{axes}) \cdot \mu_1(c_1 - \rho_1)$ and $y_2(\tilde{n}) = \tilde{n}_2 - T^p(\tilde{n}, \text{axes}) \cdot \mu_2(c_2 - \rho_2)$. Hence, the depletion time scales as follows: $T^p(a \cdot \tilde{n}, 0) = a \cdot T^p(\tilde{n}, 0)$, $a \geq 0$.

Let $0 < \zeta < 1$ be fixed, and $x > 0$. We now have the following upper bound for all initial states n with $n_1 + n_2 = x$:

$$\begin{aligned} \int_0^D (n_1^{*D}(t) + n_2^{*D}(t)) dt &= \min_{n(t)} \text{ s.t. (13)-(16)} \int_0^D (n_1(t) + n_2(t)) dt \leq \int_0^D (n_1^p(t) + n_2^p(t)) dt \\ &\leq \sup_{0 \leq t \leq D} \{n_1^p(t) + n_2^p(t)\} \cdot T^p(n, 0) \leq x \cdot \zeta \cdot (1 - \zeta) \cdot H(x). \end{aligned} \quad (22)$$

Here the function $H(x)$ is defined as

$$H(x) := \frac{\beta}{\zeta \cdot (1 - \zeta)} \cdot \sup_{l: l_1 + l_2 = x} \{T^p(l, 0)\},$$

with the constant

$$\beta := \max\left(1 + \frac{\lambda_1 + \lambda_2 - (\mu_1 c_1 + \mu_2 c_2)}{\mu_1 c_1 - \lambda_1}, 1 + \frac{\lambda_1 + \lambda_2 - (\mu_1 c_1 + \mu_2 c_2)}{\mu_2 c_2 - \lambda_2}, 1\right),$$

so that for all initial states n with $n_1 + n_2 = x$ it holds that $\sup_{0 \leq t \leq D} \{n_1^p(t) + n_2^p(t)\} = \max(x + T^p(n, \text{axes}) \cdot (\lambda_1 + \lambda_2 - (\mu_1 c_1 + \mu_2 c_2)), x) \leq \beta \cdot x$.

From (21) it easily follows that $T^p(l, 0)$ is continuous in l . Hence $\sup_{l: l_1 + l_2 = x} T^p(l, 0) < \infty$ and in particular $H(x) < \infty$ for all $x > 0$. Assume $D \geq H(x)$ (in particular, $D \geq (1 - \zeta) \cdot H(x)$). Hence, it follows from (22) that

$$\tau(x) := \arg \min_{t \geq 0} \{n_1^{*D}(t) + n_2^{*D}(t) \leq x \cdot \zeta\} \leq (1 - \zeta) \cdot H(x), \quad (23)$$

for all initial states n with $n_1 + n_2 = x$. From continuity of $n^{*D}(t)$ it follows that $n_1^{*D}(\tau(x)) + n_2^{*D}(\tau(x)) = x \cdot \zeta$.

If $n^{*D}(0) = (n_1, n_2)$, then $n^{*D}(\sum_{m=1}^{\infty} \tau((n_1 + n_2)\zeta^{m-1})) = (0, 0)$. Note that $H(a \cdot x) = a \cdot H(x)$, $a \geq 0$. Together with (23) it follows that $\sum_{m=1}^{\infty} \tau((n_1 + n_2)\zeta^{m-1}) \leq \sum_{m=1}^{\infty} \zeta^{m-1} (1 - \zeta) \cdot H(n_1 + n_2) = H(n_1 + n_2) < \infty$. Hence, relation (19) holds. \square

For the stochastic model we know that it is never optimal to serve class 2 exclusively when also work of class 1 is present. In the fluid control model this is true as well, as is stated in the lemma below. The proof may be found in Appendix C.

Lemma 4.3 *Assume (4) and (5) are satisfied, $\mu_1 \geq \mu_2$, and $c_1 + c_2 > 1$. Then, for any policy $\tilde{\pi}$ that allows $u_2^{\tilde{\pi}}(t) > 0$ when $n_1^{\tilde{\pi}}(t) > 0$, there exists a modified policy π , with $u_2^{\pi}(t) = 0$ whenever $n_1^{\pi}(t) > 0$, that does not worse than $\tilde{\pi}$, i.e., $n_1^{\pi}(t) + n_2^{\pi}(t) \leq n_1^{\tilde{\pi}}(t) + n_2^{\tilde{\pi}}(t)$, for all $t \geq 0$.*

In case $\mu_1 \leq c_1 \mu_1 + c_2 \mu_2$, the control that serves both classes in parallel whenever possible is optimal, i.e., $u_c^*(t) = 1$ when $n_1(t), n_2(t) > 0$, and $u_c^*(t) = \min(\frac{\rho_i}{c_j}, 1)$, $u_i^*(t) = 1 - u_c^*(t)$ when $n_j(t) = 0$ and $n_i(t) > 0$, for $i \neq j$, $i, j = 1, 2$. This follows from the fact that the above described policy minimizes the time to empty the system, while at the same time, it maximizes the departure rate at any moment in time. We do not include a formal proof of this fact, since the main objective of this section is to investigate close-to-optimal policies for parameter choices that did not allow us to exactly determine the optimal policy for the stochastic model. (Proposition 3.3 discusses an optimal stochastic policy when $\mu_1 \leq c_1 \mu_1 + c_2 \mu_2$.)

In the remainder of this section we concentrate on the case $\mu_1 > c_1 \mu_1 + c_2 \mu_2$, for which the following lemma enables us to prove that an optimal policy in the fluid control model can be characterized by a switching curve.

Lemma 4.4 *Assume (4) and (5) are satisfied, $\mu_1 > c_1 \mu_1 + c_2 \mu_2$, and $c_1 + c_2 > 1$. Consider a trajectory starting in $\tilde{n} \in \{n : n_1 > 0, n_2 \geq 0\}$ with the following properties: (i) first class 1 is served exclusively during a contiguous period, and then (ii) we switch to serving both classes simultaneously during another contiguous period. Let \hat{n} be the end point of this trajectory.*

Then the trajectory described above minimizes $n_1(t) + n_2(t)$ at all times (until reaching \hat{n}), among all trajectories that move from \tilde{n} to \hat{n} without coinciding with the $n_1 = 0$ axis.

Proof: Since we consider only trajectories from \tilde{n} to \hat{n} that do not coincide with the $n_1 = 0$ axis, by Lemma 4.3 we can focus on paths that do not spend any time serving class 2 individually. Denote by U_1 (U_c) the cumulative amount of time spent on serving class 1 individually (classes 1 and 2 in parallel). The net change in the amount of fluid in the two classes can be written as

$$\begin{aligned} \hat{n}_1 - \tilde{n}_1 &= (\lambda_1 - \mu_1)U_1 + (\lambda_1 - c_1 \mu_1)U_c, \\ \hat{n}_2 - \tilde{n}_2 &= \lambda_2 U_1 + (\lambda_2 - c_2 \mu_2)U_c. \end{aligned}$$

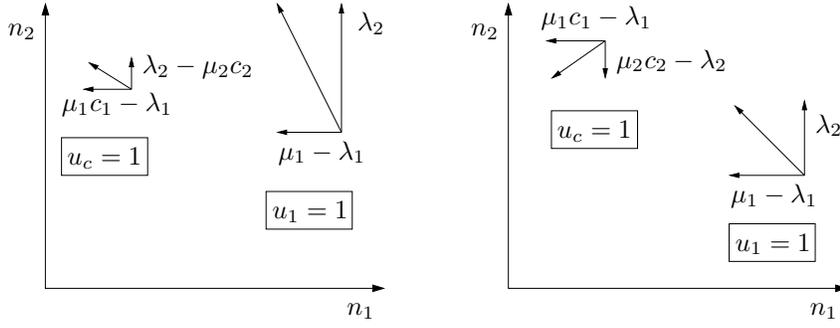


Figure 3: Drift vectors for $\rho_1 < c_1$ and $\rho_2 > c_2$ (left), and $\rho_1 < c_1$ and $\rho_2 < c_2$ (right), respectively.

Under the necessary stability conditions (4) and (5) this has a unique solution for U_1 and U_c . Hence, all trajectories spend the same cumulative amount of time serving both classes in parallel as well as serving class 1 individually.

The rate at which the total amount of fluid decreases when $n_1(t) > 0$ is given by $\frac{d(n_1(t)+n_2(t))}{dt} = \lambda_1 + \lambda_2 - u_1(t)\mu_1 - u_c(t)(\mu_1 c_1 + \mu_2 c_2)$. Since $\mu_1 > \mu_1 c_1 + \mu_2 c_2$, first serving only class 1 initially maximizes the rate at which $n_1(t) + n_2(t)$ decreases. Hence, this minimizes $n_1(t) + n_2(t)$ at all times (until reaching \hat{n}). \square

For the fluid control model we can now determine optimal policies. To do that, we distinguish between whether $\rho_1 < c_1$ or $\rho_1 \geq c_1$. Note that, cf. Bellman's principle of optimality, we only need to consider policies that base their actions on the current state $n(t)$, because of the infinite horizon and the fact that the parameters do not depend on the current time t .

4.1.1 Case $\rho_1 < c_1$

When $\rho_1 < c_1$, a necessary condition for the system to drain in finite time is $\rho_2 < 1 - \frac{\rho_1}{c_1}(1 - c_2)$ (see Lemma 4.1). Depending on ρ_2 and c_2 , the drifts are as in Figure 3. In Proposition 4.5 we describe an optimal fluid control, which is characterized by a linear switching curve. In Figure 4 the corresponding trajectory is shown. In order to state the proposition it is convenient to define

$$\alpha := \max\left(0, \frac{c_2 - \rho_2}{c_1 - \rho_1} + \frac{c_1}{c_1 + c_2 - 1} \times \frac{1 - \rho_2 - \frac{\rho_1}{c_1}(1 - c_2)}{c_1 - \rho_1} \times \frac{\mu_1 - c_1 \mu_1 - c_2 \mu_2}{\mu_2}\right). \quad (24)$$

Note that under the conditions of Proposition 4.5, it holds that $\alpha > \frac{c_2 - \rho_2}{c_1 - \rho_1}$.

Proposition 4.5 *Let $\mu_1 > \mu_1 c_1 + \mu_2 c_2$ and $c_1 + c_2 > 1$. Assume $\rho_1 < c_1$ and $\rho_2 < 1 - \frac{\rho_1}{c_1}(1 - c_2)$. An optimal control $u^*(t)$ in the fluid control model is*

- $u_1^*(t) = 1$, if $n_2(t) < \alpha \frac{\mu_2}{\mu_1} n_1(t)$.
- $u_c^*(t) = 1$, if $n_2(t) \geq \alpha \frac{\mu_2}{\mu_1} n_1(t)$ and $n_1(t) > 0$.
- $u_c^*(t) = \frac{\rho_1}{c_1}$ and $u_2^*(t) = 1 - \frac{\rho_1}{c_1}$, if $n_1(t) = 0$.

Proof: If $n_1(t) > 0$, when searching for an optimal control, by Lemma 4.3 we only need to consider controls with $u_2(t) = 0$ and $u_1(t) + u_c(t) = 1$. Hence, from $\frac{dn_1(t)}{dt} = \lambda_1 - u_1(t)\mu_1 - u_c(t)\mu_1 c_1$, and the fact that $\rho_1 < c_1 < 1$, class 1 remains empty once it hits zero. So $\frac{dn_1(t)}{dt} = 0$, or equivalently, $\rho_1 - u_1(t) - u_c(t)c_1 = 0$ when $n_1(t) = 0$.

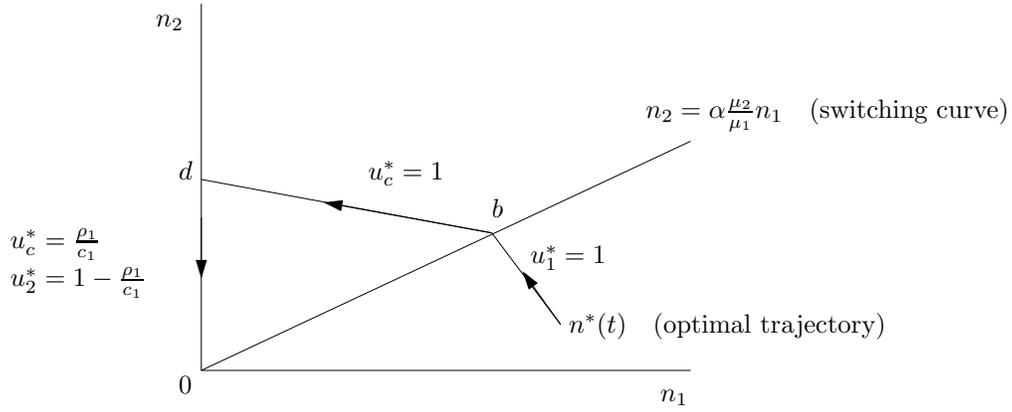


Figure 4: Optimal trajectory of the fluid control model when $\rho_1 < c_1$.

We can now determine an optimal allocation for points with $n_1(t) = 0$. Class 1 is kept empty, hence an optimal fluid control will maximize the departure rate of class 2. We should therefore maximize $u_2(t)\mu_2 + u_c(t)\mu_2c_2$ given that $\rho_1 - u_1(t) - u_c(t)c_1 = 0$, $u_1(t) + u_2(t) + u_c(t) = 1$ and $u_j(t) \geq 0$. Solving this we obtain

$$u_c^*(t) = \frac{\rho_1}{c_1}, \quad u_1^*(t) = 0 \quad \text{and} \quad u_2^*(t) = 1 - \frac{\rho_1}{c_1}.$$

So when $n_1(t) = 0$, class 1 is kept empty by serving both classes in parallel a fraction $\frac{\rho_1}{c_1}$ of time. The remaining capacity is given to class 2, see Figure 4.

Now assume we start at time $t = 0$ in $n(0) = n = (n_1, n_2)$ with $n_1 > 0$ and $n_2 \geq 0$. At some point an optimal trajectory will hit the $n_1=0$ axis for the first time. This point will be denoted by $d = (0, d_2)$, see Figure 4. Note that the path from n to d that first serves class 1 individually and at some point switches to serving both classes in parallel, is always feasible (see the drift vectors in Figure 3). Hence, by Lemma 4.4 this path is also an optimal path from n to d . The turning point where the switch occurs is denoted by $b = (b_1, b_2)$, see again Figure 4. We can calculate the costs corresponding to a certain turning point b . Let $T(x, y)$ be the time it takes to go from point x to y in the plane. We have $T(n, b) = \frac{n_1 - b_1}{\mu_1 - \lambda_1}$, $T(b, d) = \frac{b_1}{\mu_1 c_1 - \lambda_1}$,

$$T(d, 0) = \frac{d_2}{u_2\mu_2 + u_c\mu_2c_2 - \lambda_2} = \frac{d_2}{\mu_2 - \mu_2\frac{\rho_1}{c_1}(1 - c_2) - \lambda_2},$$

with $d_2 = b_2 + T(b, d)(\lambda_2 - \mu_2c_2)$ and $b_2 = n_2 + T(n, b)\lambda_2$. Let $K_n(b_1) = \int_0^\infty (n_1(t) + n_2(t))dt$ be the cost of the fluid trajectory going from n to the origin when the turning point is $b = (b_1, b_2)$. Note that $b_2 = n_2 + \frac{n_1 - b_1}{\mu_1 - \lambda_1}\lambda_2$, hence b_2 is uniquely determined by b_1 and n . We have

$$K_n(b_1) = T(n, b) \left(\frac{n_1 + b_1}{2} + \frac{n_2 + b_2}{2} \right) + T(b, d) \left(\frac{b_1}{2} + \frac{b_2 + d_2}{2} \right) + T(d, 0) \frac{d_2}{2}. \quad (25)$$

It can be checked that the function $K_n(b_1)$ is a quadratic function in b_1 and when minimizing the costs in (25), the optimal turning point b lies on the line $b_2 = \alpha \frac{\mu_2}{\mu_1} b_1$. Hence, if $n_2(t) < \alpha \frac{\mu_2}{\mu_1} n_1(t)$, then $u_1^*(t) = 1$, and if $n_2(t) \geq \alpha \frac{\mu_2}{\mu_1} n_1(t)$ and $n_1(t) > 0$, then $u_c^*(t) = 1$. This completes the characterization of an optimal control. \square

4.1.2 Case $\rho_1 \geq c_1$

When $\rho_1 \geq c_1$, the necessary stability condition is $\rho_2 < c_2$ and $\rho_1 < 1 - \frac{\rho_2}{c_2}(1 - c_1)$ (see (4) and (5)). Hence $\frac{\rho_2}{1 - \rho_1} \leq \frac{c_2 - \rho_2}{\rho_1 - c_1}$ and the drifts are as in the left picture in Figure 5. When

$\rho_1 \geq 1 - \frac{\rho_2}{c_2}(1 - c_1)$, the system is unstable which corresponds to the picture on the right in Figure 5. An optimal fluid policy is described in the next proposition, and in Figure 6 the corresponding trajectory is shown.

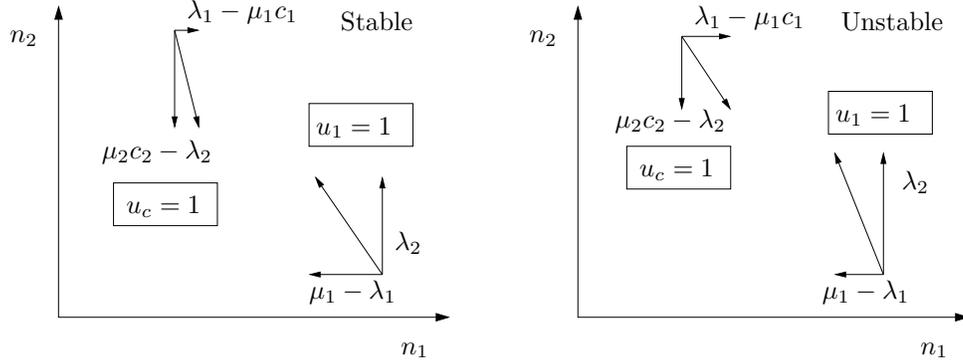


Figure 5: Vectors for $\rho_1 \geq c_1$ and $\rho_2 < c_2$. Left figure: $\rho_1 < 1 - \frac{\rho_2}{c_2}(1 - c_1)$ and hence there are policies that give a stable system. Right figure: $\rho_1 > 1 - \frac{\rho_2}{c_2}(1 - c_1)$ and hence unstable.

Proposition 4.6 *Let $\mu_1 > \mu_1 c_1 + \mu_2 c_2$ and $c_1 + c_2 > 1$. Assume $\rho_1 \geq c_1$, $\rho_2 < c_2$ and $\rho_1 < 1 - \frac{\rho_2}{c_2}(1 - c_1)$. An optimal policy in the fluid control model is to give priority to class 1, i.e.,*

- $u_1^*(t) = 1$ if $n_1(t) > 0$.
- $u_c^*(t) = \frac{1 - \rho_1}{1 - c_1}$ and $u_1^*(t) = \frac{\rho_1 - c_1}{1 - c_1}$ if $n_1(t) = 0$.

