# Understanding XCP: Equilibrium and Fairness

Steven H. Low*, Lachlan L. H. Andrew†, Bartek P. Wydrowski*
*CS and EE, California Institute of Technology
†ARC Special Research Centre for Ultra-Broadband Information Networks, University of Melbourne

6 May, 2004

## Abstract

We prove that the XCP equilibrium solves a constrained max-min fairness problem by identifying it with the unique solution of a hierarchy of optimization problems. This is the same set of problems solved by the standard max-min fair allocation, but XCP solves them under an additional constraint. We describe an algorithm to compute this equilibrium and derive a lower and upper bound on link utilization. While XCP reduces to max-min allocation at a single link, its behavior in a network can be very different. We illustrate that the additional constraint can cause flows to receive an arbitrarily small fraction of their max-min fair allocations. We confirm these results using ns2 simulations.

## 1 Introduction

TCP congestion control [1] has prevented severe congestion while the Internet underwent explosive growth during the last decade. However, the algorithm has shown serious difficulties as the network continues to scale in size and capacity [2, 3]. This has motivated several recent enhancements [4–8]. (See [6] for extensive references.) XCP [8] is distinctive in that it requires explicit communication between the traffic sources and the network. Moreover, unlike the other proposals which set the flow rates according to the *sum* of congestion measures at the links of their paths, XCP sets them according to the *minimum* "available capacity" in their paths. This has the same flavor as MaxNet [9, 10] which sets flow rates according to the *maximum* of congestion measures in their paths. In this paper, we reverse engineer XCP to understand its equilibrium properties.

In Section 2, we present a deterministic fluid model of a general XCP network with multiple links and multiple flows, and summarize our notation. In Section 3, we analyze the equilibrium rates of XCP. We show that the network queues are empty in equilibrium, agreeing with the simulation results of [8]. We prove the existence and uniqueness of XCP equilibrium rates by identifying them with the unique solution to a hierarchy of optimization problems. This is the same set of problems solved by the standard max-min fair allocation, but XCP solves them under an additional constraint. While XCP reduces to max-min allocation at a single link, its behavior in a network can be very different. We describe an algorithm to compute this equilibrium and derive upper and lower bounds on link utilization.

In Section 4, we use these bounds to investigate the impact of the choice of protocol parameters on link utilization under the additional constraint. We show that flows can receive an arbitrarily small fraction of their max-min fair allocations. Specifically, with a max-min fair allocation, as long as a link is a bottleneck for some (not necessarily all) flows that pass through it, it will be fully utilized. Under XCP, this is no longer true: when the majority of flows using a link are bottlenecked at other links, the remaining flows may not make efficient use of the residual bandwidth. With the values suggested in [8] however link utilization is at least 80% at any link. XCP has a "shuffling parameter" $\gamma > 0$ to prevent the network from settling into an unfair state [8]. We show that, given any network topology, we can choose $\gamma$ sufficiently small so that the resulting allocation is close to max-min fairness. For any fixed $\gamma > 0$, however, there are topologies in which some flow rates can be far away from their max-min allocations.

These properties and the accuracy of our algorithm are verified by ns2 simulations in Section 5. We conclude in Section 6 with limitations of this work.

## 2 Model

Consider a network with $L$ links shared by $N$ flows. Sources are indexed by $i = 1, \ldots, N$, links by $l = 1, \ldots, L$ and packets by $k$. Let $R$ be the $L \times N$ routing matrix: $R_{li} = 1$ if flow $i$ uses link $l$ and 0 otherwise. Let $L(i)$ be the set of links in the path of flow $i$:

$$L(i) \quad := \quad \{\, l \mid R_{li} = 1 \,\}$$

and $I(l)$ be the set of flows that use link $l$:

$$I(l) \quad := \quad \{\, i \mid R_{li} = 1 \,\}$$

Note that $l \in L(i) \Leftrightarrow i \in I(l)$.

We will present a continuous-time fluid model of XCP. For flows $i$, define the following variables:

- $w_i(t)$: window size at time $t$, in packets.
- $\tau_i$: round-trip propagation (and fixed processing) delay.
- $T_i(t)$: round-trip time (RTT) at time $t$.
- $x_i(t) := w_i(t)/T_i(t)$: flow rate at time $t$.

For links $l$, define the following variables:

- $c_l$: capacity, in packets/sec.
- $b_l(t)$: backlog at time $t$, in packets.
- $y_l(t) := \sum_i R_{li}x_i(t)$: aggregate input rate at link $l$ at time $t$. In equilibrium, we sometimes write $y_l(x)$ to emphasize the dependence on equilibrium rates $x$.

XCP divides time into control intervals of duration $d$, which is also used as a time scaling parameter. The original paper [8] sets $d$ to the mean RTT of all the flows at a link, making $d$ time-varying and potentially different at different links. For simplicity, we assume $d$ to be a global constant in our model.

To simplify notation, we assume all packets have the same size of 1 unit. We use "flow" and "source" interchangeably.

## 2.1 XCP description

In this subsection, we summarize the XCP algorithm. See [8] for the design rationale and a detailed description. We ignore feedback delay in our model because we are interested in equilibrium properties in this paper.

For each packet, XCP generates a feedback signal prescribing a change in window size. Let $\tilde{H}_{lk}(t)$ be the feedback generated by link $l$ for packet $k$ at time $t$. The acknowledgment for packet $k$ received by its source contains in its header the smallest feedback $\min_l H_{lk}(t)$ generated by links along its path. The source adds this quantity to its current window size.[1] We now describe how to compute the feedback.

Let

$$\phi_l(t) \quad = \quad \alpha d(c_l - y_l(t)) - \beta b_l(t)$$

where $\alpha$, $\beta > 0$ are constants, $c_l$ is the link capacity, $y_l(t)$ is the aggregate input rate, and $b_l(t)$ is the backlog at time $t$. Let $\phi_l^+(t) = \max(\phi_l(t), 0)$ and $\phi_l^-(t) = \max(-\phi_l(t), 0)$. The feedback on the $k$th packet at link $l$ is

$$\tilde{H}_{lk}(t) \quad = \quad \tilde{p}_{lk}(t) - \tilde{n}_{lk}(t)$$

where $\tilde{p}_{lk}(t)$ and $\tilde{n}_{lk}(t)$ are the increase and decrease components respectively:

$$\tilde{p}_{lk}(t) \quad = \quad (h_l(t) + \phi_l^+(t))\frac{\tilde{T}_k(t)}{d}\frac{\tilde{T}_k(t)/\tilde{w}_k(t)}{\sum_{j=1}^{K_l(t)}\tilde{T}_j(t)/\tilde{w}_j(t)} \qquad (1)$$

$$\tilde{n}_{lk}(t) \quad = \quad (h_l(t) + \phi_l^-(t))\frac{\tilde{T}_k(t)}{d}\frac{1}{K_l(t)} \qquad (2)$$

[1] In practice, the window size has a lower bound of 1 packet, but for notational simplicity, we ignore this.

where $\tilde{T}_k(t)$ and $\tilde{w}_k(t)$ are the round-trip time and window size, respectively, of the flow which transmitted packet $k$, and $K_l(t)$ is the total number of packets seen by link $l$ over the time interval $(t-d, t]$. Here

$$h_l(t) \quad = \quad \max(0, \gamma d y_l(t) - |\phi_l(t)|)$$

is a "traffic shuffling" term with $\gamma \geq 0$ a constant. (Note that we are using the definition of $\gamma$ from the appendix of [8], which differs by a factor of $d$ from that used in the corresponding equation in [8].)

## 2.2 Dynamic model

We now translate the per-packet feedback $\tilde{H}_{lk}(t)$ into per-flow feedback. Let $H_{li}(t)$ be the feedback generated by link $l$ for flow $i$ at time $t$. In general, a quantity with a tilde ( $\tilde{}$ ) pertains to a packet while the corresponding variable without a tilde pertains to a flow.

Substituting $\tilde{x}_k(t) = \tilde{w}_k(t)/\tilde{T}_k(t)$ in (1) gives

$$\tilde{p}_{lk}(t) \quad = \quad \frac{\tilde{T}_k(t)}{\tilde{x}_k(t)}\frac{h_l(t) + \phi_l^+(t)}{d\sum_{i=j}^{K_l(t)} 1/\tilde{x}_j(t)}. \qquad (3)$$

$K_l(t)$ is the total number of packets arriving at link $l$ in period $(t-d, t]$. For simplicity, we *assume* that

$$K_l(t) \quad = \quad y_l(t)d \;=\; d\sum_i R_{li}x_i(t)$$

Of these packets, we *assume* that $R_{li}x_i(t)d$ packets are from flow $i$. Hence

$$\sum_{j=1}^{K_l(t)} \frac{1}{\tilde{x}_j(t)} \quad = \quad \sum_{i=1}^{N} R_{li}x_i(t)d \cdot \frac{1}{x_i(t)} \quad = \quad N_l d$$

Thus the per-packet feedback (3) becomes per-flow feedback

$$p_{li}(t) \quad = \quad \frac{T_i(t)}{d^2}\frac{h_l(t) + \phi_l^+(t)}{N_l x_i(t)}$$

Using $K_l(t) = y_l(t)d$ again, the per-packet feedback (2) becomes

$$n_{li}(t) \quad = \quad \frac{T_i(t)}{d^2}\frac{h_l(t) + \phi_l^-(t)}{y_l(t)}$$

The feedback *per packet* to flow $i$ from link $l$ is then

$$H_{li}(t) \quad = \quad \frac{T_i(t)}{d^2}\left(\frac{h_l(t) + \phi_l^+(t)}{N_l x_i(t)} - \frac{h_l(t) + \phi_l^-(t)}{y_l(t)}\right)$$

If flow $i$ does not use link $l$, then set $H_{li}(t) = \infty$.

Let $H_i(t) = \min_{l \in L(i)} H_{li}(t)$ be the minimum feedback along $i$'s path. Since source $i$ receives $x_i(t)$ feedback packets per unit time (assuming every packet carries control information and is acknowledged), its window evolves according to:

$$\dot{w}_i(t) \quad = \quad x_i(t) \cdot H_i(t)$$

Substituting $x_i(t) = w_i(t)/T_i(t)$, we have

$$\dot{w}_i(t) = \frac{w_i(t)}{d^2} \min_{l \in L(i)} \left( \frac{h_l(t) + \phi_l^+(t)}{N_l x_i(t)} - \frac{h_l(t) + \phi_l^-(t)}{y_l(t)} \right)$$

**Remark:** The pseudo code in [8] and the NS-2 implementation contain "residual" terms not described in the text of [8]. These use the feedback from upstream links to modulate the positive and negative components $\tilde{p}_{lk}(t)$ and $\tilde{n}_{lk}(t)$ to prevent excessive positive or negative feedback in each control period. However, it can be proved (see a forthcoming paper) that the modulation of the positive component has no effect on the XCP equilibrium at all. The modulation of the negative component also has no effect on the equilibrium if the average rate of flows bottlenecked at upstream links is significant (at least half that of flows bottlenecked at link $l$ itself). Otherwise, the negative feedback is limited and the resulting link utilization is slightly increased (by around 4% in Scenario 1 of Section 5). Since these residual terms seem to impact primarily on dynamic rather than equilibrium properties, for simplicity, we ignore them in this paper.

