Queueing Systems manuscript No.
(will be inserted by the editor)

Optimal size-based opportunistic scheduler for wireless systems

Samuli Aalto · Aleksi Penttinen · Pasi Lassila · Prajwal Osti

Abstract Modern wireless cellular systems are able to utilize the opportunistic scheduling gain originating from the variability in the users' channel conditions. By favoring users with good instantaneous channel conditions, the service capacity of the system can be increased with the number of users. On the other hand, for service systems with fixed service capacity, the system performance can be optimized by utilizing the size information. Combining the advantages of size-based scheduling with opportunistic scheduling gain has proven to be a challenging task. In this paper, we consider scheduling of data traffic (finite-size elastic flows) in wireless cellular systems. Assuming that the channel conditions for different users are independent and identically distributed, we show how to optimally combine opportunistic and size-based scheduling in the transient setting with all flows available at time 0. More specifically, by utilizing the time scale separation assumption, we develop a recursive algorithm that produces the optimal long-run service rate vectors within the corresponding capacity regions. We also prove that the optimal operating policy applies the SRPT-FM principle, i.e., the shortest flow is served with the highest rate of the optimal rate vector, the second shortest with the second highest rate etc. Moreover, we determine explicitly how to implement the optimal rate vectors in the actual time slot level opportunistic scheduler. In addition to the transient setting, we explore the dynamic case with randomly arriving flows under illustrative channel scenarios by simulations. Interestingly, the scheduling policy that is optimal for the transient setting can be improved in the dynamic case under high traffic load by applying a rate-based priority scheduler that breaks the ties based on the SRPT principle.

Keywords Opportunistic scheduling · SRPT · mean delay · capacity region

S. Aalto, A. Penttinen, P. Lassila, P. Osti
Aalto University, Finland
E-mail: pasi.lassila@tkk.fi

1 Introduction

We consider optimal scheduling of downlink data traffic in modern wireless cellular systems. The traffic consists of finite-size flows that are transmitted to users in a single cell area. We assume that the scheduler operates in a very fast time scale of milliseconds transmitting in each time slot to exactly one user. From the flow level performance point of view, a scheduler is optimal if it minimizes the mean flow delay (that is, the time from the start of the transmission to its completion). The *static version* of the optimal scheduling problem considers a transient system where all the flows in the system are available at time 0 and no new flows arrive. Minimizing the mean delay corresponds to the minimization of the total completion time. In the *dynamic version*, new flows arrive randomly and the objective is to minimize the long-run mean delay (or, the steady-state mean number of flows, as well, by Little's result).

A fundamental phenomenon in all wireless systems is that the users perceive time-varying channel conditions due to fast fading phenomena. Roughly said, the better the instantaneous conditions, the higher the transmission rate, which can be exploited by *channel-aware schedulers* that get information on the changes in these conditions. This has led to *opportunistic scheduling* schemes, such as the Proportional Fair (PF) scheduler, that favor users with good channel conditions [4,11,21,14]. With opportunistic schedulers, the service capacity is no longer fixed but increases as the number users increases, which corresponds to a queueing system with scalable service capacity (or, briefly, a scalable queue).

If the scheduler is not channel-aware, then, from the scheduling point of view, the system boils down to a single-server queue with a fixed service capacity: Each user is served with its own average transmission rate. In such a case, the best that a *size-based scheduler* can do, which is aware of the remaining flow sizes, is to transmit to the user whose flow has the smallest remaining transmission time (which is the remaining size divided by the average transmission rate). This corresponds to the Shortest-Remaining-Processing-Time (SRPT) discipline, which is known to be optimal for the single-server queues [18].

The idea of SRPT is to minimize the delay by getting rid of flows as soon as possible. However, with fewer flows part of the opportunistic gain is lost, which is the basic dilemma. Combining the advantages of size-based scheduling with opportunistic scheduling gain has proven a challenging task. Two different modeling approaches have been presented to solve the problem. A direct approach [20,10,12,3] is based on *time slot level models*, while the other approach [17,2] utilizes a time scale separation argument, which leads to *flow level models*. A good baseline is provided by the PF scheduler that does not utilize size information at all.

The *time scale separation* assumption [7,9,16,13,6,17,2] implies that, at the flow level, the flows experience the time-average transmission rates produced by the time slot level opportunistic scheduler. The assumption is justified if the channel conditions can be modeled by a stationary and ergodic

process and the flow transmission times are much larger than the time scale where the scheduler operates. For a given number of flows in the system, the set of achievable rate vectors that the time slot level scheduler can support is characterized by the notion of the *capacity region*. Assuming further that the channel conditions for different users are independent and identically distributed (up to a user-specific coefficient) results in symmetric capacity regions [7]. The solution to the optimization problem for the flow level model gives the optimal rate vectors within the corresponding capacity regions. How to implement these optimal rate vectors in the time slot level opportunistic scheduler needs still to be solved separately.

Using a time slot level model and assuming that the channel conditions in different time slots are independent and identically distributed, Tsybakov [20] formulated the optimal scheduling problem for the transient system (without arrivals) as a dynamic program, which can be solved numerically. However, the dynamic program does not allow to extract any structural properties of the optimal policy. Hu et al. [10] developed heuristic algorithms that combine opportunistic scheduling with size-based information, and derived bounds for their flow level performance in the transient setting. Ayesta et al. [3] considered a slightly different problem without exact size information. They applied the theory of restless bandits to bring in more structure for the (nearly) optimal time slot level scheduler for the transient system. The approach, however, requires a Markovian description of the system and is therefore limited to geometric flow size distributions with the memoryless property. In addition, their Potential Improvement (PI) scheduler reduces to the PF scheduler in the symmetric setting.

Sadiq and de Veciana [17] applied the time scale separation argument and considered the optimal scheduling problem for the transient system with a flow level model. They were able to derive the optimal rate vectors for *nested polymatroids* that are compact, convex, coordinate-convex, and symmetric capacity regions, and showed that the optimal policy applies the SRPT-FM (SRPT-Fastest-Machine) principle, i.e., the shortest flow is served with the highest rate of the optimal rate vector, the second shortest with the second highest rate etc. In their proof, Sadiq and de Veciana utilized the known optimality result of the SRPT-FM discipline for the heterogeneous multi-server queues [15]. While the result gives a complete structural characterization for the optimal operating policy, its applicability (as such) is rather restricted, since, as Sadiq and de Veciana note, capacity regions are nested polymatroids only in some very special cases. However, the result can be utilized to find out an optimistic bound for the flow level performance whenever capacity regions are nested and convex.

In [2], we also made the time scale separation assumption and considered the optimal scheduling problem under opportunistic scheduling gain in the transient setting. We focused on the situation, where the capacity regions are compact and symmetric (including, e.g., all nested polymatroids). We gave a condition under which (i) the optimal rate vector does not depend on the sizes of the flows as long as their order (in size) remains the same and (ii) the

optimal policy applies the SRPT-FM principle. We also developed a recursive algorithm to determine both the optimal rate vector and the minimum mean delay. Unfortunately, our condition is rather implicit. In addition, we did not consider the problem of implementing the optimal rate vectors in the time slot level scheduler.

The optimal scheduling problem in the dynamic setting with new flow arrivals for the systems with opportunistic scheduling gain is still unsolved. Only some heuristic algorithms have been proposed and experimented with that try to combine opportunistic scheduling gain with size-based scheduling [10,13,12,3,17]. In this context, even stability may be an issue and, indeed, the only available theoretical results are limited to stability analysis of some families of policies (α-fair and priority based), see [7,9,8,1].

In this paper, we prove that the implicit condition given in [2] is, indeed, satisfied under the assumption that the channel conditions for different users are independent and identically distributed (up to a user-specific coefficient). As a result, we have a recursive algorithm that produces the optimal flow level rate vectors that minimize the mean delay in the transient setting. While the original recursive algorithm developed in [2] requires an explicit representation of the capacity regions, we develop another recursive algorithm that avoids this requirement by directly utilizing the time slot level channel model. Moreover, we also determine explicitly how to implement the optimal rate vectors in the time slot level opportunistic scheduler. Thus, we are able to combine opportunistic and size-based scheduling optimally in the time slot level (as long as the transient system is concerned), which is the main theoretical contribution of the paper.

In addition to the static setting, we also explore the dynamic setting by comparing the following opportunistic schedulers under some channel scenarios: (i) PF, the baseline policy that allows an explicit performance analysis in the symmetric setting [7]; (ii) SRPT-P, a priority based policy, for which the priority is based on the channel conditions and ties are broken according to the SRPT discipline [12]; (iii) SRPT-OPS(k), a policy that applies PF for (at most) k users with the shortest remaining transmission times [17]; and (iv) TR-OPT, the optimal policy for the transient system (developed in this paper). Since the theoretic analysis is still beyond mathematical tractability, we rely on simulations when comparing the policies in the dynamic setting. Interestingly, the scheduling policy TR-OPT, which is optimal for the transient case, can be improved in the dynamic setting under high traffic load by applying the rate-based priority scheduler SRPT-P.

The rest of the paper is organized as follows. In Section 2, we introduce the time slot level scheduler model and the corresponding flow level capacity regions. The flow level operating policies together with the time scale separation assumption are presented in Section 3. Section 4 includes the main theoretical results of the paper, which are related to the optimal scheduling problem in the transient setting. The theoretical results are illustrated in Section 5 by numerical examples including certain channel scenarios. In Section 6, we explore the dynamic setting based on simulations performed with two different

simulators (one based directly on the time slot level model, and the other that utilizes the time scale separation assumption). Section 7 concludes the paper and also discusses some future research directions.