The proof of Proposition 4.6 below does not give much insight into the result. Therefore, we first provide some intuition for the fact that the control u^* as defined above, is optimal when $\rho_1 > c_1$: Using Lemma 4.4 it can be argued that as long as $n_1(t) > 0$, an optimal action is $u_1^*(t) = 1$. Hence, once $n_1(t) = 0$, this optimal control will keep class 1 empty ($\rho_1 < 1$). An optimal fluid control will now choose allocations $u_j^*(t)$ such that the departure rate for class 2, $u_2(t)\mu_2 + u_c(t)\mu_2 c_2$, is maximized subject to $u_1(t) + u_c(t)c_1 = \rho_1$, $u_1(t) + u_2(t) + u_c(t) = 1$ and $u_j(t) \geq 0$. The unique solution to this is $u_2^*(t) = 0$, $u_1^*(t) = \frac{\rho_1 - c_1}{1 - c_1}$ and $u_c^*(t) = \frac{1 - \rho_1}{1 - c_1}$ when $n_1(t) = 0$.

Proof of Proposition 4.6: Consider the control $u^*(t)$ as defined in Proposition 4.6. The corresponding trajectory is denoted by $n^*(t)$. Its cost-to-go function is defined as $K_{(t,n)} := \int_t^\infty (n_1^*(s) + n_2^*(s)|n(t) = n) ds = \int_0^\infty (n_1^*(s) + n_2^*(s)|n(0) = n) ds$, for $n = (n_1, n_2) \in \mathbb{R}_+^2$. Hence, we can drop the dependence on t , and write K_n for the cost-to-go starting in state n . A sufficient condition for optimality of $u^*(t)$ is that its cost-to-go function K_n satisfies the ‘‘Hamilton-Jacobi-Bellman’’ partial differential equation

$$0 = \min_{u \text{ s.t. (14)-(16)}} \left(n_1 + n_2 + \frac{\partial K_n}{\partial n_1} \cdot (\lambda_1 - \mu_1(u_1 + c_1 u_c)) + \frac{\partial K_n}{\partial n_2} \cdot (\lambda_2 - \mu_2(u_2 + c_2 u_c)) \right), \quad (26)$$

for all $n_1, n_2 \geq 0$, and that u^* is a corresponding minimizing action, [12, Section 5.5]. In the remainder of the proof we show that this is indeed satisfied.

The cost-to-go function is easily derived. Let $d = (0, d_2)$ denote the point where the trajectory $n^*(t)$ hits the vertical axis, see Figure 6. Hence, $d_2 = n_2 + n_1 \frac{\lambda_2}{\mu_1 - \lambda_1}$ and also $d_2 = T(d, 0) \cdot$

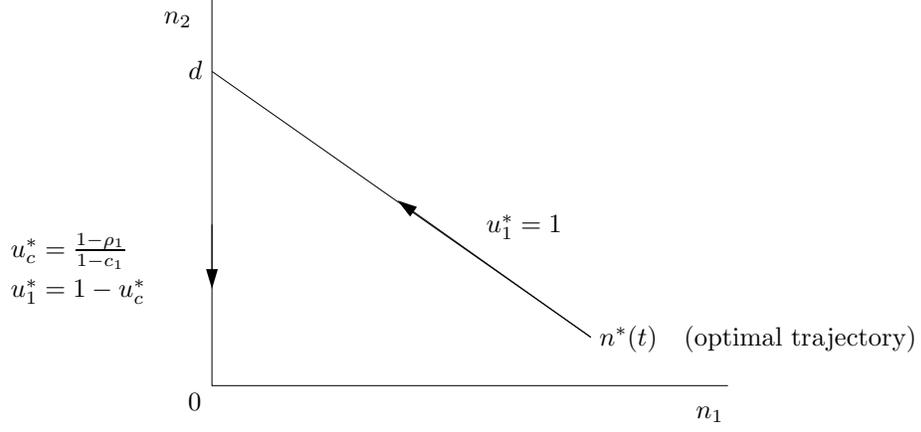


Figure 6: Optimal trajectory of the fluid control model when $\rho_1 \geq c_1$.

$(\mu_2 c_2 \frac{1-\rho_1}{1-c_1} - \lambda_2) = T(d, 0) \cdot \mu_2 \frac{c_2}{1-c_1} \cdot (1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))$, with $T(d, 0)$ the time it takes to move from d to 0. So

$$\begin{aligned} K_n &= T(n, d) \left(\frac{n_1}{2} + \frac{n_2 + d_2}{2} \right) + T(d, 0) \frac{d_2}{2} \\ &= \frac{n_1}{\mu_1 - \lambda_1} \left(\frac{n_1}{2} + \frac{2n_2 + n_1 \frac{\lambda_2}{\mu_1 - \lambda_1}}{2} \right) + \frac{(n_2 + n_1 \frac{\lambda_2}{\mu_1 - \lambda_1})^2}{2\mu_2 \frac{c_2}{1-c_1} \cdot (1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))}. \end{aligned}$$

In the Hamilton-Jacobi-Bellman equation we are interested in the function

$$\begin{aligned} &n_1 + n_2 + \frac{\partial K_n}{\partial n_1} \cdot (\lambda_1 - \mu_1(u_1 + c_1 u_c)) + \frac{\partial K_n}{\partial n_2} \cdot (\lambda_2 - \mu_2(u_2 + c_2 u_c)) \\ &= n_1 + n_2 + (\lambda_1 - \mu_1(u_1 + c_1 u_c)) \cdot \\ &\quad \left(\frac{n_1}{\mu_1 - \lambda_1} + \frac{1}{\mu_1 - \lambda_1} \left(n_1 \frac{\lambda_2}{\mu_1 - \lambda_1} + n_2 \right) + \frac{\frac{\lambda_2}{\mu_1 - \lambda_1}}{\mu_2 \frac{c_2}{1-c_1} \cdot (1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))} \left(n_1 \frac{\lambda_2}{\mu_1 - \lambda_1} + n_2 \right) \right) \\ &\quad + (\lambda_2 - \mu_2(u_2 + c_2 u_c)) \cdot \left(\frac{n_1}{\mu_1 - \lambda_1} + \frac{1}{\mu_2 \frac{c_2}{1-c_1} \cdot (1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))} \left(n_1 \frac{\lambda_2}{\mu_1 - \lambda_1} + n_2 \right) \right) \\ &= n_1 + n_2 + \frac{1}{\mu_1 - \lambda_1} \cdot \left(\rho_1(a_{11}n_1 + a_{12}n_2) + \rho_2(a_{21}n_1 + a_{22}n_2) \right. \\ &\quad \left. - n_1(u_1 a_{11} + u_c(c_1 a_{11} + c_2 a_{21}) + u_2 a_{21}) - n_2(u_1 a_{12} + u_c(c_1 a_{12} + c_2 a_{22}) + u_2 a_{22}) \right), \quad (27) \end{aligned}$$

where

$$\begin{aligned} a_{11} &= \mu_1 + \frac{\lambda_2}{1 - \rho_1} \left(1 + \frac{\frac{\rho_2}{c_2}(1 - c_1)}{(1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))} \right) = \mu_1 + \mu_2 \frac{\rho_2}{1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1)}, \\ a_{12} &= \mu_1 + \mu_1 \frac{\frac{\rho_2}{c_2}(1 - c_1)}{(1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))} = \mu_1 \frac{1 - \rho_1}{1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1)}, \\ a_{21} &= \mu_2 + \mu_2 \frac{\frac{\rho_2}{c_2}(1 - c_1)}{1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1)} = \mu_2 \frac{1 - \rho_1}{1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1)}, \\ a_{22} &= \frac{\mu_1 - \lambda_1}{\frac{c_2}{1-c_1}(1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))} = \mu_1 \frac{(1 - \rho_1) \frac{1-c_1}{c_2}}{(1 - \rho_1 - \frac{\rho_2}{c_2}(1 - c_1))}. \end{aligned}$$

Elementary calculation shows that for $u_1 = 1$, and $u_2 = u_c = 0$, equation (27) is equal to zero. In addition, under the conditions as stated in Proposition 4.6, it holds that $a_{11} > c_1 a_{11} + c_2 a_{21} > a_{21}$

and $a_{12} = c_1 a_{12} + c_2 a_{22} > a_{22}$. Hence, when $n_1(t) > 0$, the minimizing action is $u_1(t) = 1, u_2(t) = 0, u_c(t) = 0$, which is indeed prescribed by the control strategy u^* .

When $n_1 = 0$, equation (27) is equal to

$$n_2 + \frac{1}{\mu_1 - \lambda_1} \left(n_2(\rho_1 a_{12} + \rho_2 a_{22}) - n_2(u_1 a_{12} + u_c(c_1 a_{12} + c_2 a_{22}) + u_2 a_{22}) \right). \quad (28)$$

Again simple calculations show that this is equal to 0 for all u with $u_1 + u_c = 1$ and $u_2 = 0$. Besides $u_1 + u_2 + u_c \leq 1$, we have the restriction $u_1 + c_1 u_c \leq \rho_1$ (because $n_1 = 0$). Since $a_{12} = c_1 a_{12} + c_2 a_{22} > a_{22}$, any control with $u_1 + u_c = 1$ and $u_2 = 0$ such that $u_1 + c_1 u_c \leq \rho_1$, will minimize (28). The control $u_1^*(t) = \frac{\rho_1 - c_1}{1 - c_1}, u_c^*(t) = \frac{1 - \rho_1}{1 - c_1}$ and $u_2^*(t) = 0$ is therefore indeed a minimizing action. \square

4.2 Asymptotic fluid optimality

In this section we discuss the theoretical foundations that justify the use of optimal controls in the fluid model as proxies for optimal policies in the stochastic model. In particular, we prove that under a *fluid scaling*, the stochastic processes of the numbers of users under certain switching curve policies, converge to the optimal fluid trajectory $n^*(t)$ as determined in Section 4.1. Using the latter, we then show that these switching curve policies are asymptotically fluid optimal in the stochastic model. The terminology used in this section is motivated by [2, 17, 23, 25].

On a common probability space we construct different realizations of the processes, depending on the initial state. To be precise, for a given policy π we let $N_i^{\pi,r}(t)$ denote the number of class- i users at time t when the initial state equals $N_i^r(0) = r n_i, i = 1, 2$, with $r \in \mathbb{N}$. All processes $N^{\pi,r}(t)$ share the same sequences of arrivals and service requirements.

For a given policy π , denote by $T_0^{\pi,r}(t)$ the cumulative amount of time during the interval $(0, t)$ that neither class is served, by $T_i^{\pi,r}(t)$ the cumulative amount of time that was spent on serving class i individually, $i = 1, 2$, and by $T_c^{\pi,r}(t)$ the cumulative amount of time that was spent on serving classes 1 and 2 in parallel. Then, $T_0^{\pi,r}(t) + T_1^{\pi,r}(t) + T_2^{\pi,r}(t) + T_c^{\pi,r}(t) = t$, and

$$N_i^{\pi,r}(t) = r n_i + E_i(t) - F_i(T_i^{\pi,r}(t)) - F_{c,i}(T_c^{\pi,r}(t)), \quad i = 1, 2, \quad (29)$$

with $E_i(t)$ a Poisson process with rate λ_i , $F_i(\cdot)$ a Poisson process with rate μ_i and $F_{c,i}(\cdot)$ a Poisson process with rate $c_i \mu_i$, [15].

We will be interested in the processes under the fluid scaling, i.e., both time and space are scaled linearly with the parameter r :

$$\bar{N}_i^{\pi,r}(t) := \frac{N_i^{\pi,r}(rt)}{r} \quad \text{and} \quad \bar{T}_j^{\pi,r}(t) := \frac{T_j^{\pi,r}(rt)}{r}.$$

Limit points for $\bar{N}_i^{\pi,r}(t)$ and $\bar{T}_j^{\pi,r}(t)$ are described in the next lemma.

Lemma 4.7 *For almost all sample paths ω and any sequence r_k , there exists a subsequence r_{k_l} such that*

$$\begin{aligned} \lim_{l \rightarrow \infty} \bar{N}_i^{\pi,r_{k_l}}(t) &= \bar{N}_i^{\pi}(t), \quad i = 1, 2, \quad u.o.c., \\ \lim_{l \rightarrow \infty} \bar{T}_j^{\pi,r_{k_l}}(t) &= \bar{T}_j^{\pi}(t), \quad j = 0, 1, 2, c, \quad u.o.c., \end{aligned}$$

with $(\bar{N}^{\pi}, \bar{T}^{\pi})$ a continuous function. In addition, $(\bar{N}^{\pi}, \bar{T}^{\pi})$ satisfies, for $i = 1, 2, j = 0, 1, 2, c$,

$$\bar{N}_i^{\pi}(t) = n_i + \lambda_i t - \mu_i \bar{T}_i^{\pi}(t) - \mu_i c_i \bar{T}_c^{\pi}(t), \quad (30)$$

$\bar{N}_i^{\pi}(t) \geq 0, \bar{T}_j^{\pi}(0) = 0, \bar{T}_0^{\pi}(t) + \bar{T}_1^{\pi}(t) + \bar{T}_2^{\pi}(t) + \bar{T}_c^{\pi}(t) = t$, and $\bar{T}_j^{\pi}(t)$ are non-decreasing and Lipschitz continuous functions.

The notation u.o.c. stands for uniform convergence on compact sets. We call the processes $\overline{T}_j^\pi(t), j = 1, 2, c$, and $\overline{N}_i^\pi(t), i = 1, 2$, (as obtained in Lemma 4.7) fluid limits for initial fluid level n and policy π .

Proof of Lemma 4.7: Making use of (29) and the fact that $\overline{T}_j^{\pi,r}(t), j = 1, 2, c$, is Lipschitz continuous with a constant less than or equal to 1, the proof follows similarly as that of [14, Theorem 4.1], see also [26, Proposition 10.3.3 and 10.3.4]. Note that the Poisson assumptions are in fact not needed for the result of this lemma to hold. \square

As cost in the stochastic model we take $\mathbb{E}\left(\int_0^D (N_1^{\pi,r}(t) + N_2^{\pi,r}(t))dt\right)$. As $r \rightarrow \infty$, this will tend to infinity. In order to obtain a non-trivial limit we divide the cost by r^2 and consider a horizon that grows linearly in r . So we are interested in

$$\mathbb{E}\left(\int_0^{r \cdot D} \frac{N_1^{\pi,r}(t) + N_2^{\pi,r}(t)}{r^2} dt\right) = \mathbb{E}\left(\int_0^D \frac{N_1^{\pi,r}(rt) + N_2^{\pi,r}(rt)}{r} dt\right) = \mathbb{E}\left(\int_0^D (\overline{N}_1^{\pi,r}(t) + \overline{N}_2^{\pi,r}(t))dt\right). \quad (31)$$

Our goal is to find policies that minimize the cost (31) as $r \rightarrow \infty$. We have the following lower bound.

Lemma 4.8 *For any policy π we have*

$$\liminf_{r \rightarrow \infty} \mathbb{E}\left(\int_0^D (\overline{N}_1^{\pi,r}(t) + \overline{N}_2^{\pi,r}(t))dt\right) \geq \int_0^\infty (n_1^*(t) + n_2^*(t))dt, \quad \text{whenever } D \geq H(n_1 + n_2),$$

and where $n^*(t)$ represents an optimal solution of (18) for initial state n and $H(\cdot)$ is as defined in Lemma 4.2.

Proof: By applying Fatou's lemma, we obtain

$$\begin{aligned} \liminf_{r \rightarrow \infty} \mathbb{E}\left(\int_0^D (\overline{N}_1^{\pi,r}(t) + \overline{N}_2^{\pi,r}(t))dt\right) &\geq \mathbb{E}\left(\liminf_{r \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi,r}(t) + \overline{N}_2^{\pi,r}(t))dt\right) \\ &= \mathbb{E}\left(\lim_{k \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi,r_k}(t) + \overline{N}_2^{\pi,r_k}(t))dt\right), \end{aligned}$$

with the subsequence r_k (possibly depending on the sample path ω) corresponding to the lim inf-sequence. Lemma 4.7 states that for almost all sample paths ω , there exists a subsequence r_{k_l} of r_k such that $\lim_{l \rightarrow \infty} \overline{N}_1^{\pi,r_{k_l}}(t) + \overline{N}_2^{\pi,r_{k_l}}(t) = \overline{N}_1^\pi(t) + \overline{N}_2^\pi(t)$, u.o.c., with $\overline{N}_i^\pi(t)$ a fluid limit for initial fluid level n and policy π . Note that a fluid limit is an admissible trajectory for the fluid control problem. When we consider a finite horizon $D \geq H(n_1 + n_2)$, we know from Lemma 4.2 that n^* solves the finite-horizon minimization problem, and hence

$$\begin{aligned} \lim_{l \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi,r_{k_l}}(t) + \overline{N}_2^{\pi,r_{k_l}}(t))dt &= \int_0^D \lim_{l \rightarrow \infty} (\overline{N}_1^{\pi,r_{k_l}}(t) + \overline{N}_2^{\pi,r_{k_l}}(t))dt = \int_0^D (\overline{N}_1^\pi(t) + \overline{N}_2^\pi(t))dt \\ &\geq \min_{n(t) \text{ s.t. (13)-(16)}} \int_0^D (n_1(t) + n_2(t))dt = \int_0^D (n_1^*(t) + n_2^*(t))dt = \int_0^\infty (n_1^*(t) + n_2^*(t))dt, \end{aligned}$$

where in the first step we used uniform convergence of the functions $\overline{N}_i^{\pi,r_{k_l}}(t), i = 1, 2$, on $[0, D]$, in order to interchange the limit and the integral. This proves the lemma. \square

We say that a policy is asymptotically fluid optimal when the lower bound is obtained, i.e., when the scaled cost converges to the cost of the optimal trajectory in the fluid model. In the remainder of this section we characterize asymptotically fluid optimal policies.

Definition 4.9 A policy π^* is called asymptotically fluid optimal if

$$\lim_{r \rightarrow \infty} \mathbb{E} \left(\int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt \right) = \int_0^\infty (n_1^*(t) + n_2^*(t)) dt, \quad \text{with } D \geq H(n_1 + n_2),$$

where $n^*(t)$ is an optimal solution of (18) for initial state n and $H(\cdot)$ is as defined in Lemma 4.2.

4.2.1 Case $\rho_1 < c_1$

In this section we consider the case $\rho_1 < c_1$. In Proposition 4.5 we found that an optimal switching curve for the fluid control problem was given by $h(n_1) = \alpha \frac{\mu_2}{\mu_1} n_1$. In the following lemma we show that under this switching curve, the fluid scaled processes of the original stochastic model have a unique limit, which is described by the optimal trajectory of the fluid control model. Using this fact, in Proposition 4.11 we show that the policy with switching curve $h(n_1) = \alpha \frac{\mu_2}{\mu_1} n_1$ is asymptotically fluid optimal.

