A General Per-Flow Service Curve for GPS

Almut Burchard
Department of Mathematics
University of Toronto

Jörg Liebeherr
Department of ECE
University of Toronto

***Abstract*—Generalized Processor Sharing (GPS), which provides the theoretical underpinnings for fair packet scheduling algorithms, has been studied extensively. However, a tight formulation of the available service to a flow only exists for traffic that is regulated by affine arrival envelopes and a constant-rate link. In this paper, we show that the universal service curve by Parekh and Gallager can be extended to concave arrival envelopes and links with time-variable capacity. We also dispense with the previously existing assumption of a stable system.**

I. Introduction

Expressions for the available service at a link with Generalized Processor Sharing scheduling derived in [4], [5] are widely credited for the first formulation of a service curve in the network calculus. The so-called *universal service curve* for a link with rate R and GPS scheduling derived in [4], [5] is given by

$$\mathcal{S}(t) = \max_{M \subseteq \mathcal{N}} \frac{Rt - \sum_{j \in M}(\sigma_j + \rho_j\, t)}{\sum_{j \notin M} \phi_j}\,, \tag{1}$$

where $\mathcal{N}$ is the set of flows with weights $(\phi_j)_{j\in\mathcal{N}}$. The universal service curve yields a (strict) per-flow service curve $\mathcal{S}_i(t) = \phi_i \mathcal{S}(t)$ for a flow $i \in \mathcal{N}$. The service curves are derived under the assumptions that *(i)* the arrival traffic of each flow $j \in \mathcal{N}$ in a time interval of length τ is bounded by an affine arrival envelope $\mathcal{E}_j(\tau) = \sigma_j + \rho_j \tau$, and *(ii)* the system is stable in the sense that the total average arrival rate does not exceed the link capacity ($\sum_{j\in\mathcal{N}} \rho_j \le R$). For general scenarios, where the $\mathcal{E}_j$ are not necessarily affine and the system may be unstable ($\sum_{j\in\mathcal{N}} \rho_j > R$), a pessimistic estimate for the available service can be given by the minimum guaranteed rate $\frac{\phi}{\sum_{j\in\mathcal{N}} \phi_j} R$. This estimate can be somewhat improved by using knowledge of the arrival envelopes [3], [6], [7]. If the envelopes of the arrivals in Eq. (1) are replaced by envelopes for the departures, a generalization to non-affine envelopes is easily achieved. As pointed out in [2, Sec. IV.C], since a departure envelope of a flow can be expressed as a min-plus deconvolution of its arrival envelope and per-flow service curve, this only results in implicit expressions for (minimum) per-flow service curves.

In this paper, we provide the following extensions to the per-flow (strict) service curves obtained from the universal service curve in Eq. (1):

- Arrival envelopes can be arbitrary concave functions;
- The link may have a time-variable capacity;
- The link need not be stable.

These relaxations are achieved by generalizing the concepts of *feasible ordering* in [4] and *feasible partition* in [8]. Note that the extension to time-variable service rates enables the computation of the available service for hierarchical schedulers [1]. We will show that the derived service curve is best-possible.

In Sec. II we state the main result. We provide a brief description of max-min fairness in Section III, and then introduce the key notion of feasible subsets. This notion is used in Sec. IV to derive backlog and output bounds. In Sec. V, we prove the main result, Theorem 1. Sec. VI discusses GPS for general monotone arrival and service processes. We conclude the paper in Sec. VII.

II. Statement of the main result

Let A and D denote the arrival and departure processes for a flow or an aggregate of flows arriving at a service element. (Arrivals and departures for different flows will be distinguished by subscripts). The backlog is denoted by $B(t) = A(t) - D(t)$. The cumulative service process of the element will be described by a function $C(t)$. The arrivals in a half-open interval $[s,t)$ are denoted by $A(s,t) := A(t) - A(s)$, and correspondingly for the departures and the service. We always assume that arrival, departure, and service processes are nondecreasing and left-continuous, with $A(t) = D(t) = C(t) = 0$ for $t \le 0$, and $D(t) \le A(t)$ for all t.

We say that the service element is *workconserving*, if $D(s,t) = C(s,t)$ on every interval that contains no idle period, and $D(s,t) \le C(s,t)$ otherwise. An important example is the constant-rate link, $C(t) = Rt$, which serves traffic at the constant rate R whenever the backlog is positive. In case the service process is given as a time-varying rate $\dot{C}(t)$, then the service element is workconserving if the departure rate satisfies $\dot{D}(t) = \dot{C}(t)$ whenever there is a backlog at t.

Throughout this paper, we consider a finite set $\mathcal{N}$ of flows arriving to a service element. Each flow $j \in \mathcal{N}$ is associated with a positive weight $\phi_j > 0$.

Definition 1. A *Generalized Processor Sharing (GPS)* scheduler is a workconserving scheduling algorithm which ensures that for any $0 \le s < t$ and any flow $i \in \mathcal{N}$ that is backlogged on the entire interval (s,t), the departures satisfy

$$\frac{D_i(s,t)}{\phi_i} \ge \frac{D_j(s,t)}{\phi_j} \quad \text{for all } j \in \mathcal{N}\,. \tag{2}$$

Our main result provides a lower bound on $D_i(s,t)$ in terms of the parameters of the scheduler, the service process, and the traffic arriving to each of the flows $j \in \mathcal{N}$.