2 Time slot level scheduler model and capacity region

We consider downlink data transmission in a wireless cellular system, where the base station always transmits to a single user within a time slot, as in 1xEV-DO systems (with the time slot equal to 1.67 ms). The traffic consists of elastic flows (corresponding roughly to file transfers) that the users are downloading through the base station to their mobile terminals. In the rest of the paper we refer to flows, users and terminals interchangeably.

Assume now that there is a fixed number of users, n, in the system. The users (and the corresponding flows and terminals) are indexed by $i = 1, \ldots, n$. The *instantaneous rate* (in bps) of user i varies over time according to some stationary and ergodic process $R_i(t)$ taking values in $\mathcal{R} \subset \mathbb{R}_+$. Let R_i refer to the corresponding steady-state variable. We assume that the instantaneous rate processes of different users are independent and identically distributed[1] with the mean denoted by $\gamma_1 = E[R_1(t)]$. The *opportunistic gain* γ_n is defined by

$$\gamma_n = E[\max\{R_1(t), \ldots, R_n(t)\}]. \tag{1}$$

Thus, we have

$$\begin{aligned}\gamma_n &= \int_0^\infty P\{\max\{R_1(t), \ldots, R_n(t)\} > r\}\,\mathrm{d}r \\ &= \int_0^\infty (1 - P\{R_1(t) \le r\}^n)\,\mathrm{d}r.\end{aligned}$$

To simplify the presentation in this section, we assume (as in [7,9,12]) that $\mathcal{R}$ is a finite set. However, the theoretical results given in this paper do not require this restrictive assumption. Finally, let $R(t) = (R_1(t), \ldots, R_n(t))$ denote the corresponding vector process taking values in $\mathcal{R}^n \subset \mathbb{R}_+^n$.

We assume that the scheduling decision is based on the instantaneous rate vector $R(t)$. More precisely said, we consider *opportunistic schedulers* π characterized by the probabilities $p_i^\pi(\mathbf{r})$ with which flow i is scheduled in state $\mathbf{r} \in \mathcal{R}^n$ satisfying

$$\sum_{i=1}^n p_i^\pi(\mathbf{r}) \le 1.$$

Note that there may be a positive probability that the scheduler is idle even if there are flows in the system. Let $\mathit{\Pi}_n$ denote the family of such schedulers.

[1] In fact, by adjusting the corresponding flow sizes, we could, as well, have assumed that, for each user i, there is a constant a_i such that processes $a_i R_i(t)$ are independent and identically distributed, cf. [7,17].

In addition, we assume that the traffic sources are *saturated* at the time slot level [14,5,19], i.e., if time slot t is scheduled to user i, the user will receive data with rate $R_i(t)$ in this time slot. Thus, the long-term throughput for user i under scheduler $\pi \in \Pi_n$ is as follows:

$$\theta_i^\pi = \sum_{\mathbf{r} \in \mathcal{R}^n} r_i p_i^\pi(\mathbf{r}) P\{R(t) = \mathbf{r}\}. \tag{2}$$

The (opportunistic) *capacity region* is defined by all feasible throughput vectors,

$$\mathcal{C}_n = \{(\theta_1^\pi, \ldots, \theta_n^\pi) \in \mathbb{R}_+^n : \pi \in \Pi_n\}.$$

The capacity region $\mathcal{C}_n$ has the following three properties:

(i) $\mathcal{C}_n$ is a *compact* subset of $\mathbb{R}_+^n$, i.e., $\mathcal{C}_n$ is closed and bounded;
(ii) $\mathcal{C}_n$ is *symmetric*, i.e., if $\mathbf{c} \in \mathcal{C}_n$, then any permutation $\tilde{\mathbf{c}}$ of its components also lies in $\mathcal{C}_n$;
(iii) $\mathcal{C}_n$ is *convex*.

For $n = 1$, we clearly have

$$\mathcal{C}_1 = [0, \gamma_1].$$

In addition, for $n > 1$, the capacity regions generated by this model are naturally *nested*, i.e., for all $k = 2, \ldots, n$ and $\mathbf{c} \in \mathcal{C}_k$,

$$(c_1, \ldots, c_{k-1}) \in \mathcal{C}_{k-1}.$$

An important subfamily of the opportunistic schedulers Π_n consists of *weight-based schedulers* Π_n^{w}. A weight-based scheduler $\pi \in \Pi_n^{\mathrm{w}}$ is defined by a vector $\mathbf{w} = (w_1, \ldots, w_n)$ of weights $w_i \geq 0$ for which $w_1 + \ldots + w_n > 0$. Scheduler π allocates time slot t to a user i^* for which

$$w_{i^*} R_{i^*}(t) = \max_{i=1,\ldots,n} w_i R_i(t),$$

augmented with a suitable tie-breaking rule.

A well-known fact (see, e.g., [7]) is that any non-dominated feasible throughput vector $\mathbf{c} \in \mathcal{C}_n$ can be achieved by some weight-based scheduler $\pi \in \Pi_n^{\mathrm{w}}$ so that $\mathbf{c} = (\theta_1^\pi, \ldots, \theta_n^\pi)$.

We furthermore note that, in this symmetric setting, the throughput vector produced by the PF scheduler equals the throughput vector of the weight-based scheduler that applies an equal weight $w_i = 1$ for all flows i (see, e.g., [7]). Thus,

$$(\theta_1^{\mathrm{PF}}, \ldots, \theta_n^{\mathrm{PF}}) = (\gamma_n/n, \ldots, \gamma_n/n). \tag{3}$$

The following observation related to the weight-based schedulers is needed later on.

Proposition 1 *Assume that $w_i \geq 0$ for all $i = 1, \ldots, n$. Then*

$$\max_{\mathbf{c} \in \mathcal{C}_n} \sum_{i=1}^n w_i c_i = E[\max_{i=1,\ldots,n} w_i R_i].$$

Proof Note first that

$$\begin{aligned}\max_{\mathbf{c}\in\mathcal{C}_n}\sum_{i=1}^{n} w_i c_i &= \max_{\pi\in\Pi_n}\sum_{i=1}^{n} w_i\theta_i^{\pi}\\ &= \max_{\pi\in\Pi_n}\sum_{i=1}^{n} w_i \sum_{\mathbf{r}\in\mathcal{R}^n} r_i p_i^{\pi}(\mathbf{r}) P\{R(t)=\mathbf{r}\}\\ &= \max_{\pi\in\Pi_n}\sum_{\mathbf{r}\in\mathcal{R}^n}\sum_{i=1}^{n} w_i r_i p_i^{\pi}(\mathbf{r}) P\{R(t)=\mathbf{r}\}.\end{aligned}$$

Since maximizing the required expression in each state separately leads to a feasible policy, we have

$$\begin{aligned}\max_{\mathbf{c}\in\mathcal{C}_n}\sum_{i=1}^{n} w_i c_i &= \max_{\pi\in\Pi_n}\sum_{\mathbf{r}\in\mathcal{R}^n}\sum_{i=1}^{n} w_i r_i p_i^{\pi}(\mathbf{r}) P\{R(t)=\mathbf{r}\}\\ &= \sum_{\mathbf{r}\in\mathcal{R}^n}\left(\max_{\pi\in\Pi_n}\sum_{i=1}^{n} w_i r_i p_i^{\pi}(\mathbf{r})\right) P\{R(t)=\mathbf{r}\}\\ &= \sum_{\mathbf{r}\in\mathcal{R}^n}\left(\max_{i=1,\ldots,n} w_i r_i\right) P\{R(t)=\mathbf{r}\}\\ &= E[\max_{i=1,\ldots,n} w_i R_i(t)],\end{aligned}$$

which completes the proof. □

3 Time scale separation and the flow level operating model

As in [7,9,17], the flow level performance analysis and optimization is carried out under the time scale separation assumption given below.

While the actual opportunistic scheduler operates in a very fast time scale (typically milliseconds), the flow level performance is related to a much slower time scale. With the current transmission rates, the file transfers typically take from seconds to minutes. Thus, we may reasonably assume that, as long as the number of flows remains unchanged, the flows are served in continuous time and with fixed rates that correspond to some operating point in the capacity region, which is called the *time scale separation assumption*.

At the flow level, the flows no longer look like infinite traffic sources but they have finite sizes. Assume that there are n flows in the system, say at time 0, with sizes s_i, $i = 1, \ldots, n$. Without loss of generality, we may assume that

$$s_1 \geq s_2 \geq \ldots \geq s_n.$$

According to the time scale separation assumption, these flows are served with rates c_i determined by the *operating point* $\mathbf{c} = (c_1, \ldots, c_n) \in \mathcal{C}_n$. The operating point is *implemented* with the time slot level scheduler $\pi \in \Pi_n$ such that $c_i = \theta_i^{\pi}$ for all i.

We say that the operating point $\mathbf{c} \in \mathcal{C}_n$ applies the *SRPT-FM principle* if

$$c_1 \leq c_2 \leq \ldots \leq c_n.$$

In other words, the shortest flow is served with the highest rate of the operating point, the second shortest with the second highest rate etc. Note that such an operating point preserves the order of the flow sizes for the whole period until the number of flows changes, i.e., for any $t \in [0, s_n/c_n)$ (assuming no new flow arrivals),

$$s_1(t) \geq s_2(t) \geq \ldots \geq s_n(t),$$

where $s_i(t)$ denotes the size of flow i at time t.

Below we consider some operating points applying the SRPT-FM principle, which are used later on.