Lemma 4.10 Assume $c_1 + c_2 > 1$, $\rho_1 < c_1$ and $\rho_2 < 1 - \frac{\rho_1}{c_1}(1 - c_2)$. Denote by π^* the policy with switching curve $h(N_1) = \alpha \frac{\mu_2}{\mu_1} N_1$, with α as defined in (24). The functions $\overline{T}_j^{\pi^*}(t)$ are differentiable almost everywhere, and for each regular point t it holds that

$$\frac{d\overline{T}_1^{\pi^*}(t)}{dt} = 1, \quad \text{if } \overline{N}_2^{\pi^*}(t) < \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(t), \quad (32)$$

$$\frac{d\overline{T}_c^{\pi^*}(t)}{dt} = 1, \quad \text{if } \overline{N}_2^{\pi^*}(t) \geq \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(t) \text{ and } \overline{N}_1^{\pi^*}(t) > 0, \quad (33)$$

$$\frac{d\overline{T}_c^{\pi^*}(t)}{dt} = \frac{\rho_1}{c_1} \text{ and } \frac{d\overline{T}_2^{\pi^*}(t)}{dt} = 1 - \frac{\rho_1}{c_1}, \quad \text{if } \overline{N}_1^{\pi^*}(t) = 0 \text{ and } \overline{N}_2^{\pi^*}(t) > 0, \quad (34)$$

and $\frac{d\overline{T}_1^{\pi^*}(t)}{dt} + \frac{d\overline{T}_2^{\pi^*}(t)}{dt} + \frac{d\overline{T}_c^{\pi^*}(t)}{dt} + \frac{d\overline{T}_0^{\pi^*}(t)}{dt} = 1$.

In particular, $\overline{N}^{\pi^*}(t)$ is uniquely determined by

$$\overline{N}^{\pi^*}(t) = n^*(t), \quad (35)$$

with $n^*(t)$ the trajectory corresponding to the control $u^*(t)$ as defined in Proposition 4.5.

Proof: Let $\overline{N}_i^{\pi^*}(t)$, $i = 1, 2$, $\overline{T}_j^{\pi^*}(t)$, $j = 1, 2, c, 0$, be a fluid limit of policy π^* . So the functions $\overline{N}_i^{\pi^*}(t)$, $i = 1, 2$, satisfy (30), and the functions $\overline{T}_j^{\pi^*}(t)$, $j = 0, 1, 2, c$, are absolutely continuous (follows from Lipschitz continuity), and hence are differentiable almost everywhere. Fix a sample path ω such that there is a subsequence r_k with $\lim_{k \rightarrow \infty} \overline{N}_i^{\pi^*,r_k}(t) = \overline{N}_i^{\pi^*}(t)$, $i = 1, 2$, u.o.c., and $\lim_{k \rightarrow \infty} \overline{T}_j^{\pi^*,r_k}(t) = \overline{T}_j^{\pi^*}(t)$, $j = 1, 2, c$, u.o.c.. Further, let $t > 0$ be a regular point of $\overline{T}_j^{\pi^*}(t)$ for all $j = 0, 1, 2, c$.

First assume $\overline{N}_2^{\pi^*}(t) < \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(t)$. Then there is an $\epsilon > 0$ such that $\overline{N}_2^{\pi^*}(s) < \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(s)$ for $s \in [t - \epsilon, t + \epsilon]$. By the uniform convergence of $\overline{N}_i^{\pi^*,r_k}(t)$ to $\overline{N}_i^{\pi^*}(t)$, $i = 1, 2$, on $[t - \epsilon, t + \epsilon]$, we have $\overline{N}_2^{\pi^*,r_k}(r_k s) < \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*,r_k}(r_k s)$ for all r_k large enough and $s \in [t - \epsilon, t + \epsilon]$. Hence, under policy π^* , in the interval $[r_k(t - \epsilon), r_k(t + \epsilon)]$ class 1 is served and we obtain $\overline{T}_1^{\pi^*,r_k}(t + \epsilon) - \overline{T}_1^{\pi^*,r_k}(t - \epsilon) = 2\epsilon$.

Letting $r_k \rightarrow \infty$ and $\epsilon \downarrow 0$ we obtain $\frac{d\overline{T}_1^{\pi^*}(t)}{dt} = 1$.

Now assume $\overline{N}_2^{\pi^*}(t) > \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(t)$ and $\overline{N}_1^{\pi^*}(t) > 0$. Then there is an ϵ such that $\overline{N}_2^{\pi^*,r_k}(r_k s) > \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*,r_k}(r_k s)$ and $\overline{N}_1^{\pi^*,r_k}(r_k s) > 0$ for all r_k large enough and $s \in [t - \epsilon, t + \epsilon]$. Under policy π^* , in this interval both classes are served in parallel, hence $\frac{d\overline{T}_c^{\pi^*}(t)}{dt} = 1$.

Assume $\overline{N}_2^{\pi^*}(t) = \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(t)$ and $\overline{N}_1^{\pi^*}(t) > 0$. Then there is an ϵ such that $N_1^{\pi^*, r_k}(r_k s) > 0$ for all r_k large enough and $s \in [t - \epsilon, t + \epsilon]$. In this interval, class 2 is never served individually, so $\frac{d\overline{T}_1^{\pi^*}(s)}{ds} + \frac{d\overline{T}_c^{\pi^*}(s)}{ds} = 1$, for any regular point $s \in [t - \epsilon, t + \epsilon]$. Together with (30), we obtain

$$\begin{aligned}
\alpha \frac{\mu_2}{\mu_1} \frac{d\overline{N}_1^{\pi^*}(s)}{ds} - \frac{d\overline{N}_2^{\pi^*}(s)}{ds} &= \mu_2 \left(-\alpha \cdot (-\rho_1 + \frac{d\overline{T}_1^{\pi^*}(s)}{ds} + c_1 \frac{d\overline{T}_c^{\pi^*}(s)}{ds}) - \rho_2 + c_2 \frac{d\overline{T}_c^{\pi^*}(s)}{ds} \right) \\
&< \mu_2 \left(-\frac{c_2 - \rho_2}{c_1 - \rho_1} \cdot (-\rho_1 + \frac{d\overline{T}_1^{\pi^*}(s)}{ds} + c_1 \frac{d\overline{T}_c^{\pi^*}(s)}{ds}) - \rho_2 + c_2 \frac{d\overline{T}_c^{\pi^*}(s)}{ds} \right) \\
&= \mu_2 \left(-\frac{c_2 - \rho_2}{c_1 - \rho_1} \cdot (c_1 - \rho_1 + \frac{d\overline{T}_1^{\pi^*}(s)}{ds} (1 - c_1)) - \rho_2 + c_2 - c_2 \frac{d\overline{T}_1^{\pi^*}(s)}{ds} \right) \\
&= -\frac{d\overline{T}_1^{\pi^*}(s)}{ds} \cdot \frac{\mu_2}{c_1 - \rho_1} \cdot \left((c_2 - \rho_2)(1 - c_1) + c_2(c_1 - \rho_1) \right) \\
&= -\frac{d\overline{T}_1^{\pi^*}(s)}{ds} \cdot \frac{\mu_2 c_2}{c_1 - \rho_1} \cdot \left(1 - \rho_1 - \frac{\rho_2}{c_2} (1 - c_1) \right) \\
&\leq 0,
\end{aligned} \tag{36}$$

whenever $s \in [t - \epsilon, t + \epsilon]$ is a regular point. Here we used that $c_1 + c_2 > 1$, $\rho_1 < c_1 \leq 1$, $\rho_2 < 1 - \frac{\rho_1}{c_1}(1 - c_2)$, $\alpha > \frac{c_2 - \rho_2}{c_1 - \rho_1}$ and $\frac{d\overline{T}_1^{\pi^*}(s)}{ds} + \frac{d\overline{T}_c^{\pi^*}(s)}{ds} = 1$. Equation (36) implies that if at a certain time \overline{N}^{π^*} lies below the switching curve, then it moves towards the switching curve and if \overline{N}^{π^*} lies on or above the switching curve, it will move away from (and above) the switching curve. Since at time t we are in a state on the switching curve, we have $\overline{N}_2^{\pi^*}(s) < \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(s)$ for $s \in [t - \epsilon, t)$ and $\overline{N}_2^{\pi^*}(s) > \alpha \frac{\mu_2}{\mu_1} \overline{N}_1^{\pi^*}(s)$ for $s \in (t, t + \epsilon]$. Note that $\frac{d\overline{T}_1^{\pi^*}(t-)}{dt} = 1$ and $\frac{d\overline{T}_c^{\pi^*}(t-)}{dt} = 0$, while $\frac{d\overline{T}_1^{\pi^*}(t+)}{dt} = 0$ and $\frac{d\overline{T}_c^{\pi^*}(t+)}{dt} = 1$, so that the point t itself is not a regular point.

Finally assume $\overline{N}_1^{\pi^*}(t) = 0$ and $\overline{N}_2^{\pi^*}(t) > 0$. Then there is an $\epsilon > 0$ such that $\overline{N}_2^{\pi^*}(s) > \alpha \frac{\mu_1}{\mu_2} \overline{N}_1^{\pi^*}(s)$ for $s \in [t - \epsilon, t + \epsilon]$ and hence $N_2^{\pi^*, r_k}(r_k s) > \alpha \frac{\mu_1}{\mu_2} N_1^{\pi^*, r_k}(r_k s)$ for all r_k large enough and $s \in [t - \epsilon, t + \epsilon]$. Under policy π^* , in this interval class 1 is not served individually, hence $\frac{d\overline{T}_1^{\pi^*}(t)}{dt} = 0$. From (30) we then have

$$\frac{d\overline{N}_1^{\pi^*}(t)}{dt} = \lambda_1 - \mu_1 c_1 \frac{d\overline{T}_c^{\pi^*}(t)}{dt}. \tag{37}$$

Note that if $\overline{N}_1^{\pi^*}(t + \delta) > 0$, for all $0 < \delta < \Delta$, then $\frac{d\overline{T}_c^{\pi^*}(t + \delta)}{dt} = 1$. Since $\rho_1 < c_1$, from (37) we see that class 1 will stay empty, and thus $\frac{d\overline{T}_c^{\pi^*}(t)}{dt} = \frac{\rho_1}{c_1}$. We conclude that (32)–(34) are satisfied for each fluid limit $\overline{T}^{\pi^*}(t)$.

From (30) and (32)–(34) it follows that $\overline{N}_i^{\pi^*}(t)$ is uniquely determined. Using the correspondence $u_j^*(t) = \frac{d\overline{T}_j^{\pi^*}(t)}{dt}$, $j = 1, 2, c$, it follows from Proposition 4.5 that $\overline{N}^{\pi^*}(t) = n^*(t)$, with n^* as defined in Proposition 4.5. \square

In the next proposition it is stated that the linear switching curve provides a policy that is asymptotically fluid optimal for the original stochastic model.

Proposition 4.11 *Let $\mu_1 > \mu_1 c_1 + \mu_2 c_2$ and $c_1 + c_2 > 1$. If $\rho_1 < c_1$ and $\rho_2 < 1 - \frac{\rho_1}{c_1}(1 - c_2)$, then the policy π^* with switching curve $h(N_1) = \alpha \frac{\mu_2}{\mu_1} N_1$ is asymptotically fluid optimal, with α as defined in (24).*

Proof: Let r_k be a subsequence such that

$$\liminf_{r \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt = \lim_{k \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi^*,r_k}(t) + \overline{N}_2^{\pi^*,r_k}(t)) dt.$$

From Lemma 4.7 it follows that for almost all ω there exists a subsequence r_{k_l} of r_k such that $\lim_{l \rightarrow \infty} \overline{N}^{\pi^*,r_{k_l}}(t) = \overline{N}^{\pi^*}(t)$, u.o.c.. Since every fluid limit $\overline{N}^{\pi^*}(t)$ coincides with the optimal fluid control solution $n^*(t)$ (see (35)) we obtain $\lim_{l \rightarrow \infty} \overline{N}_i^{\pi^*,r_{k_l}}(t) = n_i^*(t)$, $i = 1, 2$. Since the functions $\overline{N}_i^{\pi^*,r_{k_l}}(t)$, $i = 1, 2$, converge uniformly on $[0, D]$, we can interchange the limit and the integral, so that

$$\liminf_{r \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt = \lim_{l \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi^*,r_{k_l}}(t) + \overline{N}_2^{\pi^*,r_{k_l}}(t)) dt = \int_0^D (n_1^*(t) + n_2^*(t)) dt.$$

The same holds for the lim sup and we can conclude that

$$\lim_{r \rightarrow \infty} \int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt = \int_0^D (n_1^*(t) + n_2^*(t)) dt. \quad (38)$$

We also have that $\int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt$ is uniformly integrable. This follows from the same argument as in the proof of [14, Lemma 4.5]. Here we state it briefly. Note that $\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t) \leq \frac{n_1^r + n_2^r}{r} + \frac{E_1(rt) + E_2(rt)}{r}$, with $E_i(\cdot)$ a Poisson process with rate λ_i . Since $\lim_{r \rightarrow \infty} \frac{E_1(rt) + E_2(rt)}{r} = (\lambda_1 + \lambda_2)t$ almost surely (see Lemma 4.7) and $\mathbb{E}(\frac{E_1(rt) + E_2(rt)}{r}) = (\lambda_1 + \lambda_2)t$, we obtain from [9, Theorem 3.6] that $\frac{E_1(rt) + E_2(rt)}{r}$ is uniformly integrable. Since $D < \infty$, uniform integrability of $\int_0^D (\frac{E_1(rt) + E_2(rt)}{r}) dt$ follows as well. Hence, by definition of uniform integrability it is immediate that $\int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt$ is uniformly integrable.

We then obtain

$$\begin{aligned} \lim_{r \rightarrow \infty} \mathbb{E} \left(\int_0^D (\overline{N}_1^{\pi^*,r}(t) + \overline{N}_2^{\pi^*,r}(t)) dt \right) &= \mathbb{E} \left(\int_0^D (n_1^*(t) + n_2^*(t)) dt \right) \\ &= \int_0^D (n_1^*(t) + n_2^*(t)) dt, \end{aligned} \quad (39)$$

where in the first step we used uniform integrability together with equation (38) to interchange the limit and expectation (see [9, Theorem 3.5]). Hence, policy π^* is asymptotically fluid optimal. \square

4.2.2 Case $\rho_1 > c_1$

In this section we consider the case $\rho_1 > c_1$. In Proposition 4.6 we found that for the fluid control problem it is optimal to give class 1 priority whenever present. A straightforward translation of this policy to the original stochastic model would be to give preemptive priority to class-1 users. However, the stability conditions under this policy are $\rho_1 + \rho_2 < 1$, which are more stringent than the necessary stability conditions as given in (4) and (5). Hence, a more precise interpretation of this fluid control is needed to avoid an unstable system.

Note that in the fluid control model, the policy that gives class 1 priority can keep the system stable under (4) and (5), since on the vertical axis it partly serves class 1 individually and partly serves both classes in parallel. This suggests that for the stochastic model we should serve both classes 1 and 2 in parallel when the process moves close to the vertical axis. So there is a switching curve in the original model that lies close to the vertical axis such that it is non-observable in the fluid limit. In the next conjecture we state that a policy with a switching

curve of the shape $h(N_1) = e^{N_1/\gamma}$ is an asymptotically fluid optimal policy when $\gamma > 0$ is large enough. We make no claim for small γ .

Proving the conjecture requires use of martingale equations similar to the proof of Lemma A.3 in [17], which we were unable to establish here. Following the conjecture, we sketch a proof that relies on these martingale equations.

Conjecture 4.12 *Let $\mu_1 > \mu_1 c_1 + \mu_2 c_2$ and $c_1 + c_2 > 1$. Assume $\rho_1 > c_1$, $\rho_2 < c_2$ and $\rho_1 < 1 - \frac{\rho_2}{c_2}(1 - c_1)$. The policy π^* with switching curve $h(N_1) = e^{N_1/\gamma}$ is asymptotically fluid optimal for $\gamma > 0$ large enough.*

The fluid limit of policy π^ is uniquely determined by $\overline{N}^{\pi^*}(t) = n^*(t)$, with $n^*(t)$ the trajectory corresponding to the control $u^*(t)$ as defined in Proposition 4.6.*

Sketch of proof: As noted above, the conjecture can be proved using martingale equations similar to those in the proof of Lemma A.3 in [17]. We believe such equations are valid, but were unable to formalize this. The following sketch of proof can be turned into a formal proof if the validity of these equations is proved.

Let $\overline{N}_i^{\pi^*}(t)$, $i = 1, 2$, $\overline{T}_j^{\pi^*}(t)$, $j = 1, 2, c, 0$, be a fluid limit of policy π^* . The function $\overline{T}_j^{\pi^*}(\cdot)$ is absolutely continuous. Our conjecture is proved if we can show that for each regular point t , the derivatives satisfy:

$$\frac{d\overline{T}_1^{\pi^*}(t)}{dt} = 1, \quad \text{if } \overline{N}_1^{\pi^*}(t) > 0, \quad (40)$$

$$\frac{d\overline{T}_1^{\pi^*}(t)}{dt} = \frac{\rho_1 - c_1}{1 - c_1} \quad \text{and} \quad \frac{d\overline{T}_c^{\pi^*}(t)}{dt} = \frac{1 - \rho_1}{1 - c_1}, \quad \text{if } \overline{N}_1^{\pi^*}(t) = 0, \quad (41)$$

for γ large enough. We expect that this result can be obtained by using the same techniques as in [17, Theorem 7.1] and the following correspondence between the processes ξ_t , x_t and $T(t)$ in [17], and our equivalents: $\xi_t^3 = N_1^{\pi^*}(t)$, $\xi_t^1 = N_2^{\pi^*}(t)$, $x_t^3 = \overline{N}_1^{\pi^*}(t)$, $x_t^1 = \overline{N}_2^{\pi^*}(t)$, $T_{01}(t) = \overline{T}_1^{\pi^*}(t)$ and $T_{11}(t) = \overline{T}_c^{\pi^*}(t)$, and mapping our parameters $c_1, c_2, \mu_1, \mu_2, \lambda_1$ and λ_2 , such that the drifts in the interior of Figure 4 in [17] correspond to the drifts in our Figure 5. Note that the drifts on the boundaries cannot be matched, but this does not influence the fluid analysis. Assuming martingale equations similar to those in the proof of [17, Lemma A.3] can be verified, we can establish equations similar to (A.12) and (A.13) in [17], i.e., $\frac{d\overline{T}_1^{\pi^*}(t)}{dt} = 1$ if $\overline{N}_1^{\pi^*}(t) > 0$, and $\frac{d\overline{T}_1^{\pi^*}(t)}{dt} + \frac{d\overline{T}_c^{\pi^*}(t)}{dt} = 1$ if $\overline{N}_1^{\pi^*}(t) = 0$. Since $\frac{1}{\mu_1} \frac{d\overline{N}_1^{\pi^*}(t)}{dt} = \rho_1 - \frac{d\overline{T}_1^{\pi^*}(t)}{dt} - c_1 \frac{d\overline{T}_c^{\pi^*}(t)}{dt}$ and $\rho_1 < 1 = \frac{d\overline{T}_1^{\pi^*}(t)}{dt}$ when $\overline{N}_1^{\pi^*}(t) > 0$, class 1 remains empty once it hits zero. Hence, $\frac{d\overline{T}_1^{\pi^*}(t)}{dt} + c_1 \frac{d\overline{T}_c^{\pi^*}(t)}{dt} = \rho_1$, which implies (41).

Note that $\frac{d\overline{T}_2^{\pi^*}(t)}{dt} = 0$, i.e., that the unscaled process does not stay long on the vertical axis. Hence, any capacity lost when serving class 2 individually, is negligible under fluid scaling.