In summary, an XCP network is described by the following set of equations:

$$\dot{w}_i(t) = \frac{w_i(t)}{d^2} \min_{l \in L(i)} F_{li}(t) \tag{4a}$$

$$\dot{b}_l(t) = \begin{cases} y_l(t) - c_l & \text{if } b_l(t) > 0 \\ \max(y_l(t) - c_l, 0) & \text{if } b_l(t) = 0 \end{cases} \tag{4b}$$

where

$$F_{li}(t) = \frac{h_l(t) + \phi_l^+(t)}{N_l x_i(t)} - \frac{h_l(t) + \phi_l^-(t)}{y_l(t)} \tag{5a}$$

$$\phi_l(t) = \alpha d(c_l - y_l(t)) - \beta b_l(t) \tag{5b}$$

$$h_l(t) = \max(\gamma d y_l(t) - |\phi_l(t)|, 0) \tag{5c}$$

$$x_i(t) = \frac{w_i(t)}{T_i(t)} \tag{5d}$$

$$y_l(t) = \sum_i R_{li} x_i(t) \tag{5e}$$

$$T_i(t) = \tau_i + \sum_l R_{li} \frac{b_l(t)}{c_l} \tag{5f}$$

Here, $\alpha > 0$, $\beta \geq 0$, $\gamma \geq 0$ are constants, and $\phi_l^+(t) = \max(\phi_l(t), 0)$, $\phi_l^-(t) = \max(-\phi_l(t), 0)$. Standard XCP uses $\alpha = 0.4$, $\beta = 0.226$ and $\gamma = 0.1$. We will study the behavior of the general model, which includes this as a special case. As we will see below, the qualitative properties, such as existence and uniqueness of equilibrium rates and their fairness properties, do not depend on specific values of these parameters (as long as $\gamma > 0$).

# 3 Equilibrium rates

This section characterizes the equilibrium of XCP and describes an algorithm to compute it; the next considers the implications of these results on utilization and fairness.

Equations (4)–(5) describe the evolution of the window vector $w(t) = (w_i(t)$, for all $i)$ and the backlog vector $b(t) = (b_l(t)$, for all $l)$. A pair of rate and backlog vectors $(x, b)$, with window vector $w$ given by $w_i = x_i(\tau_i + \sum_l R_{li} b_l / c_l)$, is said to be in *equilibrium* if both $\dot{w}(t) = 0$ and $\dot{b}(t) = 0$.

We start by defining a bottleneck link and other notation for XCP equilibrium. In general quantities without $t$ dependence denote equilibrium quantities, e.g., $w_i, T_i, x_i, F_{li}$, etc.

**Definition 1** *A link $l$ is said to be a* bottleneck *for source $i$ with respect to (w.r.t.) $x$ if $F_{li}$ is minimum among all the links that $i$ uses, i.e., $F_{li} = \min_{m \in L(i)} F_{mi}$. In this case, source $i$ is said to be* bottlenecked *at link $l$ w.r.t. $x$.*

By definition, every source $i$ has a bottleneck. Lemma 1 below implies that $F_{li} = 0$ in equilibrium at a bottleneck $l$.

We distinguish between links that are bottlenecks and those that are not. Let $L_1(i)$ be the set of links that are bottlenecks for source $i$ w.r.t a given equilibrium rate $x$:

$$L_1(i) \ := \ \{ \ l \in L(i) \mid F_{li} = \min_{m \in L(i)} F_{mi} \ \}$$

and $L_0(i) := L(i) \setminus L_1(i)$ be the set of links in $i$'s path that are not bottlenecks for source $i$ w.r.t $x$. We also distinguish between sources that are bottleneck locally and those that are not. Let $I_1(l)$ be the set of sources bottlenecked at link $l$ w.r.t. a given equilibrium rate $x$:

$$I_1(l) \ := \ \{ \ i \in I(l) \mid F_{li} = \min_{m \in L(i)} F_{mi} \ \}$$

and $I_0(l) := I(l) \setminus I_1(l)$ be the set of sources bottlenecked elsewhere. Let $N_l := |I(l)|$ be the number of sources at link $l$, $N_{l0} := |I_0(l)|$, and $N_{l1} := |I_1(l)|$. Let $\rho_l := N_{l0}/N_l$ be the fraction of flows through link $l$ which are not bottlenecked at link $l$, and $\sigma_l := y_{l0}/c_l$ be the fraction of the link's capacity consumed by such flows. Note that while $L(i)$, $I(l)$, and $N_l$ depend only on the routing matrix $R$, $L_1(i)$, $L_0(i)$, $I_1(l)$, $I_0(l)$, $N_{l0}$, $N_{l1}$, $\rho_l$ and $\sigma_l$ depend also on the equilibrium rate $x$ through $F_{li}$.

From (4) and the definition of $I_0(l)$, we have

**Lemma 1** *The rate and backlog vector $(x, b)$ is in equilibrium if and only if*

1. *for all $l$, $y_l \leq c_l$ with equality if $b_l > 0$ and*
2. *for all $i$, $\min_{l \in L(i)} F_{li} = 0$.*

*Moreover,*

3. *if $i \in I_0(l)$ and $j \in I_1(l)$ then $F_{li} > 0$ and $F_{lj} = 0$.*
4. *if $I_1(l) \neq \emptyset$ then $h_l = 0$ implies $\phi_l = 0$.*

**Proof:** Parts 1 to 3 are immediate. To see part 4, note that $h_l = 0$ implies

$$F_{li} = \frac{\phi_l^+}{N_l x_i} - \frac{\phi_l^-}{y_l}$$

By part 2, $F_{li} = 0$ for all $i \in I_1(l)$. Since at most one of $\phi_l^+$ and $\phi_l^-$ can be nonzero, $\phi_l^+ = \phi_l^- = 0$, whence $\phi_l = 0$. □

## 3.1 The need for bandwidth shuffling

Without bandwidth shuffling, XCP would have $\gamma = 0$, giving $h_l(t) = 0$ for all $l$ and $t$. In particular, $h_l = 0$ in equilibrium.

**Theorem 2** *Suppose $\gamma = 0$. Then $(x, b)$ with $x_i = w_i/\tau_i$ is an equilibrium if and only if*

1. *for all $l$, $y_l \le c_l$ and $b_l = 0$, and*
2. *for all $i$, there exists $l \in L(i)$ with $y_l = c_l$.*

**Proof:** The first condition in the theorem implies that for all $l$, $\phi_l \ge 0$. Combined with $h_l = 0$, this implies $F_{li} \ge 0$ for all $i$. The second condition then implies that for all $i$, $\min_{l\in L(i)} F_{li} = 0$. Hence, the conditions in the theorem are sufficient, by (4) and the first part of Lemma 1.

For necessity, there are two cases. If $I_1(l) \neq \emptyset$ then $\phi_l = 0$ by the second part of Lemma 1, and (5b) implies $y_l = c_l$ and $b_l = 0$, since $y_l \le c_l$, $b_l \ge 0$, and $\beta > 0$. Otherwise $l \in \bigcap_i L_0(i)$ and $F_{li} > 0$ by definition of $L_0(i)$. This implies $\phi_l^+ > 0$, and hence $y_l < c_l$, $b_l = 0$ in equilibrium. □

**Remark:** Without bandwidth shuffling, any (possibly unfair) boundary point of the set $\{x | Rx \le c\}$ would be an equilibrium. These are exactly the rates $x$ which maximize aggregate throughput. This is why XCP uses $\gamma > 0$ [8].

The rest of the paper considers the more complicated case of $\gamma > 0$.

## 3.2 $\gamma > 0$ case: main results

This subsection provides a conceptually simple characterization and uses it to prove the existence and uniqueness of XCP equilibrium. In the next subsection, we provide an iterative algorithm to compute this equilibrium.

From (4)–(5) and Lemma 1, $(x, b)$ is an XCP equilibrium if and only if

1. For all $l$, $y_l \le c_l$ with equality if $b_l > 0$.
2. For all sources $i$, $\min_{l\in L(i)} F_{li} = 0$.

Using (5a), condition 2 becomes: for all $i$, for all $l \in L(i)$,

$$x_i \le \frac{y_l}{N_l} \frac{h_l + \phi_l^+}{h_l + \phi_l^-} =: r^l \qquad (6)$$

with equality for some $l \in L(i)$. Hence for links $l$ with $I_1(l) \neq \emptyset$, all flows $i \in I_1(l)$ that are bottlenecked at link $l$ must have the *common* rate $r^l$. This has important implications as we will see below.

Several of the results will use the following technical lemma, which is proved in Appendix A.

**Lemma 2** *For all $l$*

1. *$x_i < x_j = r^l$ if $i \in I_0(l)$ and $j \in I_1(l)$.*
2. *$\sigma_l < \rho_l$ if $I_0(l) \neq \emptyset$.*
3. *$r^l \ge y_l/N_l$ with equality if and only if $I_0(l) = \emptyset$.*
4. *$h_l > 0$ if $I_1(l) \neq \emptyset$.*
5. *$y_l/c_l \ge \sigma_l$ with equality if and only if $I_1(l) = \emptyset$.*

Unlike in the $\gamma = 0$ case, we characterize the equilibrium backlogs and rates separately. The following result says that the equilibrium queue under XCP is zero. This originates from the definition of $\phi_l$ in (5b), which is nonnegative in equilibrium. The same property is used in REM [11] to drive the queue to zero, or more generally, to a target value.

**Theorem 3** *In equilibrium, $b_l = 0$ and $\phi_l \ge 0$ for all $l$.*

**Proof:** Links can be of three types: (a) $I_1(l) \neq \emptyset$, $I_0(l) = \emptyset$, (b) $I_1(l) = \emptyset$, $I_0(l) \neq \emptyset$, and (c) $I_1(l) \neq \emptyset$, $I_0(l) \neq \emptyset$. Each of these will be considered in turn.

Type (a) links are bottlenecks for all flows passing through them, i.e., links $l$ where (6) holds with equality for all $i \in I(l)$. Since all flows have common rate $r^l$, $y_l = N_l r^l$, whence equality in (6) implies $\phi_l^+ = \phi_l^-$. Thus $\phi_l = 0$, and (5b) implies $y_l = c_l$ and $b_l = 0$, i.e., they share the link capacity fully and equally, with no queueing delay.

Type (b) links are not bottlenecks for any of the flows they carry. Hence, for all $i \in I(l)$,

$$x_i < \frac{y_l}{N_l} \frac{h_l + \phi_l^+}{h_l + \phi_l^-}$$

Multiplying both sides by $R_{li}$ and summing over $i$, we have

$$y_l < \frac{y_l}{N_l} \frac{h_l + \phi_l^+}{h_l + \phi_l^-} \cdot \sum_i R_{li}$$

Hence

$$\frac{h_l + \phi_l^+}{h_l + \phi_l^-} > 1$$

Since both numerators and denominators are positive, $\phi_l^+ > \phi_l^-$. This implies $\phi_l > 0$ whence $y_l < c_l$ and $b_l = 0$.

