Available online at www.sciencedirect.com
ScienceDirect
Performance Evaluation 65 (2008) 586–605
PERFORMANCE EVALUATION
An International Journal
www.elsevier.com/locate/peva

Slowdown in the $M/M/1$ discriminatory processor-sharing queue☆

Sing-Kong Cheung[a], Bara Kim[b,*], Jeongsim Kim[c]

[a] *University of Twente, Department of Applied Mathematics, Stochastic Operations Research Group, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[b] *Korea University, Department of Mathematics, Telecommunication Mathematics Research Center, Anam-dong, Seongbuk-ku, Seoul 136-713, Republic of Korea*
[c] *Chungbuk National University, Department of Mathematics Education, 12, Gaeshin-dong, Heungduk-ku, Cheongju, Chungbuk, 361-763, Republic of Korea*

Received 25 January 2007; received in revised form 14 November 2007; accepted 19 November 2007
Available online 28 November 2007

Abstract

We consider a queue with multiple K job classes, Poisson arrivals, and exponentially distributed required service times in which a single processor serves according to the discriminatory processor-sharing (DPS) discipline. For this queue, we obtain the first and second moments of the slowdown, which is a measure for queueing fairness. We then provide numerical examples and discuss aspects of the slowdown in the DPS queue.

Keywords: Slowdown; Fairness; $M/M/1$ queue; Discriminatory processor-sharing; Egalitarian processor-sharing

1. Introduction

The discriminatory processor-sharing (DPS) discipline is of considerable interest and importance in computer and communication systems, as a convenient paradigm for modeling heterogeneous bandwidth sharing. The DPS discipline was introduced by Kleinrock [11] under the name priority processor-sharing. In the DPS service discipline for a single processor system with K job classes, all jobs present in the system are simultaneously served according to the set of weights $\{\alpha_i > 0, i = 1, \ldots, K\}$. If there are n_i class i jobs present in the system, then each class i job receives a fraction $\frac{\alpha_i}{\sum_{j=1}^{K} \alpha_j n_j}$ of the service capacity. When all α_i are equal, the DPS discipline reduces to the egalitarian processor-sharing (EPS) discipline. Under DPS it is possible to give preferential treatment to one or more job classes at the expense of others, by choosing a certain set of DPS weights. By appropriate choice of the weights, we may enable Quality-of-Service differentiation among different job classes.

☆ For the first author this work has been funded by the cooperation agreement between the Korea Science and Engineering Foundation (KOSEF), and the Netherlands Organization for Scientific Research (NWO). For the second author this research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment). For the second and the third authors this work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-312-C00470).
* Corresponding author.
E-mail address: bara@korea.ac.kr (B. Kim).

0166-5316/$ - see front matter
doi:10.1016/j.peva.2007.11.001

Exact analysis of DPS has proven to be more difficult as compared with the EPS discipline (e.g., see [4]). The available results for DPS are remarkably sparse compared to EPS. Kleinrock [11] and O'Donovan [13] obtained the conditional mean sojourn times for the $M/M/1$ DPS queue and the $M/G/1$ DPS queue, respectively. Fayolle et al. [6] proved that the expression for the conditional mean sojourn times obtained by Kleinrock [11] and O'Donovan [13] contains an error, and showed that the conditional mean sojourn times satisfy a system of integro-differential equations for the $M/G/1$ DPS queue. In the case of exponentially distributed required service times (job sizes), Fayolle et al. [6] derived closed-form expressions for the conditional mean sojourn times and obtained the unconditional mean sojourn times from a system of linear equations. Rege and Sengupta [14] obtained the higher moments of the queue length distribution from linear equations for the case of exponential service requirements; this result was extended to phase-type required service times by Van Kessel et al. [8]. Kim and Kim [9] found the higher moments of the sojourn times in the $M/M/1$ DPS queue as a solution of linear equations. For a recent survey on DPS queues, we refer to Altman et al. [1].

In this paper we investigate the so-called slowdown measures in the $M/M/1$ DPS queue. The slowdown is a way to measure how fairly jobs are treated by a service discipline (e.g. see [2,7,16]), and the mean slowdown is often used as a measure of system performance as opposed to the more traditional mean sojourn time. In general, it is desirable that a job's sojourn time should be correlated with its size; that is, we would like small jobs to have small sojourn times. The slowdown of a job is defined as sojourn time divided by job size, which eliminates the dominating effect of large jobs in the sojourn time measure.

It is well known that the mean slowdown in the $M/G/1$ EPS queue is $1/(1-\rho)$, which only depends on the offered load ρ. It is independent of the job size and it is also insensitive to the service requirement distributions. Therefore, EPS is often considered as a "fair" service discipline. Further support for the observation that EPS is fair has recently been provided in Cheung et al. [5]. They obtained insensitive upper bounds for all moments of the conditional sojourn time in the $M/G/1$ EPS queue, which immediately give upper bounds for moments of slowdown. The bounds are tight in some appropriate senses and they only depend on ρ and the job size.

The mean slowdown in the $M/G/1$ DPS queue depends on the job size, the job (class) type, and furthermore, the service requirement distributions. In the case of exponential service requirements, we obtain the first and second moments of the slowdown. The results can be used for numerical computation of the moments of conditional and unconditional slowdown. However, the expressions are too complicated to give any insight directly. We provide numerical examples and give some insights into the behavior of the slowdown measures.

It is a priori not clear how the unfairness depends on the job size and the weights. The jobs with the smallest DPS weight (we call this the lowest priority class) are obviously treated unfairly under DPS compared to EPS. In particular, short jobs of the lowest priority class are treated the most unfairly. It is also intuitively clear that the jobs with the largest DPS weight are treated better under DPS than under EPS, in terms of the slowdown measure.

More interestingly, when the DPS model has three or more job classes, then it is not immediately clear how the jobs of the "middle classes" (classes with weights in between the largest and lowest weights) are treated. Depending on the parameters, it is possible that the middle class jobs are always treated better or worse under DPS than under EPS. However, in some specific scenario settings, sometimes the middle class jobs are treated better under DPS than under EPS, and sometimes worse under DPS than under EPS.

The paper is organized as follows. In Section 2, we give a short review of the $M/M/1$ DPS results which are used in the current paper. In Section 3, we obtain the first and second moments of the slowdown. In Section 4, we provide numerical examples and give some insights into the behavior of the slowdown measures. Finally, in Section 5, we provide a conclusion.

2. Preliminaries

We consider a DPS queue with K job classes. Class i jobs arrive in a Poisson stream with rate λ_i, and have exponentially distributed required service times with mean μ_i^{-1}, for $i = 1, \ldots, K$. We denote the load of class i jobs by $\rho_i = \frac{\lambda_i}{\mu_i}$, and we assume that the total offered load $\rho \equiv \sum_{i=1}^{K} \rho_i$ is less than 1. In the DPS service discipline, all jobs present in the system are simultaneously served according to the set of weights $\{\alpha_i > 0, i = 1, \ldots, K\}$. If there are n_i class i jobs present in the system, then each class i job receives a fraction $\frac{\alpha_i}{\sum_{j=1}^{K} \alpha_j n_j}$ of the service capacity.

In the following subsection we give a short review of the $M/M/1$ DPS results which are used in the current paper. For the proofs we refer the reader to Rege and Sengupta [14] and Kim and Kim [9].

2.1. Moments of the number of jobs

Let $N_i,\ i = 1, \ldots, K$, be the number of class i jobs in the system at steady state. We let $Q(z_1, \ldots, z_K)$ denote the joint probability generating function of the number of each class job in the system at steady state:

$$Q(z_1, \ldots, z_K) \equiv \mathbb{E}\left(z_1^{N_1} \cdots z_K^{N_K}\right),$$

and define the following moments:

$$L_j^1 \equiv \left.\frac{\partial}{\partial z_j} Q(z_1, \ldots, z_K)\right|_{z_1=\cdots=z_K=1},$$
$$L_{jk}^2 \equiv \left.\frac{\partial^2}{\partial z_j \partial z_k} Q(z_1, \ldots, z_K)\right|_{z_1=\cdots=z_K=1},$$

for $j, k = 1, \ldots, K$. Note that L_j^1 is the mean number of class j jobs at steady state.

By Eq. (16) of Rege and Sengupta [14], we have the system of linear equations for $L_l^1,\ l = 1, \ldots, K$:

$$L_l^1 - \sum_{j=1}^{K} \alpha_j \frac{\lambda_j L_l^1 + \lambda_l L_j^1}{\alpha_j \mu_j + \alpha_l \mu_l} = \frac{\lambda_l}{\mu_l}. \tag{2.1}$$

Solving the system of K linear equations (2.1) yields $L_l^1,\ l = 1, \ldots, K$. Further, we have a system of $\frac{K(K+1)}{2}$ equations for $L_{jk}^2,\ 1 \le j \le k \le K$, by Eq. (17) of Rege and Sengupta [14] and the fact that $L_{jk}^2 = L_{kj}^2$. The linear simultaneous equations for L_{jk}^2 are

$$L_{jk}^2 - \sum_{i=1}^{K} \alpha_i \frac{\lambda_j L_{ki}^2 + \lambda_k L_{ij}^2 + \lambda_i L_{jk}^2}{\alpha_j \mu_j + \alpha_k \mu_k + \alpha_i \mu_i} = (\alpha_j + \alpha_k) \frac{\lambda_j L_k^1 + \lambda_k L_j^1}{\alpha_j \mu_j + \alpha_k \mu_k},\quad 1 \le j \le k \le K,$$

where L_i^1 in the right-hand side is obtained by (2.1).

2.2. Moments of the sojourn time

We denote $T_i(x)$ as the steady state sojourn time of a class i job with required service time (job size) x. We note that $T_i(x)$ can also be interpreted as the time necessary for a class i job whose required service time is greater than x to attain service x. Let us tag a class i job with required service time greater than x. When the tagged job attains service x, let $N_{ij}(x)$ denote the number of class j jobs in the system, $j = 1, \ldots, K$ (excluding the tagged job). We introduce the following joint transform:

$$R_{ix}(s; z_1, \ldots, z_K) \equiv \mathbb{E}\left(\mathrm{e}^{-sT_i(x)}\ z_1^{N_{i1}(x)} \cdots z_K^{N_{iK}(x)}\right),$$

which is defined for $|z_i| \le 1, i = 1, \ldots, K$, and $\mathrm{Re}(s) \ge 0$.

