Heavy-traffic analysis of the M/PH/1 discriminatory processor sharing queue with phase-dependent weights[*]

I.M. Verloop[a], U. Ayesta[b,c], R. Núñez-Queija[a,d]

[a]CWI, Amsterdam, The Netherlands
[b]BCAM - Basque Center for Applied Mathematics, Derio, Spain
[c]LAAS-CNRS, Université de Toulouse, France
[d]University of Amsterdam, The Netherlands

ABSTRACT

We analyze a generalization of the Discriminatory Processor Sharing (DPS) queue in a heavy-traffic setting. Customers present in the system are served simultaneously at rates controlled by a vector of weights. We assume phase-type distributed service requirements and allow that customers have different weights in various phases of their service.

We establish a state-space collapse for the queue length vector in heavy traffic. The result shows that in the limit, the queue length vector is the product of an exponentially distributed random variable and a deterministic vector. This generalizes a previous result by [12] who considered a DPS queue with exponentially distributed service requirements. We finally discuss some implications for residual service requirements and monotonicity properties in the ordinary DPS model.

1. INTRODUCTION

The Discriminatory Processor Sharing (DPS) model, introduced in [11], is a versatile generalization of the celebrated (Egalitarian) Processor Sharing (PS) model. DPS allows class-based differentiation by assigning different weights to customers of different classes. As new customers join the system and others leave after having completed their service requirement, the actual resource allocation to each customer fluctuates dynamically over time.

The literature devoted to the analysis of DPS has been significantly extended over the past decade, as renewed interest in DPS arose due to its relevance in communication networks with distributed control, in particular the Internet [2]. An extensive survey of the DPS literature can be found in [1]. The seminal paper [6] provided the first analysis of the mean sojourn time conditioned on the service requirement. As a by-product, the mean queue lengths of the various classes were shown to depend on the *entire* service requirement distributions, of all customer classes. This as opposed to the egalitarian PS model, where the marginal queue lengths have a geometric distribution that only depends on the average loads of all classes. Although not strictly insensitive towards higher moments of service requirement distributions, the DPS model was shown to have finite mean queue lengths irrespective of any higher-order characteristics [3].

Several papers have analyzed the discriminatory processor sharing model assuming overload conditions with general service requirement distributions. In [2] the authors determine the queue length growth rates of the standard DPS model. Extensions to bandwidth-sharing networks [5] and a framework similar to ours [4], have been obtained more recently. In these references the *transient* behavior of the queue lengths is studied under overload conditions, while we investigate the convergence of the (scaled) *steady-state* distribution as the critical load is approached.

In the present paper, we assume that customers have phase-type service requirement distributions and we allow for different weights in various phases of their service. This extension allows for example to incorporate sophisticated scheduling techniques that give preferential treatment to customers that are likely to be close to service completion, thus reducing the numbers of customers in the system and their mean response times, cf. [13]. Similar generalizations of DPS were previously considered by [4, 7, 8]. The analysis in [7] assumes heavy traffic conditions and finite second moments of the service times. Through appropriate choices for a quite general functional of the queue length process, [7] determines the heavy-traffic distributions of the marginal queue lengths and response times by using a direct relationship with critical Crump-Mode-Jagers branching processes. Our results are complementary: On one hand we mainly concentrate on the queue lengths, and on the other hand we study the *joint* queue length distribution. Doing so, we establish a *state-space collapse* for the queue length vector in heavy traffic. The reduction of dimensionality of a multi-variate stochastic process under asymptotic (heavy-traffic) scaling has been demonstrated previously in other queueing models, see for example [9, 14].

Our work is inspired by the heavy-traffic analysis of the traditional DPS model with exponential service requirements in [12], where it was used that the variability of the queue length vector is of a lower order than the mean queue lengths. In [10] it was indicated that a similar approach as in [12] can be used for DPS with phase-type distributions as well. Here

we follow a different and more direct approach, by investigating the joint probability generating function of the queue lengths in a heavy-traffic setting.