To proceed, we need some more notation. An *envelope* for an arrival function A is a nondecreasing function $\mathcal{E}$ such that

$$A(s,t) \leq \mathcal{E}(t-s) \quad \text{for all } 0 \leq s \leq t\,.$$

We also say that the arrivals *comply* to $\mathcal{E}$ and write $A \lesssim \mathcal{E}$. By convention we set $\mathcal{E}(\tau) = 0$ if $\tau \leq 0$. Without loss of generality, envelopes can be taken to be subadditive.

A nondecreasing function $\mathcal{S}$ is a *strict service curve* for a flow at a service element if $D(s,t) \geq \mathcal{S}(t-s)$ whenever the flow is backlogged on the entire interval (s,t). By convention, $\mathcal{S}(\tau) = 0$ for $\tau \leq 0$. Without loss of generality, a strict service curve may be taken to be superadditive and nonnegative.

For the special case of a workconserving service element, a function $\mathcal{C}$ is a strict service curve if

$$C(s,t) \geq \mathcal{C}(t-s) \quad \text{for all } 0 \leq s \leq t\,.$$

We say that the service process *complies* to $\mathcal{C}$ and write $C \gtrsim \mathcal{C}$. In particular, $\mathcal{C}(t) = Rt$ is a strict service curve for the workconserving link with constant rate R.

Theorem 1 (Leftover service curve)**.** *Let $\mathcal{N}$ be a finite set of flows arriving to a GPS scheduler, as in Definition 1. Assume that $C \gtrsim \mathcal{C}$. Fix $i \in \mathcal{N}$. For each $j \in \mathcal{N} \setminus \{i\}$, let $\mathcal{E}_j$ be an envelope with $A_j \lesssim \mathcal{E}_j$. If $\mathcal{C}$ is convex and each $\mathcal{E}_j$ is concave in t, then*

$$\mathcal{S}_i(t) := \max_{M \subseteq \mathcal{N}\setminus\{i\}} \frac{\phi_i}{\sum_{j\notin M}\phi_j}\Big(\mathcal{C}(t) - \sum_{j\in M}\mathcal{E}_j(t)\Big) \tag{3}$$

is the best-possible strict service curve for flow i.

We refer to $\mathcal{S}_i$ as the *leftover service curve* available to flow i under GPS. Note that there are no hypotheses on the arrivals from flow i. If no envelope is available for some flow $j \in \mathcal{N}$, a conservative estimate can be obtained by setting $\mathcal{E}_j(t) = +\infty$ for all $t > 0$.

By construction, $\mathcal{S}_i$ is nonnegative, nondecreasing, and convex in t, with $\mathcal{S}_i(0) = 0$ and $\frac{\phi_i}{\sum_{j\in\mathcal{N}}\phi_j}\mathcal{C} \leq \mathcal{S}_i \leq \mathcal{C}$. We will show that $\mathcal{S}_i(t)$ equals the service that flow i receives in a scenario where it is backlogged on $(0,t)$, the flows $j \neq i$ are greedy ($A_j = \mathcal{E}_j$), and the service element is lazy ($C = \mathcal{C}$), see Lemma 6.

Eq. (3) and the definition of the GPS scheduler are reminiscent of expressions for max-min fairness. In the proof of the theorem we will exploit this connection. The convexity and concavity assumptions will play an important role.

III. Max-min fairness and feasible subsets

Let $\mathcal{N}$ be a collection of players. As in Sec. II, let $(\phi_j)_{j\in\mathcal{N}}$ be positive weights. Each player $j \in \mathcal{N}$ requests a nonnegative share x_j of a resource X. An allocation $(y_j)_{j\in\mathcal{N}}$ with $0 \leq y_j \leq x_j$ for $j \in \mathcal{N}$ is *max-min fair*, if $\sum_{j\in\mathcal{N}} y_j = \min\big\{\sum_{j\in\mathcal{N}} x_j, X\big\}$, and for each $i \in \mathcal{N}$ with $y_i < x_i$

$$\frac{y_i}{\phi_i} \geq \frac{y_j}{\phi_j} \quad \text{for all } j \in \mathcal{N}\,. \tag{4}$$

Here, y_i represents the share allocated to player i. The first condition requires the allocation to be *waste-free*, that is, the entire resource must be used unless the requests of all players are satisfied. Eq. (4) specifies that small requests are satisfied in full while large requests are served in proportion to their weights (ϕ_j). It is known that these conditions uniquely determine the allocation. Explicitly, $y_i = \min\{x_i, \phi_i f\}$ with

$$f := \max_{M\subset\mathcal{N}} \frac{X - \sum_{j\in M} x_j}{\sum_{j\notin M}\phi_j}\,. \tag{5}$$

The value f is called the *fair share* associated with the allocation problem. By convention, for $M = \mathcal{N}$ the fraction takes the value $-\infty$ if the numerator is negative and $+\infty$ otherwise. The maximum is attained by the set of satisfied players,

$$M_{\text{sat}} := \big\{ j \in \mathcal{N} \;\big|\; x_j \leq \phi_j f \big\}\,. \tag{6}$$

Clearly, the fair share is nonnegative and jointly convex in x_j and X. It is nondecreasing in X and nonincreasing in each x_j. Its value is finite if and only if $\sum_{i\in\mathcal{N}} x_i > X$, and it satisfies the lower bound $f \geq \frac{X}{\sum_{j\in\mathcal{N}}\phi_j}$.