To find the first and second moments of the sojourn time of class i jobs with required service time x, we define the following moments:

$$M_{ix}^0 \equiv \left.\frac{\partial}{\partial s} R_{ix}(s; z_1, \ldots, z_K)\right|_{s=0,\ z_1=\cdots=z_K=1},$$
$$M_{ix}^j \equiv \left.\frac{\partial}{\partial z_j} R_{ix}(s; z_1, \ldots, z_K)\right|_{s=0,\ z_1=\cdots=z_K=1},$$

and

$$M_{ix}^{00} \equiv \left.\frac{\partial^2}{\partial s^2} R_{ix}(s; z_1, \ldots, z_K)\right|_{s=0,\ z_1=\cdots=z_K=1},$$

$$M_{ix}^{0j} \equiv \frac{\partial^2}{\partial s \partial z_j} R_{ix}(s; z_1, \ldots, z_K)\Big|_{s=0,\ z_1=\cdots=z_K=1},$$
$$M_{ix}^{jk} \equiv \frac{\partial^2}{\partial z_j \partial z_k} R_{ix}(s; z_1, \ldots, z_K)\Big|_{s=0,\ z_1=\cdots=z_K=1},$$

where $i, j, k = 1, \ldots, K$. We note that $-M_{ix}^0$ and M_{ix}^{00} are the first and second moments of the sojourn time of class i jobs with required service time x, respectively, i.e., $\mathbb{E}(T_i(x)) = -M_{ix}^0$ and $\mathbb{E}\left(T_i^2(x)\right) = M_{ix}^{00}$.

Kim and Kim [9] derived the following system of first-order linear differential equations (see Eq. (20) in Kim and Kim [9]):

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{E}[T_i(x)] = \frac{1}{\alpha_i}\boldsymbol{\alpha}^t \boldsymbol{m}_i(x) + 1, \tag{2.2}$$

where

$$\boldsymbol{\alpha} = [\alpha_1\ \alpha_2\ \cdots\ \alpha_K]^t, \qquad \boldsymbol{m}_i(x) = [M_{ix}^1\ M_{ix}^2\ \cdots\ M_{ix}^K]^t.$$

Here, and subsequently, the superscript t denotes the transpose of a vector. The vector function $\boldsymbol{m}_i(x)$ satisfies (see Eq. (21) in Kim and Kim [9])

$$\frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{m}_i(x) = \frac{1}{\alpha_i} B \boldsymbol{m}_i(x) + \boldsymbol{\lambda}, \tag{2.3}$$

where

$$B = \boldsymbol{\lambda}\boldsymbol{\alpha}^t - \mathrm{diag}(\alpha_1\mu_1, \ldots, \alpha_K\mu_K), \quad \boldsymbol{\lambda} = [\lambda_1\ \lambda_2\ \cdots\ \lambda_K]^t.$$

Further,

$$\mathbb{E}[T_i(0)] = 0,\ i = 1, \ldots, K, \tag{2.4}$$

and

$$\boldsymbol{m}_i(0) = [L_1^1\ \cdots\ L_K^1]^t \equiv \boldsymbol{L}^1, \tag{2.5}$$

by the PASTA property.

Kim and Kim [9] also derived the following system of $\frac{(K+1)(K+2)}{2}$ first-order linear differential equations (see Eqs. (24)–(29) in Kim and Kim [9]):

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{E}[T_i^2(x)] = \frac{2}{\alpha_i}\boldsymbol{\alpha}^t \boldsymbol{y}_i(x) + 2\mathbb{E}[T_i(x)], \tag{2.6}$$

where $\boldsymbol{y}_i(x) = -[M_{ix}^{01}\ \cdots\ M_{ix}^{0K}]^t$, with

$$\boldsymbol{y}_i(0) = [0\ \cdots\ 0]^t. \tag{2.7}$$

Further,

$$\frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{y}_i(x) = \frac{1}{\alpha_i} Z_i(x)\boldsymbol{\alpha} + \frac{1}{\alpha_i} B \boldsymbol{y}_i(x) + \mathbb{E}[T_i(x)]\boldsymbol{\lambda} + \left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\boldsymbol{m}_i(x), \tag{2.8}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}Z_i(x) = \frac{1}{\alpha_i}BZ_i(x) + \frac{1}{\alpha_i}Z_i(x)B^t + \left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\boldsymbol{m}_i(x)\boldsymbol{\lambda}^t + \boldsymbol{\lambda}m_i(x)^t\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right), \tag{2.9}$$

where

$$Z_i(x) = \begin{bmatrix} M_{ix}^{11} & \cdots & M_{ix}^{1K} \\ \vdots & \ldots & \vdots \\ M_{ix}^{K1} & \cdots & M_{ix}^{KK} \end{bmatrix},$$

with

$$Z_i(0) = \begin{bmatrix} L_{11}^2 & \cdots & L_{1K}^2 \\ \vdots & \ldots & \vdots \\ L_{K1}^2 & \cdots & L_{KK}^2 \end{bmatrix} \equiv L^2. \tag{2.10}$$

3. Slowdown

The unconditional slowdown of a class i job, S_i, is defined as the sojourn time divided by the job size. The conditional slowdown of a class i job whose size is x, $S_i(x)$, is given by $\frac{T_i(x)}{x}$. In this section we obtain the mean of the conditional and unconditional slowdown, i.e., $\mathbb{E}[S_i(x)]$ and $\mathbb{E}[S_i]$. Further, we express the second moment of the conditional slowdown in terms of a Laplace transform, and then obtain the second moment of the unconditional slowdown. In what follows, when we speak of moments of the unconditional slowdown we will omit the adjective "unconditional", if no confusion arises.

We begin with the following lemma.

Lemma 1. (a) *The matrix* $B = \boldsymbol{\lambda}\boldsymbol{\alpha}^t - \mathrm{diag}(\alpha_1\mu_1, \ldots, \alpha_K\mu_K)$ *is diagonalizable and the eigenvalues of* B*, say* κ_j*,* $j = 1, \ldots, K$*, are all negative. Further,* B *can be written as*

$$B = [\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_K]\,\mathrm{diag}(\kappa_1, \ldots, \kappa_K)\begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_K \end{bmatrix}, \tag{3.1}$$

with real right eigenvectors $\boldsymbol{v}_j$ *and real left eigenvectors* $\boldsymbol{u}_j$ *satisfying* $\boldsymbol{u}_j\boldsymbol{v}_k = \delta_{jk}$.

(b) $\boldsymbol{\alpha}^t(-B)^{-1} = \frac{1}{1-\rho}\,[\mu_1^{-1} \cdots \mu_K^{-1}]$

(c) $B^{-1}\boldsymbol{\lambda} = \frac{-1}{1-\rho}\,[\frac{\rho_1}{\alpha_1} \cdots \frac{\rho_K}{\alpha_K}]^t$

(d) $\boldsymbol{\alpha}^t B^{-1}\boldsymbol{\lambda} = \frac{-\rho}{1-\rho}$.

Proof. Letting $D = \mathrm{diag}(d_1, \ldots, d_K)$ with $d_i = \frac{\sqrt{\alpha_i}}{\sqrt{\lambda_i}}$, $i = 1, \ldots, K$, yields

$$DBD^{-1} = (D\boldsymbol{\lambda})(\boldsymbol{\alpha}^t D^{-1}) - \mathrm{diag}(\alpha_1\mu_1, \ldots, \alpha_K\mu_K).$$

Since $D\boldsymbol{\lambda} = \left[\ \sqrt{\alpha_1\lambda_1} \ \cdots \ \sqrt{\alpha_K\lambda_K}\ \right]^t$ and $\boldsymbol{\alpha}^t D^{-1} = \left[\ \sqrt{\alpha_1\lambda_1} \ \cdots \ \sqrt{\alpha_K\lambda_K}\ \right]$, we have

$$DBD^{-1} = (D\boldsymbol{\lambda})(D\boldsymbol{\lambda})^t - \mathrm{diag}(\alpha_1\mu_1, \ldots, \alpha_K\mu_K),$$

which is symmetric. This implies that the matrix B is diagonalizable, and B can be written as (3.1). Further, since DBD^{-1} is a real symmetric matrix, the eigenvalues κ_j, $j = 1, \ldots, K$, of B, are all real, and the eigenvectors $\boldsymbol{v}_j$ and $\boldsymbol{u}_j$ can always be taken to be real. Note that

$$[\mu_1^{-1} \cdots \mu_K^{-1}]B = -(1-\rho)\boldsymbol{\alpha}^t \ < \ \boldsymbol{0}, \tag{3.2}$$

where $\mathbf{0}$ denotes a K-dimensional row vector with its components equal to zero, and the inequality between two vectors is interpreted componentwise. Therefore eigenvalues κ_j, $j = 1, \ldots, K$, are all negative (see Theorem 2.6 on page 46 in [15]), and the proof of (a) is complete. (b) follows from (3.2). (c) follows from the identity

$$B\left[\frac{\rho_1}{\alpha_1} \cdots \frac{\rho_K}{\alpha_K}\right]^t = (\rho - 1)\boldsymbol{\lambda}.$$

Finally, (d) is immediate from (b) or (c). □

3.1. First moment

The mean of the conditional slowdown for class i job whose size is x, $\mathbb{E}[S_i(x)]$, is given by $\mathbb{E}[S_i(x)] = \frac{\mathbb{E}[T_i(x)]}{x}$.

Theorem 1. *The means of the conditional sojourn time and the conditional slowdown for a class i job whose size is x are given by*

$$\mathbb{E}[T_i(x)] = \frac{1}{1-\rho}x + a - \alpha_i b + \sum_{j=1}^{K}(\alpha_i\xi_j - \eta_j)\mathrm{e}^{\frac{\kappa_j}{\alpha_i}x}, \tag{3.3}$$

$$\mathbb{E}[S_i(x)] = \frac{1}{1-\rho} + \frac{a-\alpha_i b}{x} + \sum_{j=1}^{K}(\alpha_i\xi_j - \eta_j)\frac{\mathrm{e}^{\frac{\kappa_j}{\alpha_i}x}}{x}, \tag{3.4}$$

where κ_j, $j = 1, \ldots, K$, are eigenvalues of B,

$$a = \frac{1}{1-\rho}\,[\mu_1^{-1} \cdots \mu_K^{-1}]\boldsymbol{L}^1,$$

$$b = \frac{1}{(1-\rho)^2}\,[\mu_1^{-1} \cdots \mu_K^{-1}]\left[\frac{\rho_1}{\alpha_1} \cdots \frac{\rho_K}{\alpha_K}\right]^t,$$

and

$$\eta_j = \frac{1}{1-\rho}\left([\mu_1^{-1} \cdots \mu_K^{-1}]\boldsymbol{v}_j\right)(\boldsymbol{u}_j L^1),$$

$$\xi_j = \frac{1}{(1-\rho)^2}\left([\mu_1^{-1} \cdots \mu_K^{-1}]\boldsymbol{v}_j\right)\left(\boldsymbol{u}_j\left[\frac{\rho_1}{\alpha_1} \cdots \frac{\rho_K}{\alpha_K}\right]^t\right),$$

with $\boldsymbol{v}_j$ and $\boldsymbol{u}_j$ given in Lemma 1.