(i) Operating point

$$\mathbf{c}^{\mathrm{PF}} = (\gamma_n/n, \ldots, \gamma_n/n)$$

corresponds to the PF scheduler, see (3). The operating point may be implemented (at the time slot level) by a weight-based scheduler with weight vector

$$\mathbf{w}^{\mathrm{PF}} = (1, \ldots, 1).$$

(ii) Operating point

$$\mathbf{c}^{\mathrm{SRPT-OPS}(k)} = \begin{cases} (0, \ldots, 0, \gamma_k/k, \ldots, \gamma_k/k), & k < n, \\ (\gamma_n/n, \ldots, \gamma_n/n), & k \geq n, \end{cases}$$

with $\min\{n, k\}$ non-zero elements refers to the SRPT-OPS(k) scheduler [17] that applies the PF scheduling discipline among the $\min\{n, k\}$ flows with the shortest remaining flow sizes. The operating point may be implemented by a weight-based scheduler with weight vector

$$\mathbf{w}^{\mathrm{SRPT-OPS}(k)} = \begin{cases} (0, \ldots, 0, 1, \ldots, 1), & k < n, \\ (1, \ldots, 1), & k \geq n, \end{cases}$$

which has $\min\{n, k\}$ non-zero elements.

(iii) Operating point

$$\mathbf{c}^{\mathrm{SRPT-P}} = (\theta_1^{\mathrm{SRPT-P}}, \ldots, \theta_n^{\mathrm{SRPT-P}}),$$

where

$$\theta_i^{\mathrm{SRPT-P}} = \sum_{r \in \mathcal{R}} r P\{R_1 \leq r\}^{i-1} P\{R_1 = r\} P\{R_1 < r\}^{n-i},$$

corresponds to the SRPT-P scheduler [12], which is a priority based policy, for which the priority is based on the instantaneous rate and ties are broken according to the SRPT discipline.

(iv) Operating point

$$\mathbf{c}^{\mathrm{TR-OPT}} = (c_1^*, \ldots, c_n^*),$$

where c_i^* are defined recursively in Section 4, is the optimal operating point for the transient system, see Theorem 1. The operating point may be implemented by a weight-based scheduler with weight vector

$$\mathbf{w}^{\mathrm{TR-OPT}} = (G_1^*, \ldots, G_n^*),$$

where G_i^* are defined recursively in Section 4, see Theorem 2.

4 Optimal trade-off in the transient system

Our main results in this section are given in Theorems 1 and 2. We show that the optimal operating points that minimize mean delay in the transient system are given by the rate vectors $\mathbf{c}_k^*$ defined recursively below in (4). These operating points apply the SRPT-FM principle, and they are not depending on the actual sizes of the flows (but only on their relative order). In addition, we show that these optimal operating points can be implemented by a sequence of weight-based schedulers with the weights $w_k = G_k^*$ defined also in (4). While the recursive algorithm given in (4) and originally presented in [2], requires an explicit representation of the capacity regions, we develop another recursive algorithm (6) that avoids this requirement by directly utilizing the time slot level channel model.

4.1 Transient system

Assume now that, at time 0, there are n flows in the system with sizes s_i (in bits). Without loss of generality, the (original) flows are indexed in such a way that

$$s_1 \geq s_2 \geq \ldots \geq s_n.$$

In this section, we do not allow any further arrivals but consider a *transient system* that remains busy only until the completion of all original file transfers.

As mentioned above, the operator of the system chooses at time 0 a rate vector $\mathbf{c} = (c_1, \ldots, c_n) \in \mathcal{C}_n$. From that on, each flow i is served with rate c_i until the number of flows changes, and a new rate vector is chosen. We assume that when choosing the rate vector the operator is aware of the (remaining) sizes of the flows. Recall that a rate vector $\mathbf{c} \in \mathcal{C}_n$ is achieved by an opportunistic scheduler $\pi \in \mathit{\Pi}_n$ such that $\theta_i^\pi = c_i$ for all $i = 1, \ldots, n$. Note also that it is only the flow level operator (operating in a much slower time scale) that needs to be aware of the remaining sizes of flows, not the opportunistic scheduler (operating in the fast time scale of time slots).

4.2 Problem formulation

An *operating policy* ϕ is defined by a sequence of rate vectors, $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$, where $\mathbf{c}_k = (c_{k1}, \ldots, c_{kk}) \in \mathcal{C}_k$ for all $k = 1, \ldots, n$.[2] In other words, the operator applies the rate vector $\mathbf{c}_k$ when there are k flows in the system (called hereafter *phase* k). It is assumed that when a flow completes, the remaining $k-1$ flows are *re-indexed* in such a way that the remaining sizes $s_{k-1,i}$ again satisfy

$$s_{k-1,1} \geq s_{k-1,2} \geq \ldots \geq s_{k-1,k-1}.$$

Thus, in the next phase, the longest flow is served with rate $c_{k-1,1}$, the second longest with rate $c_{k-1,2}$, etc. Let Φ_n denote the family of all operating policies,

$$\Phi_n = \{\phi = (\mathbf{c}_1, \ldots, \mathbf{c}_n) : \mathbf{c}_k \in \mathcal{C}_k \text{ for all } k\}.$$

Note that any operating policy $\phi = (\mathbf{c}_1, \ldots, \mathbf{c}_n)$ is implemented at the time slot level by choosing a sequence of opportunistic schedulers $(\pi_1, \ldots, \pi_n)$ such that $\pi_k \in \Pi_k$ for all $k = 1, \ldots, n$ and

$$\theta_i^{\pi_k} = c_{ki}$$

for all $i = 1, \ldots, k$.

The *total completion time* [15] for operating policy $\phi \in \Phi_n$ is defined as

$$T^\phi = \sum_{k=1}^{n} k\, T_k^\phi,$$

where T_k^ϕ refers to the length of phase k for policy ϕ. Note that the *mean delay* of a flow is now given by T^ϕ/n.

In this section, we consider the scheduling problem in which the optimal operating policy minimizes the total completion time (or the mean delay, as well) in the transient system. Let ϕ^* denote such an optimal policy. Thus,

$$T^{\phi^*} = \min_{\phi \in \Phi_n} T^\phi.$$

In fact, it is even more important to find the optimal schedulers $(\pi_1^*, \ldots, \pi_n^*)$ that implement the optimal policy ϕ^* at the time slot level.

There is a clear trade-off between size-based operating policies and the opportunistic gain. The number of flows is reduced most effectively by choosing the rate vector that gives the highest maximum rate to the shortest flow, but the service capacity is maximized by choosing the rate vector that gives the highest total rate. The optimal operating policy should find the best possible compromise between these two extremes.

[2] In the present paper, it is natural to define operating policies ϕ as a sequence of rate vectors $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ due to the time scale separation assumption made in Section 3. For more generic capacity regions, the optimal control in continuous time would, of course, be a relevant question.

4.3 Optimal solution

Let $g_1, \ldots, g_n$ be a sequence of functions with $g_k(\mathbf{c}_k)$ defined on $\mathcal{C}_k$ for all k, $G_1^*, \ldots, G_n^*$ a sequence of scalars, and $\mathbf{c}_1^*, \ldots, \mathbf{c}_n^*$ a sequence of rate vectors with $\mathbf{c}_k^* \in \mathcal{C}_k$ for all k. These sequences are defined recursively as in [2, Equation (1)]:

$$\begin{cases} g_1(c_1) = \frac{1}{c_1}, \\ G_1^* = g_1(c_1^*) = \min_{c_1 \in \mathcal{C}_1} g_1(c_1), \end{cases}$$
$$\begin{cases} g_k(\mathbf{c}_k) = \frac{1}{c_{kk}} \left(k - \sum_{i=1}^{k-1} G_i^* c_{ki} \right), \\ G_k^* = g_k(\mathbf{c}_k^*) = \min_{\mathbf{c}_k \in \mathcal{C}_k} g_k(\mathbf{c}_k), \ k = 2, \ldots, n. \end{cases} \tag{4}$$

Note that the existence of the minimum values G_k^* is guaranteed by the compactness of capacity regions $\mathcal{C}_k$ (Property (i) mentioned in Section 2). In particular, we have

$$G_1^* = \min_{c \in [0, \gamma_1]} \frac{1}{c} = \frac{1}{\gamma_1}, \quad c_1^* = \gamma_1.$$

Coefficients G_k^* have the following interpretation. Consider two flows with equal (unit) sizes $s_1 = s_2 = 1$. In this case, the minimum completion time clearly satisfies

$$\begin{aligned} \min_{\phi \in \Phi_2} T^\phi &= \min_{\mathbf{c}_2 \in \mathcal{C}_2} \left(\frac{2}{c_{22}} + \left(1 - \frac{c_{21}}{c_{22}} \right) \frac{1}{c_1^*} \right) \\ &= \min_{\mathbf{c}_2 \in \mathcal{C}_2} \left(\frac{1}{c_{22}} \left(2 - \frac{c_{21}}{c_1^*} \right) \right) + \frac{1}{c_1^*} \\ &= \min_{\mathbf{c}_2 \in \mathcal{C}_2} \left(\frac{1}{c_{22}} \left(2 - G_1^* c_{21} \right) \right) + G_1^* \\ &= G_2^* + G_1^*. \end{aligned}$$

By a dynamic programming argument, this can be generalized to any number of unit-size flows n (for which $s_k = 1$ for all $k = 1, \ldots, n$),

$$\min_{\phi \in \Phi_n} T^\phi = \sum_{k=1}^{n} G_k^*.$$

Thus, G_k^* can be interpreted as the cost per unit size for flow k under the optimal operating policy. Theorem 1 reveals that this interpretation is justified also in the general case with non-equal flow sizes.