Type (c) links are bottleneck links for some but not all of the flows using them. From (6), we have

$$\frac{h_l + \phi_l^+}{h_l + \phi_l^-} = \frac{r^l}{y_l/N_l} > 1$$

where the inequality follows from Lemma 2. As for type (b) links, this implies $\phi_l > 0$, $y_l < c_l$ and $b_l = 0$. □

We next characterize the equilibrium rates of XCP. Define $g_l$ as

$$g_l(x) \quad := \quad \frac{\gamma y_l^2}{N_l\left[(\gamma+\alpha)y_l - \alpha c_l\right]}$$

where $y_l = \sum_i R_{li} x_i$. Since $g_l(x)$ depends on $x$ only through $y_l$, we will abuse notation and also write $g_l(y_l)$ or $g_l(y_l(x))$. Define the *feasible set* of source rates $x$ to be

$$X_0 := \left\{ x \in \Re_+^N \,|\, g_l(y_l) < 0 \text{ or } x_i \le g_l(y_l),\ \forall l, i \in I(l) \right\} \tag{7}$$

where $\Re_+$ denotes the set of nonnegative real numbers. We will later show that the XCP equilibrium must be in $X_0$. Note that $x \in X_0$ implies

$$Rx \quad \le \quad c$$

To see this, multiply both sides of the inequality in (7) by $R_{li}$ and sum over $i$ to get

$$y_l \quad = \quad \sum_i R_{li} x_i \ \le\ \frac{\gamma y_l^2}{(\gamma+\alpha)y_l - \alpha c_l}$$

Rearranging the above inequality yields $y_l \le c_l$. The converse may not be true, i.e., $X_0$ may be a strict subset of $\{x | Rx \le c\}$.

Our main result is to prove the existence and uniqueness of XCP equilibrium in a general network, and that this equilibrium solves a *constrained* max-min fairness problem.

**Definition 4** *A rate vector $x^* \in X_0$ is* constrained max-min fair *if for any other feasible $x \in X_0$, $x_i > x_i^*$ implies that there is a $j$ with $x_j < x_j^*$ and $x_j^* \le x_i^*$.*

Intuitively, a constrained max-min fair vector $x^*$ is such that it is not possible to increase a component $x_i^*$ without reducing another smaller or equal component $x_j^*$. This differs from standard max-min fairness only in that the feasible set $X_0$ is a subset of $\{x|Rx \le c\}$ [12]. This restriction has important ramifications, as we will see in the next section.

We will prove constructively that the unique XCP equilibrium is constrained max-min fair by identifying it with the solution of a hierarchy of optimization problems over the feasible set $X_0$: it maximizes the smallest source rates in $X_0$, and then maximizes the second smallest rates over all rates that solve the first problem, and so on. These maximization problems are defined inductively, following the idea of [13].

Let $L_0 = \emptyset$ and $I_0 = \emptyset$. The sets $(L_0, I_0, X_0)$ define the first problem $\mathbf{P}_1$, whose solution is described by the sets $(L_1, I_1, X_1)$. These sets in turn define the second problem $\mathbf{P}_2$, and so on. To simplify notation, let

$$\overline{L}_n := \bigcup_{m \le n} L_m \qquad \overline{I}_n := \bigcup_{m \le n} I_m$$

Given sets $(X_0, L_0, I_0), \ldots, (X_{n-1}, L_{n-1}, I_{n-1})$, if $\overline{I}_{n-1}$ contains all flows, then we stop. Otherwise, we define problem $\mathbf{P}_n$ and its solution $L_n, I_n, X_n$, $n \ge 1$, as follows.

$$\mathbf{P}_n\textbf{:} \qquad \max_{x \in X_{n-1}} \min_{i \notin \overline{I}_{n-1}} \ x_i \tag{8}$$

Let

$$r_n \quad := \quad \min_{l \notin \overline{L}_{n-1}} \max_{x \in X_{n-1}} \ g_l(x) \tag{9}$$

$$L_n \quad := \quad \{ \text{ minimizing } l \text{ in (9) } \} \tag{10}$$

$$I_n \quad := \quad \bigcup_{l \in L_n} I(l) \setminus \overline{I}_n \tag{11}$$

$$X_n \quad := \quad \left\{ x \in X_{n-1} \,\middle|\, x_i \left\{ \begin{array}{ll} = r_n, & \forall i \in I_n \\ > r_n, & \forall i \notin \overline{I}_n \end{array} \right. \right\} \tag{12}$$

A few important properties are immediate from these definitions. First, the rates $r_n$ are monotonic:

$$\min_l \frac{c_l}{N_l} \ = \ r_1 < r_2 < \cdots < r_n \tag{13}$$

Second, $L_n$ and $I_n$ are nonempty; moreover they are disjoint from $\overline{L}_{n-1}$ and $\overline{I}_{n-1}$, respectively. Hence $\overline{I}_n$ will eventually contain all the flows and there are only a finite number of problems $\mathbf{P}_n$. Finally, $X_n$ are *strictly* nested:

$$X_0 \supsetneq X_1 \supsetneq \cdots \supsetneq X_n$$

Indeed it will become clear that $X_n$ is exactly the set of solutions to problem $\mathbf{P}_n$, i.e., $X_1$ is the set of feasible rates $x \in X_0$ whose smallest rates are maximized, $X_2$ is a subset of $X_1$ whose second smallest rates are also maximized, and so on. We prove below that if $\mathbf{P}_{n^*}$ is the last problem, then $X_{n^*}$ is a singleton that solves all problems $\mathbf{P}_1, \ldots, \mathbf{P}_{n^*}$.

To contrast XCP equilibrium with the standard max-min fair allocation, we derive a “bottleneck” characterization that is analogous to that for max-min fairness; see the beginning of Section 4.

**Lemma 3** *Suppose $x$ is the XCP equilibrium rate vector. Link $l$ is a bottleneck for source $i \in I(l)$ w.r.t. $x$ if and only if*

1. *$x_i = g_l(x)$, and*
2. *$x_i \ge x_j$ for all $j \in I(l)$.*

**Proof:** Suppose link $l$ is a bottleneck link for source $i$ w.r.t. equilibrium $x$. Then Lemma 1(2) implies that $F_{li} = 0$, i.e., equality holds in (6). Since $\phi_l \ge 0$ by Theorem 3 and $h_l > 0$ by Lemma 2, (5c) becomes $h_l = \gamma d y_l - \phi_l$. Thus from (6)

$$x_i \quad = \quad r^l \ = \ \frac{y_l}{N_l} \frac{\gamma d y_l}{\gamma d y_l - \phi_l} \ = \ g_l(x) \tag{14}$$

proving the first condition. Condition (6) then implies the second condition.

Conversely, suppose the two conditions are satisfied. If $h_l = 0$, then $F_{li} = 0$ from (5a). Lemma 1(2) then implies $F_{li}$ is the minimum among links in source $i$'s path, i.e., link $l$ is a bottleneck. On the other hand, if $h_l > 0$, then, as above, $\phi_l \geq 0$ and $h_l = \gamma d y_l - \phi_l$. Then $x_i = g_l(x)$ is equivalent to $F_{li} = 0$, proving that $l$ is a bottleneck. □

Motivated by this lemma, we call link $l$ a *nonbottleneck w.r.t.* $x$ if either $g_l(x) < 0$ or $x_i < g_l(x)$ for all $i \in I(l)$.

Our main result is

**Theorem 5** *The problems* $\mathbf{P}_n$ *are well-defined and have a unique solution. Moreover, the following are equivalent:*

1. $x^*$ *is an XCP equilibrium.*
2. $x^*$ *is the unique rate vector that solves all the problems* $\mathbf{P}_n$.
3. $x^*$ *is constrained max-min fair.*
4. $x^* \in X_0$ *and every flow has a bottleneck w.r.t.* $x^*$*, i.e., for all* $i$*, there is an* $l \in L(i)$ *such that* $x_i^* = g_l(x^*)$ *and* $x_i^* \geq x_j^*$ *for all* $j \in I(l)$.

Before presenting our proof, we derive a (centralized) algorithm to compute the XCP equilibrium.

## 3.3 Algorithm for computing equilibrium

The equilibrium rates of XCP can be found using an algorithm analogous to that of [12] for max-min fairness. However, because the constraint on the link throughput in (6) depends on the aggregate flow rate through $h_l$ and $\phi_l$, some extra bookkeeping is required.

**Theorem 6** *The utilization of a bottleneck link* $l$ *satisfies*

$$\frac{y_l}{c_l} = \frac{\alpha + (\gamma+\alpha)\sigma_l + \sqrt{(\alpha - (\gamma+\alpha)\sigma_l)^2 + 4\alpha\gamma\sigma_l(1-\rho_l)}}{2(\gamma\rho_l + \alpha)} \tag{15}$$

*The rates of all sources* $i \in I_1(l)$ *bottlenecked at* $l$ *satisfy*

$$r^l = \frac{c_l}{N_l}\frac{[\Xi_l - \gamma\sigma_l\rho_l] + \sqrt{[\Xi_l + \gamma\sigma_l\rho_l]^2 - 4\alpha\gamma\sigma_l(\rho_l - \sigma_l)}}{2(1-\rho_l)(\gamma\rho_l + \alpha)} \tag{16}$$

*where*

$$\Xi_l = (\gamma\sigma_l + \alpha)(1 - \sigma_l) - \gamma\sigma_l(\rho_l - \sigma_l) \tag{17}$$

**Proof:** Substituting $r^l = (y_l - y_{l0})/(N_l - N_{l0})$ into (14) and solving the resulting quadratic equation gives

$$y_l = \frac{\alpha c_l + (\gamma + \alpha)y_{l0} \pm \sqrt{(\alpha c_l + (\gamma+\alpha)y_{l0})^2 - K}}{2(\gamma N_{l0}/N_l + \alpha)} \tag{18}$$

where $K = 4\alpha c_l y_{l0}(\gamma N_{l0}/N_l + \alpha)$. By Lemma 8 in Appendix A, only the larger solution of (18) satisfies part 5 of Lemma 2, and is a valid equilibrium. Rearranging the term in the square root gives (15).

To obtain (16), instead substitute $y_l = (N_l - N_{l0})r^l + y_{l0}$ into (14), giving

$$A_l\left(\frac{r^l N_l}{c_l}\right)^2 + B_l\frac{r^l N_l}{c_l} + C_l = 0, \tag{19}$$

where

$$\begin{aligned} A_l &= (1-\rho_l)(\gamma\rho_l + \alpha) \\ B_l &= 2\gamma\sigma_l\rho_l - \gamma\sigma_l - \alpha(1-\sigma_l) \\ C_l &= -\gamma\sigma_l^2. \end{aligned}$$

Since $y_l$ is increasing in $r^l$, it is again only the larger root which represents the XCP equilibrium. Thus

$$\frac{r^l N_l}{c_l} = \frac{[\Xi_l - \gamma\sigma_l\rho_l] + \sqrt{[\Xi_l - \gamma\sigma_l\rho_l]^2 - 4A_lC_l}}{2(1-\rho_l)(\gamma\rho_l + \alpha)},$$

where $\Xi_l$ is given in (17). Rearranging the expression in the square root gives (16). □

Note that the right-hand side of (16) depends on the rate vector $x$ throughput $\rho_l$ and $\sigma_l$. Hence it is not an explicit formula for the throughput of a general flow. However it says that the common "bottleneck" rate at each link $l$ depends on the rate vector $x$ only through $y_{l0}$ and $N_{l0}$ that are bottlenecked elsewhere. These are source rates smaller than the "bottleneck" rate at link $l$, by Lemma 1. This motivates an algorithm similar to the max-min algorithm of [12] that calculates the throughput $x_i$ of each flow in increasing order, without the need for recourse to simulation.