Proof. Integrating (2.3) and using (2.5), we have

$$\begin{aligned}\boldsymbol{m}_i(x) &= \mathrm{e}^{\frac{1}{\alpha_i}Bx}\boldsymbol{L}^1 + \mathrm{e}^{\frac{1}{\alpha_i}Bx}\int_0^x \mathrm{e}^{-\frac{1}{\alpha_i}Bw}\mathrm{d}w\ \boldsymbol{\lambda}\\ &= \mathrm{e}^{\frac{1}{\alpha_i}Bx}\boldsymbol{L}^1 + \alpha_i B^{-1}\left(\mathrm{e}^{\frac{1}{\alpha_i}Bx} - I\right)\boldsymbol{\lambda}.\end{aligned} \tag{3.5}$$

Similarly, by (2.2) and (2.4) together with (3.5), we have

$$\begin{aligned}\mathbb{E}[T_i(x)] &= \frac{1}{\alpha_i}\boldsymbol{\alpha}^t\int_0^x \mathrm{e}^{\frac{1}{\alpha_i}Bw}\mathrm{d}w\ \boldsymbol{L}^1 + \boldsymbol{\alpha}^t B^{-1}\int_0^x\left(\mathrm{e}^{\frac{1}{\alpha_i}Bw} - I\right)\mathrm{d}w\ \boldsymbol{\lambda} + x\\ &= \left(1 - \boldsymbol{\alpha}^t B^{-1}\boldsymbol{\lambda}\right)x + \boldsymbol{\alpha}^t B^{-1}\left(\mathrm{e}^{\frac{1}{\alpha_i}Bx} - I\right)\boldsymbol{L}^1 + \alpha_i\boldsymbol{\alpha}^t B^{-1}\left(\mathrm{e}^{\frac{1}{\alpha_i}Bx} - I\right)B^{-1}\boldsymbol{\lambda}.\end{aligned}$$

By Lemma 1,

$$\begin{aligned}\mathbb{E}[T_i(x)] &= \frac{1}{1-\rho}x - \frac{1}{1-\rho}[\mu_1^{-1}\cdots\mu_K^{-1}]\left(\mathrm{e}^{\frac{1}{\alpha_i}Bx} - I\right)\boldsymbol{L}^1 \\ &\quad + \frac{\alpha_i}{(1-\rho)^2}[\mu_1^{-1}\cdots\mu_K^{-1}]\left(\mathrm{e}^{\frac{1}{\alpha_i}Bx} - I\right)\left[\frac{\rho_1}{\alpha_1}\cdots\frac{\rho_K}{\alpha_K}\right]^t \\ &= \frac{1}{1-\rho}x + \frac{1}{1-\rho}[\mu_1^{-1}\cdots\mu_K^{-1}]\boldsymbol{L}^1 - \frac{\alpha_i}{(1-\rho)^2}[\mu_1^{-1}\cdots\mu_K^{-1}]\left[\frac{\rho_1}{\alpha_1}\cdots\frac{\rho_K}{\alpha_K}\right]^t \\ &\quad + \sum_{j=1}^{K}(-\eta_j + \alpha_i\xi_j)\mathrm{e}^{\frac{\kappa_j}{\alpha_i}x},\end{aligned}$$

where

$$\eta_j = \frac{1}{1-\rho}\left([\mu_1^{-1}\cdots\mu_K^{-1}]\boldsymbol{v}_j\right)(\boldsymbol{u}_j\boldsymbol{L}^1),$$
$$\xi_j = \frac{1}{(1-\rho)^2}\left([\mu_1^{-1}\cdots\mu_K^{-1}]\boldsymbol{v}_j\right)\left(\boldsymbol{u}_j\left[\frac{\rho_1}{\alpha_1}\cdots\frac{\rho_K}{\alpha_K}\right]^t\right).$$

Hence (3.3) is obtained, and (3.4) is immediate from $\mathbb{E}[S_i(x)] = \frac{\mathbb{E}[T_i(x)]}{x}$. □

Remark. 1. An explicit expression for the conditional mean sojourn time was also obtained by Fayolle et al. [6] in a similar form to (3.3):

$$\mathbb{E}[T_i(x)] = \frac{1}{1-\rho}x + \sum_{j=1}^{\tilde{K}}\frac{e_j - \alpha_i d_j s_j}{s_j^2}\left(1 - \mathrm{e}^{\frac{s_j}{\alpha_i}x}\right), \quad i = 1,\ldots,K,$$

where $\tilde{K}$ is the number of distinct elements in the vector $(\alpha_1\mu_1,\ldots,\alpha_K\mu_K)$, and s_j, $j = 1,\ldots,\tilde{K}$, are the $\tilde{K}$ distinct roots of

$$\sum_{j=1}^{K}\frac{\lambda_j\alpha_j}{\mu_j\alpha_j + s} = 1.$$

(They showed that $\sum_{j=1}^{K}\frac{\lambda_j\alpha_j}{\mu_j\alpha_j+s} = 1$ has exactly $\tilde{K}$ distinct roots.) Further, d_j and e_j, $j = 1,\ldots,\tilde{K}$, are given by

$$d_j = \frac{\prod_{k=1}^{\tilde{K}}(\alpha_k\mu_k + s_j)}{s_j\prod_{k\neq j}^{\tilde{K}}(s_j - s_k)}, \quad j = 1,\ldots,\tilde{K},$$
$$e_j = \frac{\left[\sum_{k=1}^{K}\lambda_k\alpha_k^2/(\mu_k^2\alpha_k^2 - s_j^2)\right]\left[\prod_{k=1}^{\tilde{K}}(\mu_k^2\alpha_k^2 - s_j^2)\right]}{\prod_{k\neq j}^{\tilde{K}}(s_k^2 - s_j^2)}, \quad j = 1,\ldots,\tilde{K},$$

with the assumption that $\alpha_j\mu_j$, $j = 1,\ldots,\tilde{K}$, are distinct.

2. We can rewrite (3.4) as

$$\mathbb{E}[S_i(x)] = \frac{1}{1-\rho} + \sum_{j=1}^{K}(\eta_j - \alpha_i\xi_j)\frac{1 - \mathrm{e}^{\frac{\kappa_j}{\alpha_i}x}}{x}. \tag{3.6}$$

Recall that $\kappa_j < 0$ for all j. (3.4) is suited for investigation when $x \to \infty$, and (3.6) is suited when $x \to 0$.

The mean slowdown for class i jobs in the $M/M/1$ DPS queue is given by

$$\mathbb{E}S_i = \int_0^\infty \mathbb{E}[S_i(x)]\mu_i \mathrm{e}^{-\mu_i x}\mathrm{d}x.$$

In the following theorem we give an expression for $\mathbb{E}S_i$.

Theorem 2. *The mean slowdown $\mathbb{E}S_i$ for class i jobs is given by*

$$\mathbb{E}S_i = \frac{1}{1-\rho} + \mu_i \sum_{j=1}^{K} (\eta_j - \alpha_i \xi_j) \log\left(1 - \frac{\kappa_j}{\alpha_i \mu_i}\right),$$

where η_j and ξ_j are given in Theorem 1.

Proof. Let $\widetilde{S}_i(s)$, $s > 0$, be the Laplace transform (LT) of $\mathbb{E}[S_i(x)]$, i.e.,

$$\widetilde{S}_i(s) \equiv \int_0^\infty \mathrm{e}^{-sx}\mathbb{E}[S_i(x)]\mathrm{d}x.$$

Taking Laplace transforms (LTs) in (3.6) yields

$$\widetilde{S}_i(s) = \frac{1}{1-\rho}\frac{1}{s} + \sum_{j=1}^{K} (\eta_j - \alpha_i \xi_j) \int_0^\infty \mathrm{e}^{-sx} \frac{1 - \mathrm{e}^{\frac{\kappa_j}{\alpha_i}x}}{x}\mathrm{d}x.$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_0^\infty \mathrm{e}^{-sx}\frac{1-\mathrm{e}^{\frac{\kappa_j}{\alpha_i}x}}{x}\mathrm{d}x = \int_0^\infty \mathrm{e}^{-sx}\left(\mathrm{e}^{\frac{\kappa_j}{\alpha_i}x} - 1\right)\mathrm{d}x = \frac{1}{s - \frac{\kappa_j}{\alpha_i}} - \frac{1}{s},$$

we have

$$\int_0^\infty \mathrm{e}^{-sx}\frac{1-\mathrm{e}^{\frac{\kappa_j}{\alpha_i}x}}{x}\mathrm{d}x = \log\left(\frac{s - \frac{\kappa_j}{\alpha_i}}{s}\right) = \log\left(1 - \frac{\kappa_j}{\alpha_i s}\right).$$

Therefore

$$\widetilde{S}_i(s) = \frac{1}{1-\rho}\frac{1}{s} + \sum_{j=1}^{K} (\eta_j - \alpha_i \xi_j) \log\left(1 - \frac{\kappa_j}{\alpha_i s}\right).$$

Since $\mathbb{E}S_i = \mu_i \widetilde{S}_i(\mu_i)$, the proof is complete. □

3.2. Second moment

Define the LT of $\mathbb{E}[S_i^2(x)]$ by

$$\widetilde{G}_i(s) \equiv \int_0^\infty \mathrm{e}^{-sx}\mathbb{E}[S_i^2(x)]\mathrm{d}x;$$

hence it follows that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) = \int_0^\infty \mathrm{e}^{-sx}\mathbb{E}[T_i^2(x)]\mathrm{d}x.$$

The above equation is expressed as follows.