In [2], we proved that, if the sequence $G_1^*, \ldots, G_n^*$ is stricly increasing, i.e.,

$$G_1^* < G_2^* < \ldots < G_n^*, \tag{5}$$

then $\mathbf{c}_k^*$ are optimal operating points to minimize the mean delay in the transient system. To have a self-contained paper, the result is repeated below, see Theorem 1. However, in [2], we worked with an abstract capacity region model (without any connection to the time slot level schedulers) so that we were not

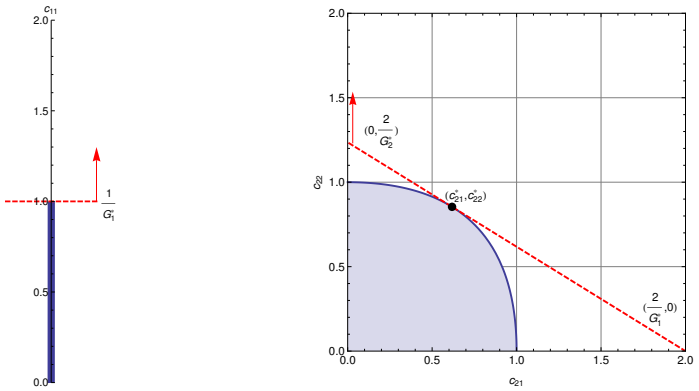

Fig. 1 Example with i.i.d. Exp(1) channels (Scenario A in Section 5.1). $G_1^* = 1$ as it is simply the inverse of the average rate. For two flows, G_2^* results from an optimization problem over the capacity set: the contours of the objective are lines that cross the c_{21}-axis at $2/G_1^*$ and c_{22}-axis at $2/g_2(\mathbf{c}_2)$ (dashed line represents the contour at the optimum). As G_1^* is fixed from the previous step, it remains to minimize $g_2(\mathbf{c}_2)$, which results in G_2^* such that $G_2^* > G_1^*$. This is also in line with the result for generic capacity regions (given in [2, p. 189]) saying that, if $c_{21} + c_{22} < 2/G_1^*$ for all $\mathbf{c}_2 \in \mathcal{C}_2$, then $G_2^* > G_1^*$.

able to give more explicit conditions for the optimality result. Now we have the connection (defined in Sections 2 and 3) that allows us to prove that the implicit condition (5) is, indeed, satisfied in our opportunistic scheduler model. This will be done in Proposition 3 below. An example is given in Fig. 1 that illustrates the result in the special case that $n = 2$.

To prove Proposition 3, we need the following auxiliary results, i.e., Proposition 2 and Corollary 1.

Proposition 2 *For all* $k = 1, \ldots, n$,

$$c_{kk}^* > 0, \qquad G_k^* > 0, \qquad \max_{\mathbf{c}_k \in \mathcal{C}_k} \sum_{i=1}^{k} G_i^* c_{ki} = \sum_{i=1}^{k} G_i^* c_{ki}^* = k.$$

Proof 1° For $k = 1$, we have

$$c_{11}^* = c_1^* = \gamma_1 > 0, \qquad G_1^* = \frac{1}{\gamma_1} > 0, \qquad G_1^* c_{11}^* = \frac{\gamma_1}{\gamma_1} = 1.$$

2° Let then $k \geq 2$. Assume now that the claim is true for $k - 1$. We will show that it is also true for k.

Let $\mathbf{c}_k \in \mathcal{C}_k$. Since the capacity regions are nested (as mentioned in Section 2), we have

$$(c_{k1}, \ldots, c_{k,k-1}) \in \mathcal{C}_{k-1}$$

implying, by the induction assumption, that

$$\sum_{i=1}^{k-1} G_i^* c_{ki} \leq k - 1.$$

Thus,

$$g_k(\mathbf{c}_k) = \frac{1}{c_{kk}} \left(k - \sum_{i=1}^{k-1} G_i^* c_{ki} \right) \geq \frac{1}{c_{kk}}.$$

We see that $g_k(\mathbf{c}_k) \to \infty$ as $c_{kk} \to 0$. It follows that

$$c_{kk}^* > 0$$

and

$$G_k^* \geq \frac{1}{c_{kk}^*} > 0.$$

In addition, by definition,

$$G_k^* = g_k(\mathbf{c}_k^*) = \frac{1}{c_{kk}^*} \left(k - \sum_{i=1}^{k-1} G_i^* c_{ki}^* \right),$$

implying that

$$\sum_{i=1}^{k} G_i^* c_{ki}^* = k.$$

On the other hand, again by definition, we have, for all $\mathbf{c}_k \in \mathcal{C}_k$,

$$G_k^* \leq g_k(\mathbf{c}_k) = \frac{1}{c_{kk}} \left(k - \sum_{i=1}^{k-1} G_i^* c_{ki} \right).$$

Thus, for all $\mathbf{c}_k \in \mathcal{C}_k$,

$$\sum_{i=1}^{k} G_i^* c_{ki} \leq k = \sum_{i=1}^{k} G_i^* c_{ki}^*,$$

which completes the proof. □

Combining Propositions 1 and 2 gives immediately the following corollary.

Corollary 1 *For all* $k = 1, \ldots, n$,

$$E[\max_{i=1,\ldots,k} G_i^* R_i] = k.$$

Since the rate processes $R_i(t)$ are assumed to be independent, we have

$$\begin{aligned} E[\max_{i=1,\ldots,k} G_i^* R_i] &= \int_0^\infty P\{\max_{i=1,\ldots,k} G_i^* R_i > r\} \, \mathrm{d}r \\ &= \int_0^\infty \left(1 - \prod_{i=1}^{k} P\{G_i^* R_i \leq r\} \right) \mathrm{d}r. \end{aligned}$$

By combining this with Corollary 1, we get an alternative recursion to determine the sequence $G_1^*, \ldots, G_n^*$:

$$\begin{aligned} &G_1^* = \tfrac{1}{\gamma_1}, \\ &\begin{cases} f_k(a) = \displaystyle\int_0^\infty \left(1 - \prod_{i=1}^{k-1} P\{G_i^* R_i \le r\} P\{aR_k \le r\}\right) \mathrm{d}r, \\ G_k^* = f_k^{-1}(k), \ k = 2, \ldots, n. \end{cases} \end{aligned} \tag{6}$$

Note that function $f_k(a)$ is clearly strictly increasing with

$$f_k(0) = \int_0^\infty \left(1 - \prod_{i=1}^{k-1} P\{G_i^* R_i \le r\}\right) \mathrm{d}r = E[\max_{i=1,\ldots,k-1} G_i^* R_i] = k - 1$$

and $f_k(a) \to \infty$ as $a \to \infty$. Thus, the inverse function $f_k^{-1}(b)$ is well defined on $[k-1, \infty)$ (including value k).

The advantage of recursion (6) is that it utilizes directly the time slot level model for the rate processes $R_i(t)$, while the original recursion (4) requires that the flow level capacity regions $\mathcal{C}_k$ be determined.

Next we show the important result that the G_k^* sequence is strictly increasing in our opportunistic scheduler model, which is based on i.i.d. rate processes $R_i(t)$ for separate users i.[3]

Proposition 3 *For the opportunistic scheduler capacity regions defined in Section 2, we have*

$$G_1^* < G_2^* < \ldots < G_n^*.$$

Proof 1° Define $G_0^* = 0$. Then

$$G_1^* = \frac{1}{\gamma_1} > 0 = G_0^*.$$

2° Denote $X_0 = 0$ and, for all $k = 1, \ldots, n-1$,

$$X_k = \max_{i=1,\ldots,k} G_i^* R_i.$$

Note that $X_0, X_1, \ldots, X_{n-1}$ is a non-decreasing sequence of random variables. In addition, for all $k = 0, 1, \ldots, n-1$, define functions $h_k(a)$, $a \in \mathbb{R}$, as follows:

$$h_k(a) = E[(aR_{k+1} - X_k) 1_{\{aR_{k+1} > X_k\}}].$$

Note that $h_k(a)$ is non-decreasing as a function of a for all k. Furthermore, by definition and Corollary 1, we have, for all $k = 0, 1, \ldots, n-1$,

$$\begin{aligned} h_k(G_{k+1}^*) &= E[(G_{k+1}^* R_{k+1} - X_k) 1_{\{G_{k+1}^* R_{k+1} > X_k\}}] \\ &= E[X_{k+1} - X_k] = 1. \end{aligned}$$

[3] An example is given in [2, p. 189] demonstrating that the result is not necessarily true for more generic capacity regions.