Different from Eq. (3), the maximum in Eq. (5) ranges over *all* subsets $M \subset \mathcal{N}$. The two formulas are related as follows.

Lemma 1. *Let $M \subset \mathcal{N}$ be a non-empty subset, and $i \in M$. Then either*

$$\frac{x_i}{\phi_i} \;\leq\; \frac{X - \sum_{j\in M\setminus\{i\}} x_j}{\sum_{j\notin M\setminus\{i\}}\phi_j} \;\leq\; \frac{X - \sum_{j\in M} x_j}{\sum_{j\notin M}\phi_j}$$

or both inequalities are reversed.

Proof. If $M = \mathcal{N}$, then the inequalities hold if and only if $\sum_{j\in\mathcal{N}} x_j \leq X$. Otherwise, set $x_i' := X - \sum_{j\in M} x_j$ and $\phi_i' := \sum_{j\notin M}\phi_j > 0$, and write

$$\begin{aligned}\frac{X - \sum_{j\in M\setminus\{i\}} x_j}{\sum_{j\notin M\setminus\{i\}}\phi_j} &= \frac{x_i + x_i'}{\phi_i + \phi_i'}\\ &= \lambda\frac{x_i}{\phi_i} + (1-\lambda)\frac{x_i'}{\phi_i'}\\ &= \lambda\frac{x_i}{\phi_i} + (1-\lambda)\frac{X - \sum_{j\in M} x_j}{\sum_{j\notin M}\phi_j}\,,\end{aligned}$$

where $\lambda = \frac{\phi_i}{\phi_i + \phi_i'}$ lies strictly between 0 and 1. Therefore either both inequalities hold, or both fail. □

As a consequence of the lemma, the fair allocation to flow i can also be computed by $y_i = \min\{x_i, f_i\}$, where

$$f_i := \max_{M\subset\mathcal{N}\setminus\{i\}} \frac{\phi_i}{\sum_{j\notin M}\phi_j}\Big(X - \sum_{j\in M} x_j\Big)\,. \tag{7}$$

We next consider the impact that a subset of requests can have on a max-min fair allocation.

Definition 2. Let $M \subset \mathcal{N}$, and $X > 0$. A collection of requests $(x_j)_{j\in M}$ is *feasible*, if

$$\max_{j\in M}\frac{x_j}{\phi_j} \leq \frac{X - \sum_{j\in M} x_j}{\sum_{j\notin M}\phi_j}\,. \tag{8}$$

In that case, M is called a *feasible subset* of $\mathcal{N}$ for the data $(\phi_j)_{j\in\mathcal{N}}$, $(x_j)_{j\in M}$, and X.

Feasibility of $(x_j)_{j\in M}$ means that $M_{\rm sat}$, the set of satisfied players from Eq. (6), contains M, regardless of the values in the set $(x_j)_{j\notin M}$. Conversely, for any set of requests $(x_j)_{j\in\mathcal{N}}$, the corresponding subset $M_{\rm sat}$ is feasible. By way of examples, a single request x_i is feasible if $x_i \le \frac{\phi_i}{\sum_{j\in\mathcal{N}}\phi_j}X$. A full set of requests $(x_j)_{j\in\mathcal{N}}$ is feasible if $\sum_{j\in\mathcal{N}} x_j \le X$.

Remark. Feasible subsets are closely related to the notion of feasible orderings introduced in [4, Sec. V.C]. By definition, a *feasible ordering* ("$\prec$") is a total order on $\mathcal{N}$ with the property that

$$\frac{x_k}{\phi_k} < \frac{X-\sum_{j\prec k} x_j}{\sum_{j\succeq k}\phi_j} \quad \text{for all } k\in\mathcal{N}\,.$$

One can verify that for any feasible ordering, the downsets $M_k := \{j \mid j \preceq k\}$ are feasible subsets. Feasible subsets are also downsets for the partial order induced by the feasible partition defined in [8].

The next lemma will be used to construct chains of feasible subsets. In the case where $M=\mathcal{N}$ and $\sum_{j\in\mathcal{N}} x_j < X$, it implies that orderings of $\mathcal{N}$ along which the fraction $\frac{x_j}{\phi_j}$ is nondecreasing are feasible. This recovers Lemma 5 in [4]. We note in passing that there exist other feasible orderings where $\frac{x_j}{\phi_j}$ is not monotone.

Lemma 2. *Let $(x_j)_{j\in M}$ be a feasible subset for a resource $X>0$ and $k\in M$. Then, $M\setminus\{k\}$ is feasible if*

$$\frac{x_k}{\phi_k} = \max_{j\in M}\frac{x_j}{\phi_j}\,.$$

Proof. By the maximality of k and the feasibility of M,

$$\max_{j\in M\setminus\{k\}}\frac{x_j}{\phi_j} \le \frac{x_k}{\phi_k} \le \frac{X-\sum_{j\in M}x_j}{\sum_{j\notin M}\phi_j}\,.$$

It follows from Lemma 1 that

$$\frac{x_k}{\phi_k} \le \frac{X-\sum_{j\in M\setminus\{k\}}x_j}{\phi_k+\sum_{j\notin M}\phi_j}\,.$$

Thus $M\setminus\{k\}$ is feasible. □

Let $(y_j)_{j\in\mathcal{N}}$ be the max-min fair allocation of a resource X resulting from requests $(x_j)_{j\in\mathcal{N}}$. Denote by $\bar y_i := x_i - y_i$ the *unmet demand* of player i. In terms of the fair share from Eq. (5), the unmet demand is given by $\bar y_i = [x_i-\phi_i f]_+$. Here, we have used the notation $[x]_+ = \max\{x,0\}$. The waste-free property of the allocation is equivalent to $\sum_{j\in\mathcal{N}}\bar y_j = \big[\sum_{j\in\mathcal{N}}x_j - X\big]_+$. The unmet demand satisfies the following useful inequalities.