1 Set $\bar{I}_0 \leftarrow \emptyset$, $\bar{L}_0 \leftarrow \emptyset$, $\sigma_l(0) \leftarrow 0$, $\rho_l(0) \leftarrow 0$ for all $l$, $n \leftarrow 1$
2 **repeat**
- 2.1 For each link, $l \notin L_{n-1}$ find $r^l(n)$ from (16) using $\sigma_l(n-1)$ and $\rho_l(n-1)$ from rates already allocated
- 2.2 Set $r_n \leftarrow \min_j r^j(n)$
- 2.3 Set $L_n \leftarrow \{l : r^l(n) = r_n\}$
- 2.4 **foreach** $l \in L_n$
  - 2.4.1 Set $r^l \leftarrow r_n$
  - 2.4.2 For each flow $i \in I(l) \setminus \bar{I}_n$, set $x_i \leftarrow r^l$

  **endfor**
- 2.5 Set $I_n \leftarrow \bigcup_{j \in L_n} I(j) \setminus \bar{I}_n$
- 2.6 Set $\bar{I}_n \leftarrow \bar{I}_{n-1} \cup I_n$
- 2.7 Set $\bar{L}_n \leftarrow \bar{L}_{n-1} \cup L_n$
- 2.8 **foreach** $l \in \bigcup_{i \notin \bar{I}_n} L(i)$
  - 2.8.1 Set $\sigma_l(n) \leftarrow \sigma_l(n-1) + \sum_{i \in I_n} R_{li}x_i/c_l$
  - 2.8.2 Set $\rho_l(n) \leftarrow \rho_l(n-1) + \sum_{i \in I_n} R_{li}/N_l$

  **endfor**
- 2.9 Set $n \leftarrow n+1$

**until** $\bar{I}_n = \{$all flows$\}$

This solves each of the optimization problems, $\mathbf{P}_n$, in turn. The key is that, by keeping track of the used capacity of each link, $\sigma_l(n)$ and $\rho_l(n)$, it can compute the maximization in (9) in closed form. For each $l$, the values $\sigma_l(n)$ and $\rho_l(n)$ vary during the algorithm. For the algorithm to be correct, they must have the right values when link $l$ is the minimum in step 2.2. This occurs as long as the link rates are allocated in increasing order, as is guaranteed by the following theorem, proved in Appendix C.

**Theorem 7** *The above algorithm calculates the correct equilibrium rates of XCP.*

If $\gamma = 0$ then (16) reduces to $r^l = (c_l - y_{l0})/(N_l - N_{l0})$, and hence the algorithm reduces to the algorithm in [12] to compute the max-min fair allocation. This suggests that, given any topology specified by the routing matrix $R$ and link capacity vector $c$, one can choose $\gamma > 0$ to be sufficiently small so that the equilibrium of (4) is close to max-min fair. On the other hand, with small $\gamma$, the convergence of individual rates to fairness can be very slow. We will return to this point in Section 4.

## 3.4 $\gamma > 0$ case: proofs and intuitions

In this subsection we prove Theorem 5. The proofs of lemmas are relegated to Appendix B. We start with a simple observation that greatly simplifies the solution of $\mathbf{P}_n$.

**Lemma 4** *Suppose $X_n$ is nonempty. The maximization in (9) can be taken over $x \in X_n$ that have equal $x_i$ for $i \notin \overline{I}_n$.*

In view of Lemma 4, we can replace $X_n$ in (12), for $n \geq 1$, by their subsets:

$$\hat{X}_n := \left\{ x \in X_{n-1} \middle| x_i = \begin{cases} r_n, & \forall i \in I_n \\ r_n + \epsilon, & \forall i \notin \overline{I}_n, \epsilon > 0 \end{cases} \right\} \quad (20)$$

and use them instead of $X_{n-1}$ in computing $r_n$:

$$r_n := \min_{l \notin \overline{L}_{n-1}} \max_{x \in \hat{X}_{n-1}} g_l(x)$$

This greatly reduces the complexity of (9) from maximizing over $n$-vectors $x \in X_n$ to over a scalar $\epsilon > 0$.

Denote an $x \in \hat{X}_n$ by $x(\epsilon; n)$, with

$$x_i(\epsilon; n) = \begin{cases} r_m, & i \in I_m, m \leq n \\ r_n + \epsilon, & i \notin \overline{I}_n \end{cases} \quad (21)$$

and let $x(0; n) := \lim_{\epsilon \to 0} x(\epsilon; n)$. Note that $x(0; n)$, $n \geq 1$, is not in $X_n$ according to definition (12), though it is in $X_0$. We will see in Lemma 6 below that $x(0, n)$ plays an important role in the proof of Theorem 5. The vector $x(\epsilon; n)$ induces link flows

$$\begin{aligned} y_l(\epsilon; n) &= \sum_i R_{li} x_i(\epsilon; n) \\ &= \sum_{m=1}^{n} r_m \sum_{i \in I_m} R_{li} + r_n \sum_{i \notin \overline{I}_n} R_{li} + \epsilon \sum_{i \notin \overline{I}_n} R_{li} \quad (22) \end{aligned}$$

This motivates the following main technical lemma.

**Lemma 5** *Given any scalars $z \geq 0$, $\eta > 0$, and $s \geq 0$, define*

$$\begin{aligned} \hat{g}_l(\epsilon) &:= \hat{g}_l(\epsilon; z, \eta) := g_l(z + \eta\epsilon) \\ &= \frac{\gamma(z + \eta\epsilon)^2}{N_l((\gamma + \alpha)(z + \eta\epsilon) - \alpha c_l)} \quad (23) \end{aligned}$$

*for some $N_l \geq 1$, $\alpha > 0$, $\gamma > 0$ and $c_l > 0$.*

1. *If either $\hat{g}_l(0) < 0$ or $s \leq \hat{g}_l(0)$ then there exists a unique $\epsilon_l \geq 0$ such that $s + \eta\epsilon_l = \hat{g}_l(\epsilon_l)$, where $\epsilon_l = 0$ if and only if $s = \hat{g}_l(0)$.*
2. *Moreover, over $\{\epsilon \,|\, \hat{g}_l(\epsilon) > 0\}$, $s + \eta\epsilon \leq \hat{g}_l(\epsilon)$ if and only if $\epsilon \leq \epsilon_l$.*

For later reference, we will denote the mapping from $(z, \eta, s)$ to the unique $\epsilon_l$ in Lemma 5 by

$$G_l(z, \eta, s) = \epsilon_l \quad (24)$$

This function is used in the rest of the proof.

Lemma 5 implies that if link $l$ is a bottleneck for some source $i$ with respect to an $x \in X_0$, then the rate of source $i$ cannot be increased without violating the feasibility constraint in (7). For instance, let $n \geq 1$ be such that $l \in L_n$. Setting $z = y_l(0; n-1)$, $s = r_{n-1}$ and $\eta = \sum_{i \notin \overline{I}_{n-1}} R_{li}$ gives $\epsilon_l = r_n - r_{n-1}$ and $l$ is a bottleneck for all $i \in I_n$ w.r.t. $x(\epsilon_l; n-1)$. Lemma 5(b) then implies that rates greater than $r_n$ are infeasible at link $l$.

The next lemma implies that all links $l \in L_n$ are bottlenecks w.r.t. all $x \in \hat{X}_n$, and all links $l \notin \overline{L}_n$ are nonbottlenecks w.r.t. $x(0; n)$. In particular, this implies that $X_n$ are nonempty.

**Lemma 6** *For each $n \geq 1$,*

1. *if $l \in L_n$, then $x_i = g_l(x)$ for all $i \in I_n$ w.r.t. all $x \in \hat{X}_n$.*
2. *if $l \notin \overline{L}_n$, then either $g_l(x(0; n)) < 0$ or $x_i(0; n) < g_l(x(0; n))$ for all $i \in I(l)$.*

Lemmas 5 and 6 suggest the following abstract algorithm to compute the solution of problems $\mathbf{P}_n$, analogous to the one of the previous section to find the XCP equilibrium. At the start of iteration $n+1$, all links $l \in L_m$, $m \leq n$, are bottlenecks for some sources $i \in I_m$ w.r.t. all $x \in X_m$. A source $i \in \overline{I}_n$ passes through at least one bottleneck $l \in \overline{L}_n$, and hence its rate cannot be raised further without violating the constraint in (7). All links $l \notin \overline{L}_n$ are nonbottleneck links w.r.t. $x(0; n)$ defined in (21). Sources $i \notin \overline{I}_n$ pass through only these nonbottlenecks, and hence their rates can be increased further, starting from $r_n$. At each nonbottleneck link $l$, nonbottlenecked sources $i \notin \overline{L}_n$ can raise their rates to (using (24), (21), and (22))

$$r_n + G_l \left( \sum_{m=1}^{n} r_m \sum_{i \in I_m} R_{li} + r_n \sum_{i \notin \overline{I}_n} R_{li},\ \sum_{i \notin \overline{I}_n} R_{li},\ r_n \right)$$

to make $l$ a bottleneck. The smallest of these rates, smallest over $l \notin \overline{L}_n$, is $r_{n+1}$. It is assigned to all previously nonbottlenecked sources going through the new bottleneck links, and is the optimal objective value for problem $\mathbf{P}_n$. These new bottleneck links are collected into $L_{n+1}$, the newly bottlenecked sources into $I_{n+1}$ and their rates into $\hat{X}_{n+1}$ (or $X_{n+1}$). The other non-minimizing links remain nonbottleneck w.r.t. the new rates $x \in \hat{X}_{n+1}$, and the cycle repeats, until all sources are assigned their bottleneck rates.

The solution of each problem $\mathbf{P}_n$ fixes the components $x_i, i \in I_n$, to be rate $r_n$, until all components have been assigned. Hence, if $\mathbf{P}_{n^*}$ is the last problem, then $X_{n^*} = \{x^*\}$ is a singleton.

The above discussion is summarized in the following lemma, which justifies Theorem 5.

**Lemma 7** *The problems $\mathbf{P}_n$ are well-defined. $X_n$ is exactly the set of solutions to problem $\mathbf{P}_n$. There is a unique solution to the hierarchy of problems.*

We now prove Theorem 5.

**Proof (Theorem 5) :** Lemma 7 implies that $\mathbf{P}_n$ are well-defined and have a unique solution. It is clear that characterizations 2 and 3 are equivalent, i.e., $x^*$ is the unique solution to the hierarchy of problems $\mathbf{P}_n$ if and only if it is constrained max-min fair. We will first prove the equivalence of characterizations 3 and 4, and then that of 1 and 4. We will use the equivalent definition of bottleneck links in Lemma 3.

**Equivalence of characterizations 3 and 4:** We will prove that $x^*$ is constrained max-min fair if and only if both $x^* \in X_0$ and every flow $i$ has a bottleneck link w.r.t. $x^*$, i.e., for all $i$, there is an $l \in L(i)$ such that $x_i^* = g_l(x)$ and $x_i^* \geq x_j^*$ for all $j \in I(l)$. The proof follows the same approach as the corresponding result for standard max-min fairness; see [12]. The difference is in the use of Lemma 5 because of the more complicated feasible set $X_0$.

