Multipath TCP: Analysis and Design

Qiuyu Peng, Anwar Walid, Steven H. Low

***Abstract*—Multi-path TCP (MP-TCP) has the potential to greatly improve application performance by using multiple paths transparently. We propose a fluid model for a large class of MP-TCP algorithms and identify design criteria that guarantee the existence, uniqueness, and stability of system equilibrium. We clarify how algorithm parameters impact TCP-friendliness, responsiveness, and window oscillation and demonstrate an inevitable tradeoff among these properties. We discuss the implications of these properties on the behavior of existing algorithms and motivate a new design that generalizes existing algorithms and strikes a good balance among TCP-friendliness, responsiveness, and window oscillation. We illustrate our analysis and the behavior of the new algorithm using ns2 simulations.**

I. Introduction

Traditional TCP uses a single path through the network even though multiple paths are usually available in today's communication infrastructure; e.g., most smart phones are enabled with both cellular and WiFi access, and servers in data centers are connected to multiple routers. Multi-path TCP (MP-TCP) has the potential to greatly improve application performance by using multiple paths transparently. It is being standardized by the MP-TCP Working Group of the Internet Engineering Task Force (IETF) [2]. In this paper we present a fluid model of MP-TCP and study how protocol parameters affect structural properties such as the existence, uniqueness and stability of equilibrium, the tradeoffs among TCP friendliness, responsiveness and window oscillation. These properties motivate a new algorithm that generalizes existing MP-TCP algorithms.

Various congestion control algorithms have been proposed as an extension of TCP NewReno for MP-TCP. A straightforward extension is to run TCP NewReno on each subpath, e.g. [3], [4]. This algorithm however can be highly unfriendly when it shares a path with a single-path TCP user. This motivates the Coupled algorithm which is fair because it has the same underlying utility function as TCP NewReno, e.g. [5], [6]. It is found in [7] however that the Coupled algorithm respond slowly in a dynamic network environment. A different algorithm is proposed in [7] (which we refer to as the Max algorithm) which is more responsive than the Coupled algorithm and still reasonably friendly to single-path TCP user. See [8] for more references to early work on multi-path congestion control.

Our goal is to develop structural understanding of MP-TCP algorithms so that we can systematically tradeoff different properties such as TCP friendliness, responsiveness, and window oscillation that can be detrimental to applications that require a steady throughput. For single-path TCP, one can associate a strictly concave utility function with each source so that the congestion control algorithm implicitly solves a network utility maximization problem [8]–[10]. The convexity of this underlying utility maximization guarantees the existence, uniqueness, and stability of most single-path TCP algorithms. For MP-TCP, it will be shown that the utility maximization interpretation fails to hold in general, necessitating the need for a different approach to understanding the equilibrium properties of MP-TCP algorithms. Moreover the relations among different performance metrics, such as fairness, responsiveness and window oscillation, need to be clarified.

The main contributions of this paper are three-fold. First we present a fluid model that covers a broad class of MP-TCP algorithms and identify the exact property that allows an algorithm to have an underlying utility function. This implies that some MP-TCP algorithms, e.g., the Max algorithm [7], has no associated utility function. We prove conditions on protocol parameters that guarantee the existence and uniqueness of the equilibrium, and its asymptotical stability. Indeed algorithms that fail to satisfy these conditions, e.g. the Coupled algorithm, can be unstable and can have multiple equilibria as shown in [7]. Second we clarify how protocol parameters impact TCP friendliness, responsiveness, and window oscillation and demonstrate the inevitable tradeoff among these properties. Finally, based on our understanding of the design space, we propose a new MP-TCP algorithm that generalizes existing algorithms and strikes a good balance among these properties. These results are illustrated using ns2 simulations.

We now summarize our proposed MP-TCP algorithm. Each source s has a set of routes r. Each route r maintains a congestion window w_r and measures its round-trip time τ_r. The window adaptation is as follows:

- For each ACK on route $r \in s$,

$$w_r \leftarrow w_r + \frac{x_r}{\tau_r \left(\sum x_k\right)^2} \left(\frac{1+\alpha_r}{2}\right) \left(\frac{4+\alpha_r}{5}\right) \tag{1}$$

- For each packet loss on route $r \in s$,

$$w_r \leftarrow \max\left\{ w_r \left(1 - \frac{1}{2}\alpha_r\right), 1 \right\} \tag{2}$$

where $x_r := w_r/\tau_r$ and $\alpha_r := \frac{\max\{x_k\}}{x_r}$.

The rest of the paper is structured as follows. In Section II we develop a fluid model for MP-TCP and use it to model existing algorithms. In Section III we prove several structural properties, focusing on design criteria that determine the existence, uniqueness, and stability of system equilibrium, TCP-friendliness, responsiveness, window oscillation, and an

A preliminary version has appeared in [1].

Q. Peng and S. H. Low are with California Institute of Technology, Pasadena, CA, 91125, USA (Email:{qpeng,slow}@caltech.edu)

A. Walid is with Bell Laboratories, Murray Hill, NJ 07974, USA (Email: anwar@research.bell-labs.com)

inevitable tradeoff among these properties. In section IV we discuss the implications of these properties on the existing algorithms. This motivates our new MP-TCP algorithm and we explain our design rationale. In Section V we use ns2 simulations to compare the performance of the proposed algorithm with the existing algorithms. We conclude in Section VI.