2. MAIN RESULT

Our model can be viewed as a service station with J customer types. Customers arrive according to a Poisson arrival process with rate λ, and an arriving customer is of type i with probability p_{0i}. Customers of type i have an exponentially distributed service requirement with mean $\frac{1}{\mu_i}$. After service completion, customers of type i become of type j with probability p_{ij}, and leave the system with probability $p_{i0} := 1 - \sum_{j=1}^{J} p_{ij}$. We denote the number of type-j customers in stationarity by Q_j. The J customer types share a common resource of capacity one. There are positive weights $g_1, \ldots, g_J$ associated with each of the types. Whenever there are q_i type-i customers, $i = 1, \ldots, J$, present in the system, each type-j customer is served at rate

$$\frac{g_j}{\sum_{i=1}^{J} g_i q_i}, \quad j = 1, \ldots, J.$$

Note that, indeed, this framework essentially represents the M/PH/1 DPS queue with J service phases, allowing customers in different service phases (i.e., customers of different types in the above description) to have different weights. In fact, since customers may start in each of the J phases, one can also view it as a *multi-class* M/PH/1 DPS queue having phase-dependent weights, with at most J classes, and each class having a distribution with at most J phases. The fact that all phase-type distributions are composed with the same set of exponential phases is no restriction, since the number J is arbitrary and phases need not mutually communicate. To avoid confusion, we stick to the terminology of types used in the model description above.

The expected remaining service requirement until departure for a customer that is now of type i, denoted by $\mathbb{E}R_i$, satisfies $\mathbb{E}(R_i) = \frac{1}{\mu_i} + \sum_{j=1}^{J} p_{ij} \mathbb{E}(R_j)$. Note that this includes service in all subsequent stages as the customer changes from one type to another. Let $\mathbb{E}(\bar{R}) = (\mathbb{E}(R_1), \ldots, \mathbb{E}(R_J))^T$ and let P be a $J \times J$ matrix with $P = (p_{ij})$, $i, j = 1, \ldots, J$. Since P is a sub-stochastic matrix, $(I - P)^{-1}$ is well defined and we can write

$$\mathbb{E}(\bar{R}) = (I - P)^{-1} \bar{m}, \quad \text{with} \quad \bar{m} = (1/\mu_1, \ldots, 1/\mu_J)^T.$$

Define the total traffic load by

$$\rho := \lambda \sum_{j=1}^{J} p_{0j} \mathbb{E}(R_j).$$

Let γ_i represent the expected number of times a customer is of type i during its stay at the service station. Hence, $\gamma_1, \ldots, \gamma_J$, satisfy the following equations

$$\gamma_i = p_{0i} + \sum_{j=1}^{J} \gamma_j p_{ji}, \quad i = 1, \ldots, J,$$

i.e., $\bar{\gamma}^T = \bar{p}_0^T (I - P)^{-1}$, with $\bar{\gamma} = (\gamma_1, \ldots, \gamma_J)^T$ and $\bar{p}_0 = (p_{01}, \ldots, p_{0J})^T$. Note that $\frac{\gamma_i}{\mu_i}$ represents the expected cumulative amount of service a customer requires while being of type i during its stay at the station. We denote the load corresponding to customers while they are of type i, by

$$\rho_i := \lambda \frac{\gamma_i}{\mu_i}.$$

Hence, for the total traffic load ρ we may equivalently write

$$\rho = \lambda \bar{p}_0^T \mathbb{E}(\bar{R}) = \lambda \bar{p}_0^T (I - P)^{-1} \bar{m} = \lambda \bar{\gamma}^T \bar{m} = \sum_{j=1}^{J} \rho_j.$$

Our main result shows that the steady-state distribution of the multi-dimensional queue length process takes a rather simple form when the system is near saturation, i.e., $\rho \uparrow 1$, which is commonly referred to as the heavy-traffic regime. This regime can be accomplished by fixing the $\bar{p}_0, P$ and $\bar{m}$, and letting

$$\lambda \uparrow \hat{\lambda} := \frac{1}{\bar{p}_0^T (I - P)^{-1} \bar{m}},$$

since then $\rho = \lambda \bar{p}_0^T (I - P)^{-1} \bar{m} \uparrow 1$. In heavy traffic, we denote by

$$\hat{\rho}_i = \hat{\lambda} \frac{\gamma_i}{\mu_i}$$

the load corresponding to customers while they are of type i ($\sum_{j=1}^{J} \hat{\rho}_j = 1$).

We will now state our main result, which establishes a state-space collapse for the queue length vector in the heavy-traffic regime.