Suppose $x^* \in X_0$ and every flow $i$ has a bottleneck link w.r.t. $x^*$. If $x^*$ is not constrained max-min fair, then there exists another $x \in X_0$ such that $x_i > x_i^*$ for some $i$, and if $x_j^* \leq x_i^*$ then $x_j \geq x_j^*$. We will derive a contradiction. Let $l \in L(i)$ be a bottleneck for $i$ w.r.t. $x^*$. Then $x_i^* \geq x_j^*$ for all $j \in I(l)$, and hence $\epsilon_j := x_j - x_j^* \geq 0$ for all $j \in I(l)$, with $\epsilon_i > 0$. Write the link flow $y_l(x)$ due to rates $x$ in terms of $\epsilon_j$ and the link flow $y_l(x^*)$ due to $x^*$:

$$
\begin{aligned}
y_l(x) &= \sum_j R_{lj}(x_j^* + \epsilon_j) \\
&= y_l(x^*) + \sum_j R_{lj}\epsilon_j
\end{aligned}
$$

Let the scalar $\epsilon$ be the average $\epsilon_j$:

$$
\epsilon = \frac{\sum_j R_{lj}\epsilon_j}{\sum_j R_{lj}} > 0
$$

Then the rate vector $\hat{x}$ defined by $\hat{x}_j = x_j^* + \epsilon$ if $j \in I(l)$ and $\hat{x}_j = x_j^*$ otherwise induces the same flow rate at link $l$ as $x$ does:

$$
y_l(\hat{x}) = y_l(x^*) + \epsilon \cdot \sum_j R_{lj} = y_l(x) \tag{25}
$$

Since $\epsilon \leq \max_j \epsilon_j$, and $x$ is feasible, we must have, for all $j \in I(l)$,

$$
\begin{aligned}
\hat{x}_j &\leq \max_{j \in I(l)} x_j \\
&\leq g_l(y_l(x)) \\
&= g_l(y_l(\hat{x}))
\end{aligned} \tag{26}
$$

where the last equality follows from (25). Hence $\hat{x}$ is also feasible. But link $l$ is a bottleneck for source $i$ w.r.t. $x^*$, i.e.,

$$
x_i^* = g_l(y_l(x^*)) > 0 \tag{27}
$$

and so applying Lemma 5(1) with $z = y_l(x^*)$, $s = x_i^*$ and $\eta = \sum_j R_{lj}$ gives $\epsilon_l = 0$. Since $\hat{x}_j = x_j^* + \epsilon$ for all $j \in I(l)$, with $\epsilon > \epsilon_l$, Lemma 5(2) and (25) imply that

$$
\hat{x}_i > g_l(y_l(\hat{x}))
$$

contradicting (26). Hence $x^*$ is constrained max-min fair.

Conversely, let $x^* \in X_0$ be constrained max-min fair. If there is a source $i$ that has no bottleneck link w.r.t. $x^*$, then for all $l \in L(i)$, either $g_l(y_l(x^*)) < 0$ or $x_i^* < g_l(y_l(x^*))$ for all $i \in I(l)$. Lemma 5 then implies that there exists a unique scalar $\epsilon_i > 0$, given by

$$
\epsilon_i = \min_{l \in L(i)} G_l\left(y_l(x^*), 1, x_i^*\right)
$$

such that $x \in X_0$, given by $x_i = x_i^* + \epsilon_i$ and $x_j = x_j^*$ for $j \neq i$, strictly increases component $i$ without having to reduce other components $j$, contradicting the fact that $x^*$ is constrained max-min fair.

**Equivalence of characterizations 1 and 4:** We will prove that a vector $x^* \in \Re_+^N$ is an XCP equilibrium if and only if $x^* \in X_0$ and every flow $i$ has a bottleneck link w.r.t. $x^*$.

The discussion at the beginning of Section 3.2 shows that $x^*$ is an XCP equilibrium if and only if, for all $i$, (6) holds for all $l \in L(i)$, with equality for some $l \in L(i)$. This, with (14), establishes $x^* \in X_0$. As observed after Definition 1, every flow has a bottleneck by definition.

To show characterization 4 implies characterization 1, it suffices to show that the characterization in Lemma 3 implies statements 1. and 2. at the start of Section 3.2. The discussion after (7), and setting $b = 0$, establishes Statement 1. This shows $\phi_l^- = 0$ for all $l$. If $x_i \leq g_l(x)$ then (5c) and (5a) give $F_{li} \geq 0$, with equality when $x_i = g_l(x)$. Otherwise, $g_l(x) < 0$ giving $h = 0$ and, by (5a), $F_{li} \geq 0$. □

# 4 Utilization and fairness

In this section, we discuss some implications of the results in Section 3 on link utilization and fairness of the equilibrium rates. Theorem 5 shows that XCP equilibrium is constrained max-min fair. It is instructive to compare the XCP equilibrium with the (standard) max-min fair allocation and a class of algorithms proposed in [13].

It is proved in [13] that a (standard) max-min fair rate vector $x^*$ is the unique solution of the same hierarchy of problems $\mathbf{P}_n$ (8)–(12) defined in Section 3, except that the feasible set $X_0$ in (7) is replaced with the superset

$$\overline{X}_0 \quad := \quad \{\, x \in \Re^N_+ \mid Rx \le c \,\} \tag{28}$$

The key feature that results from this much simpler feasible set $\overline{X}_0$ is that the bottleneck links under a max-min fair allocation are all fully utilized. Indeed, a rate vector $x^* \in \overline{X}_0$ is max-min fair if and only if, for every source $i$, there is a link $l \in L(i)$ in its path such that [12]

1. $y_l(x^*) = c_l$
2. $x^*_i \ge x^*_j$ for all $j \in I(l)$,

From Theorem 5, condition 1 is replaced with the fixed point equation $x^*_i = g_l(y_l(x^*))$ for XCP equilibrium. The simpler condition for max-min fairness has several implications.

First it allows a much simpler proof of max-min fair vector as the unique solution of the problems $\mathbf{P}_n$; see [13]. Second the (centralized) algorithm to compute the max-min fair rate vector (see [13, 12]) is simpler than that in Section 3.3 for the constrained max-min fair vector. Third, and most importantly, the XCP equilibrium can underutilize link capacities and deviate by an arbitrarily large factor from the max-min fair allocation, as we illustrate below.

Max-min fairness is generalized in [13] by restricting the feasible set to a (strict) subset of $\overline{X}_0$ in (28). Like XCP, the restriction is specified as additional constraints on source rates $x_i$ and link flows $y_l$. An example is that, in addition to being in $\overline{X}_0$, a feasible rate vector $x$ must also satisfy

$$x_i \quad \le \quad \frac{1}{s^2 c_l}(c_l - y_l)^2 \quad \forall i, \forall l \in L(i)$$

This is motivated by an explicit design objective of trading off full link utilization for the ability to accommodate random rate fluctuations. If the standard deviation of the rate of source $i$ is $sx_i$, then it is shown in [13] that the standard deviation of the link flow $y_l$ is less than the spare capacity $c_l - y_l$, so that overshoot is avoided, i.e., $y_l(t) \le c_l$ for all $t$ in the absence of feedback delay. An alternative additional constraint in [13] is

$$x_i \quad \le \quad \frac{f_i}{e_l}(c_l - y_l) \quad \forall i, \forall l \in L(i)$$

This is again motivated by an explicit design objective: the link parameter $e_l$ controls utilization and source parameter $f_i$ controls fairness, akin to XCP's efficiency and fairness controllers. A distributed algorithm to compute the equilibrium rates is also provided in [13], and its convergence proved. Like XCP, explicit feedback is required: each link $l$ feeds back the spare capacity $c_l - y_l(t)$ to sources that go through this link. Sources adjust their individual rates based on feedback on its path in a way that is distributed, yet avoids overshoot.

We now illustrate the effect of the additional constraint (7) in XCP on link utilization and fairness.

As we explained in the proof of Theorem 3, there are three types of links. The first type are bottlenecks for all the flows that go through that link. All links of this type, such as all $l \in L_1$ in problem $\mathbf{P}_1$, are fully utilized, $y_l = c_l$. The second type are bottlenecks for none of the flows that go through that link. They are underutilized, $y_l < c_l$, because the flow rates going through the link are constrained elsewhere. The third type are bottlenecks for some, but not all, of the flows that go through the link. In contrast to the standard max-min fair allocation, these links are also underutilized, $y_l < c_l$. We can bound the utilization of these partial bottlenecks.

**Theorem 8** *If $l \in L_1(i)$ for some $i$ then*

$$\frac{\alpha}{\gamma\rho_l + \alpha} \;\le\; \frac{y_l}{c_l} \;\le\; 1 - \frac{\gamma\sigma_l(\rho_l - \sigma_l)}{\gamma\rho_l + \alpha}$$

**Proof:** Noting that $\rho_l < 1$ (and that $2(\gamma\rho_l + \alpha) > 0$), removing the last term from the square root in (15) gives the lower bound:

$$\begin{aligned}\frac{y_l}{c_l} \;&\ge\; \frac{\alpha + (\gamma+\alpha)\sigma_l + |(\alpha - (\gamma+\alpha)\sigma_l)|}{2(\gamma\rho_l + \alpha)} \\ &\ge\; \frac{\alpha}{\gamma\rho_l + \alpha}\end{aligned} \tag{29}$$

where the second inequality is an equality if $\sigma \le \alpha/(\gamma+\alpha)$.

To derive the upper bound, first note that $\rho_l \ge \sigma_l$ from Lemma 2(2). Since $2(1-\rho_l)(\gamma\rho_l+\alpha) > 0$ and $\Xi \ge 0$, removing the last term from the square root of (16) yields

$$\frac{r^l N_l}{c_l} \le \frac{1-\sigma_l}{1-\rho_l} - \frac{\gamma\sigma_l(\rho_l - \sigma_l)}{(1-\rho_l)(\gamma\rho_l + \alpha)} \tag{30}$$

Multiplying both sides by $(1-\rho_l)$ and adding $\sigma_l$ lead to the upper bound on utilization

$$\frac{y_l}{c_l} \le 1 - \frac{\gamma\sigma_l(\rho_l - \sigma_l)}{\gamma\rho_l + \alpha} \tag{31}$$

□

Substituting either $\gamma = 0$ or $\rho_l = \sigma_l$ into either the exact expressions (15) and (16) or the upper and lower bounds (29) and (30) gives full utilization as in the max-min case: $y_l = c_l$ and $r^l = c_l(1-\sigma_l)/(N_l(1-\rho_l))$. This shows that XCP could be made to approach max-min fairness if the bandwidth shuffling were reduced.

On the other hand, link utilization could be arbitrarily low if $\alpha$ and $\gamma$ had been chosen poorly. With the values suggested in [8] however the utilization is at least 80%. Consider a network of two links. Link 1 has $c_1 = 1$ and carries $N_1$ flows, while link 2 has $c_2 = 1 + \gamma/\alpha$ and carries $N_2 = N_1 + 1$ flows, consisting of all the traffic on link 1 plus one other flow. As $N_1 \to \infty$ we get $\rho_2 \to 1$. This gives $\sigma_2 = \alpha/(\gamma + \alpha)$ in the limit. Thus, both terms in the square root of (15) go to zero, and (29) becomes tight, and $y_l/c_l \to 0$ as $\gamma/\alpha \to \infty$. However, with $\alpha = 0.4$ and $\gamma = 0.1$ [8], (15) gives $y_l/c_l = 0.8$.