Now, let $k = 1, \ldots, n-1$, and assume that, for all $j = 0, 1, \ldots, k-1$,

$$G^*_{j+1} > G^*_j.$$

(i) If the steady-state rate distribution is deterministic (i.e., $P\{R_1 = \gamma_1\} = 1$), then clearly $P\{G^*_k R_{k+1} > X_k\} = 0$ and

$$h_k(G^*_k) = E[(G^*_k R_{k+1} - X_k) 1_{\{G^*_k R_{k+1} > X_k\}}] = 0.$$

(ii) Assume now that the steady-state rate distribution is non-deterministic. In this case, by the induction assumption,

$$\begin{aligned} &P\{G^*_k R_{k+1} > X_k > X_{k-1}\} \\ &\quad \geq \; P\{R_{k+1} > R_k \geq R_{k-1} \geq \ldots \geq R_1\} \; > \; 0, \end{aligned}$$

which implies that

$$\begin{aligned} h_k(G^*_k) &= E[(G^*_k R_{k+1} - X_k) 1_{\{G^*_k R_{k+1} > X_k\}}] \\ &= E[(G^*_k R_{k+1} - X_k) 1_{\{G^*_k R_{k+1} > X_k > X_{k-1}\}}] \\ &\quad + E[(G^*_k R_{k+1} - X_k) 1_{\{G^*_k R_{k+1} > X_k = X_{k-1}\}}] \\ &< E[(G^*_k R_{k+1} - X_{k-1}) 1_{\{G^*_k R_{k+1} > X_k\}}] \\ &\leq E[(G^*_k R_{k+1} - X_{k-1}) 1_{\{G^*_k R_{k+1} > X_{k-1}\}}] \\ &= E[(G^*_k R_k - X_{k-1}) 1_{\{G^*_k R_k > X_{k-1}\}}] \\ &= h_{k-1}(G^*_k) \; = \; 1. \end{aligned}$$

Thus, we have $h_k(G^*_k) < 1$ in both cases ((i) and (ii)). Since $h_k(G^*_{k+1}) = 1$ and $h_k(a)$ is a non-decreasing function, we conclude that

$$G^*_{k+1} > G^*_k,$$

which completes the proof. □

Another monotonicity result is given below (cf. [2, Proposition 1]).

Proposition 4 *For all $k = 2, \ldots, n$,*

$$c^*_{k,1} \leq c^*_{k,2} \leq \ldots \leq c^*_{k,k}.$$

Proof 1° Let $k \in \{2, \ldots, n\}$ and $j \in \{1, \ldots, k-2\}$. In addition, let $\tilde{\mathbf{c}}^*_k$ denote the modification of $\mathbf{c}^*_k$ where the service rates $c^*_{k,j}$ and $c^*_{k,j+1}$ have changed their places,

$$\tilde{\mathbf{c}}^*_k = (c^*_{k,1}, \ldots, c^*_{k,j-1}, c^*_{k,j+1}, c^*_{k,j}, c^*_{k,j+2}, \ldots, c^*_{k,k}).$$

Note that $\tilde{\mathbf{c}}^*_k \in \mathcal{C}_k$, since $\mathcal{C}_k$ is symmetric (Property (ii) mentioned in Section 2). Now

$$\begin{aligned} &c^*_{kk} \left(g_k(\mathbf{c}^*_k) - g_k(\tilde{\mathbf{c}}^*_k) \right) \\ &\quad = \; c^*_{k,j+1} G^*_j + c^*_{k,j} G^*_{j+1} - c^*_{k,j} G^*_j - c^*_{k,j+1} G^*_{j+1} \\ &\quad = \; (c^*_{k,j} - c^*_{k,j+1})(G^*_{j+1} - G^*_j). \end{aligned}$$

Since $c^*_{kk}\,(g_k(\mathbf{c}^*_k) - g_k(\tilde{\mathbf{c}}^*_k)) \le 0$ by definition, and $G^*_{j+1} - G^*_j > 0$ by Proposition 3, we conclude that

$$c^*_{k,j} - c^*_{k,j+1} \le 0.$$

2° Consider now the remaining case where $k \in \{2,\ldots,n\}$ and $j = k-1$. If $c^*_{k,k-1} = 0$, then $c^*_{k,k-1} - c^*_{kk} \le 0$ for sure. Thus, we may assume that $c^*_{k,k-1} > 0$. Now, let $\tilde{\mathbf{c}}^*_k$ denote the modification of $\mathbf{c}^*_k$ where the service rates $c^*_{k,k-1}$ and $c^*_{k,k}$ have changed their places,

$$\tilde{\mathbf{c}}^*_k = (c^*_{k,1},\ldots,c^*_{k,k-2},c^*_{k,k},c^*_{k,k-1}).$$

Note again that $\tilde{\mathbf{c}}^*_k \in \mathcal{C}_k$, since $\mathcal{C}_k$ is symmetric. Now

$$\begin{aligned}
&c^*_{kk}c^*_{k,k-1}\,(g_k(\mathbf{c}^*_k) - g_k(\tilde{\mathbf{c}}^*_k)) \\
&= c^*_{k,k-1}\left(k - \sum_{j=1}^{k-2} c^*_{kj}G^*_j - c^*_{k,k-1}G^*_{k-1}\right) \\
&\quad - c^*_{kk}\left(k - \sum_{j=1}^{k-2} c^*_{kj}G^*_j - c^*_{kk}G^*_{k-1}\right) \\
&= (c^*_{k,k-1} - c^*_{kk})\left(k - \sum_{j=1}^{k-1} c^*_{kj}G^*_j - c^*_{kk}G^*_{k-1}\right) \\
&= (c^*_{k,k-1} - c^*_{kk})\left(c^*_{kk}G^*_k - c^*_{kk}G^*_{k-1}\right) \\
&= c^*_{kk}(c^*_{k,k-1} - c^*_{kk})(G^*_k - G^*_{k-1}).
\end{aligned}$$

Since $c^*_{kk}c^*_{k,k-1}\,(g_k(\mathbf{c}^*_k) - g_k(\tilde{\mathbf{c}}^*_k)) \le 0$ by definition, and $G^*_k - G^*_{k-1} > 0$ by Proposition 3, we conclude that

$$c^*_{k,k-1} - c^*_{kk} \le 0,$$

which completes the proof. □

Due to Proposition 3, the optimality result given in Theorem 1 below follows now from [2, Theorem 1]. For completeness, however, we give the proof also in this paper. In addition, we are now able to give the implementation of the optimal flow level policy at the time slot level, see Theorem 2 below.

Theorem 1 *For all flow sizes $s_1 \ge s_2 \ge \ldots \ge s_n$, the optimal operating policy is $\phi^* = (\mathbf{c}^*_1,\ldots,\mathbf{c}^*_n)$. The minimum total completion time T^{ϕ^*} satisfies*

$$T^{\phi^*} = \sum_{k=1}^{n} G^*_k s_k.$$

In addition, for all $k = 2,\ldots,n$,

$$c^*_{k,1} \le c^*_{k,2} \le \ldots \le c^*_{k,k},$$

i.e., the optimal policy applies the SRPT-FM principle.

Proof The monotonicity of the optimal rates is already shown in Proposition 4. The optimality result itself is proved below by induction.

1° For $n = 1$, the result is clearly true:

$$T^{\phi^*} = \frac{s_1}{\gamma_1} = \min_{c_1 \in \mathcal{C}_1} \frac{s_1}{c_1} = \min_{\phi} T^{\phi}.$$

In addition, $G_1^* = 1/\gamma_1$ so that $T^{\phi^*} = G_1^* s_1$ as claimed.

2° Assume now that $n \geq 2$ and the result is true for all values $1, \ldots, n-1$. We will show that it is also true for value n.

It follows from the induction assumption that the optimal policy applies rate vectors $\mathbf{c}_k^*$ for all $k = 1, \ldots, n-1$. Thus, for any policy $\phi = (\mathbf{c}_1, \ldots, \mathbf{c}_n) \in \Phi_n$, the modified policy $\tilde{\phi} = (\mathbf{c}_1^*, \ldots, \mathbf{c}_{n-1}^*, \mathbf{c}_n) \in \Phi_n$ results in a smaller total completion time so that

$$\begin{aligned}
T^{\phi} &\geq T^{\tilde{\phi}} \\
&= n\, T_n^{\tilde{\phi}} + \sum_{k=1}^{n-1} k\, T_k^{\tilde{\phi}} \\
&= n\, T_n^{\tilde{\phi}} + \sum_{k=1}^{n-1} G_k^* \left(s_{i(k)} - T_n^{\tilde{\phi}} c_{n,i(k)} \right) \\
&= n\, \frac{s_{i(n)}}{c_{n,i(n)}} + \sum_{k=1}^{n-1} G_k^* \left(s_{i(k)} - \frac{s_{i(n)}}{c_{n,i(n)}}\, c_{n,i(k)} \right) \\
&= \frac{s_{i(n)}}{c_{n,i(n)}} \left(n - \sum_{k=1}^{n-1} G_k^* c_{n,i(k)} \right) + \sum_{k=1}^{n-1} G_k^* s_{i(k)} \\
&= g_n((c_{n,i(1)}, \ldots, c_{n,i(n)}))s_{i(n)} + \sum_{k=1}^{n-1} G_k^* s_{i(k)},
\end{aligned}$$

where $i(k)$ refers to the original index of the flow that completes at the end of phase k under policy $\tilde{\phi}$. Note that $(c_{n,i(1)}, \ldots, c_{n,i(n)}) \in \mathcal{C}_n$, since $\mathbf{c}_n = (c_{n1}, \ldots, c_{nn}) \in \mathcal{C}_n$ and $\mathcal{C}_n$ is symmetric (Property (ii)). Thus,

$$g_n((c_{n,i(1)}, \ldots, c_{n,i(n)})) \geq G_n^*$$

implying that

$$T^{\phi} \geq \sum_{k=1}^{n} G_k^* s_{i(k)} \geq \sum_{k=1}^{n} G_k^* s_k,$$

where the latter inequality follows from the facts that $s_1 \geq \ldots \geq s_n$ (by assumption) and $G_1^* < \ldots < G_n^*$ (by Proposition 3).