Theorem 3. *We have*

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) = \frac{2}{\alpha_i^2}\frac{1}{s}\sum_{j=1}^{K} c_j \left\{ \boldsymbol{\alpha}^t \sum_{k=1}^{K} \frac{\boldsymbol{v}_k \boldsymbol{u}_k L^2 \boldsymbol{u}_j^t}{(s - \frac{\kappa_k}{\alpha_i})(s - \frac{\kappa_j + \kappa_k}{\alpha_i})} \right.$$

$$
\begin{aligned}
&+\boldsymbol{\alpha}^t \sum_{k=1}^{K}\sum_{m=1}^{K} \frac{\boldsymbol{v}_k\boldsymbol{u}_k(I+\frac{1}{\alpha_i}\mathrm{diag}(\alpha_1,\ldots,\alpha_K))\boldsymbol{v}_m\boldsymbol{u}_m(\boldsymbol{\lambda}+s\boldsymbol{L}^1)(\boldsymbol{\lambda}^t\boldsymbol{u}_j^t)}{s(s-\frac{\kappa_k}{\alpha_i})(s-\frac{\kappa_j+\kappa_k}{\alpha_i})(s-\frac{\kappa_m}{\alpha_i})} \\
&+\boldsymbol{\alpha}^t \sum_{k=1}^{K}\sum_{m=1}^{K} \frac{\boldsymbol{v}_k\boldsymbol{u}_k\boldsymbol{\lambda}(\boldsymbol{\lambda}+s\boldsymbol{L}^1)^t\boldsymbol{u}_m^t\boldsymbol{v}_m^t(I+\frac{1}{\alpha_i}\mathrm{diag}(\alpha_1,\ldots,\alpha_K))\boldsymbol{u}_j^t}{s(s-\frac{\kappa_k}{\alpha_i})(s-\frac{\kappa_j+\kappa_k}{\alpha_i})(s-\frac{\kappa_m}{\alpha_i})} \Bigg\} \\
&+\frac{2}{s^3}\left(1+\frac{1}{\alpha_i}\sum_{j=1}^{K}\frac{\boldsymbol{\alpha}^t\boldsymbol{v}_j\boldsymbol{u}_j(\boldsymbol{\lambda}+s\boldsymbol{L}^1)}{s-\frac{\kappa_j}{\alpha_i}}\right)\left(\frac{1}{\alpha_i}\sum_{k=1}^{K}\frac{\boldsymbol{\alpha}^t\boldsymbol{v}_k\boldsymbol{u}_k\boldsymbol{\lambda}}{s-\frac{\kappa_k}{\alpha_i}}+1\right) \\
&+\frac{2}{\alpha_i}\sum_{j=1}^{K}\sum_{k=1}^{K}\frac{\boldsymbol{\alpha}^t\boldsymbol{v}_j\boldsymbol{u}_j(I+\frac{1}{\alpha_i}\mathrm{diag}(\alpha_1,\ldots,\alpha_K))\boldsymbol{v}_k\boldsymbol{u}_k(\boldsymbol{\lambda}+s\boldsymbol{L}^1)}{s^2(s-\frac{\kappa_j}{\alpha_i})(s-\frac{\kappa_k}{\alpha_i})}, \qquad (3.7)
\end{aligned}
$$

where $[c_1\ \cdots\ c_K]=\boldsymbol{\alpha}^t[\boldsymbol{v}_1\ \cdots\ \boldsymbol{v}_K]$.

Proof. Let $\widetilde{T}_i(s)$, $s>0$, be the LT of $\mathbb{E}[T_i(x)]$, i.e.,

$$\widetilde{T}_i(s)\equiv\int_0^\infty \mathrm{e}^{-sx}\mathbb{E}[T_i(x)]\mathrm{d}x.$$

Taking LTs in (2.2), we readily obtain

$$\widetilde{T}_i(s)=\frac{1}{s}\left(\frac{1}{\alpha_i}\boldsymbol{\alpha}^t\widetilde{\boldsymbol{m}}_i(s)+\frac{1}{s}\right), \qquad (3.8)$$

where $\widetilde{\boldsymbol{m}}_i(s)$ is the LT of $\boldsymbol{m}_i(x)$. Similarly, from (2.3), we have

$$s\widetilde{\boldsymbol{m}}_i(s)-\boldsymbol{m}_i(0)=\frac{1}{\alpha_i}B\widetilde{\boldsymbol{m}}_i(s)+\frac{1}{s}\boldsymbol{\lambda}.$$

Since $\boldsymbol{m}_i(0)=\boldsymbol{L}^1$, the equation becomes

$$\widetilde{\boldsymbol{m}}_i(s)=\alpha_i(\alpha_i sI-B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda}+\boldsymbol{L}^1\right), \qquad (3.9)$$

and substitution into (3.8) yields

$$\widetilde{T}_i(s)=\frac{1}{s^2}+\frac{1}{s^2}\boldsymbol{\alpha}^t(\alpha_i sI-B)^{-1}\left(\boldsymbol{\lambda}+s\boldsymbol{L}^1\right). \qquad (3.10)$$

Taking LTs in (2.6) and using (3.10) leads to

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s)=\frac{1}{s}\frac{2}{\alpha_i}\boldsymbol{\alpha}^t\widetilde{\boldsymbol{y}}_i(s)+2\left(\frac{1}{s^3}+\frac{1}{s^3}\boldsymbol{\alpha}^t(\alpha_i sI-B)^{-1}(\boldsymbol{\lambda}+s\boldsymbol{L}^1)\right), \qquad (3.11)$$

where $\widetilde{\boldsymbol{y}}_i(s)$ is the LT of $\boldsymbol{y}_i(x)$. Taking LTs in (2.8) and using (3.9), (3.10), and (2.7), we obtain

$$
\begin{aligned}
\frac{1}{\alpha_i}(\alpha_i sI-B)\widetilde{\boldsymbol{y}}_i(s)=&\frac{1}{\alpha_i}\widetilde{Z}_i(s)\boldsymbol{\alpha}+\left(\frac{1}{s^2}+\frac{1}{s^2}\boldsymbol{\alpha}^t(\alpha_i sI-B)^{-1}(\boldsymbol{\lambda}+s\boldsymbol{L}^1)\right)\boldsymbol{\lambda} \\
&+\left(I+\frac{1}{\alpha_i}\mathrm{diag}(\alpha_1,\ldots,\alpha_K)\right)\alpha_i(\alpha_i sI-B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda}+\boldsymbol{L}^1\right),
\end{aligned}
$$

or, equivalently,

$$
\begin{aligned}
\widetilde{\boldsymbol{y}}_i(s)=&(\alpha_i sI-B)^{-1}\widetilde{Z}_i(s)\boldsymbol{\alpha}+\alpha_i\left(\frac{1}{s^2}+\frac{1}{s^2}\boldsymbol{\alpha}^t(\alpha_i sI-B)^{-1}(\boldsymbol{\lambda}+s\boldsymbol{L}^1)\right)(\alpha_i sI-B)^{-1}\boldsymbol{\lambda} \\
&+\alpha_i(\alpha_i sI-B)^{-1}\left(I+\frac{1}{\alpha_i}\mathrm{diag}(\alpha_1,\ldots,\alpha_K)\right)\alpha_i(\alpha_i sI-B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda}+\boldsymbol{L}^1\right), \qquad (3.12)
\end{aligned}
$$

where $\widetilde{Z}_i(s)$ is the LT of $Z_i(x)$. Substitution of (3.12) into (3.11) yields

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) &= \frac{2}{\alpha_i}\frac{1}{s}\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\widetilde{Z}_i(s)\boldsymbol{\alpha} \\ &\quad + 2\left(\frac{1}{s^3} + \frac{1}{s^3}\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}(\boldsymbol{\lambda} + s\boldsymbol{L}^1)\right)\left(\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\boldsymbol{\lambda} + 1\right) \\ &\quad + \frac{2}{s^2}\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\alpha_i(\alpha_i sI - B)^{-1}(\boldsymbol{\lambda} + s\boldsymbol{L}^1). \end{aligned} \tag{3.13}$$

We now need to investigate $\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\widetilde{Z}_i(s)\boldsymbol{\alpha}$. Taking LTs in (2.9) and using (3.9) leads to

$$\begin{aligned}s\widetilde{Z}_i(s) - Z_i(0) &= \frac{1}{\alpha_i}B\widetilde{Z}_i(s) + \frac{1}{\alpha_i}\widetilde{Z}_i(s)B^t + \left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\alpha_i(\alpha_i sI - B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)\boldsymbol{\lambda}^t \\ &\quad + \boldsymbol{\lambda}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)^t\alpha_i(\alpha_i sI - B^t)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right), \end{aligned}$$

where $Z_i(0) = L^2$, see (2.10). Postmultiplying the above by $\boldsymbol{u}_j^t$ yields

$$\begin{aligned}\frac{1}{\alpha_i}((\alpha_i s - \kappa_j)I - B)\widetilde{Z}_i(s)\boldsymbol{u}_j^t &= L^2\boldsymbol{u}_j^t + \left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\alpha_i(\alpha_i sI - B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)(\boldsymbol{\lambda}^t\boldsymbol{u}_j^t) \\ &\quad + \boldsymbol{\lambda}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)^t\alpha_i(\alpha_i sI - B^t)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\boldsymbol{u}_j^t, \end{aligned}$$

so

$$\begin{aligned}\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\widetilde{Z}_i(s)\boldsymbol{u}_j^t &= \alpha_i\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\left((\alpha_i s - \kappa_j)I - B\right)^{-1}L^2\boldsymbol{u}_j^t \\ &\quad + \alpha_i^2\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}((\alpha_i s - \kappa_j)I - B)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right) \\ &\quad \times(\alpha_i sI - B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)(\boldsymbol{\lambda}^t\boldsymbol{u}_j^t) + \alpha_i^2\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}((\alpha_i s - \kappa_j)I - B)^{-1} \\ &\quad \times\boldsymbol{\lambda}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)^t(\alpha_i sI - B^t)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\boldsymbol{u}_j^t. \end{aligned} \tag{3.14}$$