Lemma 3. *Let $(\bar y_j)_{j\in\mathcal{N}}$ be the unmet demands in the max-min fair allocation of a resource X resulting from requests $(x_j)_{j\in\mathcal{N}}$, and let $(\bar y'_j)_{j\in\mathcal{N}}$ be defined accordingly from X' and $(x'_j)_{j\in\mathcal{N}}$. Then*

$$\sum_{j\in\mathcal{N}}|\bar y_j-\bar y'_j| \le \sum_{j\in\mathcal{N}}|x_j-x'_j| + |X-X'|\,. \tag{9}$$

Moreover, we have the monotonicity property

$$\left.\begin{array}{l} x_j\le x'_j \text{ for all } j\in\mathcal{N}\\ X\ge X'\end{array}\right\} \implies \bar y_j \le \bar y'_j \text{ for all } j\in\mathcal{N}\,.$$

Proof. We start with the second claim. Fix $i\in\mathcal{N}$. By definition, $\bar y_i = [x_i - f_i]_+$, and correspondingly for $\bar y'_i$. It is apparent from Eq. (7) that f_i is nondecreasing in X and nonincreasing in the variables x_j for $j\ne i$. This proves monotonicity.

For Eq. (9), let $(x_j)_{j\in\mathcal{N}}$ and $(x'_j)_{j\in\mathcal{N}}$ be as in the statement of the lemma. Denote by $(\bar z_j)_{j\in\mathcal{N}}$ the unmet demand resulting from the requests $\min(\{x_j,x'_j\})_{j\in\mathcal{N}}$ for the resource $\max\{X,X'\}$, and by $(\bar w_j)_{j\in\mathcal{N}}$ be the unmet demand resulting from requests $\max(\{x_j,x'_j\})_{j\in\mathcal{N}}$ for the resource $\min\{X,X'\}$. By monotonicity,

$$\bar z_j \le \bar y_j \le \bar w_j \quad \text{for all } \ j\in\mathcal{N}\,,$$

and likewise for $\bar y'_j$. Therefore

$$\begin{aligned}\sum_{j\in\mathcal{N}}|\bar y_j-\bar y'_j| &\le \sum_{j\in\mathcal{N}}(\bar w_j-\bar z_j)\\ &= \Big[\sum_{j\in\mathcal{N}}\max\{x_j,x'_j\}-\min\{X,X'\}\Big]_+\\ &\qquad -\Big[\sum_{j\in\mathcal{N}}\min\{x_j,x'_j\}-\max\{X,X'\}\Big]_+\\ &\le \sum_{j\in\mathcal{N}}|x_j-x'_j| + |X-X'|\,,\end{aligned}$$

where the second step used the waste-free property. □

The lemma implies that the max-min fair allocation for a fixed value of X, viewed as a mapping $(x_j)_{j\in\mathcal{N}}\mapsto(\bar y_j)_{j\in\mathcal{N}}$, contracts the ℓ^1-distance and preserves the natural order.

IV. Performance bounds

Next we apply the results of Sec. III to the allocation of service at a GPS scheduler. Henceforth, x_j expresses the fraction of the service process C allocated to flow j. The following theorem says that the aggregate cumulative departures from a feasible subset $(x_j)_{j\in M}$ are at least as large as if each flow $j\in M$ were allocated a dedicated link with service process x_jC. Note that no assumption is made on busy periods.

Theorem 2 (Departures)**.** *Let $(A_j(t))_{j\in\mathcal{N}}$ be arrivals from a set of flows to a GPS scheduler with service process $C(t)$. Fix $M\subset\mathcal{N}$, and let $(x_j)_{j\in M}$ be a feasible subset of requests for a resource $X=1$. Then for all $t\ge 0$,*

$$\sum_{j\in M}D_j(t) \ge \sum_{j\in M}\inf_{s\le t}\big\{A_j(s)+x_jC(s,t)\big\}\,. \tag{10}$$

Proof. We proceed by induction on the number of elements of M. When $M=\emptyset$, there is nothing to show.

For the inductive step, let $M\subset\mathcal{N}$ be a non-empty feasible subset, and suppose the claim has already been established for its proper feasible subsets. Choose $k\in M$ to maximize the ratio $\frac{x_j}{\phi_j}$. By Lemma 2, $M \setminus \{k\}$ is feasible. The inductive hypothesis implies that, for all $t \geq 0$,

$$\sum_{j\in M\setminus\{k\}} D_j(t) \geq \sum_{j\in M\setminus\{k\}} \inf_{r\leq t}\big\{A_j(r) + x_j C(r,t)\big\}\,. \quad (11)$$

Fix $t > 0$ and let s be the start of the busy period for flow k that contains t. If $D_k(s,t) \geq x_k C(s,t)$, then

$$D_k(t) \geq A_k(s) + x_k C(s,t)\,,$$

since $D_k(s) = A_k(s)$. Eq. (10) follows by adding Eq. (11).