Consider then policy $\phi^* = (\mathbf{c}_1^*, \ldots, \mathbf{c}_n^*)$ and let $i^*(k)$ denote the original index of the flow that completes at the end of phase k under this policy ϕ^*. It

follows from Proposition 4 that $i^*(k) = k$ for all k. Thus,

$$\begin{aligned} T^{\phi^*} &= n\,T_n^{\phi^*} + \sum_{k=1}^{n-1} k\,T_k^{\phi^*} \\ &= n\,\frac{s_n}{c_{nn}^*} + \sum_{k=1}^{n-1} G_k^* \left(s_k - \frac{s_n}{c_{nn}^*}\,c_{nk}^* \right) \\ &= \frac{s_n}{c_{nn}^*} \left(n - \sum_{k=1}^{n-1} G_k^* c_{nk}^* \right) + \sum_{k=1}^{n-1} G_k^* s_k \\ &= g_n(\mathbf{c}_n^*) s_n + \sum_{k=1}^{n-1} G_k^* s_k \\ &= \sum_{k=1}^{n} G_k^* s_k \end{aligned}$$

so that $T^{\phi} \geq T^{\phi^*}$ for any $\phi \in \Phi_n$. □

Theorem 2 *The optimal operating policy $\phi^* = (\mathbf{c}_1^*, \ldots, \mathbf{c}_n^*)$ can be implemented by a sequence of weight-based schedulers that, in phase k, use the weight vector $\mathbf{w}_k = (G_1^*, \ldots, G_k^*)$ and break the ties by giving the time slot t to the flow with the highest index i^* such that*

$$G_{i^*}^* R_{i^*}(t) = \max\{G_1^* R_1(t), \ldots, G_k^* R_k(t)\}.$$

Proof Let $k = 1, \ldots, n$. For $k = 1$, the result is trivially true. Thus, we may assume that $k > 1$.

Let π denote the weight-based scheduler based on weight vector $\mathbf{w}_k = (G_1^*, \ldots, G_k^*)$. In addition, let $\mathbf{c}^{\pi} = (\theta_1^{\pi}, \ldots, \theta_k^{\pi}) \in \mathcal{C}_k$ denote the corresponding throughput vector. Note that, for all $i = 1, \ldots, k$,

$$\theta_i^{\pi} = E[R_i \prod_{j=1}^{i-1} 1_{\{G_i^* R_i \geq G_j^* R_j\}} \prod_{j=i+1}^{k} 1_{\{G_i^* R_i > G_j^* R_j\}}].$$

Below we will show that

$$g_k(\mathbf{c}^{\pi}) = \min_{\mathbf{c}_k \in \mathcal{C}_k} g_k(\mathbf{c}_k) = G_k^*,$$

which, by (4), implies that the claim is true.

First, by Proposition 1, we have

$$E[\max_{i=1,\ldots,k} G_i^* R_i] = \max_{\mathbf{c} \in \mathcal{C}_k} \sum_{i=1}^{k} G_i^* c_i \geq \sum_{i=1}^{k} G_i^* \theta_i^{\pi}. \tag{7}$$

On the other hand, by definition,

$$\begin{aligned}\sum_{i=1}^{k} G_i^* \theta_i^\pi &= \sum_{i=1}^{k} G_i^* E[R_i \prod_{j=1}^{i-1} 1_{\{G_i^* R_i \geq G_j^* R_j\}} \prod_{j=i+1}^{k} 1_{\{G_i^* R_i > G_j^* R_j\}}] \\ &= E[\sum_{i=1}^{k} G_i^* R_i \prod_{j=1}^{i-1} 1_{\{G_i^* R_i \geq G_j^* R_j\}} \prod_{j=i+1}^{k} 1_{\{G_i^* R_i > G_j^* R_j\}}] \\ &= E[\max_{i=1,\ldots,k} G_i^* R_i] \qquad (8)\end{aligned}$$

Equations (7) and (8), together with Proposition 2, imply that

$$\sum_{i=1}^{k} G_i^* \theta_i^\pi = \max_{\mathbf{c} \in \mathcal{C}_k} \sum_{i=1}^{k} G_i^* c_i = k.$$

Thus,

$$g_k(\mathbf{c}^\pi) = \frac{1}{\theta_k^\pi} \left(k - \sum_{i=1}^{k-1} G_i^* \theta_i^\pi \right) = \frac{1}{\theta_k^\pi} \left(\sum_{i=1}^{k} G_i^* \theta_i^\pi - \sum_{i=1}^{k-1} G_i^* \theta_i^\pi \right) = G_k^*,$$

which completes the proof. □

5 Numerical examples for the transient system

In this section, we present numerical examples for three different channel scenarios and illustrate the differences in the flow level performance of the transient system when policies mentioned in Section 3 are applied. Scenarios A and B reflect two extreme cases: in Scenario A, the opportunistic gain grows without limits as the number of users increases, while it is rather restricted and bounded in Scenario B. The third one, Scenario C, lies somewhere between the two and is motivated by a practical system.

5.1 Scenario A: Exponential channel

In the first channel scenario, we assume Rayleigh fading and linear data rate to SINR dependency. In this model, the rate process is exponential, $R_i \sim \text{Exp}(1)$, and the opportunistic gain γ_k in phase k is the maximum of k independent Exp(1) random variables, which equals the harmonic number H_k,

$$\gamma_k = H_k = \sum_{i=1}^{k} \frac{1}{i}.$$

Note that the opportunistic gain for this model is unbounded, $\gamma_k \to \infty$ as $k \to \infty$. Recursion (6) is now based on

$$f_k(a) = \int_0^\infty \left(1 - \left(1 - e^{-\frac{r}{a}}\right) \prod_{i=1}^{k-1} \left(1 - e^{-\frac{r}{G_i^*}}\right) \right) \mathrm{d}r$$

from which optimal weights G_k^* can be computed numerically utilizing a simple line search. The corresponding optimal flow level operating point $\mathbf{c}_k^*$ can then be computed numerically as follows:

$$c_{ki}^* = \int_0^\infty r e^{-r} \prod_{j \neq i} \left(1 - e^{-\frac{G_i^*}{G_j^*} r} \right) \mathrm{d}r.$$

5.2 Scenario B: Two-state channel

In the second scenario, the instantaneous rate process has only two discrete values, $r_{\min}$ and $r_{\max}$, with $r_{\min} < r_{\max}$. Let $p = P\{R_i = r_{\max}\}$. In all forthcoming numerical examples, we use values $r_{\min} = 1$, $r_{\max} = 10$ and $p = 1/2$ (to have a significant difference in the rates).

The opportunistic gain γ_k in phase k is now

$$\gamma_k = r_{\min} (1-p)^k + r_{\max} (1 - (1-p)^k),$$

which clearly converges to $r_{\max}$ as $k \to \infty$.

To compute G_k^*, we define $m(a)$, $a > 0$, as the maximum $m \in \{0, \ldots, k-1\}$ for which $G_m^* r_{\max} < a r_{\min}$, with the convention $G_0^* = 0$. Recursion (6) is now based on

$$f_k(a) = a \left(r_{\min} (1-p)^{k-m(a)} + r_{\max} p \right) + r_{\max} \sum_{i=m(a)+1}^{k-1} G_i^* p (1-p)^{k-i},$$

allowing a numerical solution for G_k^*. Once G_k^* is determined, let $m_k = m(G_k^*)$. The optimal rate vector $\mathbf{c}_k^*$ is obtained in a straightforward fashion:

$$c_{ki}^* = \begin{cases} 0, & i \leq m_k, \\ r_{\max} p (1-p)^{k-i}, & m_k < i < k, \\ r_{\min} (1-p)^{k-m_k} + r_{\max} p, & i = k. \end{cases}$$

Note that the SRPT-P policy results in a rate vector which is the same as above but with $m_k \equiv 0$. Indeed, the only difference between the SRPT-P and the optimal policy is that in the optimum only the $k - m_k$ shortest flows are served with non-zero rates whereas SRPT-P serves (in principle) all. Naturally, the "cut point" and thus the difference between the policies depends on the parameters of the model, i.e., $r_{\min}$, $r_{\max}$, and p.

5.3 Scenario C: HDR channel

The last channel model example is motivated by the empirical distribution for the achievable rates in an actual HDR system, see [4]. The model has been used, e.g., in [7]. In this case, there are 11 rates, given below in kbps,

$$r = \{38.4, 76.8, 102.6, 153.6, 204.8, 307.2, 614.4, 921.6, 1228.8, 1843.2, 2457.6\},$$

with the respective probabilities

$$p = \{0, 0.01, 0.04, 0.08, 0.15, 0.24, 0.18, 0.09, 0.12, 0.05, 0.04\}.$$

The optimal weights G_k^* and the corresponding optimal flow level operating points $\mathbf{c}_k^*$ can be computed similarly to the two-state model but we omit the computational details for brevity.

5.4 Performance comparison at flow level

To compare the performance of different policies in the transient system, we consider the three channel scenarios introduced above. The considered policies ϕ are the optimal policy TR-OPT together with PF and SRPT-P introduced in Section 3. As PF represents our baseline policy, in the results below we always study the performance relative to PF. The performance is measured by the total completion time under the time-scale separation assumption, i.e., based on the flow level model.

First, the total completion time of the different policies ϕ relative to the total completion time under PF, i.e., the ratio

$$\frac{T^{\phi}}{T^{\mathrm{PF}}},$$

is shown in Fig. 2 (left), which depicts the performance ratio as a function of n, the initial number of flows in the system, for 1000 randomly selected initial sizes drawn from an exponential distribution. In Scenario A, the rate distribution is continuous and thus PF and SRPT-P are identical policies, but the optimal policy TR-OPT produces a significant benefit over PF. In Scenario B, the benefit from the optimal policy TR-OPT is more significant (close to 50% for large n), and interestingly it gives numerically identical performance to SRPT-P. However, this is a result of the rather large difference between $r_{\min}$ and $r_{\max}$, recall the discussion on the relation of SRPT-P with the optimal policy in Section 5.2. Finally, in Scenario C the benefit of the optimal policy TR-OPT is somewhere between Scenarios A and B, and SRPT-P is somewhat inferior to the TR-OPT policy.