From (30), (40) and (41) it follows that $\overline{N}_i^{\pi^*}(t)$ is uniquely determined. Using the correspondence $u_j^*(t) = \frac{d\overline{T}_j^{\pi^*}(t)}{dt}$, $j = 1, 2, c$, it follows from Proposition 4.6 that $\overline{N}^{\pi^*}(t) = n^*(t)$, with $n^*(t)$ the trajectory corresponding to the control $u^*(t)$ as defined in Proposition 4.6. The proof of the conjecture can then be completed similar as in the proof of Proposition 4.11. \square

4.3 Exponential switching curves

When $\rho_1 < c_1$, an asymptotically fluid optimal policy can be characterized by a linear switching curve and the slope of this curve has been exactly determined. When $\rho_1 > c_1$, Conjecture 4.12

states that an exponential switching curve $h(N_1) = e^{N_1/\gamma}$ is asymptotically fluid optimal for any γ that is *large enough*. The purpose of this section is to determine a reasonable *rule of thumb* for the choice of γ . This rule of thumb should avoid possible poor performance that may result when choosing very small values of γ (for which Conjecture 4.12 is inconclusive).

An asymptotically fluid optimal policy π^* satisfies

$$\begin{aligned} \mathbb{E}\left(\int_0^{r \cdot D} (N_1^{\pi^*,r}(t) + N_2^{\pi^*,r}(t))dt\right) &= r^2 \cdot \mathbb{E}\left(\int_0^D (\bar{N}_1^{\pi^*,r}(t) + \bar{N}_2^{\pi^*,r}(t))dt\right) \\ &= r^2 \int_0^\infty (n_1^*(t) + n_2^*(t))dt + o(r^2). \end{aligned}$$

Hence, one way to determine a reasonable value for γ is by choosing that value for γ that minimizes the next order term, $o(r^2)$. For the discrete-time version of our model, it is possible to find an estimate of this term under exponential switching curves, using the techniques of [18]. Consider a discrete-time system with Bernoulli arrivals. In an interval of length Δ , a class- i user arrives with probability $\lambda_i \Delta$, and it leaves the system with probability $\mu_i s_i \Delta$, $s \in S$ (with S the capacity region as defined in Section 2). We are interested in policies with exponential switching curves, i.e., $s = (1, 0)$ if $N_2 < e^{N_1/\gamma}$, and $s = (c_1, c_2)$ if $N_2 \geq e^{N_1/\gamma}$ and $N_1 > 0$. When $\Delta \rightarrow 0$, this approximates the continuous-time system with Poisson arrivals and exponential distributed service requirements. (The user departure rate in the discrete model is $\mu_i s_i$, which is equal to the user departure rate in the stochastic model.) For a given parameter γ of the switching curve, denote the state at time k by $N_i^\gamma(k)$, $i = 1, 2$.

Following the reasoning in [18] we consider different realizations of the queue length process, indexed by a superscript $r \in \mathbb{N}$. We take as initial point $n^r = (\gamma \ln[rn_2], [rn_2])$ and as time horizon $r \cdot D$ for some fixed D with $n_2 > D$. We then write $\mathbb{E}(\sum_{k=0}^{r \cdot D} N_1^{\gamma,r}(k) + N_2^{\gamma,r}(k)) = \sum_{k=1}^4 V_k^\gamma(n^r)$ with

$$\begin{aligned} V_1^\gamma(n^r) &= \sum_{k=0}^{r \cdot D} \mathbb{E}(N_1^{\gamma,r}(k)), \\ V_2^\gamma(n^r) &= \sum_{k=0}^{r \cdot D} (n_2^r + k(\lambda_2 - \frac{1 - \rho_1}{1 - c_1} \mu_2 c_2)), \\ V_3^\gamma(n^r) &= \sum_{k=0}^{r \cdot D} \frac{\mu_2 c_2}{\mu_1 (1 - c_1)} (n_1^r - \mathbb{E}(N_1^{\gamma,r}(k))), \\ V_4^\gamma(n^r) &= \sum_{k=0}^{r \cdot D} \mu_2 \frac{c_1 + c_2 - 1}{1 - c_1} \mathbb{E}(v_k^{\gamma,r}), \end{aligned}$$

where $v_k^{\gamma,r} = \sum_{m=0}^{k-1} \mathbf{1}_{(N_1^{\gamma,r}(m)=0)}$ is the number of times the process serves class 2 individually. Since $c_1 < \rho_1 < 1$ and $\rho_1 < 1 - \frac{\rho_2}{c_2}(1 - c_1)$, we can use the large-deviation results in [18] to show that

$$\begin{aligned} V_1^\gamma(n^r) &= Dr\gamma \ln(r) + O(r), \\ V_2^\gamma(n^r) &= r^2 \int_0^\infty (n_1^*(t) + n_2^*(t))dt + O(r), \\ V_3^\gamma(n^r) &= O(r) \\ V_4^\gamma(n^r) &= \mu_2 \frac{c_1 + c_2 - 1}{1 - c_1} \cdot r^{2-\beta(\Delta)\gamma+o(1)}, \end{aligned}$$

as $r \rightarrow \infty$, with $\beta(\Delta) = \ln(\frac{\rho_1}{c_1} \frac{1 - \mu_1 c_1 \Delta}{1 - \lambda_1 \Delta})$, and $n^*(t)$ the optimal fluid trajectory corresponding to initial state $(0, n_2)$. As explained in [18], setting the value of γ larger than $1/\beta(\Delta)$ gives good

second-order asymptotics. The condition that γ should be large enough is natural: Setting the value of γ near 0 would almost everywhere give priority to class 1, a strategy that we know can perform poorly. In our numerical experiments in Section 6.2 (for the continuous-time setting), this is indeed observed for large loads, as the performance severely degrades for small values of γ (see Figure 12). For large values of γ , performance is also suboptimal: When $\gamma > 1/\beta(\Delta)$, the second-order term is given by $D\gamma r \ln r$, so that it is not attractive to choose γ too large either. However, performance turns out to be less sensitive to small changes in γ for large values of γ , see also our numerical experiments in Section 6.2.

Letting $\Delta \rightarrow 0$, we have $\lim_{\Delta \rightarrow 0} \beta(\Delta) = \ln(\frac{\rho_1}{c_1})$. Choosing as estimate $\gamma = 1/\ln(\frac{\rho_1}{c_1})$ in the continuous-time system proves to be a reasonable *rule of thumb* in all our experiments, see Section 6.2.

5 Heavy-traffic regime

One of the goals of this paper is to describe policies that approximate optimal policies rather well (in cases where an optimal policy could not be determined explicitly). In Section 4 we did so by considering a simpler (fluid) model that only took into account the mean drifts. We proved that certain policies are asymptotically fluid optimal, and therefore are potentially close to optimal in the original stochastic model as well. In this section we discuss another approach to obtain policies that are in some sense approximately optimal: We review optimality results available in the literature for a heavy-traffic regime (when the system is close to saturation). These results can be used as approximations for the original system when the load is rather high, however, there is no guarantee for the performance of these policies in moderately-loaded systems. Therefore, in Section 6 we numerically compare (under moderate load conditions) the performance of the policies that are optimal in heavy traffic with our asymptotically fluid optimal policies. Note that both of these policies are motivated by a certain asymptotic regime, and beforehand it is unclear how well they perform outside these regimes.

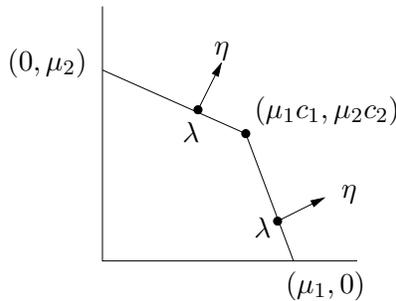


Figure 7: Stability set.

We know the system can be kept stable when (4) and (5) are satisfied. Equivalently, we may say that the system can be kept stable when the vector (λ_1, λ_2) lies in the interior of the stability set as depicted in Figure 7. The system is said to be in heavy traffic when the vector (λ_1, λ_2) lies on the northeast boundary of the stability set in Figure 7. In addition, the complete resource pooling condition is satisfied if the outer normal, η , to the stability set at that λ is unique up to scaling and all its coordinates are strictly positive, i.e., (λ_1, λ_2) is such that $\lambda_i > 0$ for $i = 1, 2$ and $(\lambda_1, \lambda_2) \neq (\mu_1 c_1, \mu_2 c_2)$. More precisely, the parameters of a heavily-loaded system under the resource pooling condition correspond to one of the following two regions (see also Figure 7):

- Region A: $\rho_2 = 1 - \frac{\rho_1}{c_1}(1 - c_2)$ and $\rho_2 > c_2$, $c_1 > \rho_1$. The outer normal vector to a point in this region is $\eta = (\mu_2(1 - c_2), \mu_1 c_1)$.

- Region B: $\rho_1 = 1 - \frac{\rho_2}{c_2}(1 - c_1)$ and $\rho_1 > c_1, c_2 > \rho_2$. The outer normal vector to a point in this region is $\eta = (\mu_2 c_2, \mu_1(1 - c_1))$.

Recall that our model may be viewed as a parallel two-server model, see Remark 2.1. Policies that are in some sense asymptotically optimal in a heavy-traffic setting with complete resource pooling have been investigated in among others [5, 6, 24, 32]. In Section 5.1 we briefly state the results of Bell and Williams [5, 6]. They prove that threshold-based policies asymptotically minimize the (scaled) total number of users in heavy traffic. In Section 5.2 we recall the definition of Max-Weight policies (or $Gc\mu$ -rule) and describe the results concerning their behavior in heavy traffic as obtained by Mandelbaum and Stolyar [24, 32].

5.1 Threshold policies

Bell and Williams [5, 6] have investigated the parallel-server model (with an arbitrary number of servers and classes) with i.i.d. interarrival times and service requirements, and with FCFS as intra-class policy. (In the case of exponential service requirements, the behavior of the system is independent of the non-anticipating intra-class policy.) In this section we collect results specific for the parallel two-server model. Their model is in fact a slight variation to the model we consider in this paper. First, in their model, once a server starts serving a user, this user has to obtain its full service from this server. Secondly, their model has slightly different behavior near the boundaries: when $N_i = 1$, their model can have a departure rate of at most $\mu_i c_i$ for class i , since a single user cannot be served simultaneously by the two servers. In the model we consider, we can have a departure rate of μ_i .

Bell and Williams [5, 6] consider a sequence of parameters indexed by r , μ_i^r and λ_i^r ($\rho_i^r = \frac{\lambda_i^r}{\mu_i^r}$), with $\lambda_i^r \rightarrow \lambda_i, \mu_i^r \rightarrow \mu_i$ such that $\lambda_1, \lambda_2, \mu_1$ and μ_2 correspond either to Region A or Region B. An additional condition involves the rate at which the system converges:

$$\lim_{r \rightarrow \infty} \sqrt{r} \mu_i^r (\rho_i^r - \rho_i) = \theta_i, \quad \text{with } \theta_i \in \mathbb{R}, i = 1, 2.$$

Let $N_i^{\pi, r}(t)$ be the number of class- i users in the r -th system under policy π , and let $\hat{N}_i^{\pi, r}(t) = \frac{N_i^{\pi, r}(rt)}{\sqrt{r}}$ be the diffusion scaled number of class- i users. It is assumed that the system is initially empty. Define $\hat{J}^r(\pi) = \mathbb{E}(\int_0^\infty e^{-\xi t} (\hat{N}_1^{\pi, r}(t) + \hat{N}_2^{\pi, r}(t)) dt)$ where $\xi > 0$ is a constant. In [5, 6], a sequence of policies $\tilde{\pi}^r$ is called *asymptotically optimal in heavy traffic* when $\lim_{r \rightarrow \infty} \hat{J}^r(\tilde{\pi}^r) \leq \liminf_{r \rightarrow \infty} \hat{J}^r(\pi^r)$ for any sequence of policies π^r .

When $\mu_1 \leq c_1 \mu_1 + c_2 \mu_2$, the optimal policy is to serve both classes in parallel whenever possible. This remains valid in heavy traffic. For $\mu_1 > c_1 \mu_1 + c_2 \mu_2$ the following result holds:

Proposition 5.1 ([6]) *Assume $\mu_1 > c_1 \mu_1 + c_2 \mu_2, c_1 + c_2 > 1$ and consider a heavy-traffic setting with complete resource pooling.*

- *If (ρ_1, ρ_2) corresponds to Region A, then the policy that serves classes 1 and 2 in parallel whenever possible, is an asymptotically optimal policy in heavy traffic.*
- *If (ρ_1, ρ_2) corresponds to Region B, then the sequence of threshold policies that gives priority to class 1 when $N_1 > c \ln(\sqrt{r})$ (with c large enough), and that otherwise serves classes 1 and 2 in parallel, is asymptotically optimal in heavy traffic.*

Denote by $Th(r)$ the minimum value for the threshold Th such that the r -th system is stable under the threshold policy that serves class 1 individually when $N_1(t) > Th$ and serves classes 1 and 2 in parallel otherwise. In [34] it is shown that any threshold Th with $Th > Th(r)$, makes the r -th system stable. In addition, $Th(r)/\ln(\sqrt{r}) \rightarrow \hat{c}$ for some constant $\hat{c} > 0$. This shows that the threshold $c \cdot \ln(\sqrt{r})$ in the above proposition is of a minimum order.

In Section 6.2 we will evaluate the performance of threshold-based policies in the moderately-loaded case, and compare it with the optimal policy found numerically, and with the asymptotically fluid optimal policies as we proposed in this paper.

5.2 Max-Weight policies

In this section we summarize results on Max-Weight policies cf. [24, 32]. For a parallel-server system with K classes and L servers, the $Gc\mu$ -rule (Max-Weight policy is a special case of this) is defined as follows. When server l becomes free, it starts serving a user from class k such that $k = \arg \max_i \mu_{il} C'_i(N_i)$, and serves this user until it leaves the system. Here μ_{il} is the departure rate of class- i users when served by server l , and the function $C_i(N_i)$ can be interpreted as the cost of having N_i users in class i . In order for the heavy-traffic results to hold (which will be stated later on), the function $C_i(N_i)$ needs to satisfy certain conditions as specified in [24]. In particular, the second derivative needs to be strictly positive and continuous in $(0, \infty)$. This excludes the function $C_i(N_i) = N_i$, which would be needed to minimize the total number of users. We will focus on functions of the type $C_i(N_i) = \gamma_i N_i^{\beta+1}$, with parameters $\beta, \gamma_i > 0$, which do satisfy the conditions. These cost functions correspond to the Max-Weight policies. An important property of the Max-Weight policies is that they maintain a stable system under the necessary stability conditions [32].

Mandelbaum and Stolyar [24] consider i.i.d. inter-arrival times and service requirements. They consider a sequence of systems indexed by r , λ_i^r , with $\lambda_i^r \rightarrow \lambda_i$, and keep μ_{il} fixed. The parameters λ_i and μ_{il} are such that the system is in heavy traffic and the complete resource pooling condition is satisfied. In addition, $\lim_{r \rightarrow \infty} \sqrt{r}(\lambda_i^r - \lambda_i) = \theta_i$, with $\theta_i \in \mathbb{R}$, $i = 1, \dots, K$. The initial state converges under the diffusion scaling such that $\lim_{r \rightarrow \infty} \frac{N_i^r(0)}{\sqrt{r}} = m_i$ with $(\gamma_1 m_1^\beta, \dots, \gamma_K m_K^\beta)$ proportional to (η_1, \dots, η_K) .

The authors of [24] focus on policies with FCFS as intra-class policy, and once a user is taken for service by a server, this user cannot be served by any other server. The next proposition states the results for the Max-Weight policy, which will be denoted by MW . In particular, for a heavy traffic setting it states that the Max-Weight policy minimizes (under diffusion scaling) both the cost, $\sum_i C_i(N_i(t)) = \sum_i \gamma_i \cdot (N_i(t))^{\beta+1}$, and the “virtual” workload, $\sum_i \eta_i N_i(t)$, at all times. As before, η is the outer normal vector to the stability set.

Proposition 5.2 ([24]) *Consider a heavy-traffic setting with complete resource pooling. For any policy π that is allowed in the framework of [24] (as described above), it holds that*

$$\liminf_{r \rightarrow \infty} \sum_i \gamma_i \cdot (\hat{N}_i^{\pi,r}(t))^{\beta+1} \geq \lim_{r \rightarrow \infty} \sum_i \gamma_i \cdot (\hat{N}_i^{MW,r}(t))^{\beta+1},$$

and,

$$\liminf_{r \rightarrow \infty} \sum_i \eta_i \hat{N}_i^{\pi,r}(t) \geq \lim_{r \rightarrow \infty} \sum_i \eta_i \hat{N}_i^{MW,r}(t).$$

for all time t . In addition, the vector

$$\lim_{r \rightarrow \infty} (\gamma_1 \cdot (\hat{N}_1^{MW,r}(t))^\beta, \dots, \gamma_K \cdot (\hat{N}_K^{MW,r}(t))^\beta), \quad (42)$$

is proportional to (η_1, \dots, η_K) .

The result that the vector in (42) is proportional to (η_1, \dots, η_K) , is referred to as a state space collapse, since the dimension of the K -dimensional process decreases to one.

Note that the Max-Weight policy does not minimize the total number of users, since β must be strictly positive. However, the Max-Weight policy can be used to come close to this setting,

for example, by setting $\beta > 0$ very small and $\gamma_i = 1, i = 1, \dots, K$. An alternative option is by making use of the fact that the Max-Weight policy does minimize the (diffusion-scaled) virtual workload $\sum_i \eta_i N_i(t)$. Hence, when trying to minimize the total number of users among the Max-Weight policies, it is best to set the parameters (γ_i 's and β) such that $N_k^{MW}(t)$ is as large as possible, where k is such that $\eta_k \geq \eta_i$ for all $i \neq k$. For this reason, in [24] it is suggested that in heavy traffic a good choice for the parameters is $\beta = 1$, $\gamma_i = 1, i \neq k$ and $\gamma_k = \epsilon_k$, with $\epsilon_k > 0$ small, since the state space collapse result implies that then $N_k^{MW}(t)$ will become relatively large compared to $N_i^{MW}(t), i \neq k$.

The model that we consider in this paper, is closely matched² by the parallel two-server model as considered in [24] when taking $\mu_{11} = \mu_1 c_1, \mu_{21} = \mu_2(1 - c_2), \mu_{12} = \mu_1(1 - c_1)$ and $\mu_{22} = \mu_2 c_2$, see Remark 2.1. When $c_1 + c_2 > 1$ and $c_i \leq 1$, the corresponding Max-Weight policy is as follows:

- Serve class 1 when $N_2 < \left(\frac{\gamma_1(1-c_1)\mu_1}{\gamma_2 c_2 \mu_2} \right)^{\frac{1}{\beta}} N_1$.
- Serve classes 1 and 2 in parallel when $\left(\frac{\gamma_1(1-c_1)\mu_1}{\gamma_2 c_2 \mu_2} \right)^{\frac{1}{\beta}} N_1 \leq N_2 < \left(\frac{\gamma_1 c_1 \mu_1}{\gamma_2 (1-c_2)\mu_2} \right)^{\frac{1}{\beta}} N_1$.
- Serve class 2 when $\left(\frac{\gamma_1 c_1 \mu_1}{\gamma_2 (1-c_2)\mu_2} \right)^{\frac{1}{\beta}} N_1 \leq N_2$.