Otherwise, since flow k is backlogged on (s,t),

$$\frac{D_j(s,t)}{\phi_j} \leq \frac{D_k(s,t)}{\phi_k} < \frac{x_k}{\phi_k} C(s,t)$$

for all $j \in \mathcal{N}$ by Eq. (2). Therefore

$$\begin{aligned}\sum_{j\notin M} D_j(s,t) &< \Big(\sum_{j\notin M}\phi_j\Big)\frac{x_k}{\phi_k}C(s,t)\\ &\leq \Big(1 - \sum_{j\in M} x_j\Big)C(s,t)\,,\end{aligned}$$

where the second inequality is by the feasibility of $(x_j)_{j\in M}$. Since the scheduler is workconserving, it follows that

$$\sum_{j\in M} D_j(s,t) > \sum_{j\in M} x_j C(s,t)\,,$$

and therefore

$$\sum_{j\in M} D_j(t) > \sum_{j\in M}\big(D_j(s) + x_j C(s,t)\big)\,.$$

Clearly, $D_k(s) = A_k(s)$ by the choice of s. For the flows $j \neq k$, we use Eq. (11) at time s to obtain

$$\sum_{j\in M\setminus\{k\}}\big(D_j(s) + x_j C(s,t)\big) \geq \sum_{j\in M\setminus\{k\}} \inf_{r\leq s}\big\{A_j(r) + x_j C(r,t)\big\}\,.$$

Eq. (10) follows once we add the term for $j = k$ and extend the range of the infima to $r \leq t$. This completes the induction. ☐

In the case where $M = \{i\}$, Theorem 2 yields

$$D_i(t) \geq \inf_{s\leq t}\Big\{A_i(s) + \frac{\phi_i}{\sum_{j\in\mathcal{N}}\phi_j} C(s,t)\Big\}\,.$$

More generally, the theorem implies the following key estimates.

Corollary 1 (Backlog). *Define*

$$B_j^*(t) := \sup_{r\leq t}\{A_j(r,t) - x_j C(r,t)\}$$

for $j \in M$. Under the assumptions of Theorem 2,

$$\sum_{j\in M} B_j(t) \leq \sum_{j\in M} B_j^*(t)\,, \qquad t \geq 0\,. \quad (12)$$

Proof. Write $B_j(t) = A_j(t) - D_j(t)$ and apply Eq. (10). ☐

For later use, we note that if $A_j(t) \lesssim \sigma_j + \rho_j t$ and $C(t) \gtrsim R(t-L)$ with $\rho_j \leq x_j R$, then

$$B_j^*(t) \leq \sigma_j + \rho_j L\,, \qquad t \geq 0\,. \quad (13)$$

Corollary 1 implies Theorem 4 in [4] as follows. The assumption in [4] is that the arrivals comply to token-bucket envelopes, $A_j \lesssim \sigma_j + \rho_j t$, that the link offers a constant-rate service $C \gtrsim Rt$, and that the stability condition $\sum_{j\in\mathcal{N}}\rho_j < R$ holds. If we choose $x_j = \frac{\rho_j}{R}$, then $\sigma_j - \sigma_j^t$ equals $B_j^*(t) - B_j(t)$, where σ_j^t is defined in [4] as the sum of the filling level of the token bucket and the backlog at time t. Further, in [4] the set M is assumed to be a downset for a feasible ordering of $\mathcal{N}$. Under these assumptions, Eq. (12) reduces to the central conclusion in [4] that $\sum_{j\in M}\sigma_j^t \leq \sum_{j\in M}\sigma_j$.

Corollary 2 (Output burstiness). *Under the assumptions of Theorem 2,*

$$\sum_{j\in M} D_j(s,t) \leq \sum_{j\in M}\big(B_j^*(t) + x_j C(s,t)\big)\,, \quad 0 \leq s \leq t\,.$$

Proof. By Theorem 2,

$$\begin{aligned}\sum_{j\in M} D_j(s,t) &\leq \sum_{j\in M}\big(A_j(t) - D_j(s)\big)\\ &\leq \sum_{j\in M}\sup_{r\leq s}\big\{A_j(r,t) - x_j C(r,s)\big\}\\ &\leq \sum_{j\in M}\big(B_j^*(t) + x_j C(s,t)\big)\,.\end{aligned}$$

In the last step, we have extended the range of the supremum to $r \leq t$ and applied the definition of $B_j^*(t)$. ☐

V. The leftover service curve

Consider the definition of the leftover service curve $\mathcal{S}_i$ in Eq. (3). It follows from Lemma 1 that

$$\min\big\{\mathcal{E}_i(t), \mathcal{S}_i(t)\big\} = \min\big\{\mathcal{E}_i(t), \phi_i\mathcal{S}(t)\big\}\,,$$

where

$$\mathcal{S}(t) := \max_{M\subset\mathcal{N}} \frac{\mathcal{C}(t) - \sum_{j\in M}\mathcal{E}_j(t)}{\sum_{j\notin M}\phi_j}\,. \quad (14)$$

Note the structural similarities of Eq. (3) to Eq. (7), and of Eq. (14) to Eq. (5). In the special case where the envelopes $\mathcal{E}_j$ are affine, $\mathcal{S}$ agrees with the universal service curve in Eq. (1). The maximum in Eq. (14) is attained by

$$M^* := \{j \in \mathcal{N} \mid \mathcal{E}_j(t) \leq \mathcal{S}_j(t)\}\,, \quad (15)$$

see Eq. (6).