PROPOSITION 2.1. *When scaled with $1 - \rho$, the queue length vector has a proper limiting distribution as $(\rho_1, \ldots, \rho_J) \to (\hat{\rho}_1, \ldots, \hat{\rho}_J)$, such that $\rho \uparrow 1$,*

$$(1 - \rho)(Q_1, Q_2, \ldots, Q_J) \xrightarrow{d} X \cdot \left(\frac{\hat{\rho}_1}{g_1}, \frac{\hat{\rho}_2}{g_2}, \ldots, \frac{\hat{\rho}_J}{g_J} \right),$$

where $\xrightarrow{d}$ denotes convergence in distribution and X is an exponentially distributed random variable with mean

$$\mathbb{E}(X) = \frac{\sum_{j=1}^{J} \hat{\rho}_j \mathbb{E}(R_j)}{\sum_{j=1}^{J} \frac{\hat{\rho}_j}{g_j} \mathbb{E}(R_j)}.$$

3. SKETCH OF THE PROOF

The basis for the proof of Proposition 2.1 is a functional equation for the generating function of the joint queue length process which we derive here. Denote by $\bar{Q}$ and $\bar{q}$ the vectors $(Q_1, Q_2, \ldots, Q_J) \geq \bar{0}$ and $(q_1, q_2, \ldots, q_J) \geq \bar{0}$, respectively. The corresponding equilibrium distribution is denoted by $\pi(\bar{q}) := \mathbb{P}(\bar{Q} = \bar{q})$. For notational convenience we use the following transformation:

$$R(\bar{0}) := 0 \quad \text{and} \quad R(\bar{q}) := \frac{\pi(\bar{q})}{\sum_{j=1}^{J} g_j q_j}, \quad \text{for} \ \bar{q} \neq \bar{0}.$$

Also, let $p(\bar{z})$ and $r(\bar{z})$ denote the generating functions of $\pi(\bar{q})$ and $R(\bar{q})$, respectively, where $\bar{z} = (z_1, \ldots, z_J)$ and $|z_i| < 1$ for $i = 1, \ldots, J$:

$$p(\bar{z}) = \sum_{q_1=0}^{\infty} \cdots \sum_{q_J=0}^{\infty} z_1^{q_1} \cdot \ldots \cdot z_J^{q_J} \pi(\bar{q}),$$

$$r(\bar{z}) = \sum_{q_1=0}^{\infty} \cdots \sum_{q_J=0}^{\infty} z_1^{q_1} \cdot \ldots \cdot z_J^{q_J} R(\bar{q}).$$

Since $\pi(\bar{0}) = 1 - \rho$, it follows that

$$\sum_{i=1}^{J} g_i z_i \frac{\partial r(\bar{z})}{\partial z_i} + 1 - \rho = p(\bar{z}). \tag{1}$$

From the balance equations for $\pi(\bar{q})$ [10], we then obtain the following partial differential equation for $r(\bar{z})$:

$$\sum_{i=1}^{J} \left(\mu_i g_i (p_{i0} + \sum_{j=1}^{J} p_{ij} z_j - z_i) - \lambda g_i z_i (1 - \sum_{j=1}^{J} p_{0j} z_j) \right) \frac{\partial r}{\partial z_i}$$
$$= \lambda(1-\rho)(1 - \sum_{i=1}^{J} p_{0i} z_i). \tag{2}$$

This equation was derived in [12] for exponentially distributed service requirements. Equation (2) turns out to be very useful to analyze the joint queue length distribution in heavy traffic, because it allows for an explicit solution in that asymptotic regime as we will indicate next.

We write $\bar{s} = (s_1, \ldots, s_J)$ and use the short hand notation $\mathrm{e}^{-(1-\rho)\bar{s}} = (\mathrm{e}^{-(1-\rho)s_1}, \ldots, \mathrm{e}^{-(1-\rho)s_J})$. Our goal is to study

$$\lim_{\rho \uparrow 1} \mathbb{E}(\mathrm{e}^{-(1-\rho)s_1 Q_1} \cdot \ldots \cdot \mathrm{e}^{-(1-\rho)s_J Q_J}) = \lim_{\rho \uparrow 1} p(\mathrm{e}^{-(1-\rho)\bar{s}}).$$

From (1) it can be argued that there exists a function $\hat{r}(\cdot)$ such that

$$\lim_{\rho \uparrow 1} p(\mathrm{e}^{-(1-\rho)\bar{s}}) = \sum_{i=1}^{J} g_i \frac{\partial \hat{r}(\bar{s})}{\partial s_i}. \tag{3}$$

Taking $\bar{z}$ equal to $\mathrm{e}^{-(1-\rho)\bar{s}}$ in (2), dividing both sides by $1-\rho$ and taking the limit of $\rho \uparrow 1$, this gives

$$\sum_{i=1}^{J} g_i \left(\mu_i(s_i - \sum_{j=1}^{J} p_{ij} s_j) - \hat{\lambda} \sum_{j=1}^{J} p_{0j} s_j \right) \frac{\partial \hat{r}(\bar{s})}{\partial s_i} = 0.$$