Let $[c_1 \ \cdots \ c_K] = \boldsymbol{\alpha}^t[\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_K]$. Then $\boldsymbol{\alpha}^t = [c_1 \ \cdots \ c_K]\begin{bmatrix}\boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_K\end{bmatrix}$, and (3.14) leads to

$$\begin{aligned}\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\widetilde{Z}_i(s)\boldsymbol{\alpha} &= \sum_{j=1}^{K} c_j\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}\widetilde{Z}_i(s)\boldsymbol{u}_j^t \\ &= \sum_{j=1}^{K} c_j\left\{\alpha_i\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}((\alpha_i s - \kappa_j)I - B)^{-1}L^2\boldsymbol{u}_j^t\right. \\ &\quad + \alpha_i^2\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}((\alpha_i s - \kappa_j)I - B)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right) \\ &\quad \times(\alpha_i sI - B)^{-1}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)(\boldsymbol{\lambda}^t\boldsymbol{u}_j^t) + \alpha_i^2\boldsymbol{\alpha}^t(\alpha_i sI - B)^{-1}((\alpha_i s - \kappa_j)I - B)^{-1} \\ &\quad \left.\times\boldsymbol{\lambda}\left(\frac{1}{s}\boldsymbol{\lambda} + \boldsymbol{L}^1\right)^t(\alpha_i sI - B^t)^{-1}\left(I + \frac{1}{\alpha_i}\mathrm{diag}(\alpha_1, \ldots, \alpha_K)\right)\boldsymbol{u}_j^t\right\}. \end{aligned} \tag{3.15}$$

Finally, substitution of (3.15) into (3.13) yields (3.7). □

Remark. Note that if $\alpha_1 = \alpha_2 = \cdots = \alpha_K$ and $\mu_1 = \mu_2 = \cdots = \mu_K$, then the $M/M/1$ DPS queue is the $M/M/1$ EPS queue. By setting $\alpha_1 = \alpha_2 = \cdots = \alpha_K$ and $\mu_1 = \mu_2 = \cdots = \mu_K$, we can reproduce formula (4.3) in

Yashkov [17] for the second moment of the conditional slowdown as follows. Let $\mu = \mu_1 = \mu_2 = \cdots = \mu_K$. We may assume that $\alpha_1 = \cdots = \alpha_K = \frac{1}{\mu}$. Then

$$B = \boldsymbol{\lambda}\boldsymbol{\alpha}^t - I,$$

and the eigenvalues of B are

$$\kappa_1 = -(1-\rho), \qquad \kappa_j = -1, \qquad j = 2, \ldots, K. \tag{3.16}$$

We may choose

$$\boldsymbol{v}_1 = \rho^{-1}[\rho_1 \ \cdots \ \rho_K]^t, \qquad \boldsymbol{u}_1 = \mathbf{1}^t, \tag{3.17}$$

where $\mathbf{1}$ denotes a K-dimensional column vector with all its components equal to one. Further, we note that

$$\boldsymbol{L}^1 = \frac{1}{1-\rho}[\rho_1 \ \cdots \ \rho_K]^t, \qquad L^2 = \frac{2}{(1-\rho)^2}[\rho_1 \ \cdots \ \rho_K]^t[\rho_1 \ \cdots \ \rho_K]. \tag{3.18}$$

Since

$$\begin{aligned}
&\boldsymbol{\alpha}^t\boldsymbol{v}_k = \mu^{-1}\boldsymbol{u}_1\boldsymbol{v}_k \ = \ \mu^{-1}\delta_{k1},\\
&\boldsymbol{u}_1 L^2 \boldsymbol{u}_j^t = \frac{2\rho^2}{(1-\rho)^2}\delta_{j1},\\
&\boldsymbol{\lambda}^t\boldsymbol{u}_j^t = \mu\rho\boldsymbol{v}_1^t\boldsymbol{u}_j^t = \mu\rho\delta_{j1},\\
&\left(\boldsymbol{\lambda} + s\boldsymbol{L}^1\right)^t \boldsymbol{u}_m^t = \left(\mu\rho + \frac{s\rho}{1-\rho}\right)\boldsymbol{v}_1^t\boldsymbol{u}_m^t = \left(\mu\rho + \frac{s\rho}{1-\rho}\right)\delta_{m1},\\
&c_1 = \frac{1}{\mu},
\end{aligned}$$

substituting (3.16)–(3.18) into (3.7) leads to

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) = {} & \frac{4\rho^2}{(1-\rho)^2}\frac{1}{s(s+\mu(1-\rho))(s+2\mu(1-\rho))} + \frac{4\mu\rho^2}{1-\rho}\frac{1}{s^2(s+\mu(1-\rho))(s+2\mu(1-\rho))}\\
& + \frac{4\mu\rho^2}{1-\rho}\frac{1}{s^2(s+\mu(1-\rho))(s+2\mu(1-\rho))} + \frac{2}{1-\rho}\frac{1}{s^3}\left(1 + \frac{\mu\rho}{s+\mu(1-\rho)}\right)\\
& + \frac{4\rho}{1-\rho}\frac{1}{s^2(s+\mu(1-\rho))},
\end{aligned}$$

which is simplified to

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) = \frac{2}{(1-\rho)^2}\frac{1}{s^3} + \frac{2\rho}{(1-\rho)^2}\frac{1}{s^2(s+\mu(1-\rho))}.$$

Decomposing the above into partial fractions yields

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) = -\frac{2\rho}{\mu^2(1-\rho)^4}\frac{1}{s} + \frac{2\rho}{\mu(1-\rho)^3}\frac{1}{s^2} + \frac{2}{(1-\rho)^2}\frac{1}{s^3} + \frac{2\rho}{\mu^2(1-\rho)^4}\frac{1}{s+\mu(1-\rho)}. \tag{3.19}$$

By the inversion formula, we have

$$\mathbb{E}[T_i^2(x)] = \frac{x^2}{(1-\rho)^2} + \frac{2\rho x}{\mu(1-\rho)^3} - \frac{2\rho}{\mu^2(1-\rho)^4}(1 - \mathrm{e}^{-\mu(1-\rho)x}),$$

and

$$\mathbb{E}[S_i^2(x)] = \frac{1}{(1-\rho)^2} + \frac{2\rho}{\mu^2(1-\rho)^4}\frac{\mu(1-\rho)x - 1 + \mathrm{e}^{-\mu(1-\rho)x}}{x^2},$$

which is the same as formula (4.3) in Yashkov [17].

Now we express the second moment of the unconditional slowdown for class i jobs in the $M/M/1$ DPS queue. The second moment of the unconditional slowdown for a class i job, $\mathbb{E}[S_i^2]$, is given as

$$\mathbb{E}[S_i^2] = \int_0^\infty \mathbb{E}[S_i^2(x)]\mu_i \mathrm{e}^{-\mu_i x}\mathrm{d}x = \mu_i \widetilde{G}_i(\mu_i).$$

Let us decompose (3.7) into partial fractions

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\widetilde{G}_i(s) = \frac{\epsilon_{i1}^1}{s} + \frac{\epsilon_{i2}^1}{s^2} + \frac{\epsilon_{i3}^1}{s^3} + \sum_{j=1}^K \left(\frac{\epsilon_{ij}^2}{s - \frac{\kappa_j}{\alpha_i}} + \frac{\epsilon_{ij}^3}{(s - \frac{\kappa_j}{\alpha_i})^2} \right) + \sum_{1\le j\le k\le K} \frac{\epsilon_{ijk}^4}{s - \frac{\kappa_j+\kappa_k}{\alpha_i}}, \tag{3.20}$$

for some constants $\epsilon_{i1}^1, \epsilon_{i2}^1, \epsilon_{i3}^1, \epsilon_{ij}^2, \epsilon_{ij}^3$, $j = 1, \ldots, K$, and ϵ_{ijk}^4, $1 \le j \le k \le K$.

Theorem 4. *The second moment of the slowdown for class i jobs is*

$$\begin{aligned}\mathbb{E}[S_i^2] = {} & \frac{1}{(1-\rho)^2} + \mu_i \sum_{j=1}^K \epsilon_{ij}^2 \left\{ \left(\mu_i - \frac{\kappa_j}{\alpha_i}\right) \log\left(1 - \frac{\kappa_j}{\alpha_i\mu_i}\right) + \frac{\kappa_j}{\alpha_i} \right\} - \mu_i \sum_{j=1}^K \epsilon_{ij}^3 \log\left(1 - \frac{\kappa_j}{\alpha_i\mu_i}\right) \\ & + \mu_i \sum_{1\le j\le k\le K} \epsilon_{ijk}^4 \left\{ \left(\mu_i - \frac{\kappa_j+\kappa_k}{\alpha_i}\right) \log\left(1 - \frac{\kappa_j+\kappa_k}{\alpha_i\mu_i}\right) + \frac{\kappa_j+\kappa_k}{\alpha_i} \right\}.\end{aligned}$$

Proof. Integrating (3.20) twice, we get

$$\begin{aligned}\widetilde{G}_i(s) = {} & \epsilon_{i1}^1(s\log s - s) - \epsilon_{i2}^1 \log s + \frac{\epsilon_{i3}^1}{2}\frac{1}{s} + \sum_{j=1}^K \epsilon_{ij}^2 \left\{ \left(s - \frac{\kappa_j}{\alpha_i}\right)\log\left(s - \frac{\kappa_j}{\alpha_i}\right) - \left(s - \frac{\kappa_j}{\alpha_i}\right) \right\} \\ & - \sum_{j=1}^K \epsilon_{ij}^3 \log\left(s - \frac{\kappa_j}{\alpha_i}\right) + \sum_{1\le j\le k\le K} \epsilon_{ijk}^4 \left\{ \left(s - \frac{\kappa_j+\kappa_k}{\alpha_i}\right)\log\left(s - \frac{\kappa_j+\kappa_k}{\alpha_i}\right) - \left(s - \frac{\kappa_j+\kappa_k}{\alpha_i}\right) \right\} \\ & + C_1 s + C_2,\end{aligned} \tag{3.21}$$

for some constants C_1 and C_2. We note that

$$\log(s+a) = \log s + \frac{a}{s} + o\left(\frac{1}{s}\right), \quad \text{as } s \to \infty, \tag{3.22}$$

for all real a, where '$f(s) = o(g(s))$ as $s \to \infty$' means that $\lim_{s\to\infty} \frac{f(s)}{g(s)} = 0$. Substituting (3.22) into (3.21), after some arithmetic, we can rewrite (3.21) as