To further illustrate the differences of the policies, we examine the total service rate of the policies ϕ relative to the total service rate of PF, i.e., the ratio

$$\frac{c_{n1}^{\phi} + \ldots + c_{nn}^{\phi}}{c_{n1}^{\mathrm{PF}} + \ldots + c_{nn}^{\mathrm{PF}}}.$$

Note that since SRPT-P always gives priority to the highest instantaneous rate, the total service rate is exactly the same as that achieved by PF, only the division of the rate among the flows is different between SRPT-P and PF. Thus, we only consider the ratio for the TR-OPT policy. The results are given in Fig. 2 (right). There is a striking qualitative difference in the behavior between Scenario A and Scenarios B-C. In Scenario A, where the opportunistic

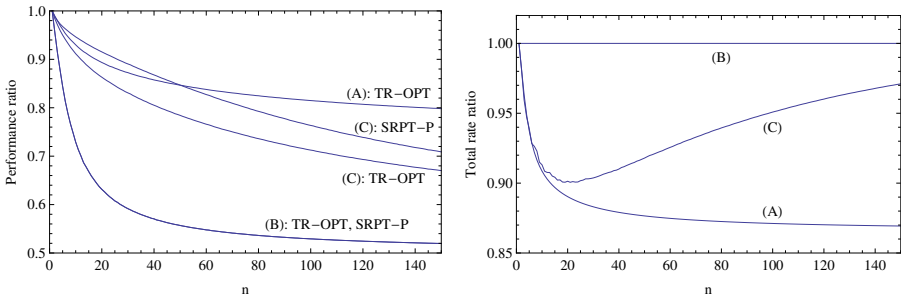

Fig. 2 The ratio of the total completion time for different policies relative to the total completion time under PF (left) and the ratio of the total service rate of the optimal policy TR-OPT relative to the total service rate with PF (right) for Scenarios A, B and C.

gain is unbounded, the ratio is less than one and appears to be decreasing, i.e., the total service rate achieved by the optimal policy TR-OPT is clearly behind that of PF. However, in Scenario B the ratio is equal to one (up to numerical accuracy) and in Scenario C the ratio appears to approach one as n increases.

5.5 Performance comparison at time slot level

The optimal time slot level scheduler for the transient problem *without assuming time-scale separation* can be solved numerically using dynamic programming, see [20,12]. In the following, we investigate how close the performance of the time-slot level implementation of the TR-OPT policy, as well as SRPT-P and PF, are to the performance of the optimal time slot level scheduler. The numerical complexity of solving the dynamic program limits our analysis to the two-state channel (Scenario B with varying rate ratio $r_{\max}/r_{\min}$) and two flows ($n = 2$). We note that the performance of the time slot level implementations of TR-OPT, SRPT-P and PF can also be obtained numerically using so-called Howard equations, cf. [12, Sect. 4].

For TR-OPT and its implementation at the time slot level, we conclude that if $r_{\max}/r_{\min} < 2$ then $G_2^* = 2G_1^*$ and the scheduler always serves the shorter flow until completion. Otherwise, if $r_{\max}/r_{\min} > 2$, it holds that

$$G_2^* = \frac{2 - p(1-p)r_{\max}/((1-p)r_{\min} + p\, r_{\max})}{(1-p)^2 r_{\min} + p\, r_{\max}} < 2G_1^*,$$

and the scheduler serves the initially shorter flow 2 in all channel states except when $R_1(t) = r_{\max}$ and $R_2(t) = r_{\min}$.

In Fig. 3, we show the ratio of the mean completion time of TR-OPT, SRPT-P and PF policies to the optimal time slot level policy as a function of the mean size of the flows, $\mathrm{E}[X]$ (initial flow sizes X, expressed in time slots, obey a geometric distribution). In the left panel, the ratio $r_{\max}/r_{\min} = 1.5$, and TR-OPT very quickly becomes as good as the optimal solution as the

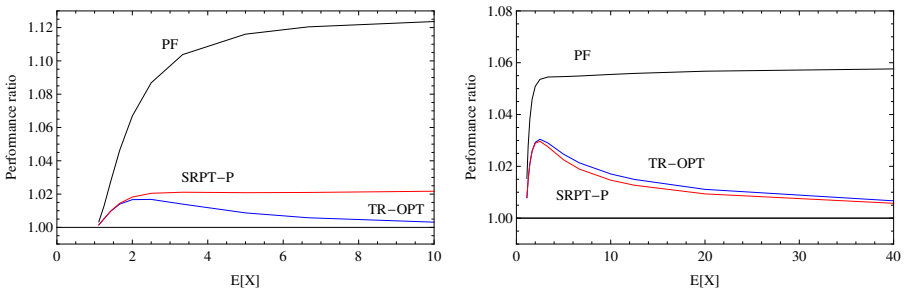

Fig. 3 The ratio of the total completion time for the time slot level implementations of TR-OPT, SRPT-P and PF relative to the total completion time of two flows under the optimal time slot level scheduler for two-state channel as a function of mean flow size with $r_{\max}/r_{\min} = 1.5$ (left) and $r_{\max}/r_{\min} = 5$ (right).

mean flow size grows, while for SRPT-P a small gap remains. In the right panel, the ratio $r_{\max}/r_{\min} = 5$ and the difference between TR-OPT and the optimal time slot level policy is now greater and somewhat larger flow sizes are needed until the difference disappears. Interestingly, in this case, SRPT-P gives practically identical performance as TR-OPT. Finally, in both cases PF performs worst and a clear performance gap remains.

Our numerical experiments demonstrate that the time scale separation assumption is not that bad even if the flow sizes are only one or two orders of magnitude greater than the length of the time slot. In addition, the difference between the truly optimal scheduler and the time slot level implementation of TR-OPT becomes marginal when the mean flow size is sufficiently large.

6 Dynamic scheduling problem

It is also interesting to investigate how the optimal policy of the transient system, TR-OPT, performs in the dynamic scheduling problem, where flows are not fixed at the beginning but arrive randomly.

To study the dynamic problem, we utilize two independent simulators. The flow level simulator (implemented in Mathematica) assumes the time scale separation. Flows arrive according to a Poisson process with rate λ and in each flow level state n they are served using a rate vector chosen from the corresponding capacity region $\mathcal{C}_n$ according to the given policy. The second simulator, referred to as the packet level simulator, is applied to the same flow level scenario, but in this case the scheduling decisions are done on the packet level: In each time slot t the rate process values $R_i(t)$ are drawn independently from their respective steady-state distributions and the policy defines which flow is served in each time slot. For example, the TR-OPT policy schedules the flow with maximum $G_i^* R_i(t)$ in time slot t. The packet level simulator was implemented in C++.

In the simulations, we considered the same channel scenarios as in Section 5, i.e., Scenarios A-C. The simulations were run for exponentially dis-

tributed flow sizes. The flow level simulations were run with $5 \cdot 10^6$ arrivals. In the packet level simulator, the time slot length and the mean file size were kept fixed and the arrival rate λ was varied to change the load. For Scenarios A and B, the time slot duration and the mean size of the flows were chosen so that the time-scale separation is reasonably well preserved, which resulted in time slot duration of 0.001 and mean size of 1. For Scenario C, the time slot duration was taken from the specifications of HDR systems and thus the time slot duration was 1.67 ms and the mean file size was 50 KB. The simulation runs in the packet level simulations also consisted of $5 \cdot 10^6$ arrivals.

In addition to the policies compared in Section 5 with PF serving as the baseline policy, we also considered the SRPT-OPS(k) policy introduced in Section 3. In the results below, we have optimized the value of k separately for each value of the load. However, the optimization was not done in an exhaustive manner over a fixed range of k, but instead, as the load was increased, the optimization was done by searching for the next local minimizer when starting from the current optimal value valid for the previous load value. Below the optimized SRPT-OPS(k) policy is briefly referred to as OPS*. The simulations with this policy were run with 10^6 arrivals.

6.1 Scenario A: Exponential channel

Scenario A, with a continuous rate distribution and infinitely growing opportunistic gain, resulted in the mean occupancies as a function of λ as depicted in Fig. 4 (left). The results from the flow level simulator are shown with continuous lines, and the results from the packet level simulations are indicated with dots. PF and SRPT-P are identical policies in this case, and even the OPS* policy resulted in almost identical performance (the difference being so small that it can not be distinguished in the figure). The TR-OPT policy is clearly getting increasingly worse as the arrival rate λ increases, indicating possible stability problems. Recall also that the total service rate in Scenario A for TR-OPT does not achieve the same as for PF, see Fig. 2 (right).

Fig. 4 (right) gives the ratio of the mean number of flows under TR-OPT relative to PF as a function of the arrival rate λ. Note that the stability limit of the corresponding constant rate non-opportunistic M/G/1 queue (in which the channel information is not utilized) is at $\lambda = 1$. From the figure we observe that at low loads, and slightly beyond the M/G/1 stability limit, TR-OPT is still able to achieve a slight gain over PF and the other policies.

6.2 Scenario B: Two-state channel

Scenario B, the two-state channel model, represents a case where the capacity is bounded from above by $r_{\max}$. The results are shown in Fig. 5. The results for the mean number of flows as a function of the load is given in the left panel and the relative performance compared with PF is shown in the right panel.