Lemma 4. *Let $\mathcal{N}$, $\mathcal{C}$, i, and $\mathcal{E}_j$ be as in Theorem 1. Given $\tau > 0$, define M^* by Eq. (15) with $t = \tau$ and $\mathcal{E}_i = +\infty$. Then*

$$x_j := \frac{\dot{\mathcal{E}}_j(\tau_-)}{\dot{\mathcal{C}}(\tau_-)}\,, \quad j \in M^*$$

defines a feasible subset for the resource $X = 1$.

Here, we used the notation $f(x_-) = \sup_{y<x} f(y)$.

Proof. By Eqs. (14) and (15), the subset of requests $x_j' := \frac{\mathcal{E}_j(\tau)}{\mathcal{C}(\tau)}$, $j \in M^*$ is feasible for $X = 1$. Since $\mathcal{E}_j(\tau) \geq \tau\dot{\mathcal{E}}_j(\tau_-)$ by concavity and $\mathcal{C}(\tau) \leq \tau\dot{\mathcal{C}}(\tau_-)$ by convexity, we have $x_j' \geq x_j$ for all $j \in M^*$. Thus, $(x_j)_{j\in M^*}$ is a feasible subset. ☐

We next consider the special case of token-bucket envelopes and latency-rate service curves. (The general proof follows immediately afterwards.)

Lemma 5. *Under the hypotheses of Theorem 1, suppose additionally that the service curve has the form $\mathcal{C}(t) = R(t-L)$, and the envelopes are given by $\mathcal{E}_j(t) = \sigma_j + \rho_j t$ for $j \in \mathcal{N} \setminus \{i\}$. Then Eq.* (3) *defines a strict service curve for flow i.*

Proof. Suppose that flow i is backlogged on some interval (s,t). We need to show that $D_i(s,t) \geq S_i(t-s)$.

Set $\tau = t - s$. Let M^* be as in Eq. (15) with τ in place of t and $\mathcal{E}_i = +\infty$, and set $x_j = \frac{\rho_j}{R}$ for $j \in M^*$. By Lemma 4, the subset of requests $(x_j)_{j\in M^*}$ is feasible for $X = 1$. By Corollary 2,

$$\sum_{j\in M^*} D_j(s,t) \leq \sum_{j\in M^*} \left\{B_j^*(t) + x_j C(s,t)\right\} .$$

Since the scheduler is workconserving, it follows that

$$\begin{aligned}\sum_{j\notin M^*} D_j(s,t) &\geq \Big(1 - \sum_{j\in M^*} x_j\Big) C(s,t) - \sum_{j\in M^*} B_j^*(t)\\ &\geq \Big(R - \sum_{j\in M^*} \rho_j\Big)(t-s-L) - \sum_{j\in M^*} \{\sigma_j + \rho_j L\}\\ &= \mathcal{C}(t-s) - \sum_{j\in M^*} \mathcal{E}_j(t-s) .\end{aligned}$$

In the first line, the coefficient of $C(s,t)$ is nonnegative by the feasibility of $(x_j)_{j\in M^*}$. In the second line, we have used that $C(t) \gtrsim R(t-L)$ and applied Eq. (13). In the last line, we have canceled the terms $\rho_j L$ and inserted the envelopes and service curves. By Eq. (2),

$$\begin{aligned}D_i(s,t) &\geq \frac{\phi_i}{\sum_{j\notin M^*}\phi_j} \sum_{j\notin M^*} D_j(s,t)\\ &\geq \frac{\phi_i}{\sum_{j\notin M^*}\phi_j}\Big(\mathcal{C}(t-s) - \sum_{j\in M^*}\mathcal{E}_j(t-s)\Big)\\ &= \mathcal{S}_i(t-s) .\end{aligned}$$

The final step used the maximality of M^* in Eq. (3). □

We are ready to tackle the main result.

Proof of Theorem 1. Given $0 \leq s < t$, set $\tau = t - s$, and fix $i \in \mathcal{N}$. For $j \in \mathcal{N} \setminus \{i\}$, consider the tangent line to the graph of $\mathcal{E}_j$ at τ, defined by $\mathcal{E}'_j(u) = \sigma_j + \rho_j u$ with

$$\rho_j := \dot{\mathcal{E}}_j(\tau_-) , \qquad \sigma_j := \mathcal{E}_j(\tau) - \rho_j \tau \geq 0 .$$

Since $\mathcal{E}_j \leq \mathcal{E}'_j$ by concavity, the arrival process A_j complies to the token-bucket envelope $\mathcal{E}'_j$. Also consider the tangent line to $\mathcal{C}$ at τ, defined by $\mathcal{C}'(u) = R(u-L)$ with

$$R := \dot{\mathcal{C}}(\tau_-) , \quad L := \tau - \tfrac{\mathcal{C}(\tau)}{R} \geq 0 .$$