Similarly, a given flow may obtain an arbitrarily small proportion of its max-min fair bandwidth for *any* $\alpha > 0$ and $\gamma > 0$. The ratio of the upper bound on XCP bandwidth (30) to the max-min fair bandwidth, $r^{l,mm} = c_l(1-\sigma_l)/(N_l(1-\rho_l))$, is minimized with respect to $\sigma_l$ when $\rho_l = 2\sigma_l - \sigma_l^2$. Substituting this value into (16) and dividing by $r^{l,mm}$ gives

$$\frac{r^l}{r^{l,mm}} = \frac{C - \frac{D}{1-\sigma_l} + \sqrt{\left[C + \frac{D}{1-\sigma_l}\right]^2 - \frac{E}{1-\sigma_l}}}{2(\gamma(2\sigma_l - \sigma_l^2) + \alpha)} \qquad (32)$$

where

$$\begin{aligned} C &= \Xi_l/(1-\sigma_l) = \gamma\sigma_l(1-\sigma_l) + \alpha, \\ D &= \gamma\sigma_l^2(2-\sigma_l) \\ E &= 4\gamma\sigma_l^2\alpha. \end{aligned}$$

Thus

$$\frac{r^l}{r^{l,mm}} = \frac{C - \frac{D}{1-\sigma_l} + \frac{D}{1-\sigma_l}\sqrt{1 + \frac{(1-\sigma_l)^2}{D^2}\left(C^2 + \frac{2CD-E}{1-\sigma_l}\right)}}{2(\gamma(2\sigma_l - \sigma_l^2) + \alpha)}$$

Applying the identity $\sqrt{1+x} \le 1 + x/2$, for $x \ge -1$, gives

$$\frac{r^l}{r^{l,mm}} \le \frac{C + \dfrac{1-\sigma_l}{2D}C^2 + \dfrac{2CD-E}{2D}}{2(\gamma\sigma_l(2-\sigma_l) + \alpha)}. \qquad (33)$$

In the limit as $\sigma_l \to 1$, the right hand side tends to 0 for any $\gamma \neq 0$. This demonstrates that, for any non-zero amount of bandwidth shuffling, XCP can be arbitrarily unfair for some topology.

Hence, although the equilibrium of (4) converges to max-min as $\gamma \to 0$, this convergence is not uniform with respect to topology. In other words, given any topology specified by $(R, c)$, we can choose $\gamma$ sufficiently small so that the resulting allocation is close to max-min fairness. However, for any fixed $\gamma > 0$, such as 0.1 used by XCP, there are topologies in which some source rates can be far away from their max-min allocations.

This behavior can be exhibited by a simple two link network: one link has capacity 1 and carries $n^2$ flows, while the other carries $n^2 - 1$ of those same flows and has capacity $(n-1)/n$. This network has $\sigma_2 = (n-1)/n$ and $\rho_2 = (n^2-1)/n^2 = 2\sigma_2 - \sigma_2^2$. Hence, $\sigma_2 \to 1$ as $n \to \infty$ and $r^2/r^{2,mm} \to 0$.

These asymptotic results will be illustrated and confirmed by simulation in the following section.

# 5 Simulation results

In this section, we present simulation results using the implementation available from [8] for NS-2 [14]. These results verify the accuracy of our algorithm in Section 3.3 and confirm our qualitative discussion in Section 4 on the utilization and fairness properties of XCP.

All sources always have packets to send. All links have equal propagation delay of $d_l$ in both directions. The variable `avg_rrt_` in the XCP implementation ($d$ in the analysis) is fixed to the maximum of all RTTs in the network, $4 \cdot d_l$. The XCP default parameters $\alpha = 0.4$, $\beta = 0.226$ and $\gamma = 0.1$ are used. All our simulations use the unmodified XCP code in NS-2 that includes the "residual" terms. Though these terms are omitted in our model, as we remark in Section 2.2, they do not impact significantly the equilibrium properties. Hence, the simulation results agree well with theoretical predictions, as we now show.

The topology used for Scenarios 1 and 2 is shown in Figure 1 and consists of two links, with $i + j$ sources traversing link L1 and $j$ sources traversing L2.

Scenario 1 investigates the utilization of L1 as the number of sources traversing L1 and L2 is changed. In the experiment $i \ge j$, with $c_1$ =155Mbps and $c_2$ =100Mbps. The utilization of L1 for a range of $i$ and $j$ is shown in Figure 2. A max-min fair allocation would result in a full utilization of L1 for all $i$ and $j$ combinations. However, as the number of sources bottlenecked at L2 increases, XCP's utilization of L1 decreases.

Since XCP's "residual" terms depend on feedback from upstream nodes, the equilibrium rates depend on the order in which links are traversed. If the direction of flow in this network were reversed, then the utilization would be 0–4% higher than for the case considered (and the theoretical predictions will be 0–4% lower than the simulation results).

Scenario 2 demonstrates that XCP can be arbitrarily unfair for some topology. Let $c_1$ =155Mbps, $c_2 = c_1(n-1)/n$, $i = n^2 - 1$ and $j = 1$. The ratio of the rate of the source traversing only L1 to the max-min fair rate is plotted in Figure 3. Indeed the unfairness increases with the number of sources in the network, confirming the theory.

Scenario 3 demonstrates the claim that as $\gamma/\alpha \to \infty$, $y_2/c_2 \to 0$. We set $c_2$ =200Mbps, $c_1 = c_2(1 + \gamma/\alpha)$, $i = 256$ and $j = 1$. The parameter $\alpha$ is varied from $0.512$ to $0.016$ and the utilisation of L1 as a function of $\gamma/\alpha$, as well as the upper and lower bounds from ( 31) and ( 29), are plotted in Figure 4.

Scenario 4 tests the rate allocation algorithm for a more complicated topology as shown in 5. The link capacities in Mbps are $c_1 = 10$, $c_3 = 8$, $c_4 = 8$, $c_5 = 7$, $c_6 = 6$ and $c_2$ is varied in this experiment. Delay is set to $d_l$ =10ms. The source rates are plotted in Figure 6. There is a good agreement between the predicted and measured rates even though the lower bandwidth delay product makes the fluid flow approximation more questionable.

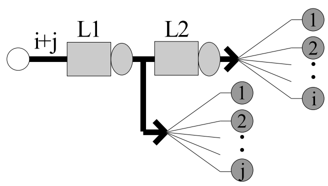


Figure 1: Topology for Scenarios 1 and 2.

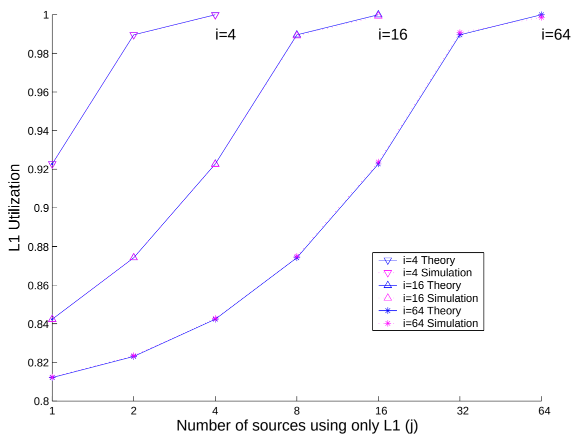


Figure 2: Scenario 1: utilization.

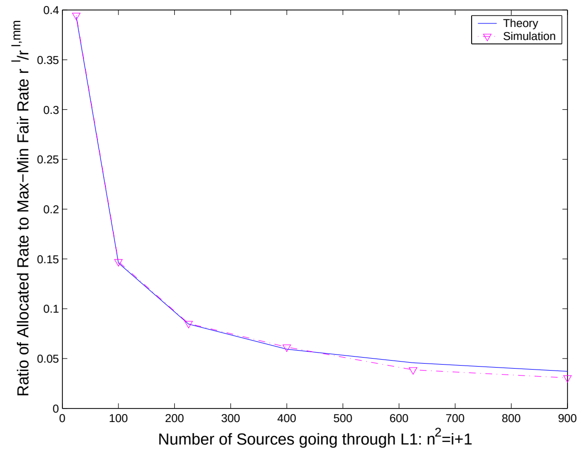


Figure 3: Scenario 2: unfairness, $r^2/r^{2,mm} \to 0$ as $n \to \infty$.

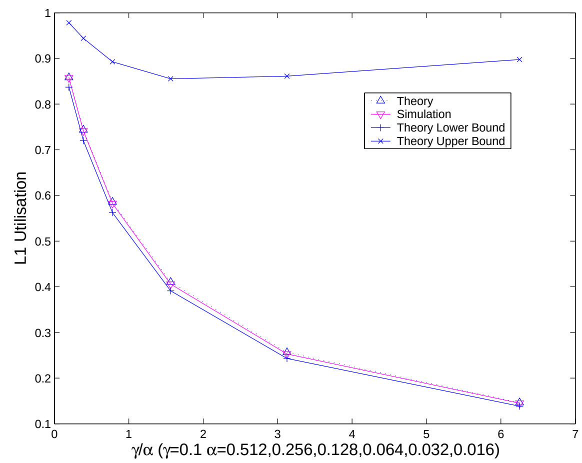


Figure 4: Scenario 3: utilisation of L1 as function of $\gamma/\alpha$.

# 6 Conclusion

We have presented a dynamic model of XCP and used it to completely characterize its equilibrium properties. We have shown that XCP clears the queues in equilibrium, and has unique equilibrium rates that solve a constrained max-min fairness problem. The additional constraint under XCP can lead to unfairness for some network topologies. XCP gives a utilization of at least 80%, but a poor choice of $\alpha$ or $\gamma$ could lead to arbitrarily low utilization. We have provided an algorithm to compute the equilibrium for general networks, and have presented simulation results to illustrate these findings.

An important question that we have not pursued is the dynamic properties of XCP, such as its stability. Even though the "residual" terms in XCP code do not seem to affect equilibrium properties drastically, they may be important in determining its dynamic properties, and hence should be taken into account in such an analysis. It is important to understand the stability of individual source rates $x_i(t)$, in addition to the aggregate rate $y_l(t)$, as studied in [8], in general networks in the presence of delay. Since equilibrium queues are zero, the usual practice of linearizing around the equilibrium needs caution at the tightest bottlenecks that have zero queue yet full utilization.

# A Valid equilibrium rate

**Proof (Lemma 2) :**

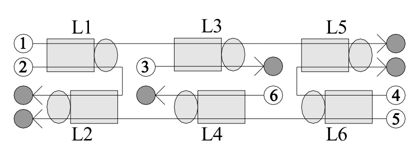


Figure 5: Scenario 3 topology.

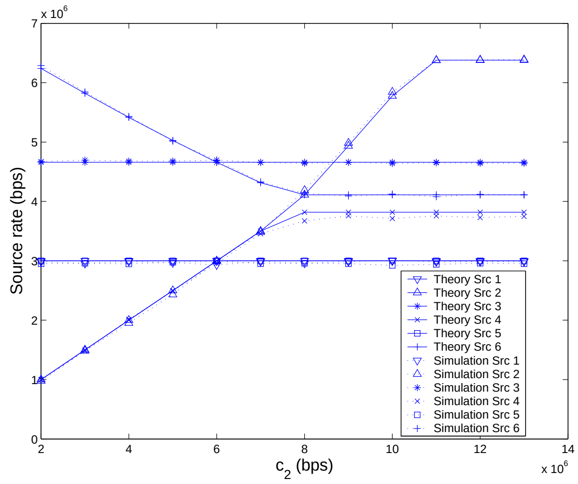


Figure 6: Scenario 4: throughputs.