Hence the Max-Weight policy has two linear switching curves. In Figure 8 these switching

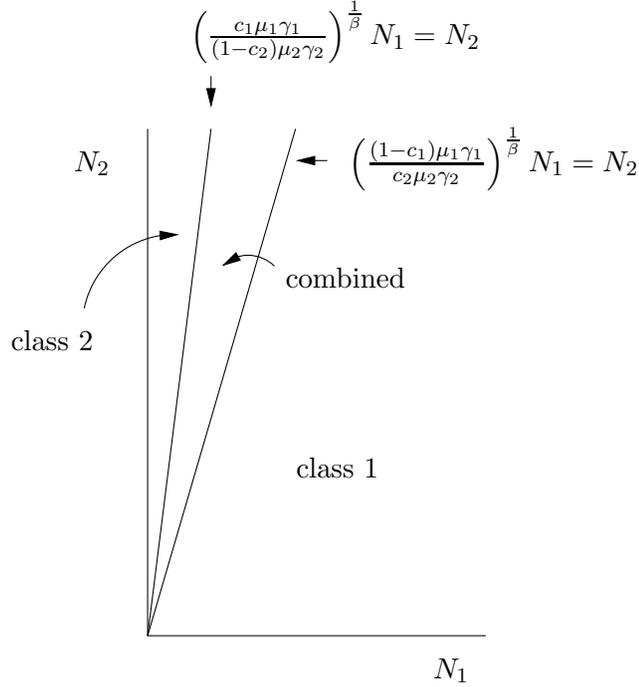


Figure 8: The Max-Weight policy.

curves are plotted. Note that in heavy traffic, the state space collapses to the line $N_2 = \left(\frac{c_1 \mu_1 \gamma_1}{(1-c_2) \mu_2 \gamma_2} \right)^{\frac{1}{\beta}} N_1$ if we are in Region A and to the line $N_2 = \left(\frac{(1-c_1) \mu_1 \gamma_1}{c_2 \mu_2 \gamma_2} \right)^{\frac{1}{\beta}} N_1$ if we are in Region B.

²Note that the two models are not exactly equivalent since the possible allocations in states with $N_i = 1$, for an $i = 1, 2$, are not identical.

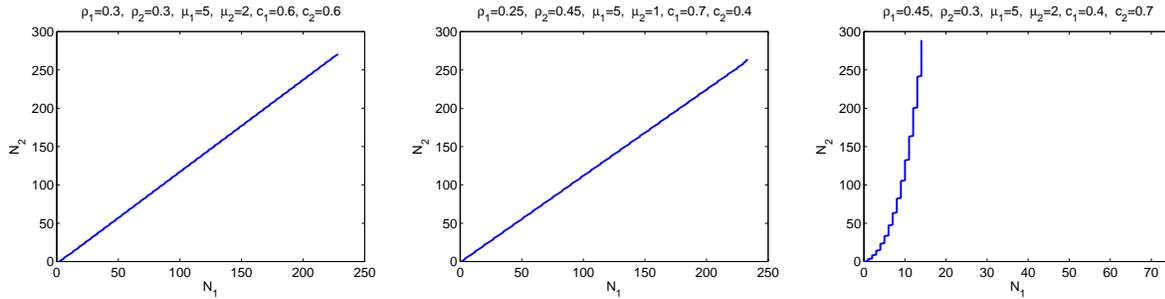


Figure 9: Optimal switching curve when a) $\rho_1 < c_1$, $\rho_2 < c_2$, b) $\rho_1 < c_1$ and $\rho_2 > c_2$ and c) $\rho_1 > c_1$ and $\rho_2 < c_2$.

In Section 6.3 we will investigate the performance of the Max-Weight policies in the moderately-loaded case, and compare it with the optimal policy found numerically, and with the asymptotically fluid optimal policies as we proposed in this paper.

6 Numerical results

An average-optimal policy for the original stochastic model can be computed numerically by value iteration using appropriate truncation of the state space. Figure 9 illustrates for various scenarios that an optimal policy is characterized by a switching curve. We note that finding these optimal curves numerically was extremely time-consuming. Figures 9 a) and b) consider the setting $\rho_1 < c_1$. We see that the switching curve is linear and coincides exactly with the asymptotically fluid optimal switching curve $h(N_1) = \alpha \frac{\mu_2}{\mu_1} N_1$ from Proposition 4.11. Figure 9 c) corresponds to a scenario with $\rho_1 > c_1$ and illustrates that then an optimal policy resembles an exponentially shaped curve, which agrees with Conjecture 4.12. In the remainder of this section we will assess the gains that can be achieved by choosing the best switching-curve policies.

6.1 Linear switching curve policies for $\rho_1 < c_1$

In Figure 10 we focus on the case $\rho_1 < c_1$ and plot the mean total number of users under policies with a linear switching curve $h(N_1) = dN_1$ (obtained by simulation). On the horizontal axis we vary the value of d . Note that $d = 0$ corresponds to always serving both classes in parallel. When the slope grows large ($d \rightarrow \infty$), the policy gives higher priority to serving class 1 exclusively (whenever present). Note that strict priority for class 1 leads to instability if $\rho_1 + \rho_2 > 1$, which can be the case even if the stability conditions (4) and (5) are met. The two graphs on the left in Figure 10 correspond to a moderately-loaded system. There we also plot the optimal policy found numerically by value iteration. We observe that when the parameter d is chosen well, the linear switching curve policy coincides with the optimal policy. The two graphs on the right in Figure 10 represent a heavily-loaded system. We did not determine the optimal policy for this parameter setting, since this is extremely time-consuming. Choosing d very large implies that the mean number of users will be large (since $\rho_1 + \rho_2 > 1$). It seems that a good choice for heavily-loaded systems is $d = 0$, i.e., always serve both classes in parallel. In a heavy-traffic setting with $\rho_1 < c_1$ (and necessarily $\rho_2 > c_2$ while $\rho_2 + \frac{\rho_1}{c_1}(1 - c_2) \rightarrow 1$) we see that the policy that always serves both classes in parallel is also the asymptotically optimal policy as found by both the fluid analysis (since then $\alpha = 0$, so the slope is equal to 0) and the heavy-traffic analysis.

In Figure 11 we repeated the experiment for different parameter choices to illustrate that the relative differences in performance between the optimal linear policy (obtained numerically by

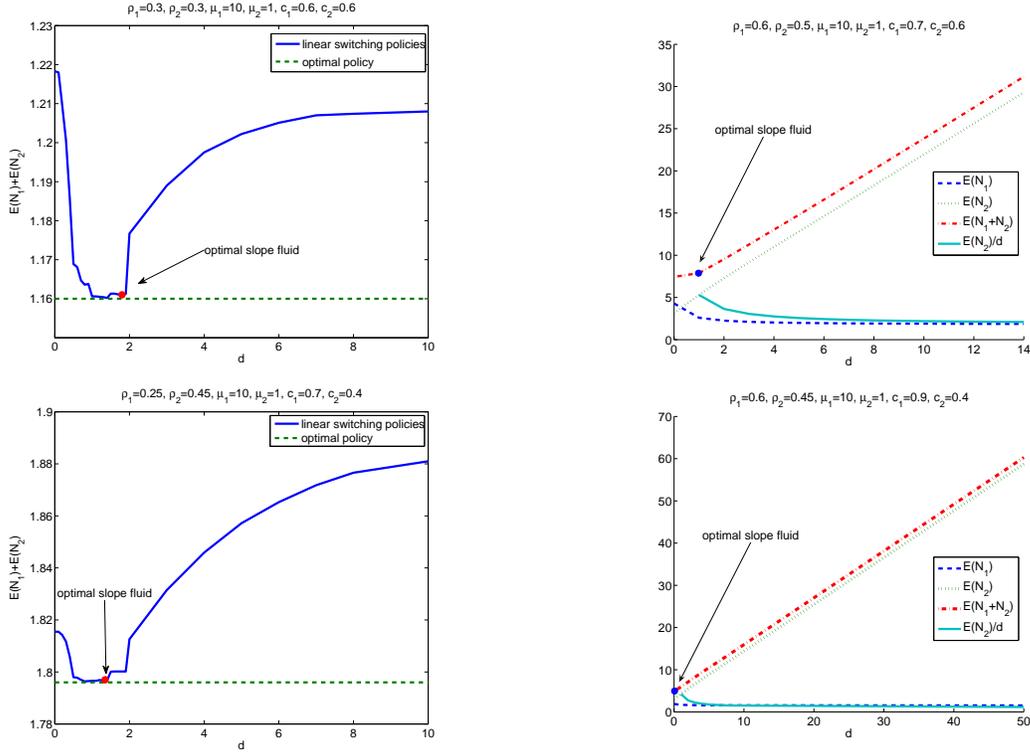


Figure 10: Mean total number of users for policies with a linear switching curve. The marker indicates the optimal slope for the fluid approximation. The two graphs on the top row correspond to cases with $\rho_1 < c_1$ and $\rho_2 < c_2$. The lower graphs have $\rho_1 < c_1$ and $\rho_2 > c_2$.

value iteration) and the strategy that maximizes the service capacity at all times (slope $d = 0$) can be quite significant.

An important observation in Figures 10 and 11 is that the asymptotically fluid optimal policy as found by the fluid analysis in Section 4 (denoted in the figures by “optimal slope fluid”) is always close to optimal and performs very well.

In the two graphs on the right in Figure 10, we observe that the total mean number of users grows linearly in d as $d \rightarrow \infty$. In the following remark we provide intuition for this effect.

Remark 6.1 Consider the policy with a linear switching curve $h(N_1) = dN_1$. If d tends to ∞ , then the system dynamics tends to a priority queue where class 1 is given preemptive priority. When $\rho_1 + \rho_2 < 1$, this policy is stable, and we indeed observe in the two graphs on the left in Figure 10 and in Figure 11 that the mean number of users will converge to a constant. However, when $\rho_1 + \rho_2 > 1$, this policy is not stable, and $\mathbb{E}(N_1 + N_2)$ will grow infinitely large as $d \rightarrow \infty$. The two graphs on the right in Figure 10 suggest that the mean number of users grows linearly in d as $d \rightarrow \infty$. This can be intuitively understood as follows.

Conditioned on $jd \leq N_2 < (j+1)d$, class 1 has as departure rate $\mu_1 c_1$ if $N_1 \leq j$, and μ_1 otherwise. For a given j , let $\pi(j)$ denote the equilibrium distribution for the process with departure rates as described above. Hence, $\pi_i(j) = \pi_0(j) \left(\frac{\rho_1}{c_1}\right)^i$ if $i \leq j$ and $\pi_i(j) = \pi_0(j) \left(\frac{\rho_1}{c_1}\right)^j \rho_1^{i-j}$ if $i > j$. If d is large, we assume that class 1 reaches equilibrium during the time that $jd \leq N_2 < (j+1)d$. Then the mean departure rate for class 2 is $\mu_2(j) := \mu_2 \pi_0(j) + \mu_2 c_2 \sum_{i=1}^j \pi_i(j)$ (when $jd \leq N_2 < (j+1)d$), since both classes are served in parallel whenever $N_2 \geq dN_1$. It can be checked that this is increasing in j , hence there exists a j^* such that $\mu_2(j^* - 1) < \lambda_2 \leq \mu_2(j^*)$ (for convenience we define $\mu_2(-1) = 0$). Note that $j^* > 0$, unless $\rho_1 + \rho_2 < 1$. Hence, if $jd \leq N_2 < (j+1)d$ with $j < j^*$, then the mean drift in class 2 is positive, and the probability that the increase in N_2 is

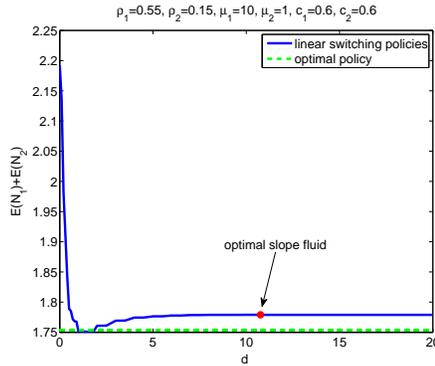


Figure 11: Mean total number of users for policies with a linear switching curve. The marker indicates the optimal slope for the fluid approximation.

$O(d)$ tends to 1 as $d \rightarrow \infty$. If $jd \leq N_2 < (j+1)d$ with $j \geq j^*$, then the mean drift in class 2 is negative. Hence, the probability that the decrement of N_2 is of order $O(d)$ tends to 1 as $d \rightarrow \infty$. It is therefore plausible that the process N_2/d will most of the time be around the level j^* .

If the region $(j^*+1)d \leq N_2$ is not reached (which is not a strong assumption, since this region will be rarely visited as $d \rightarrow \infty$), then the number of class-1 users can be upper bounded by the number of class-1 users in a system with departure rates $\mu_1 c_1$ if $N_1 \leq j^*$ and μ_1 otherwise. Since j^* does not depend on d , the upper bound for the number of class-1 users does not scale with d . For the parameters used in the graph on the top right in Figure 10, the j^* is equal to 2. We observe in the figure that $\mathbb{E}(N_2)/d$ indeed converges to $j^* = 2$ and that $\mathbb{E}(N_1)$ does not scale with d . For the parameters that belong to the graph on the bottom right in Figure 10, the j^* is equal to 1. In that case too, we observe in the figure that $\mathbb{E}(N_2)/d$ indeed converges to $j^* = 1$ and that $\mathbb{E}(N_1)$ does not scale with d .

6.2 Exponential switching curves and threshold-based policies for $\rho_1 > c_1$

In Figure 12 we consider several parameter settings with $\rho_1 > c_1$, and plot the total mean number of users under policies with switching curves of the shape $h(N_1) = e^{N_1/\gamma}$ (obtained by simulation). On the horizontal axis we vary the value of γ . Note that when γ grows large, this tends to the policy that always serves both classes in parallel. We observed that the best choice for the parameter γ , delivers virtually the same performance as the optimal policy (found numerically by value iteration). In Conjecture 4.12 it is stated that exponential switching curves are asymptotically fluid optimal. The large-deviation analysis further suggests that $\gamma = \frac{1}{\ln(\rho_1/c_1)}$ is a safe choice, see Section 4.3 (denoted in the figures by “rule of thumb”). In the three graphs on the top row in Figure 12 this corresponds to $\gamma = 8.5$. We observe that in fact the better choices for the parameter γ are smaller than 8.5. Still, the large deviations result gives a safe estimate (the policy is stable) with better performance than the capacity-maximizing strategy (serving both classes in parallel whenever possible, i.e., $\gamma \rightarrow \infty$). In the three graphs on the last row in Figure 12, the rule of thumb is equal to $\gamma = 2.5$. In this case, the rule of thumb is very close to the optimal performance. In general, in all our tests we observed that the rule of thumb for γ proves to be a reasonable choice.

Recall that when $\rho_1 > c_1$, a threshold policy is asymptotically optimal in a heavily-loaded system. That is, both classes should be served in parallel whenever the number of class-1 users is below or equal to some threshold $Th \geq 0$. When the threshold grows large, this coincides with the policy that always serves both classes in parallel. In Figure 12 we consider a moderately-loaded system. We vary the value of the threshold Th , and plot the mean total number of

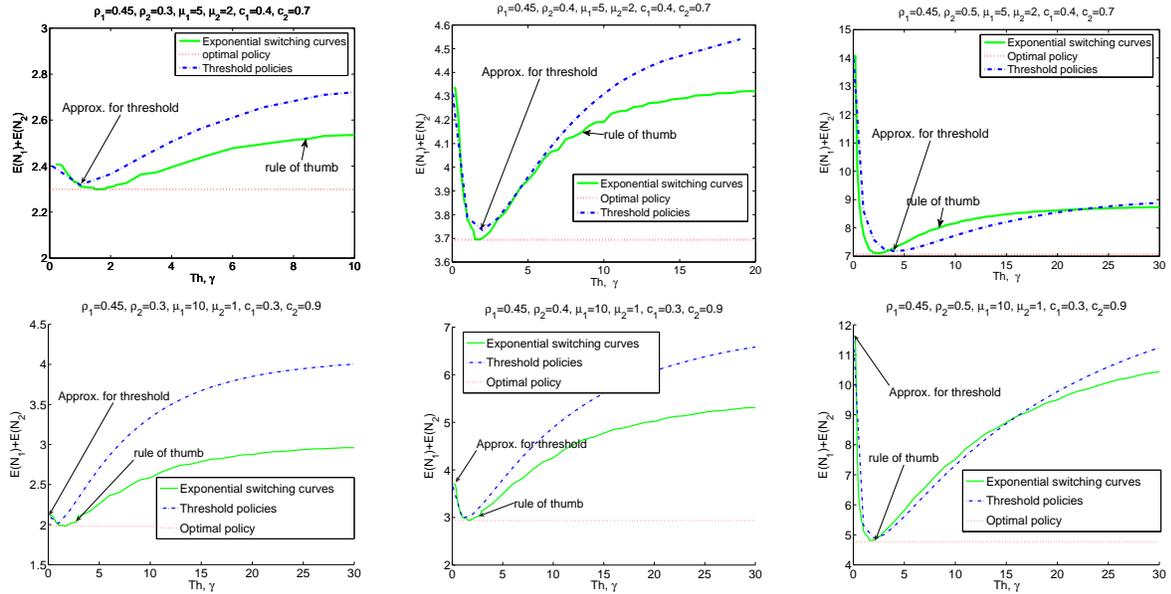


Figure 12: Mean total number of users when $\rho_1 > c_1$ and $\rho_2 < c_2$ for policies with exponential switching curves (as a function of γ), and for threshold policies (as a function of Th).

users (obtained by simulation). For certain small values of the threshold, the threshold policy performs rather well. However, when the threshold is chosen too small, the performance of the system can degrade considerably. In fact, for a system with large loads ($\rho_1 + \rho_2 > 1$), the policy with $Th = 0$ is unstable. In the two graphs on the right in Figure 12 (where still $\rho_1 + \rho_2 < 1$) we already see that the total number of users doubles when the threshold is set equal to 0. In [33] the authors propose a method to obtain estimates for the value of the threshold. For transparency of presentation, we do not describe this method here. For the settings in Figure 12 we have calculated the estimates for the threshold using their method (denoted in the figures by “Approx. for threshold”). We see that in the figures on the top, the approximation for the threshold matches exactly with the best threshold value. However, in the figures on the bottom, the approximation of the threshold is too small, which results in severe performance degradation in case of high loads ($\rho_1 = 0.45, \rho_2 = 0.5$).

In general, for the case $\rho_1 > c_1$ our fluid-based method (rule of thumb) proved to be a rather safe option, while threshold policies (using the approximation in [33]) may sometimes perform better, but can also be far from optimal. Although this is supported by a rather extensive set of experiments, it remains as a challenge to provide a theoretical basis for the robustness of fluid-based policies.

6.3 Comparison with Max-Weight policies for moderate loads

As stated in Section 5.2, Max-Weight policies can be close to optimal in a heavy-traffic setting. In this section we investigate the performance of the Max-Weight policies in a moderately-loaded system and compare this to the performance of the asymptotically fluid optimal policies as found in this paper. We need to distinguish between whether $\mu_1 c_1 + \mu_2 c_2 \geq \mu_1$ or $\mu_1 c_1 + \mu_2 c_2 < \mu_1$. We will see that in the second case the fluid-based policies can outperform the Max-Weight policies, and that the parameter choices for the Max-Weight policies as suggested by the heavy-traffic results are not necessarily a good choice in a moderately-loaded system.