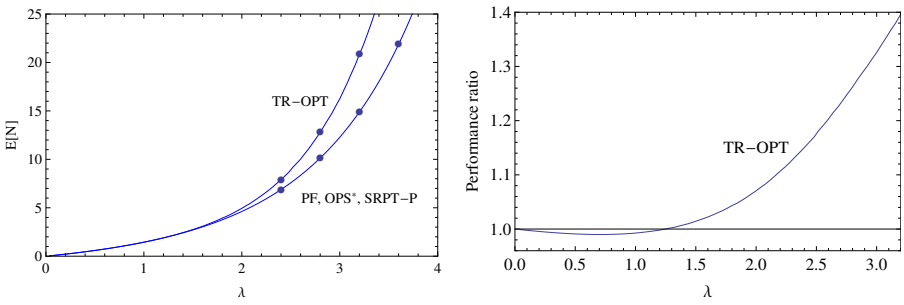

Fig. 4 Mean number of flows with different policies (left) and ratio of the mean number of flows relative to PF (right) as a function of load for Scenario A.

The load parameter is given by $\rho = \lambda \bar{X}/r_{\max}$, where $\bar{X}$ denotes the mean file size, and the necessary condition for stability reads $\rho < 1$. In this case the stability limit of the corresponding M/G/1 queue is just $\bar{r}/r_{\max} = 0.55$, where $\bar{r}$ denotes the mean rate.

Throughout the load region, the TR-OPT and SRPT-P policies yield practically identical performance, and the performance is significantly better than with PF and also clearly better than with OPS*.

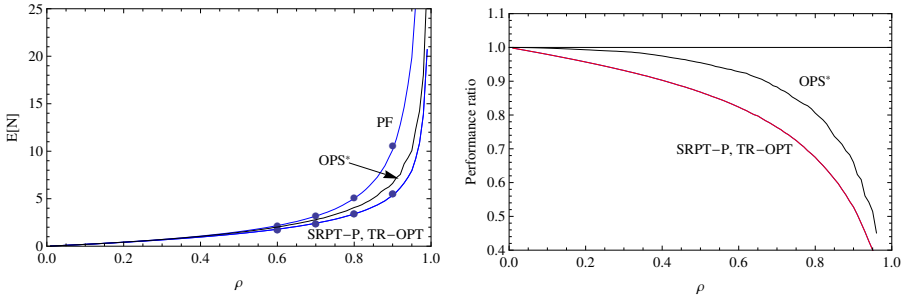

Fig. 5 Mean number of flows with different policies (left) and ratio of the mean number of flows relative to PF (right) as a function of load for Scenario B.

6.3 Scenario C: HDR channel

Finally, Scenario C, the 11-state channel model gives the results shown in Fig. 6, where the left panel gives the mean occupancy and the right panel the relative performance compared with PF, as a function of the load. Here the load is defined as in Scenario B but with $r_{\max}$ representing the maximum HDR rate and the stability limit of the M/G/1 system is approximately at 0.27.

The results from the flow level simulator are shown with continuous lines, and the results from the packet level simulations are indicated with dots. As

can be seen, the time scale separation is nicely realized even in this scenario that is based on realistic parameters.

The results confirm partly the observations from the two-state model with SRPT-P exhibiting best performance over all loads and being superior in particular at high loads, while TR-OPT performs clearly worse at high loads than in the two-state model. The OPS* policy is practically the same as PF, and its irregular behavior at high loads is a result of the local optimization of the k-parameter of the policy.

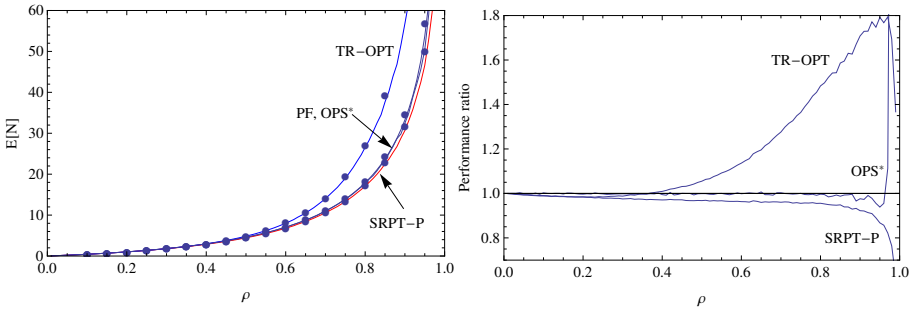

Fig. 6 Mean number of flows with different policies (left) and ratio of the mean number of flows relative to PF (right) as a function of load for Scenario C

6.4 Summary

As a synthesis for the dynamic case based on these three different scenarios, we find that the scheduling policy SRPT-P is uniformly better than the baseline policy PF or the optimized SRPT-OPS(k) policy, OPS*, and almost uniformly better than TR-OPT. Even in the cases where SRPT-P is not the best one, its difference to the best policy (TR-OPT) is rather marginal. The TR-OPT policy performs well whenever the traffic load is light or the opportunistic gain is very restricted.

7 Conclusions

In this paper we have considered the problem of combining optimally opportunistic and size-based scheduling in the context of data transmission in a wireless cellular system. Our model covers the symmetric case where the channel conditions for different users are independent and identically distributed (up to a user-specific coefficient), and our approach is based on the time scale separation assumption that allows us to define the optimal scheduling problem at the flow level. We have found the optimal trade-off in the transient setting where all the elastic traffic flows are available at time 0 and no new

flows arrive. The optimal policy is characterized both at the flow level (by giving the optimal operating points in all the phases of the service process) and at the time slot level (by giving the parameters of the optimal opportunistic scheduler). We have also explored the dynamic case with randomly arriving new flows by simulations in three very different scenarios. Interestingly, the scheduling policy TR-OPT, which is optimal for the transient system, can be improved in the dynamic case under high traffic load by applying a rate-based priority scheduler that breaks the ties based on the SRPT principle.

While we have already been considering the trade-off between opportunistic and size-based scheduling in our earlier works [12, 1, 2], in the present paper, we are, for the first time, able to combine opportunistic and size-based scheduling optimally at the time slot level. The results cover the symmetric case in the static setting with a fixed number of flows. Thus, the future challenge relates to extending the results to cover the asymmetric case with user-specific channel conditions and/or the dynamic setting with randomly arriving new flows.

References

1. Aalto, S., Lassila, P.: Flow-level stability and performance of channel-aware priority-based schedulers. In: Proceedings of NGI (2010)
2. Aalto, S., Penttinen, A., Lassila, P., Osti, P.: On the optimal trade-off between SRPT and opportunistic scheduling. In: Proceedings of ACM SIGMETRICS, pp. 185–195 (2011)
3. Ayesta, U., Erausquin, M., Jacko, P.: A modeling framework for optimizing the flow-level scheduling with time-varying channels. Performance Evaluation **67**, 1014–1029 (2010)
4. Bender, P., Black, P., Grob, M., Padovani, R., Sindhushyana, N., Viterbi, S.: CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users. IEEE Communications Magazine **38**(7), 70–77 (2000)
5. Berggren, F., Jäntti, R.: Asymptotically fair transmission scheduling over fading channels. IEEE Transactions on Wireless Communications **3**, 326–336 (2004)
6. Bonald, T., Borst, S., Hegde, N., Jonckheere, M., Proutiére, A.: Flow-level performance and capacity of wireless networks with user mobility. Queueing Systems **63**, 131–164 (2009)
7. Borst, S.: User-level performance of channel-aware scheduling algorithms in wireless data networks. IEEE/ACM Transactions on Networking **13**, 636–647 (2005)
8. Borst, S.: Flow-level performance and user mobility in wireless data networks. Philosophical Transactions of the Royal Society **366**, 2047–2058 (2008)
9. Borst, S., Jonckheere, M.: Flow-level stability of channel-aware scheduling algorithms. In: Proceedings of WiOpt (2006)
10. Hu, M., Zhang, J., Sadowsky, J.: Traffic aided opportunistic scheduling for wireless networks: algorithms and performance bounds. Computer Networks **46**, 505–518 (2004)
11. Jalali, A., Padovani, R., Pankaj, R.: Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system. In: Proceedings of IEEE VTC 2000-Spring Conference, pp. 1854–1858 (2000)
12. Lassila, P., Aalto, S.: Combining opportunistic and size-based scheduling in wireless systems. In: Proceedings of ACM MSWiM, pp. 323–332 (2008)
13. Lei, L., Lin, C.: Opportunistic scheduler evaluation using discriminatory processor sharing model. In: Proceedings of IEEE ICC, pp. 2926–2930 (2008)
14. Liu, X., Chong, E., Schroff, N.: A framework for opportunistic scheduling in wireless networks. Computer Networks **41**, 451–474 (2003)
15. Pinedo, M.: Scheduling: Theory, Algorithms and Systems, 3rd edn. Springer (2008)

16. Prakash, R., Veeravalli, V.: Centralized wireless data networks with user arrivals and departures. IEEE Transactions on Information Theory **53**, 695–713 (2007)
17. Sadiq, B., de Veciana, G.: Balancing SRPT prioritization vs opportunistic gain in wireless systems with flow dynamics. In: Proceedings of ITC-22 (2010)
18. Schrage, L.: A proof of the optimality of the shortest remaining processing time discipline. Operations Research **16**, 687–690 (1968)
19. Stolyar, A.: On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation. Operations Research **53**, 12–25 (2005)
20. Tsybakov, B.: File transmission over wireless fast fading downlink. IEEE Transactions on Information Theory **48**, 2323–2337 (2002)
21. Viswanath, P., Tse, D., Laroia, R.: Opportunistic beamforming using dumb antennas. IEEE Transactions on Information Theory **48**, 1277–1294 (2002)