1. By Lemma 6, $F_{li} > 0$ and $F_{lj} = 0$. This implies

$$\frac{h_l + \phi_l^+}{N_l x_i} \;>\; \frac{h_l + \phi_l^-}{y_l} \;=\; \frac{h_l + \phi_l^+}{N_l x_j}$$

2. If this were not the case, then the average rate of flows in $I_0(l)$ (and hence the rate of at least one such flow) would be greater than $c_l/N_l$. Part 1 would then require the total rate to exceed the capacity.
3. From part 1,

$$\begin{aligned} N_l r^l &= N_{l1} r^l + N_{l0} r^l \\ &\geq N_{l1} r^l + y_{l0} \;=\; y_l \end{aligned}$$

   with equality if and only if $y_{l0} = 0$.
4. Otherwise, $h_l = 0$ and $\phi_l = 0$ since $I_1(l) \neq \emptyset$ (by Lemma 1). Then $y_l = c_l$ and $b_l = 0$, whence $h_l = \max(\gamma d c_l - 0, 0) > 0$.
5. $\sigma_l = y_{l0}/c_l \leq y_l/c_l$ with equality if and only if $y_{l0} = y_l$. □

**Lemma 8** *The smaller solution to (18) does not give a valid equilibrium rate for link $l$.*

**Proof:** For a rate, $y_l$, to be valid requires $y_l/c_l > \sigma_l$, by Lemma 2(5).

Let $\theta = (1+\gamma/\alpha)\sigma_l$ and $\psi = (1+\rho_l\gamma/\alpha)\sigma_l \leq \theta$. Then dividing the numerator and denominator of (18) by $\alpha$ gives

$$\frac{y_l}{c_l} = \frac{1+\theta \pm \sqrt{(1+\theta)^2 - 4\psi}}{2\psi/\sigma_l}.$$

If the lower root were valid, then

$$1 + \theta - 2\psi > \sqrt{(1+\theta)^2 - 4\psi}$$

whence

$$(1+\theta)^2 - 4(1+\theta)\psi + 4\psi^2 > (1+\theta)^2 - 4\psi$$

But $0 \leq \psi \leq \theta$, which yields a contradiction. Thus the lower root is not valid. □

# B Properties of optimization problems

**Proof (Lemma 4) :** Note that $g_l(x) = g_l(y_l)$ depends on $x$ only through $y_l$. Now $x \in X_n$ are all of the form

$$x_i \;=\; \begin{cases} r_m & i \in I_m, m \leq n \\ r_n + \epsilon_i & i \notin \overline{I}_n \end{cases}$$

where $\epsilon_i > 0$. Write $x \in X_n$ as $x(\epsilon_i, i \notin \overline{I}_n)$. Hence we can write $y_l$ also as a function of $(\epsilon_i, i \notin \overline{I}_n)$:

$$\begin{aligned} & y_l\left(\epsilon_i, i \notin \overline{I}_n\right) \\ &= \sum_{m=1}^{n} r_m \sum_{i \in I_m} R_{li} + \sum_{i \notin \overline{I}_n} R_{li}(r_n + \epsilon_i) \\ &= \sum_{m=1}^{n-1} r_m \sum_{i \in I_m} R_{li} + r_n \sum_{i \notin \overline{I}_{n-1}} R_{li} + \sum_{i \notin \overline{I}_n} R_{li}\epsilon_i \qquad (34) \end{aligned}$$

Given any $(\epsilon_i, i \notin \overline{I}_n)$, define the average $\epsilon$ by

$$\epsilon \cdot \sum_{i \notin \overline{I}_{n-1}} R_{li} \;=\; \sum_{i \notin \overline{I}_n} R_{li}\epsilon_i$$

and consider the vector $(\epsilon_i = \epsilon, i \notin \overline{I}_n)$ with equal components. From (34), this vector produces the same link flow $y_l$. Moreover, $x(\epsilon) := x(\epsilon_i = \epsilon, i \notin \overline{I}_n)$ defined by this vector also satisfies $x_i(\epsilon) \leq g_l(y_l)$ for all $i \in I(l)$, and hence is in $X_n$. This is because for $i \in I_m$, $m \leq n$, $x_i(\epsilon) = r_m \leq g_l(y_l)$ since the original $x((\epsilon_i, i \notin \overline{I}_n))$ is in $X_n$. For $i \notin \overline{I}_n$,

$$x_i(\epsilon) = r_n + \epsilon \;\leq\; r_n + \max_i \epsilon_i \;\leq\; g_l(y_l)$$

where the first inequality follows because $\epsilon$ is the average of $\epsilon_i$, and the last inequality follows because the original $x((\epsilon_i, i \notin \overline{I}_n))$ is in $X_n$. Hence, if $(\epsilon_i^*, i \notin \overline{I}_n)$ achieves the maximum in (9), the vector $(\epsilon_i^* = \epsilon^*, i \notin \overline{I}_n)$ with a common value $\epsilon^*$ also achieves the maximum. □

**Proof (Lemma 5) :** Define

$$f(\epsilon) \;:=\; \hat{g}_l(\epsilon) - (s + \eta\epsilon)$$

We will show that there exists $\epsilon_l \geq 0$, with equality if and only if $s = \hat{g}_l(0)$, such that $f(\epsilon_l) = 0$ under the conditions given in the lemma. Consider the two cases separately.

**Case 1:** $s \le \hat{g}_l(0)$. If $s = \hat{g}_l(0)$ then $f(0) = 0$ and $\epsilon_l = 0$. Otherwise, $f(0) > 0$ and it remains to show that there exists a suitable $\epsilon_l > 0$. Consider

$$\begin{aligned}
\lim_{\epsilon\to\infty} f(\epsilon) &= \lim_{\epsilon\to\infty} \hat{g}_l(\epsilon) - (s + \eta\epsilon) \\
&= \lim_{\epsilon\to\infty} \frac{\gamma(z+\eta\epsilon)^2}{N_l((\gamma+\alpha)(z+\eta\epsilon) - \alpha c_l)} - (s+\eta\epsilon) \\
&= \lim_{\epsilon\to\infty} \epsilon \cdot \\
&\quad \left(\lim_{\epsilon\to\infty} \frac{\gamma(z/\epsilon+\eta)^2}{N_l((\gamma+\alpha)(z/\epsilon+\eta) - \alpha c_l/\epsilon)} - (s/\epsilon+\eta)\right) \\
&= \lim_{\epsilon\to\infty} \epsilon \cdot \eta \left(\frac{1}{N_l}\frac{\gamma}{\gamma+\alpha} - 1\right) \\
&= -\infty
\end{aligned}$$

since $N_l \ge 1$. This implies that there exists $\epsilon_1$ sufficiently large such that $f(\epsilon_1) < 0$. Since $f$ is continuous on $[0, \epsilon_1]$, there exists an $\epsilon_l \in (0, \epsilon_1)$ such that $f(\epsilon_l) = 0$. This is illustrated in Figure 7.

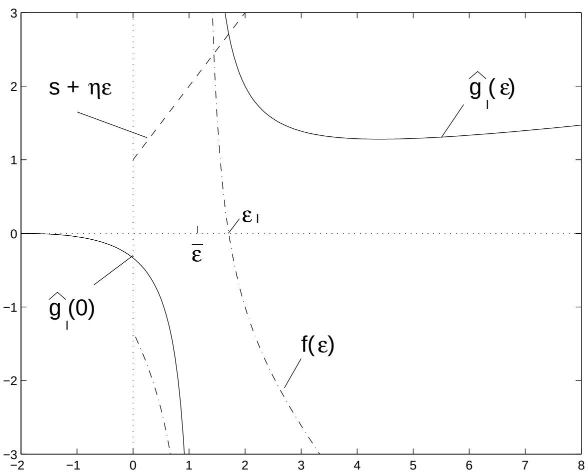


Figure 7: $\hat{g}_l(\epsilon)$ in (23) and $f(\epsilon) = \hat{g}_l(\epsilon) - (s + \eta\epsilon)$.

**Case 2:** $\hat{g}_l(0) < 0$. Then

$$z \quad < \quad \frac{\alpha c_l}{\gamma+\alpha}$$

and hence

$$\bar{\epsilon} \quad := \quad \frac{1}{\eta}\left(\frac{\alpha c_l}{\gamma+\alpha} - z\right) \ > \ 0$$

Moreover, as $\epsilon \to \bar{\epsilon}$ from above, $\hat{g}_l(\epsilon) \to +\infty$. Since $\hat{g}_l(\epsilon)$ is continuous for $\epsilon > \bar{\epsilon}$, there exists $\epsilon_0 > \bar{\epsilon}$ such that $s + \eta\epsilon_0 < \hat{g}_l(\epsilon_0)$, i.e., $f(\epsilon_0) > 0$. The same argument as in Case 1 shows that there exists an $\epsilon_1$ with $f(\epsilon_1) < 0$. Since $\hat{g}_l$ is continuous on $[\epsilon_0, \epsilon_1]$, there exists $\epsilon_l \in (\epsilon_0, \epsilon_1)$ where $f(\epsilon_l) = 0$.

We now prove that $\epsilon_l$ is unique. The argument also shows that over $\{\epsilon \mid \hat{g}_l(\epsilon) > 0\}$, $s + \eta\epsilon > \hat{g}_l(\epsilon)$ if and only if $\epsilon > \epsilon_l$. First note that $\bar{\epsilon}$ may be negative. They key observation, illustrated in Figure 7, is that

1. for $\epsilon < \bar{\epsilon}$, $\hat{g}_l(\epsilon)$ is negative, concave, and approaches $-\infty$ as $\epsilon \to \bar{\epsilon}$ from below.
2. for $\epsilon > \bar{\epsilon}$, $\hat{g}_l(\epsilon)$ is positive, convex, attains its minimum at

$$\epsilon \quad = \quad \frac{1}{\eta}\left(\frac{2\alpha c_l}{\gamma+\alpha} - z\right)$$

and approaches $+\infty$ as $\epsilon \to \bar{\epsilon}$ from above.

A moment of thought then convinces one that it suffices to show that $\hat{g}_l(\epsilon)$ grows less rapidly than $s + \eta\epsilon$, i.e.,

$$\frac{d\hat{g}_l}{d\epsilon} \quad < \quad \eta \tag{35}$$

Letting $y_l = z + \eta\epsilon$, we have

$$\frac{d\hat{g}_l}{d\epsilon} \quad = \quad \frac{\eta}{N_l} \cdot \frac{\gamma y_l((\gamma+\alpha)y_l - 2\alpha c_l)}{((\gamma+\alpha)y_l - \alpha c_l)^2}$$

Since $N_l \ge 1$, if (35) is violated, the above implies that

$$\frac{\gamma y_l((\gamma+\alpha)y_l - 2\alpha c_l)}{((\gamma+\alpha)y_l - \alpha c_l)^2} \quad \ge \quad 1$$

which implies

$$\alpha^2(y_l + c_l)^2 + \alpha\gamma y_l^2 \quad \le \quad 0$$

which is a contradiction since $c_l,\ \alpha > 0$ and $\gamma,\ y_l \ge 0$. Hence, (35) must hold and $s + \eta\epsilon > g_l(\epsilon)$ if and only if $\epsilon > \epsilon_l$ whenever $g_l(\epsilon) > 0$. $\square$

**Proof (Lemma 6) :** The first assertion follows directly from the definitions of $L_n$, $I_n$ and $\hat{X}_n$ in (10), (11) and (20), respectively. We will prove the second assertion by induction on $n$.

**Base case** $n = 1$**:** Fix any $x \in \hat{X}_1$.

Consider an $l \notin L_1 = \overline{L}_1$. Then $c_l/N_l > r_1$ by (13). Now

$$x_i(\epsilon; 1) \quad = \quad \begin{cases} r_1 & i \in I_1 \\ r_1 + \epsilon & i \notin I_1 \end{cases}$$

Hence

$$g_l(x(0;1)) \quad = \quad \frac{\gamma(N_l r_1)^2}{N_l((\gamma+\alpha)N_l r_1 - \alpha c_l)}$$

If

$$N_l r_1 \quad < \quad \frac{\alpha}{\gamma+\alpha} c_l$$

then $g_l(x(0;1)) < 0$. Otherwise, $g_l(x(0;1)) > 0$. We claim that $x_i(0,1) = r_1 < g_l(x(0;1))$ for all $i \in I(l)$. If not, then

$$r_1 \quad \ge \quad \frac{\gamma(N_l r_1)^2}{N_l((\gamma+\alpha)N_l r_1 - \alpha c_l)}$$

yielding $N_l r_1 \geq c_l$, a contradiction. Hence, if $l \notin \overline{L}_1$, then either $g_l(x(0;1)) < 0$ or $x_i(0;1) < g_l(x(0;1))$ for all $i \in I(l)$.