From this partial differential equation it can be shown that the function $\hat{r}(\bar{s})$ is constant on the $J-1$ dimensional hyperplane

$$H_c := \{\bar{s} \geq \bar{0} : \sum_{j=1}^{J} \frac{\hat{\rho}_j}{g_j} s_j = c\}, \quad c > 0.$$

Together with (3), it follows that $\lim_{\rho \uparrow 1} p(\mathrm{e}^{-(1-\rho)\bar{s}})$ depends only on $\bar{s}$ through $\sum_{j=1}^{J} \frac{\hat{\rho}_j}{g_j} s_j$. This can be used to prove Proposition 2.1, see [15] for further details.

4. THE STANDARD DPS QUEUE

The standard DPS model with several customer classes and phase-type distributed service requirements is a special case of the framework studied here. In particular, Proposition 2.1 gives a proof of the state space collapse stated in [10, Theorem 5]. It also implies several other interesting new results for the standard DPS queue, which we briefly mention here, referring to the technical report [15] for full details.

For example, we can show that the (scaled) numbers of customers in the various classes and the *remaining* service requirements of any finite subset of customers are independent in a heavy-traffic setting. In particular, the remaining service requirement of any customer is distributed according to the forward recurrence time of its service requirement.

In addition, it follows for a heavy-traffic setting that the total mean queue length in a standard DPS queue reduces as customers with lower variability (measured in terms of the mean forward recurrence time) in their service requirements obtain larger weights. This property can be understood from the standard intuition of size-based scheduling: Customers belonging to classes with highly variable service distributions are likely to have longer residual service requirements and should therefore be given lower priority.

5. REFERENCES

[1] E. Altman, K.E. Avrachenkov, and U. Ayesta. A survey on discriminatory processor sharing. *Queueing Systems*, 53:53–63, 2006.

[2] E. Altman, T. Jimenez, and D. Kofman. DPS queues with stationary ergodic service times and the performance of TCP in overload. In *Proceedings of IEEE INFOCOM*, Hong Kong, 2004.

[3] K.E. Avrachenkov, U. Ayesta, P. Brown, and R. Núñez-Queija. Discriminatory processor sharing revisited. In *Proceedings of IEEE INFOCOM*, Miami, FL, USA, 2005.

[4] A. Ben Tahar and A. Jean-Marie. The fluid limit of the multiclass processor sharing queue. *INRIA Research Report*, RR-6867, 2009.

[5] R. Egorova, S.C. Borst, and A.P. Zwart. Bandwidth-sharing networks in overload. *Performance Evaluation*, 64:978–993, 2007.

[6] G. Fayolle, I. Mitrani, and R. Iasnogorodski. Sharing a processor among many job classes. *Journal of the ACM*, 27:519–532, 1980.

[7] G. Grishechkin. On a relationship between processor-sharing queues and Crump-Mode-Jagers branching processes. *Advances in Applied Probability*, 24:653–698, 1992.

[8] M. Haviv and J. van der Wal. Mean sojourn times for phase-type discriminatory processor sharing systems. *European Journal of Operational Research*, 189:375–386, 2008.

[9] W.N. Kang, F.P. Kelly, N.H. Lee, and R.J. Williams. State space collapse and diffusion approximation for a network operating under a fair bandwidth sharing policy. *Annals of Applied Probability*, 2009. To appear.

[10] G. van Kessel, R. Núñez-Queija, and S.C. Borst. Asymptotic regimes and approximations for discriminatory processor sharing. *Performance Evaluation Review*, 32:44–46, 2004.

[11] L. Kleinrock. Time-shared systems: A theoretical treatment. *Journal of the ACM*, 14:242–261, 1967.

[12] K.M. Rege and B. Sengupta. Queue length distribution for the discriminatory processor-sharing queue. *Operations Research*, 44:653–657, 1996.

[13] R. Righter and J.G. Shanthikumar. Scheduling multiclass single server queueing systems to stochastically maximize the number of successful departures. *Probability in the Engineering and Informational Sciences*, 3:323–333, 1989.

[14] A.L. Stolyar. Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Annals of Applied Probability*, 14:1–53, 2004.

[15] I.M. Verloop, U. Ayesta, and R. Núñez-Queija. Heavy-traffic analysis of a multiple-phase network with discriminatory processor sharing. *CWI research report PNA-E0905*, 2009.