Case $\mu_1 c_1 + \mu_2 c_2 \geq \mu_1$

From Section 3.1 we know that when $\mu_1 c_1 + \mu_2 c_2 \geq \mu_1$, the policy which serves classes 1 and 2 in parallel whenever possible, stochastically minimizes the total mean number of users present in the system. Note that when $\mu_1 c_1 + \mu_2 c_2 \geq \mu_1$, the Max-Weight policy with $\gamma_1 = \gamma_2 = 1$ and β close to zero, will almost always serve both classes in parallel. Numerically we observed that the Max-Weight policy (with $\gamma_1 = \gamma_2 = 1$ and β close to zero) turns out to be very effective and nearly matches the optimal performance. For this reason, we have not included any graphs for this case.

Case $\mu_1 c_1 + \mu_2 c_2 < \mu_1$

When $\mu_1 c_1 + \mu_2 c_2 < \mu_1$, the asymptotically fluid optimal policy we proposed is described by a switching curve $h(N_1)$ (either linear or exponential), where class 1 is served in states below the switching curve, and classes 1 and 2 are served in parallel in states above the switching curve. We compare these policies with Max-Weight policies. We choose the parameters as described in Section 5.2. So we take $\gamma_1 = \gamma_2 = 1$ and $\beta = 10^{-4}$. When $\mu_1 > \mu_2$ and $\mu_1 c_1 + \mu_2 c_2 < \mu_1$, we have $\eta_1 < \eta_2$, both in Region A and in Region B of Figure 7. Hence, we will also consider Max-Weight policies with $\gamma_1 = 1$, $\gamma_2 = \epsilon_2$, $\epsilon_2 > 0$, and $\beta = 1$.

In Figures 13 and 14, we compare (by simulation) the performance of the Max-Weight policies with the performance of the best linear or exponential switching curve policies. On the horizontal axis we vary ϵ_2 and on the vertical axis we plot the total mean number of users under the various policies. First of all, we note that in both Figures 13 and 14, the Max-Weight policy with $\beta = 10^{-4}$ and $\gamma_i = 1$, $i = 1, 2$, performs rather poorly. This is not surprising, since if $\mu_1 c_1 + \mu_2 c_2 < \mu_1$, then the Max-Weight policy (with $\beta = 10^{-4}$ and $\gamma_1 = \gamma_2 = 1$) will almost always serve class 1 individually, which is far from optimal.

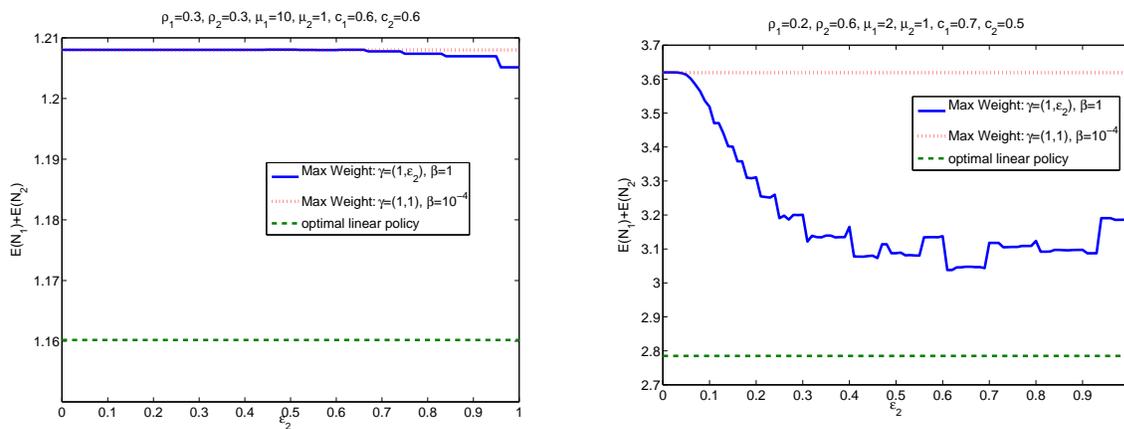


Figure 13: Mean total number of users under Max-Weight policies and under the optimal linear switching curve, with $\mu_1 c_1 + \mu_2 c_2 < \mu_1$ and $\rho_1 < c_1$.

For the parameters as in Figure 13 a), the fluid approximation suggests that if $N_2 \leq 1.8N_1$, then serve class 1, and otherwise serve both classes in parallel. Numerically, we found that this is also the best linear policy for the stochastic model. The Max-Weight policy (with $\beta = 1$ and $\gamma = (1, \epsilon_2)$) will serve class 1 most of the time, since that is the prescribed action in states such that $N_2 \leq 6\frac{2}{3\epsilon_2}N_1$. From the figure, we see that this is only 5% worse than the optimal linear policy. For the parameters as in Figure 13 b), the fluid approximation serves always classes 1 and 2 in parallel. Numerically, we found that this is also the best linear policy for the stochastic

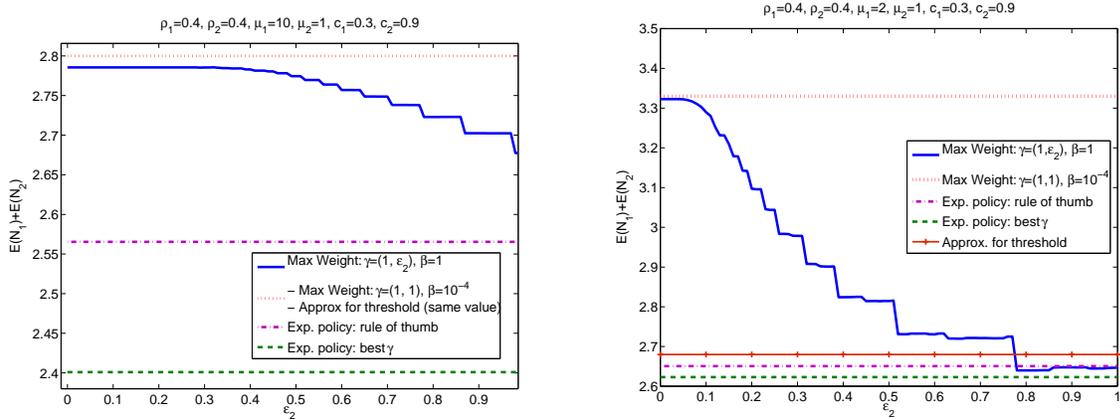


Figure 14: Mean total number of users under Max-Weight policies and under exponential switching curves, with $\mu_1 c_1 + \mu_2 c_2 < \mu_1$ and $\rho_1 > c_1$.

model. The Max-Weight policy (with $\beta = 1$ and $\gamma = (1, \epsilon_2)$) however, serves class 1 individually as soon as $N_2 \leq \frac{12}{10\epsilon_2} N_1$. These states will be visited more often when $\epsilon_2 \downarrow 0$. In the figure, the performance degrades from 15% worse ($\epsilon_2 = 1$), to 30% worse ($\epsilon_2 \downarrow 0$), compared with the optimal linear policy.

In Figure 14 the parameters are such that Conjecture 4.12 implies an exponential switching curve. We plot the performance of both the best exponential switching curve (determined numerically), and of the exponential switching curve where γ is set according to the rule of thumb, i.e., $\gamma = \frac{1}{\ln(\rho_1/c_1)} = 3.48$. For $\mu_1 = 10$, the Max-Weight policy (with $\beta = 1$ and $\gamma = (1, \epsilon_2)$) is about 15% worse compared with the best exponential policy. For $\mu_1 = 2$, it is close to optimal when $\epsilon_2 = 1$, but the performance degrades when $\epsilon_2 \downarrow 0$. Observe that in both cases the policy with an exponential switching curve where γ is chosen according to the rule of thumb, performs rather well. We have also calculated the performance of the threshold policy as suggested in [33]. For the setting of Figure 14 a) it suggests a threshold equal to 0, in which case there are approximately 2.8 users in the system. Hence, the proposed policy does not give good performance. For the setting of Figure 14 b) it suggests a threshold equal to 1, in which case there are approximately 2.68 users in the system. This is rather close to optimal.

7 Conclusion and future work

We have studied optimal policies for systems that have capacity gains when serving users in parallel. Fluid limit analysis indicates that asymptotically fluid optimal policies can be characterized by either linear or exponentially shaped switching curves. The results yield directly usable estimates for efficient policies in the stochastic setting, comparing favorably with threshold-based policies and Max-Weight policies for moderately-loaded regimes. A proof of Conjecture 4.12 seems within reach, using similar steps as in Lemma A.3 in [17]. At this stage we did not succeed to formalize the required martingale inequality used there.

Several extensions to this work are of interest. For example, it is interesting to investigate how our results change if the capacity is also favorably affected by the numbers of users within each class. For example, in wireless networks the aggregate transmission rate increases with the number of users, due to opportunistic scheduling that exploits multiuser diversity [21].

An intermediate step that is of interest on its own would be to consider our current model with

several possible service capacity vectors when serving classes in parallel. For example, if in addition to the service capacities c_1 and c_2 we can choose d_1 and d_2 that are not in the convex hull depicted in Figure 1.

A third direction of interest is to study our model with more than two classes. This could also serve as an intermediate step towards handling multiuser diversity gains as mentioned above, which is presumably more difficult to handle. These issues will be addressed in on-going and future research.

Acknowledgments

The authors are grateful to Sem Borst for valuable discussions and comments, and to the referees for their constructive reviews.

References

- [1] Ata, B., Kumar, S. (2005). Heavy traffic analysis of open processing networks with complete resource pooling: asymptotic optimality of discrete review policies. *Annals of Applied Probability* **15**, 331–391.
- [2] Bäuerle, N. (2000). Asymptotic optimality of tracking policies in stochastic networks. *Annals of Applied Probability* **10**, 1065–1083.
- [3] Bäuerle, N. (2002). Optimal control of queueing networks: an approach via fluid models. *Advances in Applied Probability* **34**, 313–328.
- [4] Bayati, M., Sharma, M., Squillante, M.S. (2006). Optimal scheduling in a multiserver stochastic network. *ACM SIGMETRICS/Performance Evaluation Review* **34**, 45–47.
- [5] Bell, S. L., Williams, R. J. (2001). Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: asymptotic optimality of a threshold policy. *Annals of Applied Probability* **11**, 608–649.
- [6] Bell, S. L., Williams, R. J. (2005). Dynamic scheduling of a parallel server system in heavy traffic with complete resource pooling: asymptotic optimality of a threshold policy. *Electronic J. of Probability* **10**, 1044–1115.
- [7] Bhardwaj, S., Williams, R.J. (2009). Diffusion approximation for a heavily loaded multi-user wireless communication system with cooperation. *Queueing Systems*. DOI: 10.1007/s11134-009-9119-8.
- [8] Bhardwaj, S., Williams, R. J., Acampora, A.S. (2007). On the performance of a two user MIMO downlink system in heavy traffic. *IEEE Transactions on Information Theory* **53**, 1851–1859.
- [9] Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley, New York.
- [10] Bonald, T., Proutière, A. (2006). Flow-level stability of utility-based allocations for non-convex rate regions. In: *Proc. CISS 2006 Conference on Information Sciences and Systems (Princeton University)*.
- [11] Borst, S.C., Leskelä, L., Jonckheere, M. (2008). Stability of parallel queueing systems with coupled service rates. *Discrete Event Dynamic Systems* **18**, 447–472.

- [12] Cesari, L. (1983). *Optimization – Theory and Applications*. Springer, New York.
- [13] Cohen, J.W., Boxma, O.J. (1983). *Boundary Value Problems in Queueing System Analysis*. North-Holland, Amsterdam.
- [14] Dai, J.G. (1995). On positive Harris recurrence of multiclass queueing networks: A unified approach via fluid limit models. *Annals of Applied Probability* **5**, 49–77.
- [15] Ethier, S.N., Kurtz, T.G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.
- [16] Fayolle, G., Iasnogorodski, R. (1979). Two coupled processors: the reduction to a Riemann-Hilbert problem. *Z. Wahr. verw. Geb.* **47**, 325–351.
- [17] Gajrat, A., Hordijk, A. (2000). Fluid approximation of a controlled multiclass tandem network. *Queueing Systems* **35**, 349–380.
- [18] Gajrat, A., Hordijk, A., Ridder, A. (2003). Large-deviations analysis of the fluid approximation for a controllable tandem queue. *Annals of Applied Probability* **13**, 1423–1448.
- [19] Harrison, J.M. (1998). Heavy traffic analysis of a system with parallel servers: asymptotic optimality of discrete-review policies. *Annals of Applied Probability* **8**, 822–848.
- [20] Harrison, J.M., López, M.J. (1999). Heavy traffic resource pooling in parallel-server systems. *Queueing systems* **33**, 339–368.
- [21] Liu, X., Chong, E., Shroff, N. (2003). A framework for opportunistic scheduling in wireless networks. *Computer Networks* **41**, 451–474.
- [22] Liu, J., Proutière, A., Yi, Y., Chiang, M., Poor, H.V. (2007). Flow-level stability of data networks with non-convex and time-varying rate regions. *ACM SIGMETRICS Performance Evaluation Review* **35**, 239–250.
- [23] Maglaras, C. (2000). Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid-scale asymptotic optimality. *Annals of Applied Probability* **10**, 897–929.
- [24] Mandelbaum, A., Stolyar, A.L. (2004). Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized $c\mu$ -rule. *Operations Research* **52**, 836–855.
- [25] Meyn, S.P. (1997). Stability and optimization of queueing networks and their fluid models. In: *Mathematics of stochastic manufacturing systems. Lectures in Applied Mathematics* **33**, 175–199.
- [26] Meyn, S.P. (2008). *Control Techniques for Complex Networks*. Cambridge University Press, New York.
- [27] Osogami, T., Harchol-Balter, M., Scheller-Wolf, A. (2005). Analysis of cycle stealing with switching times and thresholds. *Performance Evaluation* **61**, 347–369.
- [28] Osogami, T., Harchol-Balter, M., Scheller-Wolf, A., Zhang, L. (2004). Exploring Threshold-based Policies for Load Sharing. *Forty-second Annual Allerton Conference on Communication, Control, and Computing*, 1012–1021.
- [29] Puterman, M.L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, New York.

- [30] Righter, R., Shanthikumar, J.G. (1989). Scheduling multiclass single-server queueing systems to stochastically maximize the number of successful departures. *Prob. Eng. Inf. Sc.* **3**, 323–333.
- [31] Seierstad, A., Sydsæter, K. (1987). *Optimal Control Theory with Economic Applications*. North-Holland, Amsterdam.
- [32] Stolyar, A.L. (2004). MaxWeight Scheduling in a Generalized Switch: State Space Collapse and Workload Minimization in Heavy Traffic. *Annals of Applied Probability* **14**, 1–53.
- [33] Squillante, M.S., Xia, C.H., Yao, D.D., Zhang, L. (2001). Threshold-based priority policies for parallel-server systems with affinity scheduling. *Proc. of the IEEE American Control Conference* **4**, 2992–2999.
- [34] Tezcan, T. (2010). Stability analysis of N-model systems under a static priority rule using augmented fluid models. Under submission.
- [35] Verloop, I.M., Borst, S.C., Núñez-Queija, R. (2006). Delay optimization in bandwidth-sharing networks. In: *Proc. CISS 2006 Conference on Information Sciences and Systems (Princeton University)*.
- [36] Verloop, I.M., Núñez-Queija, R. (2009). Assessing the efficiency of resource allocations in bandwidth-sharing networks. *Performance Evaluation* **66**, 59–77.

Appendix A: Proof of Lemma 3.2

The proof is by induction on the time index k . For $k = 0$ the statement holds. In order to apply induction, assume it holds for $Z = V_k$. We show that it holds for $Z = V_{k+1}$ as well.

In the remainder of the appendix we show that

$$\begin{aligned}
& (\mu_1 + \mu_2)V_{k+1}(x) + \mu_1c_1V_{k+1}(x - e_1) + \mu_2c_2V_{k+1}(x - e_2) \\
& \leq \mu_1V_{k+1}(x) + \mu_2V_{k+1}(x - e_2) + (\mu_1c_1 + \mu_2c_2)V_{k+1}(x)
\end{aligned} \tag{43}$$

is indeed satisfied. The proof of

$$\begin{aligned}
& (\mu_1 + \mu_2)V_{k+1}(x) + \mu_1c_1V_{k+1}(x - e_1) + \mu_2c_2V_{k+1}(x - e_2) \\
& \leq \mu_1V_{k+1}(x - e_1) + \mu_2V_{k+1}(x) + (\mu_1c_1 + \mu_2c_2)V_{k+1}(x)
\end{aligned}$$

follows exactly the same steps, but with the role of class 1 and class 2 interchanged.

First assume $x_1 > 0$ and $x_2 = 1$. By definition of the function $V_{k+1}(\cdot)$ (see (7)), we can write

$$\begin{aligned}
& \mu_2V_{k+1}(x_1, 1) + \mu_1c_1V_{k+1}(x_1 - 1, 1) + \mu_2c_2V_{k+1}(x_1, 0) \\
& \leq \mu_2[\lambda_1V_k(x_1 + 1, 1) + \lambda_2V_k(x_1, 2) + \mu_1V_k(x_1, 1) + \mu_2V_k(x_1, 0) + (\mu_1c_1 + \mu_2c_2)V_k(x_1, 1)] \\
& \quad + \mu_1c_1[\lambda_1V_k(x_1, 1) + \lambda_2V_k(x_1 - 1, 2) + \mu_1V_k(x_1 - 1, 1) + \mu_2V_k(x_1 - 1, 0) \\
& \quad \quad + (\mu_1c_1 + \mu_2c_2)V_k(x_1 - 1, 1)] \\
& \quad + \mu_2c_2[\lambda_1V_k(x_1 + 1, 0) + \lambda_2V_k(x_1, 1) + \mu_1V_k(x_1 - 1, 0) + \mu_2V_k(x_1, 0) + (\mu_1c_1 + \mu_2c_2)V_k(x_1, 0)].
\end{aligned} \tag{44}$$

Rearranging terms in (44), gives

$$\begin{aligned}
& \lambda_1[\mu_2 V_k(x_1 + 1, 1) + \mu_1 c_1 V_k(x_1, 1) + \mu_2 c_2 V_k(x_1 + 1, 0)] \\
& + \lambda_2[\mu_2 V_k(x_1, 2) + \mu_1 c_1 V_k(x_1 - 1, 2) + \mu_2 c_2 V_k(x_1, 1)] \\
& + \mu_1[\mu_2 V_k(x_1, 1) + \mu_1 c_1 V_k(x_1 - 1, 1) + \mu_2 c_2 V_k(x_1, 0)] \\
& + (\mu_1 c_1 + \mu_2 c_2)[\mu_2 V_k(x_1, 1) + \mu_1 c_1 V_k(x_1 - 1, 1) + \mu_2 c_2 V_k(x_1, 0)] \\
& + \mu_2[\mu_2 V_k(x_1, 0) + \mu_2 c_2 V_k(x_1, 0)] \\
& + \mu_1 \mu_2[(c_1 + c_2)V_k(x_1 - 1, 0) - c_2 V_k(x_1, 0)].
\end{aligned} \tag{45}$$

Since $V_k(\cdot)$ is increasing in x_1 (see Lemma 3.1), $c_1 + c_2 \geq 1$, and since (8) holds by induction for $V_k(\cdot)$, equation (45) is less than or equal to

$$\begin{aligned}
& \lambda_1[\mu_2 V_k(x_1 + 1, 0) + (\mu_1 c_1 + \mu_2 c_2)V_k(x_1 + 1, 1)] \\
& + \lambda_2[\mu_2 V_k(x_1, 1) + (\mu_1 c_1 + \mu_2 c_2)V_k(x_1, 2)] \\
& + \mu_1[\mu_2 V_k(x_1, 0) + (\mu_1 c_1 + \mu_2 c_2)V_k(x_1, 1)] \\
& + (\mu_1 c_1 + \mu_2 c_2)[\mu_2 V_k(x_1, 1) + \mu_1 c_1 V_k(x_1 - 1, 1) + \mu_2 c_2 V_k(x_1, 0)] \\
& + \mu_2[\mu_2 V_k(x_1, 0) + \mu_2 c_2 V_k(x_1, 0)] \\
& + \mu_1 \mu_2[(c_1 - 1)V_k(x_1, 0) + V_k(x_1 - 1, 0)] \\
& = \mu_2[\lambda_1 V_k(x_1 + 1, 0) + \lambda_2 V_k(x_1, 1) + \mu_1 V_k(x_1 - 1, 0) + (\mu_2 + \mu_1 c_1 + \mu_2 c_2)V_k(x_1, 0)] \\
& + (\mu_1 c_1 + \mu_2 c_2)[\lambda_1 V_k(x_1 + 1, 1) + \lambda_2 V_k(x_1, 2) + (\mu_1 + \mu_2)V_k(x_1, 1) \\
& \quad + \mu_1 c_1 V_k(x_1 - 1, 1) + \mu_2 c_2 V_k(x_1, 0)],
\end{aligned} \tag{46}$$

where in the last step we rearranged the terms. Since (8) holds by induction for $V_k(\cdot)$, equation (46) is equal to $\mu_2 V_{k+1}(x_1, 0) + (\mu_1 c_1 + \mu_2 c_2)V_{k+1}(x_1, 1)$. Hence, (43) is proved.