Since $\mathcal{C} \geq \mathcal{C}'$ by convexity, the service process C complies to the latency-rate service curve $\mathcal{C}'$. By Lemma 5,

$$\mathcal{S}'_i := \max_{M\subseteq\mathcal{N}\setminus\{i\}} \frac{\phi_i}{\sum_{j\notin M}\phi_j}\Big(\mathcal{C}' - \sum_{j\in M}\mathcal{E}'_j\Big)$$

is a strict service curve for flow i. In particular, if flow i is backlogged on (s,t) then

$$D_i(s,t) \geq \mathcal{S}'_i(t-s) = \mathcal{S}_i(t-s) ,$$

where the equality is by the choice of $\tau = t-s$. We conclude that $\mathcal{S}_i$ is a strict service curve. By Lemma 6 below, there are scenarios where the departures saturate the service curve. Therefore $\mathcal{S}_i$ is best possible. □

Lemma 6 (The greedy/lazy scenario)**.** *In the setup of Theorem 1, let the service process be $C(t) = \mathcal{C}(t)$, and the arrival processes $A_j(t) = \mathcal{E}_j(t)$ for $j \in \mathcal{N}$ and $t \geq 0$. Then*

$$D_j(t) = \min\big\{\mathcal{E}_j(t), \mathcal{S}_j(t)\big\} , \quad j \in \mathcal{N} . \tag{16}$$

Proof. Let $t > 0$ be given. Since the scheduler is workconserving, the aggregate departures satisfy

$$\sum_{j\in\mathcal{N}} D_j(t) = \inf_{0\leq s\leq t}\Big\{\sum_{j\in\mathcal{N}} A_j(s) + C(s,t)\Big\} .$$

Inserting the assumptions on the arrival and service processes, we obtain

$$\begin{aligned}\sum_{j\in\mathcal{N}} D_j(t) &= \inf_{0\leq s\leq t}\Big\{\sum_{j\in\mathcal{N}}\mathcal{E}_j(s) + \mathcal{C}(t) - \mathcal{C}(s)\Big\}\\ &= \min\Big\{\sum_{j\in\mathcal{N}}\mathcal{E}_j(t), \mathcal{C}(t)\Big\}\\ &= \sum_{j\in\mathcal{N}} \min\big\{\mathcal{E}_j(t), \mathcal{S}_j(t)\big\} .\end{aligned} \tag{17}$$

The second step follows since the minimum is attained at $s = 0$ or $s = t$ by concavity. In the last step, we have used that $y_j = \min\{\mathcal{E}_j(t), \mathcal{S}_j(t)\}$ is a max-min fair allocation of the resource $X = \mathcal{C}(t)$, and therefore waste-free.

On the other hand, since $\mathcal{S}_j$ is a service curve for flow j,

$$\begin{aligned}D_j(t) &\geq \inf_{0\leq s\leq t}\big\{\mathcal{E}_j(s) + \mathcal{S}_j(t) - \mathcal{S}_j(s)\big\}\\ &= \min\big\{\mathcal{E}_j(t), \mathcal{S}_j(t)\big\} .\end{aligned}$$

Since this holds for every $j \in \mathcal{N}$, by Eq. (17) it holds with equality. □

Lemma 6 demonstrates that the departures from a GPS scheduler in the greedy scenario necessarily satisfy Eq. (16). For completeness of the argument, we show that these departures actually conform to Definition 1. The workconserving property follows from the waste-free property of the max-min fair allocation. It remains to verify Eq. (2) on an arbitrary interval where flow i is backlogged.

Eq. (16) yields $B_j(t) = \big[\mathcal{E}_j(t) - \mathcal{S}_j(t)\big]_+$. By concavity, the ratio $\frac{B_j(t)}{t}$ is nonincreasing in t. Therefore, if flow i is backlogged at time t, then it is backlogged for all $0 < s \leq t$. By Eq. (4), $\frac{D_i(t)}{\phi_i} \geq \frac{D_j(t)}{\phi_j}$, with equality if flow j is backlogged as well. If flow j is backlogged at time s, then $\frac{D_i(s)}{\phi_i} = \frac{D_j(s)}{\phi_j}$, and Eq. (2) follows. Otherwise, flow j is not backlogged at time s, and $D_j(s) = \mathcal{E}_j(s)$. The difference $\frac{D_i(s,t)}{\phi_i} - \frac{D_j(s,t)}{\phi_j} = \frac{\mathcal{S}_i(t)-\mathcal{S}_i(s)}{\phi_i} - \frac{\mathcal{E}_j(t)-\mathcal{E}_j(s)}{\phi_j}$ is concave in s, and nonnegative at $s = 0, t$. Therefore it is nonnegative for every $0 \leq s \leq t$, proving Eq. (2) also in this case.

VI. The Backlog Process

We briefly address the question how to describe the departures from a GPS scheduler with a general nondecreasing service process $C(t)$ and nondecreasing arrival processes $A_j(t)$, $j \in \mathcal{N}$. We will argue that the workconserving property together with Eq. (2) completely determines the backlog process, and hence the departures.

Consider once more the relation between the GPS scheduler and max-min fairness, as evidenced by Eq. (2) and Eq. (4). The departures $D_j(s,t)$ over a time interval $[s,t)$ define an allocation of the resource $X = C(s,t)$ among a set of flows $j \in \mathcal{N}$, each of which requests a share $x_j = B_j(s) + A_j(s,t)$. The backlog $B_j(t)$ plays the role of the unmet demand.