$$\begin{aligned}\widetilde{G}_i(s) = {} & \left\{ \epsilon_{i1}^1 + \sum_{j=1}^K \epsilon_{ij}^2 + \sum_{1\le j\le k\le K} \epsilon_{ijk}^4 \right\} s\log s + \left\{ C_1 - \epsilon_{i1}^1 - \sum_{j=1}^K \epsilon_{ij}^2 - \sum_{1\le j\le k\le K} \epsilon_{ijk}^4 \right\} s \\ & - \left\{ \epsilon_{i2}^1 + \sum_{j=1}^K \epsilon_{ij}^2 \frac{\kappa_j}{\alpha_i} + \sum_{j=1}^K \epsilon_{ij}^3 + \sum_{1\le j\le k\le K} \epsilon_{ijk}^4 \frac{\kappa_j+\kappa_k}{\alpha_i} \right\} \log s + C_2 \\ & + \left\{ \frac{\epsilon_{i3}^1}{2} + \sum_{j=1}^K \epsilon_{ij}^2 \left(\frac{\kappa_j}{\alpha_i}\right)^2 + \sum_{j=1}^K \epsilon_{ij}^3 \frac{\kappa_j}{\alpha_i} + \sum_{1\le j\le k\le K} \epsilon_{ijk}^4 \left(\frac{\kappa_j+\kappa_k}{\alpha_i}\right)^2 \right\} \frac{1}{s} + o\left(\frac{1}{s}\right),\end{aligned}$$

as $s \to \infty$. Since $\lim_{s\to\infty} \widetilde{G}_i(s) = 0$, the following conditions should hold:

$$\begin{aligned}\epsilon_{i1}^1 &= -\sum_{j=1}^K \epsilon_{ij}^2 - \sum_{1\le j\le k\le K} \epsilon_{ijk}^4, \\ C_1 &= 0,\end{aligned}$$

$$\epsilon^1_{i2} = -\sum_{j=1}^{K} \epsilon^2_{ij} \frac{\kappa_j}{\alpha_i} - \sum_{j=1}^{K} \epsilon^3_{ij} - \sum_{1\le j\le k\le K} \epsilon^4_{ijk} \frac{\kappa_j + \kappa_k}{\alpha_i},$$

$$C_2 = 0.$$

Substituting the conditions above into (3.21), after some arithmetic, we obtain

$$\widetilde{G}_i(s) = \frac{\epsilon^1_{i3}}{2}\frac{1}{s} + \sum_{j=1}^{K} \epsilon^2_{ij} \left\{ \left(s - \frac{\kappa_j}{\alpha_i} \right) \log\left(1 - \frac{\kappa_j}{\alpha_i s} \right) + \frac{\kappa_j}{\alpha_i} \right\} - \sum_{j=1}^{K} \epsilon^3_{ij} \log\left(1 - \frac{\kappa_j}{\alpha_i s} \right) + \sum_{1\le j\le k\le K} \epsilon^4_{ijk} \left\{ \left(s - \frac{\kappa_j + \kappa_k}{\alpha_i} \right) \log\left(1 - \frac{\kappa_j + \kappa_k}{\alpha_i s} \right) + \frac{\kappa_j + \kappa_k}{\alpha_i} \right\}. \tag{3.23}$$

By (3.20), $\epsilon^1_{i3} = s^3 \left(\frac{\mathrm{d}^2}{\mathrm{d}s^2} \widetilde{G}_i(s) \right)\Big|_{s=0}$. From this and (3.7) together with Lemma 1(d), it follows that

$$\epsilon^1_{i3} = 2\left(1 + \sum_{j=1}^{K} \frac{\boldsymbol{\alpha}^t \boldsymbol{v}_j \boldsymbol{u}_j \boldsymbol{\lambda}}{-\kappa_j} \right)^2 = 2\left(1 - \boldsymbol{\alpha}^t B^{-1} \boldsymbol{\lambda} \right)^2 = \frac{2}{(1-\rho)^2}. \tag{3.24}$$

Finally, substituting (3.24) into (3.23) and noticing that $\mathbb{E}[S_i^2] = \mu_i \widetilde{G}_i(\mu_i)$, we finish the proof. □

Consider the case when $\alpha_1 = \alpha_2 = \cdots = \alpha_K$ and $\mu_1 = \mu_2 = \cdots = \mu_K$ (i.e., EPS discipline). Then using (3.19) and Theorem 4 gives the following corollary. We can see that in the case of the $M/M/1$ EPS queue, the second moment of the slowdown is determined by only the offered load ρ.

Corollary 1. *For the $M/M/1$ EPS queue, the second moment of the slowdown S is given by*

$$\mathbb{E}[S^2] = \frac{1}{(1-\rho)^2} - \frac{2\rho}{(1-\rho)^3} + \frac{2\rho(2-\rho)\log(2-\rho)}{(1-\rho)^4}. \tag{3.25}$$

Although the results obtained in this section can be used for numerical computation of the moments of conditional and unconditional slowdown, the expressions are too complicated to give any insight directly. We provide numerical examples and give some insights into the behavior of the slowdown measures in the following section.

4. Numerical examples

In this section, we provide numerical examples to discuss aspects of the slowdown in the DPS queue. For convenience, we refer to the "highest priority class" as the class with the largest weight, and the "lowest priority class" as the class with the smallest weight. The classes with weights in between the largest and smallest weights are labelled as the "middle classes".

With the figures in this section, we can observe that

- The conditional mean slowdown of the highest priority class increases as the job size increases.
- The conditional mean slowdown of the lowest priority class decreases as the job size increases.
- It could happen that the conditional mean slowdown of the middle classes is neither increasing nor decreasing. See Figs. 4 and 5. This phenomenon was also observed in [10].

It is known that (see Remark 2 in [6]), for the $M/G/1$ DPS queue, as the job size increases to infinity, the conditional mean slowdown of each class tends to $1/(1-\rho)$, which is the same as the conditional mean slowdown of the EPS model. If the conditional mean slowdown of a job with size x is larger (resp. smaller) than $1/(1-\rho)$, then we say that this job is treated worse (resp. better) under DPS than under EPS.

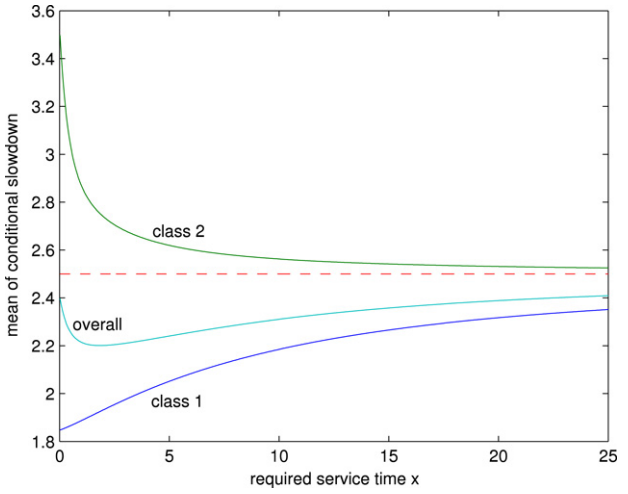

Fig. 1. Mean of the conditional slowdown for Example 1a.

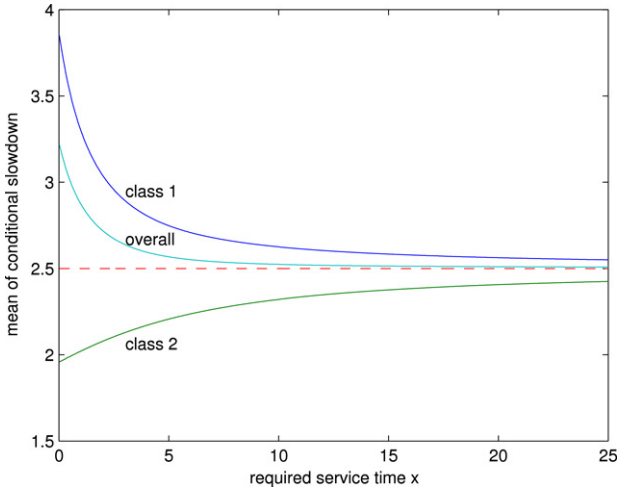

Fig. 2. Mean of the conditional slowdown for Example 1b.

4.1. Mean slowdown for a two-class DPS model

We consider the case of $K = 2$ job classes with weights α_1 and α_2. We assume equal loads of $\rho_1 = \rho_2 = 0.3$, and hence $1/(1-\rho) = 2.5$. We consider the following two DPS models.

Example 1a. We assume $\mu_1 = 2$ and $\mu_2 = 1$; hence $\lambda_1 = 0.6$ and $\lambda_2 = 0.3$. Take $\alpha_1 = 3$ and $\alpha_2 = 1$.

Example 1b. We assume $\mu_1 = 2$ and $\mu_2 = 1$; hence $\lambda_1 = 0.6$ and $\lambda_2 = 0.3$. Take $\alpha_1 = 1$ and $\alpha_2 = 3$.

Note that, in both examples, class 1 has smaller mean job size compared to class 2. It is shown in [3,10] that if $\alpha_1 \geq \alpha_2$, then DPS outperforms EPS from the viewpoint of mean number of jobs and the mean sojourn time at the steady state.

In Figs. 1 and 2, we depict the conditional mean slowdown of each class for Examples 1a and 1b, respectively, varying the required service time x. We observe that the overall conditional mean slowdown is better in Example 1a compared to Example 1b. Also it is intuitively clear that the highest priority class is always treated better under DPS than under EPS and the lowest priority class is always treated worse under DPS than under EPS, which is consistent with the figures. In addition, the conditional mean slowdown curve of each class does not cross those of other class,

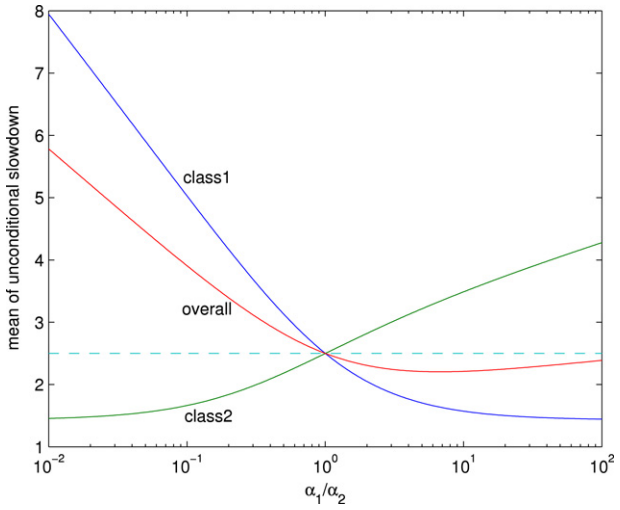

Fig. 3. Mean slowdown when $\mu_1 = 2$ and $\mu_2 = 1$.