Now assume $x_1 > 0$ and $x_2 > 1$. By definition of $V_{k+1}(\cdot)$ we can write

$$\begin{aligned}
& \mu_2 V_{k+1}(x) + \mu_1 c_1 V_{k+1}(x - e_1) + \mu_2 c_2 V_{k+1}(x - e_2) \\
& \leq \mu_2[\lambda_1 V_k(x + e_1) + \lambda_2 V_k(x + e_2) + \mu_1 V_k(x) + \mu_2 V_k(x - e_2) \\
& \quad + \mu_1 c_1 V_k(x) + \mu_2 c_2 V_k(x)] \\
& + \mu_1 c_1[\lambda_1 V_k(x) + \lambda_2 V_k(x - e_1 + e_2) + \mu_1 V_k(x - e_1) + \mu_2 V_k(x - e_1 - e_2) \\
& \quad + \mu_1 c_1 V_k(x - e_1) + \mu_2 c_2 V_k(x - e_1)] \\
& + \mu_2 c_2[\lambda_1 V_k(x + e_1 - e_2) + \lambda_2 V_k(x) + \mu_1 V_k(x - e_2) + \mu_2 V_k(x - 2e_2) \\
& \quad + \mu_1 c_1 V_k(x - e_2) + \mu_2 c_2 V_k(x - e_2)]
\end{aligned} \tag{47}$$

Rearranging terms in (47), gives

$$\begin{aligned}
& \lambda_1[\mu_2 V_k(x + e_1) + \mu_1 c_1 V_k(x) + \mu_2 c_2 V_k(x + e_1 - e_2)] \\
& + \lambda_2[\mu_2 V_k(x + e_2) + \mu_1 c_1 V_k(x - e_1 + e_2) + \mu_2 c_2 V_k(x)] \\
& + \mu_1[\mu_2 V_k(x) + \mu_1 c_1 V_k(x - e_1) + \mu_2 c_2 V_k(x - e_2)] \\
& + \mu_2[\mu_2 V_k(x - e_2) + \mu_1 c_1 V_k(x - e_1 - e_2) + \mu_2 c_2 V_k(x - 2e_2)] \\
& + (\mu_1 c_1 + \mu_2 c_2)[\mu_2 V_k(x) + \mu_1 c_1 V_k(x - e_1) + \mu_2 c_2 V_k(x - e_2)].
\end{aligned} \tag{48}$$

Since (8) holds by induction for $V_k(\cdot)$, equation (48) is less than or equal to

$$\begin{aligned}
& \lambda_1[\mu_2 V_k(x + e_1 - e_2) + \mu_1 c_1 V_k(x + e_1) + \mu_2 c_2 V_k(x + e_1)] \\
& + \lambda_2[\mu_2 V_k(x) + \mu_1 c_1 V_k(x + e_2) + \mu_2 c_2 V_k(x + e_2)] \\
& + \mu_1[\mu_2 V_k(x - e_2) + \mu_1 c_1 V_k(x) + \mu_2 c_2 V_k(x)] \\
& + \mu_2[\mu_2 V_k(x - e_2) + \mu_1 c_1 V_k(x - e_1 - e_2) + \mu_2 c_2 V_k(x - 2e_2)] \\
& + (\mu_1 c_1 + \mu_2 c_2)[\mu_2 V_k(x) + \mu_1 c_1 V_k(x - e_1) + \mu_2 c_2 V_k(x - e_2)] \\
& = \mu_2[\lambda_1 V_k(x + e_1 - e_2) + \lambda_2 V_k(x) + (\mu_1 + \mu_2) V_k(x - e_2) \\
& \quad + \mu_1 c_1 V_k(x - e_1 - e_2) + \mu_2 c_2 V_k(x - 2e_2)] \\
& \quad + (\mu_1 c_1 + \mu_2 c_2)[\lambda_1 V_k(x + e_1) + \lambda_2 V_k(x + e_2) + (\mu_1 + \mu_2) V_k(x) \\
& \quad \quad + \mu_1 c_1 V_k(x - e_1) + \mu_2 c_2 V_k(x - e_2)], \tag{49}
\end{aligned}$$

where in the last step we rearranged the terms. Since (8) holds by induction for $V_k(\cdot)$, equation (49) is equal to $\mu_2 V_{k+1}(x - e_2) + \mu_1 c_1 V_{k+1}(x) + \mu_2 c_2 V_{k+1}(x)$. Hence, (43) is proved. \square

Appendix B: Proof of Lemma 3.4

We use t^+ to denote any element in an interval $(t, t + \delta]$, for a sufficiently small $\delta > 0$. Throughout the proof we use that

$$W_i(t) > 0 \text{ implies } W_i(t^+) > 0, \text{ and that } W_i(t) = 0 \text{ implies } W_i(t^+) = 0. \tag{50}$$

This follows since the workload process $W_i(t)$, $i = 1, 2$, is right-continuous and increases only with an arrival.

Note that $S_i(t)$, $i = 1, 2$, is continuous. In order to show relation (9), we therefore consider the first time instant t such that (9) holds with equality and is violated immediately after time t . So $S_1^\pi(t) = S_1^{\tilde{\pi}}(t)$, and by (1) also $W_1^\pi(t) = W_1^{\tilde{\pi}}(t)$, while $s_1^\pi(t^+) < s_1^{\tilde{\pi}}(t^+)$, so that $S_1^\pi(t^+) < S_1^{\tilde{\pi}}(t^+)$. Since $W_1^\pi(t) = W_1^{\tilde{\pi}}(t)$, by (50) and by construction of policy π we obtain that $s_1^\pi(t^+) \geq s_1^{\tilde{\pi}}(t^+)$. This gives contradiction and hence (9) holds for all $t \geq 0$.

Let time t be the first time instant such that either (10) or (11) holds with equality and is violated immediately after time t . We will show that such a t does not exist. The remainder of the proof consists of two parts, depending on whether equation (10) or equation (11) is the first to be violated.

Part I: Assume (10) is the first equation that fails to hold, i.e., $S_1^\pi(t) + S_2^\pi(t) = S_1^{\tilde{\pi}}(t) + S_2^{\tilde{\pi}}(t)$, and by (1) also $W_1^\pi(t) + W_2^\pi(t) = W_1^{\tilde{\pi}}(t) + W_2^{\tilde{\pi}}(t)$, while $s_1^\pi(t^+) + s_2^\pi(t^+) < s_1^{\tilde{\pi}}(t^+) + s_2^{\tilde{\pi}}(t^+)$, so that $S_1^\pi(t^+) + S_2^\pi(t^+) < S_1^{\tilde{\pi}}(t^+) + S_2^{\tilde{\pi}}(t^+)$. We will show that

$$W_1^\pi(t) + W_2^\pi(t) = W_1^{\tilde{\pi}}(t) + W_2^{\tilde{\pi}}(t) \text{ implies } W_i^\pi(t) = W_i^{\tilde{\pi}}(t), \quad i = 1, 2. \tag{51}$$

By (50) and by construction of policy, $W_i^\pi(t) = W_i^{\tilde{\pi}}(t)$, $i = 1, 2$, implies that $s_1^\pi(t^+) + s_2^\pi(t^+) \geq s_1^{\tilde{\pi}}(t^+) + s_2^{\tilde{\pi}}(t^+)$, and hence we reach a contradiction. So let us prove (51).

- We first assume that there is an interval $[u, t)$ in which policy $\tilde{\pi}$ has more work in the system compared to policy π , i.e., $W_1^\pi(v) + W_2^\pi(v) < W_1^{\tilde{\pi}}(v) + W_2^{\tilde{\pi}}(v)$ for all $v \in [u, t)$, and at time t , $W_1^\pi(t) + W_2^\pi(t) = W_1^{\tilde{\pi}}(t) + W_2^{\tilde{\pi}}(t)$. We can choose this interval such that $\tilde{\pi}$ has made up for the lost capacity in one of the three ways described below. Define $M_c^{\tilde{\pi}}(u, t)$ as the cumulative amount of time that both classes are served in parallel under policy $\tilde{\pi}$ in the time interval $[u, t)$.

- (i) During the interval $[u, t)$ policy $\tilde{\pi}$ has work in the system, while policy π has an empty system.

- (ii) In the interval $[u, t)$ we have $M_c^{\tilde{\pi}}(u, t) > 0$, while policy π serves class 1 with service capacity 1. Hence $W_2^\pi(v) = 0$ and $W_1^\pi(v) > 0$, for all $v \in [u, t)$.
- (iii) In the interval $[u, t)$ we have $M_c^{\tilde{\pi}}(u, t) > 0$, while policy π serves class 2 with service capacity 1. Hence $W_1^\pi(v) = 0$ and $W_2^\pi(v) > 0$, for all $v \in [u, t)$.

Note that the three cases are mutually exclusive. We will show that (51) holds for (i), (ii) and (iii). Although not mentioned explicitly, in all three cases we use that a possible arrival at time t alters the workload in both systems in the same way. Let t^- denote any element in an interval $[t - \delta, t)$ with $\delta > 0$ sufficiently small.

In case (i) we have $W_i^\pi(t^-) = 0$ for $i = 1, 2$. Since at time t it holds that $W_1^\pi(t) + W_2^\pi(t) = W_1^{\tilde{\pi}}(t) + W_2^{\tilde{\pi}}(t)$, we obtain that $W_i^{\tilde{\pi}}(t^-) = 0$, $i = 1, 2$. Hence, we have $W_i^\pi(t) = W_i^{\tilde{\pi}}(t)$, $i = 1, 2$.

In case (ii) we have that $W_2^\pi(t^-) = 0$, hence $W_2^\pi(t) \leq W_2^{\tilde{\pi}}(t)$. From $W_1^\pi(t) + W_2^\pi(t) = W_1^{\tilde{\pi}}(t) + W_2^{\tilde{\pi}}(t)$ and $W_1^\pi(t) \leq W_1^{\tilde{\pi}}(t)$ (follows from (1) and (9)), we obtain $W_2^\pi(t) \geq W_2^{\tilde{\pi}}(t)$. Hence, $W_i^\pi(t) = W_i^{\tilde{\pi}}(t)$, $i = 1, 2$.

In case (iii) we have

$$M_c^{\tilde{\pi}}(u, t)(c_1 + c_2 - 1) = W_1^{\tilde{\pi}}(u) + W_2^{\tilde{\pi}}(u) - W_2^\pi(u), \quad (52)$$

since the total amount of additional capacity that policy $\tilde{\pi}$ gets compared to policy π in the interval $[u, t)$ (left-hand side in (52)), is equal to the difference in the total workload at time u (right-hand side in (52)). Since $W_1^\pi(u) = 0$, from (1) and (11) we obtain that $c_1 W_2^\pi(u) = (1 - c_2) W_1^\pi(u) + c_1 W_2^\pi(u) \leq (1 - c_2) W_1^{\tilde{\pi}}(u) + c_1 W_2^{\tilde{\pi}}(u)$. Rewriting, this gives $W_1^{\tilde{\pi}}(u) \leq \frac{c_1}{c_1 + c_2 - 1} (W_1^{\tilde{\pi}}(u) + W_2^{\tilde{\pi}}(u) - W_2^\pi(u)) = c_1 M_c^{\tilde{\pi}}(u, t)$. Note that $S_1^{\tilde{\pi}}(t) - S_1^{\tilde{\pi}}(u) \geq c_1 M_c^{\tilde{\pi}}(u, t)$ and $A_1(u, t^-) = 0$ (since $W_1^\pi(v) = 0$ for all $v \in [u, t)$). Together this gives $W_1^{\tilde{\pi}}(t^-) = W_1^{\tilde{\pi}}(u) + A_1(u, t^-) - (S_1^{\tilde{\pi}}(t) - S_1^{\tilde{\pi}}(u)) \leq 0$. Since we also know that $W_1^\pi(t^-) = 0$, it follows that $W_1^\pi(t) = W_1^{\tilde{\pi}}(t)$, and hence $W_2^\pi(t) = W_2^{\tilde{\pi}}(t)$.

- Now consider the case when there is an interval $[w, t]$ such that $W_1^\pi(v) + W_2^\pi(v) = W_1^{\tilde{\pi}}(v) + W_2^{\tilde{\pi}}(v)$ for all $v \in [w, t]$ and $W_1^\pi(w^-) + W_2^\pi(w^-) < W_1^{\tilde{\pi}}(w^-) + W_2^{\tilde{\pi}}(w^-)$. From the previous case, we obtain that $W_i^\pi(w) = W_i^{\tilde{\pi}}(w)$, $i = 1, 2$. Together with the fact that in the interval $[w, t]$ the total workload is equal under both policies, and by construction of policy π , it follows that $\tilde{\pi}$ did not serve class 2 individually while π serves both classes in parallel. Hence, $W_i^\pi(v) = W_i^{\tilde{\pi}}(v)$ for all $v \in [w, t]$, $i = 1, 2$.

Part II: Assume (11) is the first equation that fails to hold, i.e., $(1 - c_2)S_1^\pi(t) + c_1S_2^\pi(t) = (1 - c_2)S_1^{\tilde{\pi}}(t) + c_1S_2^{\tilde{\pi}}(t)$, and by (1) also $(1 - c_2)W_1^\pi(t) + c_1W_2^\pi(t) = (1 - c_2)W_1^{\tilde{\pi}}(t) + c_1W_2^{\tilde{\pi}}(t)$, while $(1 - c_2)s_1^\pi(t^+) + c_1s_2^\pi(t^+) < (1 - c_2)s_1^{\tilde{\pi}}(t^+) + c_1s_2^{\tilde{\pi}}(t^+)$. So that $(1 - c_2)S_1^\pi(t^+) + c_1S_2^\pi(t^+) < (1 - c_2)S_1^{\tilde{\pi}}(t^+) + c_1S_2^{\tilde{\pi}}(t^+)$. With slight abuse of notation, let $f_1(t^+), f_2(t^+), f_c(t^+), f_0(t^+)$ be the coefficients that define the capacity vector in the capacity region S under policy $\tilde{\pi}$ at time t^+ , i.e., $(s_1^{\tilde{\pi}}(t^+), s_2^{\tilde{\pi}}(t^+)) = f_1(t^+) \cdot (1, 0) + f_2(t^+) \cdot (0, 1) + f_c(t^+) \cdot (c_1, c_2) + f_0(t^+) \cdot (0, 0)$. Note that $1 = f_1(t^+) + f_2(t^+) + f_c(t^+) + f_0(t^+)$. We have the following possibilities:

- If $W_1^\pi(t) > 0$ and $W_2^\pi(t) > 0$, then by (50) and by definition of policy π we have $(s_1^\pi(t^+), s_2^\pi(t^+)) = f_1(t^+) \cdot (1, 0) + (f_c(t^+) + f_2(t^+)) \cdot (c_1, c_2) + f_0(t^+) \cdot (0, 0)$, hence $(1 - c_2)s_1^\pi(t^+) + c_1s_2^\pi(t^+) = (1 - c_2)(f_1(t^+) + c_1(f_c(t^+) + f_2(t^+))) + c_1c_2(f_c(t^+) + f_2(t^+)) = (1 - c_2)(f_1(t^+) + c_1f_c(t^+)) + c_1(f_2(t^+) + c_2f_c(t^+)) = (1 - c_2)s_1^{\tilde{\pi}}(t^+) + c_1s_2^{\tilde{\pi}}(t^+)$.
- If $W_1^\pi(t) = 0$ and $W_2^\pi(t) > 0$, then, by definition, policy π serves class 2 individually for a fraction of time $1 - f_0(t^+)$ and otherwise idles. So $(1 - c_2)s_1^\pi(t^+) + c_1s_2^\pi(t^+) = c_1(1 - f_0(t^+))$. Since $c_1 + c_2 > 1$, we have that $c_1(1 - f_0(t^+)) \geq (1 - c_2)f_1(t^+) + c_1(f_c(t^+) + f_2(t^+)) = (1 - c_2)(f_1(t^+) + c_1f_c(t^+)) + c_1(f_2(t^+) + c_2f_c(t^+)) = (1 - c_2)s_1^{\tilde{\pi}}(t^+) + c_1s_2^{\tilde{\pi}}(t^+)$.

- If $W_1^\pi(t) > 0$ and $W_2^\pi(t) = 0$, then we have $(1 - c_2)W_1^\pi(t) = (1 - c_2)W_1^{\tilde{\pi}}(t) + c_1W_2^{\tilde{\pi}}(t)$ and $W_1^\pi(t) \leq W_1^{\tilde{\pi}}(t)$ (by (9)). Hence $W_1^\pi(t) = W_1^{\tilde{\pi}}(t)$ and $0 = W_2^\pi(t) = W_2^{\tilde{\pi}}(t)$. By (50) we obtain $f_2(t^+) = 0$, so by definition of policy π , $s_i^\pi(t^+) = s_i^{\tilde{\pi}}(t^+)$, $i = 1, 2$.
- If $W_1^\pi(t) + W_2^\pi(t) = 0$, then $0 = (1 - c_2)W_1^{\tilde{\pi}}(t) + c_1W_2^{\tilde{\pi}}(t)$. By (50) we have $W_i^\pi(t^+) = W_i^{\tilde{\pi}}(t^+) = 0$, and hence $(1 - c_2)s_1^\pi(t^+) + c_1s_2^\pi(t^+) = (1 - c_2)s_1^{\tilde{\pi}}(t^+) + c_1s_2^{\tilde{\pi}}(t^+) = 0$.