**Induction hypothesis:** Suppose the second assertion holds for $n$. We will prove it for $n+1$.

**Induction:** Fix an $l \notin \overline{L}_{n+1}$. First note that $l \notin \overline{L}_n$. Consider all links $\hat{l} \notin \overline{L}_n$. By the induction hypothesis, either $g_{\hat{l}}(x(0;n)) < 0$ or $x_i(0;n) < g_{\hat{l}}(x(0;n))$ for all $i \in I(\hat{l})$. For an $i \notin \overline{I}_n$, we have from (21), for $\epsilon \geq 0$,

$$\begin{aligned} x_i(\epsilon;n) &= r_n + \epsilon \\ y_{\hat{l}}(\epsilon;n) &= z_{\hat{l}} + \eta_{\hat{l}}\epsilon_{\hat{l}} \end{aligned}$$

where, from (22),

$$\begin{aligned} z_{\hat{l}} &:= \sum_{m=1}^{n} r_m \sum_{i \in I_m} R_{\hat{l}i} + r_n \sum_{i \notin \overline{I}_n} R_{\hat{l}i} \\ \eta_{\hat{l}} &:= \sum_{i \notin \overline{I}_n} R_{\hat{l}i} \end{aligned}$$

From (23), $\hat{g}_{\hat{l}}(0) = g_{\hat{l}}(z_{\hat{l}}) = g_{\hat{l}}(x(0;n)) < 0$, and hence the induction hypothesis implies that either $\hat{g}_{\hat{l}}(0) < 0$ or $r_n < \hat{g}_{\hat{l}}(0)$. Lemma 5 then implies that, for each $\hat{l} \notin \overline{L}_n$, there exists a unique $\epsilon_{\hat{l}} > 0$ such that $r_n + \epsilon_{\hat{l}} = \hat{g}_{\hat{l}}(\epsilon_{\hat{l}}) = g_{\hat{l}}(z_{\hat{l}} + \eta_{\hat{l}}\epsilon_{\hat{l}})$. The minimum $g_{\hat{l}}(z_{\hat{l}} + \eta_{\hat{l}}\epsilon_{\hat{l}})$ over $\hat{l} \notin \overline{L}_n$ is $r_{n+1}$, and these minimizing $\hat{l}$ constitute $L_{n+1}$. All the sources $i \notin \overline{I}_n$ that go through a link in $L_{n+1}$ are assigned the common rate $r_{n+1}$, and they are collected into $X_{n+1}$.

Since $l \notin \overline{L}_n$ and $l \notin L_{n+1}$, the corresponding $\epsilon_l$ satisfies

$$r_n + \epsilon_l = g_l(z_l + \eta_l \epsilon_l) \tag{36a}$$

$$r_n + \epsilon_l > r_{n+1} = r_n + \min_{\hat{l} \notin \overline{L}_n} \epsilon_{\hat{l}} \tag{36b}$$

But $x_i(\epsilon; n+1) = r_{n+1} + \epsilon$ for all $i \notin \overline{I}_{n+1}$. Hence (36) and Lemma 5(2) (with $z = y_l(0;n)$, $s = r_n$, $\eta = \sum_{i \notin \overline{I}_n} R_{li}$) imply that either $g_l(x(0;n+1)) \leq 0$ or $x_i(0;n+1) = r_n + \min_{\hat{l} \in \overline{L}_n} \epsilon_{\hat{l}} < \hat{g}_l(\min_{\hat{l} \in \overline{L}_n} \epsilon_{\hat{l}}) = g_l(x(0;n+1))$. The proof is completed by noting that $g_l(x(0;n+1)) \neq 0$, since $y_l \geq r_1 > 0$. □

# C Correctness of rate algorithm

To establish the correctness of the algorithm in Section 3.3 to find XCP's equilibrium rates, it is sufficient to show that the calculated values satisfy the two conditions stated in Section 3.2.

Each source $i$ is assigned a rate in the same step as a particular link $l$. If $r^l$ is chosen according to (16) with the true equilibrium values of $\rho_l$ and $\sigma_l$, then $F_{li} = 0$, since that is the condition from which (16) was derived. Similarly, (16) implies $y_l \leq c_l$, as established in the proof of Theorem 8.

The correctness of the algorithm can thus be established by showing that the true equilibrium values of $\sigma_l$ and $\rho_l$ are used when the final value $r^l$ is calculated in step 2.4.1. For each link $l$, the values of $\sigma_l$ and $\rho_l$ depend only on network parameters and flows $i$ with rates $x_i < r^l$, by Lemma 2(1). That is, if the rates selected by the algorithm are such that the rate $r^l$ of each link $l$ is greater than the rates of the flows flowing through $l$ but bottlenecked elsewhere, then the rates must form an equilibrium of XCP. The theorem then results from the following lemma.

**Lemma 9** *For each $l$ and $n$ for which $r^l(n+1)$ is defined, $r^l(n+1) \geq r_n$. Moreover, if $\gamma < \alpha$, then for each $l$ and $n$ for which $r^l(n+1)$ is defined, $r^l(n+1) \geq r^l(n)$.*

**Proof:** Consider an arbitrary link, $l$, and iteration, $n$. If $l$ does not carry any flows in $\Delta(n)$, then $r^l(n+1) = r^l(n) > r_n$ as required. Consider now the case that $l$ does carry a flow in $\Delta(n)$.

Let $x = r_n N_l / c_l$ be the rate of flows allocated in iteration $n$, normalized to link $l$. Let $\rho(u) = \rho_l(n) + u$ and $\sigma(u) = \sigma_l(n) + ux$. Then $\rho(0)$ and $\sigma(0)$ are the fraction of allocated flows and allocated capacity on link $l$ before step 2.4.1 of iteration $n$, while $\rho(\delta)$ and $\sigma(\delta)$ are the values after the update, where $\delta = \rho_l(n+1) - \rho_l(n) > 0$.

Let $z(u)$ be the value of $r^l N_l / c_l$ calculated from (19) using $\rho(u)$ and $\sigma(u)$. Note that $z \geq 1$ by Lemma 2(3), since $c_l \geq y_l$. Moreover, $x < z(0)$, since if $z(0) \leq x$ then link $l$ would have been an element of $M(n)$ and $r^l(n+1)$ would not be defined. To prove the lemma, it is sufficient to show that $z(\delta) \geq x$ and that if $\gamma > \alpha$ then $dz/du \geq 0$.

Below, the argument $(u)$ will be dropped when no ambiguity can arise. Differentiating (19) with respect to $u$, and noting that $z(u)$ is the larger of the two solutions of (19), gives

$$\begin{aligned} 0 &= \frac{dA_l}{du} z^2 + 2zA_l \frac{dz}{du} + \frac{dB_l}{du} z + B_l \frac{dz}{du} + \frac{dC_l}{du} \\ &= \sqrt{B_l^2 - 4A_l C_l} \frac{dz}{du} + \frac{dA_l}{du} z^2 + \frac{dB_l}{du} z + \frac{dC_l}{du}. \end{aligned}$$

Now

$$\begin{aligned} &\frac{dA_l}{du} z^2 + \frac{dB_l}{du} z + \frac{dC_l}{du} \\ &\quad = [\gamma(1-\rho) - (\gamma\rho + \alpha)] z^2 \\ &\quad\quad + [2\gamma\sigma - \gamma x(1-\rho) + x(\gamma\rho + \alpha)] z - 2\gamma\sigma x \\ &\quad = (x - z)[z(\alpha - \gamma) + 2\gamma(z\rho - \sigma)] \end{aligned}$$

giving

$$\frac{dz}{du} = (z - x) \frac{[z(\alpha - \gamma) + 2\gamma(z\rho - \sigma)]}{\sqrt{B_l^2 - 4A_l C_l}}$$

If $\gamma < \alpha$ then $dz/du > 0$, and $z$ remains greater than $x$. This is because $2\gamma(z\rho - \sigma) \geq 0$, since $z \geq 1$ and $0 \leq \sigma \leq \rho$. This establishes the second part of the lemma.

If $\gamma \geq \alpha$ then the right hand side need not be positive. However, the second factor is bounded, and so $z$ approaches $x$ exponentially as $u$ increases, and so can never drop below $x$. In particular, $z(\delta) > x$, which establishes the first part of the lemma. $\square$

# D Acknowledgment

This work is supported by NSF, Caltech, ARO, AFOSR, and Cisco as part of the FAST Project, and supported by the Australian Research Council. We thank Dina Katabi of MIT for helpful discussions.

# References

[1] V. Jacobson, "Congestion avoidance and control," *Proceedings of SIGCOMM'88, ACM*, August 1988. An updated version is available via `ftp://ftp.ee.lbl.gov/papers/congavoid.ps.Z`.

[2] C. Hollot, V. Misra, D. Towsley, and W. Gong, "Analysis and design of controllers for AQM routers supporting TCP flows," *IEEE Transactions on Automatic Control*, vol. 47, no. 6, pp. 945–959, 2002.

[3] S. H. Low, F. Paganini, J. Wang, and J. C. Doyle, "Linear stability of TCP/RED and a scalable control," *Computer Networks Journal*, vol. 43, no. 5, pp. 633–647, 2003. `http://netlab.caltech.edu`.

[4] C. Casetti, M. Gerla, S. Mascolo, M. Sansadidi, and R. Wang, "TCP Westwood: end-to-end congestion control for wired/wireless networks," *Wireless Networks Journal*, vol. 8, pp. 467–479, 2002.

[5] S. Floyd, "HighSpeed TCP for large congestion windows." Internet draft draft-floyd-tcp-highspeed-02.txt, work in progress, `http://www.icir.org/floyd/hstcp.html`, February 2003.

[6] C. Jin, D. X. Wei, and S. H. Low, "TCP FAST: motivation, architecture, algorithms, performance," in *Proceedings of IEEE Infocom*, March 2004. `http://netlab.caltech.edu`.

[7] T. Kelly, "Scalable TCP: Improving performance in highspeed wide area networks." Submitted for publication, `http://www-lce.eng.cam.ac.uk/~ctk21/scalable/`, December 2002.

[8] D. Katabi, M. Handley, and C. Rohrs, "Congestion control for high-bandwidth delay product networks," in *Proc. ACM Sigcomm*, August 2002.

[9] B. Wydrowski and M. Zukerman, "MaxNet: A congestion control architecture for maxmin fairness," *IEEE Communications Letters*, vol. 6, pp. 512–514, November 2002.

[10] B. Wydrowski, L. L. H. Andrew, and M. Zukerman, "MaxNet: A congestion control architecture for scalable networks," *IEEE Communications Letters*, vol. 7, pp. 511–513, October 2003.

[11] S. Athuraliya, V. H. Li, S. H. Low, and Q. Yin, "REM: active queue management," *IEEE Network*, vol. 15, pp. 48–53, May/June 2001. Extended version in *Proceedings of ITC17*, Salvador, Brazil, September 2001. `http://netlab.caltech.edu`.

[12] D. Bertsekas and R. Gallager, *Data Networks*. Prentice-Hall Inc., 2nd ed. ed., 1992.

[13] E. M. Gafni and D. P. Bertsekas, "Dynamic control of session input rates in communication networks," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 1009–1016, Jan. 1984.

[14] "NS network simulator." `http://www.isi.edu/nsnam/ns/`.