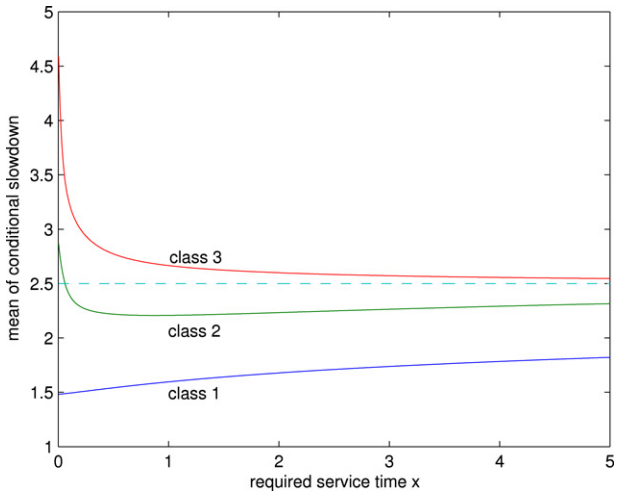

Fig. 4. Mean of the conditional slowdown for Example 2a with $\alpha_2 = 2.0$.

which also follows from the stochastic ordering result for conditional sojourn times; see Theorem 2 in Avrachenkov et al. [3]. As illustrated in Figs. 1 and 2, it is also observed that the conditional mean slowdown for the lowest priority class is much larger for small job sizes x. Short jobs of the lowest priority class are treated relatively the most unfairly, which can be explained by the so-called "ON–OFF" effect: If the ratio of weights α_1/α_2 is large, then from a class 2 point-of-view, the queue behaves almost as an ON–OFF processor-sharing queue (see Section 4.4 of [4,12]). When the number of class 1 jobs gets large, then the service process for class 2 may seem frozen (OFF period). When there are no high priority class jobs in the system, the low priority class jobs get full service capacity (ON period).

In Fig. 3, we plot the unconditional mean slowdown of each class, varying the weight ratio α_1/α_2. It is observed that in the case of $\alpha_1/\alpha_2 = 1$ (i.e., EPS discipline), the mean slowdown of a class 1 job equals that of a class 2 job. Further, the mean slowdown of a class 2 job (resp. a class 1 job) increases (resp. decreases) as the weight ratio α_1/α_2 increases, as we expect.

4.2. Mean slowdown for three-class DPS model

Now we consider the more interesting case of a DPS model with $K = 3$ job classes, with the presence of a middle class. We assume equal loads of $\rho_1 = \rho_2 = \rho_3 = 0.2$, and hence $1/(1-\rho) = 2.5$.

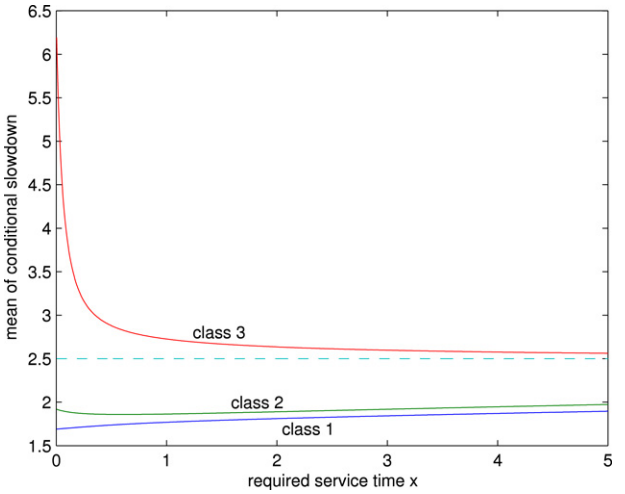

Fig. 5. Mean of the conditional slowdown for Example 2a with $\alpha_2 = 6.0$.

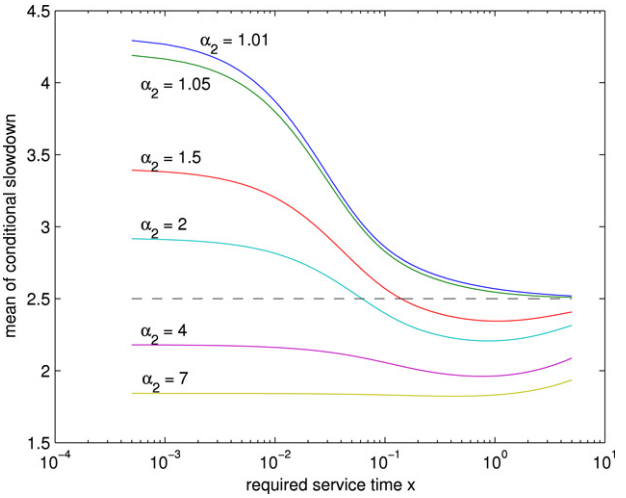

Fig. 6. Mean of the conditional slowdown for the middle class in Example 2a.

Example 2a (*See* Figs. 4–6). We assume $\mu_1 = 10, \mu_2 = 5, \mu_3 = 1$. Take $\alpha_1 = 8, \alpha_3 = 1$ and choose α_2 such that $\alpha_3 \leq \alpha_2 \leq \alpha_1$. In this case, class 1 is the highest priority class, class 2 is the middle class and class 3 is the lowest priority class.

Figs. 4 and 5 show the conditional mean slowdown of each class for different values of α_2, varying the required service time x. Fig. 6 shows the conditional mean slowdown of only the middle class for different values of α_2. We observe that if α_2 is small, then the conditional mean slowdown curve of the middle class is above the curve $1/(1-\rho)$, i.e., middle class jobs are always treated worse under DPS compared to EPS. If the weight α_2 of the middle class is moderate ($\alpha_2 = 1.5, \alpha_2 = 2.0$), then the slowdown curve of the middle class crosses the curve $1/(1-\rho)$. Sometimes the middle class job is treated worse and sometimes better under DPS compared to EPS, depending on the job size x of the middle class job. If α_2 gets larger, then the conditional mean slowdown curve of the middle class will be always below the curve $1/(1-\rho)$, i.e., middle class jobs are always treated better under DPS compared to EPS. It is observed that the ON–OFF effect experienced by class 2 jobs becomes larger when α_2 becomes smaller.

Example 2b (*See* Figs. 7 *and* 8). We assume $\mu_1 = 10, \mu_2 = 5, \mu_3 = 1$. Take $\alpha_1 = 1, \alpha_3 = 8$ and choose α_2 such that $\alpha_1 \leq \alpha_2 \leq \alpha_3$. In this case, class 3 is the highest priority class, class 2 is the middle class and class 1 is the lowest priority class.

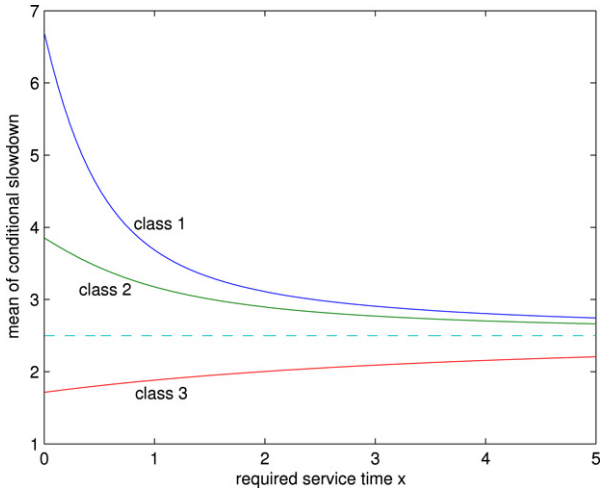

Fig. 7. Mean of the conditional slowdown for Example 2b with $\alpha_2 = 2.0$.

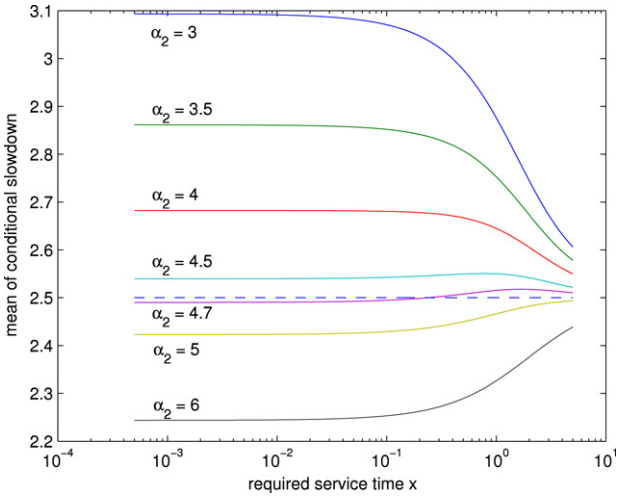

Fig. 8. Mean of the conditional slowdown for the middle class in Example 2b.

Fig. 7 shows the conditional mean slowdown of each class in the case of $\alpha_2 = 2.0$. Fig. 8 shows the conditional mean slowdown of only the middle class for different values of α_2.

As illustrated in Fig. 7, it is observed that if α_2 is small, then the conditional mean slowdown curve of the middle class is always above the curve $1/(1-\rho)$; however, the shape of the curve changes. From Fig. 8, we see that if α_2 is moderate, then the short middle class jobs have a smaller conditional mean slowdown under DPS than under EPS; however, long middle class jobs are still treated unfairly under DPS in this situation. Fig. 8 illustrates that the conditional mean slowdown curve of the middle class jobs will be always below the curve $1/(1-\rho)$, indicating that if α_2 is large, then the middle class gets served better under DPS for all job sizes. Fig. 8 also indicates that the ON–OFF effect experienced by class 2 jobs becomes larger when α_2 becomes smaller.

4.3. Variance of the slowdown for three-class DPS model

In this subsection, we investigate the conditional and unconditional variance of the slowdown for a DPS model with $K = 3$ job classes. We assume equal loads of $\rho_1 = \rho_2 = \rho_3 = 0.2$, and hence $1/(1-\rho) = 2.5$.

Example 3 (*See* Figs. 9 *and* 10). We assume $\mu_1 = 10$, $\mu_2 = 5$, $\mu_3 = 1$; hence $\lambda_1 = 2$, $\lambda_2 = 1$ and $\lambda_3 = 0.2$.