For all the four possibilities we reached a contradiction with $(1 - c_2)s_1^\pi(t^+) + c_1s_2^\pi(t^+) < (1 - c_2)s_1^{\tilde{\pi}}(t^+) + c_1s_2^{\tilde{\pi}}(t^+)$ and this concludes the proof. \square

Appendix C: Proof of Lemma 4.3

We construct policy π below. Note that $u_2^\pi(t) = 0$ when $n_1^\pi(t) > 0$.

- If $n_1^\pi(t) > 0$ and $n_2^\pi(t) > 0$, then $u_c^\pi(t) = u_2^{\tilde{\pi}}(t) + u_c^{\tilde{\pi}}(t)$, $u_1^\pi(t) = u_1^{\tilde{\pi}}(t)$ and $u_2^\pi(t) = 0$.
- If $n_1^\pi(t) = 0$ and $n_2^\pi(t) > 0$, then $u_c^\pi(t) = \min\left(u_2^{\tilde{\pi}}(t) + u_c^{\tilde{\pi}}(t), \frac{\rho_1}{c_1}\right)$, $u_1^\pi(t) = \min\left(u_1^{\tilde{\pi}}(t), \rho_1 - c_1u_c^\pi(t)\right)$ and $u_2^\pi(t) = u_c^\pi(t) + u_1^{\tilde{\pi}}(t) + u_2^{\tilde{\pi}}(t) - u_c^\pi(t) - u_1^{\tilde{\pi}}(t)$.
- If $n_1^\pi(t) > 0$ and $n_2^\pi(t) = 0$, then $u_c^\pi(t) = \min\left(u_2^{\tilde{\pi}}(t) + u_c^{\tilde{\pi}}(t), \frac{\rho_2}{c_2}\right)$, $u_1^\pi(t) = u_c^\pi(t) + u_1^{\tilde{\pi}}(t) + u_2^{\tilde{\pi}}(t) - u_c^\pi(t)$ and $u_2^\pi(t) = 0$.
- If $n_1^\pi(t) = 0$ and $n_2^\pi(t) = 0$, then take $u^\pi(t)$ such that $\rho_i = u_i^\pi(t) + c_iu_c^\pi(t)$, $i = 1, 2$.

Once $n_1^\pi(t) + n_2^\pi(t) = 0$, policy π will keep the system empty from that moment on (this is possible since the stability conditions are satisfied). Therefore, we will focus on states with $n_1^\pi(t) + n_2^\pi(t) > 0$.

For policies π and $\tilde{\pi}$, we will prove the following inequalities:

$$U_1^\pi(t) + c_1U_c^\pi(t) \geq U_1^{\tilde{\pi}}(t) + c_1U_c^{\tilde{\pi}}(t) \quad (53)$$

$$U_1^\pi(t) + U_2^\pi(t) + (c_1 + c_2)U_c^\pi(t) \geq U_1^{\tilde{\pi}}(t) + U_2^{\tilde{\pi}}(t) + (c_1 + c_2)U_c^{\tilde{\pi}}(t) \quad (54)$$

$$(1 - c_2)U_1^\pi(t) + c_1(U_2^\pi(t) + U_c^\pi(t)) \geq (1 - c_2)U_1^{\tilde{\pi}}(t) + c_1(U_2^{\tilde{\pi}}(t) + U_c^{\tilde{\pi}}(t)). \quad (55)$$

They are similar to the inequalities of the stochastic model (9)–(11) when setting $S_i(t) = U_i(t) + c_iU_c(t)$. When multiplying (53) by $\mu_1 - \mu_2 \geq 0$ and (54) by μ_2 and adding the two inequalities, we obtain $\mu_1U_1^\pi(t) + \mu_2U_2^\pi(t) + (\mu_1c_1 + \mu_2c_2)U_c^\pi(t) \geq \mu_1U_1^{\tilde{\pi}}(t) + \mu_2U_2^{\tilde{\pi}}(t) + (\mu_1c_1 + \mu_2c_2)U_c^{\tilde{\pi}}(t)$. By (13) we get $n_1^\pi(t) + n_2^\pi(t) \leq n_1^{\tilde{\pi}}(t) + n_2^{\tilde{\pi}}(t)$ for all $t \geq 0$, which was to be proved. The remainder of the appendix is devoted to the proof of inequalities (53)–(55). Throughout the proof, we consider the workload fluid processes $w_i(\cdot) = n_i(t)/\mu_i$, $i = 1, 2$.

Note that $U_j(t)$, $j = 1, 2, c$, is continuous. In order to show (53), we therefore consider the first time instant t such that (53) holds with equality and is violated immediately after time t . So $U_1^\pi(t) + c_1U_c^\pi(t) = U_1^{\tilde{\pi}}(t) + c_1U_c^{\tilde{\pi}}(t)$, and by (13) also $n_1^\pi(t) = n_1^{\tilde{\pi}}(t)$, while $u_1^\pi(t^+) + c_1u_c^\pi(t^+) < u_1^{\tilde{\pi}}(t^+) + c_1u_c^{\tilde{\pi}}(t^+)$, so $n_1^\pi(t^+) > n_1^{\tilde{\pi}}(t^+)$. Since $n_1^\pi(t^+) > 0$, by construction of policy π we obtain $u_1^\pi(t) + c_1u_c^\pi(t) \geq u_1^{\tilde{\pi}}(t) + c_1u_c^{\tilde{\pi}}(t)$, which gives contradiction. Hence (53) holds for all $t \geq 0$.

Let time t be the first time instant that either (54) or (55) holds with equality and is violated immediately after time t . The remainder of the proof consists of two parts, depending on whether equation (54) or equation (55) is the first to be violated.

Part I: Assume (54) is the first equation that fails to hold, i.e., $U_1^\pi(t) + U_2^\pi(t) + (c_1 + c_2)U_c^\pi(t) = U_1^{\tilde{\pi}}(t) + U_2^{\tilde{\pi}}(t) + (c_1 + c_2)U_c^{\tilde{\pi}}(t)$, and by (13) also $w_1^\pi(t) + w_2^\pi(t) = w_1^{\tilde{\pi}}(t) + w_2^{\tilde{\pi}}(t)$, while $u_1^\pi(t^+) + u_2^\pi(t^+) + (c_1 + c_2)u_c^\pi(t^+) < u_1^{\tilde{\pi}}(t^+) + u_2^{\tilde{\pi}}(t^+) + (c_1 + c_2)u_c^{\tilde{\pi}}(t^+)$. In what follows we use the following implication, which will be proved later on:

$$w_1^\pi(t) + w_2^\pi(t) = w_1^{\tilde{\pi}}(t) + w_2^{\tilde{\pi}}(t) \quad \text{implies} \quad w_i^\pi(t) = w_i^{\tilde{\pi}}(t), \quad i = 1, 2. \quad (56)$$

We now distinguish between three cases: (i) If $w_1^\pi(t^+) > 0$ and $w_2^\pi(t^+) > 0$, then by construction of policy π , $u_c^\pi(t^+) \geq u_c^{\tilde{\pi}}(t^+)$. (ii) If $w_1^\pi(t^+) = 0$, then $0 = w_1^\pi(t) (= w_1^{\tilde{\pi}}(t))$, since $w_1(\cdot)$ is continuous. Policy π is able to keep class 1 empty at time t^+ while $\tilde{\pi}$ might not, so we have $\rho_1 = u_1^\pi(t^+) + c_1 u_c^\pi(t^+) \geq u_1^{\tilde{\pi}}(t^+) + c_1 u_c^{\tilde{\pi}}(t^+)$. In particular, $u_c^{\tilde{\pi}}(t^+) \leq \rho_1/c_1$, and by construction of policy π , this implies $u_c^\pi(t^+) \geq u_c^{\tilde{\pi}}(t^+)$. (iii) If $w_2^\pi(t^+) = 0$, then $0 = w_2^\pi(t) (= w_2^{\tilde{\pi}}(t))$, since $w_2(\cdot)$ is continuous. In a similar fashion as in the previous item, we obtain that $u_c^\pi(t^+) \geq u_c^{\tilde{\pi}}(t^+)$. Hence, in all cases it holds that $u_c^\pi(t^+) \geq u_c^{\tilde{\pi}}(t^+)$. Together with $c_1 + c_2 \geq 1$ and $u_1^\pi(t^+) + u_2^\pi(t^+) + u_c^\pi(t^+) = u_1^{\tilde{\pi}}(t^+) + u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)$, we can conclude that $u_1^\pi(t^+) + u_2^\pi(t^+) + (c_1 + c_2)u_c^\pi(t^+) \geq u_1^{\tilde{\pi}}(t^+) + u_2^{\tilde{\pi}}(t^+) + (c_1 + c_2)u_c^{\tilde{\pi}}(t^+)$, and we reach a contradiction. It now only remains to prove that the implication in (56) is satisfied. We distinguish between the following two cases:

- Assume there is an interval $[u, t]$ in which policy $\tilde{\pi}$ has more work in the system compared to policy π , i.e., $w_1^\pi(v) + w_2^\pi(v) < w_1^{\tilde{\pi}}(v) + w_2^{\tilde{\pi}}(v)$, for all $v \in [u, t]$. If the interval is such that $w_1^\pi(v) > 0$ and $w_2^\pi(v) > 0$, for all $v \in [u, t]$, then policy $\tilde{\pi}$ can never catch up with π (by construction of policy π). Hence, we can choose the interval $[u, t]$ such that:

- (i) For all $v \in [u, t]$, $w_2^\pi(v) = 0$ and $w_1^\pi(v) > 0$.
- (ii) For all $v \in [u, t]$, $w_1^\pi(v) = 0$ and $w_2^\pi(v) > 0$.

Note that the two cases are mutually excluding. We show that (56) holds in both cases.

By continuity of $w_2^\pi(\cdot)$, in case (i) we have as well $w_2^\pi(t) = 0$. Hence, $w_1^\pi(t) = w_1^{\tilde{\pi}}(t) + w_2^{\tilde{\pi}}(t)$. By (13) and (53) we have $w_1^\pi(t) \leq w_1^{\tilde{\pi}}(t)$. Together this gives $w_2^{\tilde{\pi}}(t) = 0 (= w_2^\pi(t))$ and $w_1^{\tilde{\pi}}(t) = w_1^\pi(t)$. Hence, in case (i) relation (56) is proved.

Let $M_j^{\tilde{\pi}}(u, t) = \int_u^t u_j^{\tilde{\pi}}(s) ds$ be the cumulative amount of time that activity j occurs under policy $\tilde{\pi}$ in the time interval $[u, t]$. The total amount of additional capacity that policy $\tilde{\pi}$ gets compared with policy π in the interval $[u, t]$ is

$$\begin{aligned} & (c_1 + c_2)M_c^{\tilde{\pi}}(u, t) + M_1^{\tilde{\pi}}(u, t) + M_2^{\tilde{\pi}}(u, t) - (c_1 + c_2)M_c^\pi(u, t) - M_1^\pi(u, t) - M_2^\pi(u, t) \\ &= (c_1 + c_2 - 1)(M_c^{\tilde{\pi}}(u, t) - M_c^\pi(u, t)), \end{aligned}$$

where we used that $M_c^{\tilde{\pi}}(u, t) + M_1^{\tilde{\pi}}(u, t) + M_2^{\tilde{\pi}}(u, t) = M_c^\pi(u, t) + M_1^\pi(u, t) + M_2^\pi(u, t)$.

This is equal to the difference in the total workload at time u , so $(c_1 + c_2 - 1)(M_c^{\tilde{\pi}}(u, t) - M_c^\pi(u, t)) = w_1^{\tilde{\pi}}(u) + w_2^{\tilde{\pi}}(u) - w_1^\pi(u) - w_2^\pi(u)$. In case (ii), $w_1^\pi(u) = 0$, hence we obtain from (13) and (55) that $c_1 w_2^\pi(u) = (1 - c_2)w_1^\pi(u) + c_1 w_2^\pi(u) \leq (1 - c_2)w_1^{\tilde{\pi}}(u) + c_1 w_2^{\tilde{\pi}}(u)$. Rewriting, this gives

$$w_1^{\tilde{\pi}}(u) \leq \frac{c_1}{c_1 + c_2 - 1} (w_1^{\tilde{\pi}}(u) + w_2^{\tilde{\pi}}(u) - w_2^\pi(u)) = c_1 (M_c^{\tilde{\pi}}(u, t) - M_c^\pi(u, t)). \quad (57)$$

Note that $\rho_1(t - u) = c_1 M_c^\pi(u, t) + M_1^\pi(u, t)$ (since in case (ii) class 1 is kept empty under policy π), and $M_1^{\tilde{\pi}}(u, t) \geq M_1^\pi(u, t)$ (by definition of policy π). Together with (57) this gives

$$w_1^{\tilde{\pi}}(t) = w_1^{\tilde{\pi}}(u) + \rho_1(t - u) - c_1 M_c^{\tilde{\pi}}(u, t) - M_1^{\tilde{\pi}}(u, t) \leq 0.$$

By continuity of $w_1^\pi(\cdot)$, in case (ii) we have as well $w_1^\pi(t) = 0$. Hence it follows immediately from $w_1^\pi(t) + w_2^\pi(t) = w_1^{\tilde{\pi}}(t) + w_2^{\tilde{\pi}}(t)$ that $w_i^\pi(t) = w_i^{\tilde{\pi}}(t)$, $i = 1, 2$.

- Now consider the case when there is an interval $[v, t]$ such that $w_1^\pi(u) + w_2^\pi(u) = w_1^{\tilde{\pi}}(u) + w_2^{\tilde{\pi}}(u)$ for all $u \in [v, t]$ and $w_1^\pi(v^-) + w_2^\pi(v^-) < w_1^{\tilde{\pi}}(v^-) + w_2^{\tilde{\pi}}(v^-)$. From the previous item, we obtain that $w_i^\pi(v) = w_i^{\tilde{\pi}}(v)$, $i = 1, 2$. Together with the fact that in the interval $[v, t]$ the total workload is equal under both policies, and by construction of policy π , it follows that π does exactly the same as policy $\tilde{\pi}$. Hence, $w_i^\pi(u) = w_i^{\tilde{\pi}}(u)$ for all $u \in [v, t]$, $i = 1, 2$.

Part II: Assume (55) is the first equation that fails to hold, i.e., $(1 - c_2)U_1^\pi(t) + c_1(U_2^\pi(t) + U_c^\pi(t)) = (1 - c_2)U_1^{\tilde{\pi}}(t) + c_1(U_2^{\tilde{\pi}}(t) + U_c^{\tilde{\pi}}(t))$, and by (13) also $(1 - c_2)w_1^\pi(t) + c_1w_2^\pi(t) = (1 - c_2)w_1^{\tilde{\pi}}(t) + c_1w_2^{\tilde{\pi}}(t)$, while $(1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) < (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$. We have the following possibilities:

- If $w_1^\pi(t^+) > 0$ and $w_2^\pi(t^+) > 0$, then by definition of policy π we have $(1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) = (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$.
- If $w_1^\pi(t^+) = 0$ and $w_2^\pi(t^+) > 0$, then we distinguish between the following three cases:
 - (i) If $\rho_1 \leq c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$, then $u_1^\pi(t^+) = 0$, $u_2^\pi(t^+) = u_1^{\tilde{\pi}}(t^+) + u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+) - \frac{\rho_1}{c_1}$ and $u_c^\pi(t^+) = \frac{\rho_1}{c_1}$. Since $c_1 + c_2 > 1$, we have

$$\begin{aligned} (1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) &= c_1(u_1^{\tilde{\pi}}(t^+) + u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)) \\ &\geq (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)). \end{aligned}$$

- (ii) If $c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)) < \rho_1 \leq u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$, then $u_1^\pi(t^+) = \rho_1 - c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$, $u_2^\pi(t^+) = u_1^{\tilde{\pi}}(t^+) - \rho_1 + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$ and $u_c^\pi(t^+) = u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)$. Together with $c_1 + c_2 > 1$, we obtain

$$\begin{aligned} &(1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) \\ &= (1 - c_2)(\rho_1 - c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))) \\ &+ c_1(u_1^{\tilde{\pi}}(t^+) + u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+) - \rho_1 + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))) \\ &= (1 - c_1 - c_2)\rho_1 + c_1(c_1 + c_2)(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)) + c_1u_1^{\tilde{\pi}}(t^+) \\ &\geq (1 - c_1 - c_2)(u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))) + c_1(c_1 + c_2)(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)) + c_1u_1^{\tilde{\pi}}(t^+) \\ &= (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)). \end{aligned}$$

- (iii) If $u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)) < \rho_1$, then $u_1^\pi(t^+) = u_1^{\tilde{\pi}}(t^+)$, $u_2^\pi(t^+) = 0$ and $u_c^\pi(t^+) = u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)$. So we have $(1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) = (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$.

- If $w_1^\pi(t^+) > 0$ and $w_2^\pi(t^+) = 0$, then by continuity of $w_2^\pi(\cdot)$ we have $w_2^\pi(t) = 0$. Hence, $(1 - c_2)w_1^\pi(t) = (1 - c_2)w_1^{\tilde{\pi}}(t) + c_1w_2^{\tilde{\pi}}(t)$. Since also $w_1^\pi(t) \leq w_1^{\tilde{\pi}}(t)$, this gives $w_1^\pi(t) = w_1^{\tilde{\pi}}(t)$ and $0 = w_2^\pi(t) = w_2^{\tilde{\pi}}(t)$. Note that when $w_2^{\tilde{\pi}}(t^+) = 0$, then $u_2^{\tilde{\pi}}(t^+) + c_2u_c^{\tilde{\pi}}(t^+) = \rho_2$. If instead $w_2^{\tilde{\pi}}(t^+) > 0$, then $u_2^{\tilde{\pi}}(t^+) + c_2u_c^{\tilde{\pi}}(t^+) < u_2^\pi(t^+) + c_2u_c^\pi(t^+) = \rho_2$ (the inequality follows from $0 = w_2^\pi(t) = w_2^{\tilde{\pi}}(t)$, and the fact that policy π is able to keep class 2 empty at time t^+ , while policy $\tilde{\pi}$ is not). Hence, it holds that $u_2^{\tilde{\pi}}(t^+) + c_2u_c^{\tilde{\pi}}(t^+) \leq \rho_2$ (so also $u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+) \leq \frac{\rho_2}{c_2}$). By construction of policy π , this implies $u_c^\pi(t^+) = u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+)$, $u_1^\pi(t^+) = u_1^{\tilde{\pi}}(t^+)$ and $u_2^\pi(t^+) = 0$. Hence, $(1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) = (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$.

For all the three possibilities we reach a contradiction with $(1 - c_2)u_1^\pi(t^+) + c_1(u_2^\pi(t^+) + u_c^\pi(t^+)) < (1 - c_2)u_1^{\tilde{\pi}}(t^+) + c_1(u_2^{\tilde{\pi}}(t^+) + u_c^{\tilde{\pi}}(t^+))$ and this concludes the proof. \square