On any interval where the arrival processes $A_j(t)$ are concave and $C(t)$ is convex, the departures are given by the max-min fair allocation

$$D_j(s,t) = \min\bigl\{B_j(s) + A_j(s,t), \phi_j f\bigr\}\,, \quad j \in \mathcal{N}\,,$$

where f is defined by Eq. (5) with $x_j = B_j(s) + A_j(s,t)$ and $X = C(s,t)$. This follows by applying Lemma 6 to the time-shifted processes $A'_j(\tau) = B_j(s) + A_j(s, s+\tau)$ and $C'(\tau) = C(s, s+\tau)$, and then setting $\tau = t - s$. The backlog satisfies the difference equation

$$B_j(t) = \bigl[B_j(s) + A_j(s,t) - \phi_j f\bigr]_+\,, \quad j \in \mathcal{N}\,. \qquad (18)$$

However, Eq. (18) cannot hold for general arrival and service processes on arbitrary intervals. Flows that are backlogged at time t but are idle at an earlier time $s < t$ receive less service than indicated by Eq. (18). The underlying reason is that Eq. (2) provides no explicit service guarantees for such flows.

Since Eq. (18) is valid when s is so close to t that the set of backlogged flows remains constant from s to t, taking the limit $s \to t$ yields the differential equation

$$\dot{B}_i(t) = \dot{A}_i(t) - \frac{\phi_i}{\sum_{j \notin M(t)} \phi_j}\Bigl(\dot{C}(t) - \sum_{j \in M(t)} \dot{A}_j(t)\Bigr)\,, \quad (19)$$

so long as $B_i(t) > 0$. Here, $M(t) = \{j \in \mathcal{N} \mid B_j(t) = 0\}$ is the set of flows that are not backlogged at time t. The differential equation holds at every time t where the arrival and service processes are differentiable, except at instants where $M(t)$ changes. (If the arrival and service processes are not absolutely continuous, the differential equation should be supplemented by equations that account for their jumps and singular continuous components.)

Eq. (19) determines the backlog process on intervals where $M(t)$ is constant. These intervals in turn depend on the departures, rendering the differential equation nonlinear. Standard theorems that guarantee the existence and uniqueness of solutions for nonlinear differential equations do not apply, because the right hand side of Eq. (19) does not have the requisite continuity properties.

We construct the backlog process as follows. Given arrival and service processes $A_j(t)$ and $C(t)$, we approximate them with piecewise linear nondecreasing functions. Specifically, we consider the class of functions that are linear on intervals $(t_\ell, t_{\ell+1}]$, where the breakpoints t_ℓ form an increasing sequence with $t_0 = 0$ and $\lim t_\ell = +\infty$. Jumps are permitted at each t_ℓ. Since linear functions are simultaneously convex and concave, Lemma 6 implies that the backlog process for the approximating scenario satisfies Eq. (18) on each interval $(t_\ell, t_{\ell+1}]$. Then $B_j(t)$ and $D_j(t)$ lie again in the piecewise linear class, with at most $|\mathcal{N}|$ additional breakpoints appearing between t_ℓ and $t_{\ell+1}$ at instants where some flow ceases to be backlogged. By Lemma 3, all errors can be bounded explicitly in terms of the original discretization error. Consequently, the backlog process does not depend on the precise approximation scheme that was used in its construction.

Thanks to Lemma 3, the backlog evolves by an order-preserving family of contractions. One implication is that the backlog process at a GPS scheduler with random stationary arrival and service processes that is started with empty queues is stochastically increasing. As $t \to \infty$, the flows separate into two groups, one consisting of underloaded flows whose backlog process approaches a steady state, and the other of overloaded flows whose backlog becomes unbounded.

VII. Conclusions

We have addressed a longstanding open problem in the theory of fair queueing algorithms, and extended the strict service curve formulation for GPS schedulers by Parekh and Gallager to concave arrival envelopes and links with time-variable capacity. We show that the service curves hold under any load condition and are not limited to stable systems. With this paper, the leftover service curve formulation for GPS has a comparable degree of generality as existing leftover formulations of other classical scheduling algorithms, such as Static Priority, FIFO, and Earliest-Deadline-First.

References

[1] J. C. Bennett and H. Zhang. Hierarchical packet fair queueing algorithms. *IEEE/ACM Transactions on Networking*, 5(5):675–689, October 1997.

[2] M. Fidler. Survey of deterministic and stochastic service curve models in the network calculus. *IEEE Communications Surveys & Tutorials*, 12(1):59–86, 2010.

[3] C. Li, A. Burchard, and J. Liebeherr. A network calculus with effective bandwidth. *IEEE/ACM Transactions on Networking*, 15(6): 1442–1453, December 2007.

[4] A. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM Transactions on Networking*, 1(3):344–357, June 1993.

[5] A. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: the multiple node case. *IEEE/ACM Transactions on Networking*, 2(2):137–150, April 1994.

[6] J. Y. Qiu and E. W. Knightly. Inter-class resource sharing using statistical service envelopes. In *Proc. IEEE Infocom*, pages 1404–1411, March 1999.

[7] J. Y. Qiu, C. Cetinkaya, C. Li, and E. W. Knightly. Inter-class resource sharing using statistical service envelopes. Technical Report, 2000, *http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.37.9839&rep=rep1&type=pdf*.

[8] Z. L. Zhang, Z. Liu, and D. Towsley. Closed-form deterministic end-to-end performance bounds for the generalized processor sharing scheduling discipline. *Journal of Combinatorial Optimization*, 1(4):457–481, 1998.