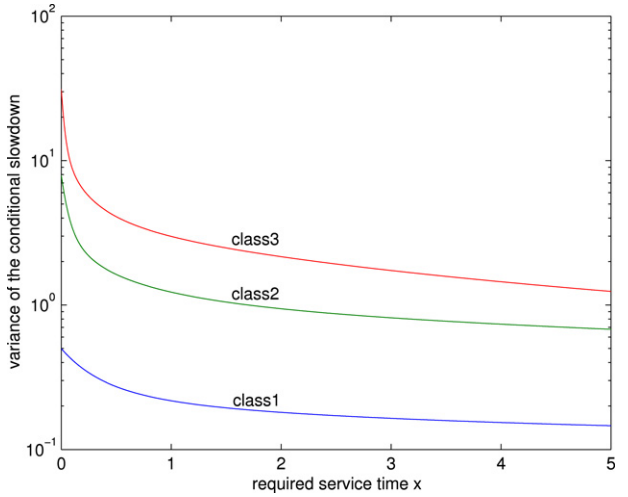

Fig. 9. Variance of the conditional slowdown in Example 3.

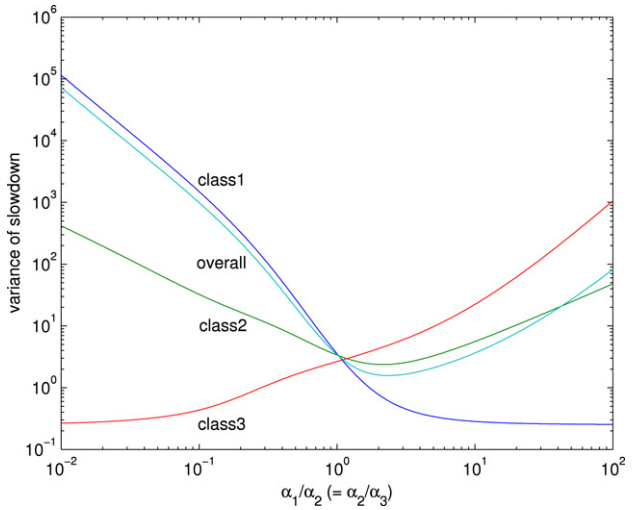

Fig. 10. Variance of slowdown in Example 3.

In Fig. 9, we plot the variance of the conditional slowdown of each class with required service time x, varying the required service time x, for the case when $\alpha_1 = 8, \alpha_2 = 2$ and $\alpha_3 = 1$. The variance of the conditional slowdown of each class decreases as the required service time x increases. Further, we observe that the smaller the weight becomes, the larger the variance of the conditional slowdown becomes.

In Fig. 10, we plot the variance of the slowdown of each class and of overall classes, varying the weight ratio α_1/α_2 $(=\alpha_2/\alpha_3)$. We observe that the variance of the slowdown of a class 3 job (resp. a class 1 job) increases (resp. decreases) as the weight ratio α_1/α_2 increases. The variance of the slowdown of overall classes will be discussed in the following subsection.

4.4. An unfairness measure

Let S denote the slowdown of an arbitrary customer in overall classes. Then

$$\mathbb{E}S = \sum_{i=1}^{K} \frac{\lambda_i}{\lambda} \mathbb{E}S_i \quad \text{and} \quad \text{Var}\, S = \sum_{i=1}^{K} \frac{\lambda_i}{\lambda} \mathbb{E}S_i^2 - (\mathbb{E}S)^2,$$

where $\lambda = \sum_{i=1}^{K} \lambda_i$. Avi-Itzhak et al. [2] introduced two unfairness measures, one of which is Var$S$; see the last paragraph above Section 3.1 in [2]. In Fig. 10, the variance of the slowdown of overall classes corresponds to the unfairness measure in Example 3.

5. Conclusion

In this paper we obtained the first and second moments of the slowdown in the $M/M/1$ queue with the discriminatory processor-sharing (DPS) service discipline. The slowdown is a measure for queueing fairness: jobs in the standard $M/G/1$ queue with egalitarian processor-sharing (EPS) have a constant mean slowdown, i.e., the mean slowdown is independent of the job size, and this reflects the fairness of the EPS service discipline. In contrast, DPS aims to differentiate the Quality-of-Service among different types of jobs, and we discuss the observation that a job of a certain size may sometimes be treated better or worse (in terms of slowdown) compared to a similar queueing model with equal weights.

How fairly jobs are treated under DPS depends on several parameters; in particular, it depends on the set of DPS weights $(\alpha_1, \ldots, \alpha_K)$ in combination with the mean job sizes $(1/\mu_1, \ldots, 1/\mu_K)$, and the job size $x > 0$ of a particular class i, $i = 1, \ldots, K$. The highest priority class with the largest DPS weight is always treated better under DPS than under EPS at the expense of other classes. The lowest priority class is always treated worse under DPS than under EPS. However, the unfairness also depends on the job size; short lowest priority jobs are generally treated the most unfairly. Short highest priority jobs are generally treated the best; these jobs benefit the most from the preemptive priority effect that the highest priority jobs observe.

When there are middle classes, i.e., classes with weights in between the largest and smallest weights, the characterization of the fairness and unfairness of the middle class is less straightforward. In the numerical examples we have observed and explained the following possible cases for the middle class:

- All jobs are always treated worse under DPS than under EPS.
- All jobs are always treated better under DPS than under EPS.
- Sometimes short jobs are treated worse and long jobs are treated better under DPS than under EPS.
- Sometimes short jobs are treated better and long jobs are treated worse under DPS than under EPS.

We also observed that the slowdown curve for the middle class jobs is generally not monotonic in the job size, unless the weight of the middle class is sufficiently close to weight of the highest or lowest priority class. The slowdown curves of the highest and lowest priority classes are increasing and decreasing, respectively.

In the paper of Avi-Itzhak et al. [2] two (un)fairness measures are proposed. The first fairness measure is defined by $\mathrm{Var}(T - \frac{\mathbb{E}T}{\mathbb{E}X}X)$, where T denotes the sojourn time and X denotes the required service time for an arbitrary customer. A second and alternative approach for defining fairness is the variance of the slowdown S. Avi-Itzhak et al. [2] mentioned that this fairness measure is very hard to compute, and they left computation of such a fairness metric as an open research topic. This paper provides an analytical computation of the latter fairness measure for the $M/M/1$ DPS queue. Our results can be used for the analysis of the fairness measure and provide tools for a numerical study of the $M/M/1$ DPS queues.

References

[1] E. Altman, K.E. Avrachenkov, U. Ayesta, A survey on discriminatory processor sharing, Queueing Systems 53 (2006) 53–63.
[2] B. Avi-Itzhak, E. Brosh, H. Levy, SQF: A slowdown queueing fairness measure, Performance Evaluation 64 (2007) 1121–1136.
[3] K. Avrachenkov, U. Ayesta, P. Brown, R. Núñez-Queija, Discriminatory processor sharing revisited, in: Proceedings of IEEE INFOCOM 2005, Miami, USA.
[4] S.-K. Cheung, J.L. van den Berg, R.J. Boucherie, Decomposing the queue length distribution of processor-sharing models into queue lengths of permanent customer queues, Performance Evaluation 62 (2005) 100–116.
[5] S.-K. Cheung, J.L. van den Berg, R.J. Boucherie, Insensitive bounds for the moments of the sojourn time distribution in the $M/G/1$ processor-sharing queue, Queueing Systems 53 (2006) 7–18.
[6] G. Fayolle, I. Mitrani, R. Iasnogorodski, Sharing a processor among many job classes, Journal of the ACM 27 (1980) 519–532.
[7] M. Harchol-Balter, K. Sigman, A. Wierman, Asymptotic convergence of scheduling policies with respect to slowdown, Performance Evaluation 49 (2002) 241–256.
[8] G. Van Kessel, R. Núñez-Queija, S.C. Borst, Asymptotic regimes and approximations for discriminatory processor sharing, ACM SIGMETRICS Performance Evaluation Review (2004) 44–46.

[9] J. Kim, B. Kim, Sojourn time distribution in the $M/M/1$ queue with discriminatory processor-sharing, Performance Evaluation 58 (2004) 341–365.
[10] B. Kim, J. Kim, Comparison of DPS and PS systems according to DPS weights, IEEE Communications Letters 10 (2006) 558–560.
[11] L. Kleinrock, Time-shared systems: A theoretical treatment, Journal of the ACM 14 (1967) 242–261.
[12] R. Núñez-Queija, Sojourn times in a processor sharing queue with service interruptions, Queueing Systems 34 (1-4) (2000) 351–386.
[13] T.M. O'Donovan, Direct solutions of $M/G/1$ processor-sharing models, Operations Research 22 (1974) 1232–1235.
[14] K.M. Rege, B. Sengupta, Queue length distribution for the discriminatory processor-sharing queue, Operations Research 44 (1996) 653–657.
[15] E. Seneta, Non-negative Matrices and Markov Chains, 2nd ed, Springer-Verlag, New York Inc., 1981.
[16] A. Wierman, M. Harchol-Balter, Classifying scheduling policies with respect to unfairness in an $M/GI/1$, in: Proceedings of ACM SIGMETRICS 2003, San Diego, pp. 238–249.
[17] S.F. Yashkov, Processor-sharing queues: Some progress in analysis, Queueing Systems 2 (1987) 1–17.

Sing-Kong Cheung received his master's degree in Econometrics and Operations Research from the Vrije Universiteit, Amsterdam, The Netherlands. He received his Ph.D. from the University of Twente, The Netherlands, in 2007 for a thesis on processor-sharing queues and resource sharing in wireless LANs. During 2006, Sing-Kong visited the Telecommunication Mathematics Research Center (TMRC) of Korea University (KU) in Seoul, South Korea, and the Korea Advanced Institute of Science and Technology (KAIST) in Daejeon, South Korea.

Bara Kim is an associate professor in the Department of Mathematics at Korea University, Seoul, Korea. He received B.S., M.S. and Ph.D. in Mathematics from Korea Advanced Institute of Science and Technology (KAIST). During 2002 he was a visiting scholar in the School of Industrial and Systems Engineering at Georgia Institute of Technology. His areas of interests include probability theory, stochastic models, applied operations research, queueing theory and their applications to the communication systems and industrial engineering.

Jeongsim Kim is an assistant professor in the Department of Mathematics Education at Chungbuk National University. She received her B.S., M.S. and Ph.D. in Mathematics from Korea University. Her research interests include probability theory, queueing theory and its applications to the communication systems.