II. Multipath TCP model

In this section we first propose a fluid model of MP-TCP and then use it to model MP-TCP algorithms in the literature. Unless otherwise specified, a boldface letter $\mathbf{x} \in \mathbb{R}^n$ denotes a vector with components x_i. We use $\mathbf{x}_{-i} := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ to denote the $n-1$ dimensional vector without x_i and $\|\mathbf{x}\|_k := (\sum x_i^k)^{1/k}$ to denote the L_k-norm of $\mathbf{x}$. Given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x} \geq \mathbf{y}$ means $x_i \geq y_i$ for all components i. A capital letter denotes a matrix or a set, depending on the context. A symmetric matrix P is said to be *positive (negative) semidefinite* if $\mathbf{x}^T P \mathbf{x} \geq 0 (\leq 0)$ for any $\mathbf{x}$, and *positive (negative) definite* if $\mathbf{x}^T P \mathbf{x} > 0 (< 0)$ for any $\mathbf{x} \neq \mathbf{0}$. For any matrix P, define $[P]^+ := (P + P^T)/2$ to be its symmetric part. Given two arbitrary matrices A and B (not necessarily symmetric), $A \succeq B$ means $[A - B]^+$ is positive semidefinite. For a vector $\mathbf{x}$, $diag\{\mathbf{x}\}$ is a diagonal matrix with entries given by $\mathbf{x}$.

A. Fluid model

Consider a network that consists of a set $L = \{1, \ldots, L\}$ of links with finite capacities c_l. The network is shared by a set $S = \{1, \ldots, S\}$ of sources. Available to source $s \in S$ is a fixed collection of routes (or paths) r. A route r consists of a set of links l. We abuse notation and use s both to denote a source and the set of routes r available to it, depending on the context. Likewise, r is used both to denote a route and the set of links l in the route. Let $R := \{r \mid r \in s, s \in S\}$ be the collection of all routes. Let $H \in \{0,1\}^{|L| \times |R|}$ be the routing matrix: $H_{lr} = 1$ if link l is in route r (denoted by '$l \in r$'), and 0 otherwise.

For each route $r \in R$, τ_r denotes its round trip time (RTT). For simplicity we assume τ_r are constants. Each source s maintains a congestion window $w_r(t)$ at time t for every route $r \in s$. Let $x_r(t) := w_r(t)/\tau_r$ represent the sending rate on route r. Each link l maintains a congestion price $p_l(t)$ at time t. Let $q_r(t) := \sum_{l \in L} H_{lr} p_l(t)$ be the aggregate price on route r. In this paper $p_l(t)$ represents the packet loss probability at link l and $q_r(t)$ represents the approximate packet loss probability on route r.

We associate three state variables $(x_r(t), w_r(t), q_r(t))$ for each route $r \in s$. Let $\mathbf{x}_s(t) := (x_r(t), r \in s)$, $\mathbf{w}_s(t) := (w_r(t), r \in s)$, $\mathbf{q}_s(t) := (q_r(t), r \in s)$. Then $(\mathbf{x}_s(t), \mathbf{w}_s(t), \mathbf{q}_s(t))$ represents the corresponding state variables for each source $s \in S$. For each link l, let $y_l(t) := \sum_{r \in R} H_{lr} x_r(t)$ be its aggregate traffic rate.

Congestion control is a distributed algorithm that adapts $\mathbf{x}(t)$ and $\mathbf{p}(t)$ in a closed loop. Motivated by the AIMD algorithm of TCP Newreno, we model MP-TCP by

$$\dot{x}_r = k_r(\mathbf{x}_s) \left(\phi_r(\mathbf{x}_s) - q_r\right)^+_{x_r} \qquad r \in s \quad s \in S \qquad (3)$$

$$\dot{p}_l = \gamma_l \left(y_l - c_l\right)^+_{p_l} \qquad l \in L, \qquad (4)$$

where $(a)^+_x = a$ for $x > 0$ and $\max\{0, a\}$ for $x \leq 0$. We omit the time t in the expression for simplicity. (3) models how sending rates are adapted in the congestion avoidance phase of TCP at each end system and (4) models how the congestion price is (often implicitly) updated at each link. The MP-TCP algorithm installed at source s is specified by (K_s, Φ_s), where $K_s(\mathbf{x}_s) := (k_r(\mathbf{x}_s), r \in s)$ and $\Phi_s(\mathbf{x}_s) := (\phi_r(\mathbf{x}_s), r \in s)$. Here $K_s(\mathbf{x}_s) \geq 0$ is a vector of positive gains that determines the dynamic property of the algorithm. $\Phi_s(\mathbf{x}_s)$ determines the equilibrium properties of the algorithm. The link algorithm is specified by γ_l, where $\gamma_l > 0$ is a positive gain that determines the dynamic property. This is a simplified model for the RED algorithm that assumes the loss probability is proportional to the backlog, and is used in, e.g., [9], [10].

B. Existing MP-TCP algorithms

We first show how to relate the fluid model (3) to the window-based MP-TCP algorithms proposed in the literature. On each route r the source increases its window at the return of each ACK. Let this increment be denoted by $I_r(\mathbf{w}_s)$ where $\mathbf{w}_s$ is the vector of window sizes on different routes of source s. The source decreases the window on route r when it sees a packet loss on route r. Let this decrement be denoted by $D_r(\mathbf{w}_s)$. Then most loss based MP-TCP algorithms take the form of the following pseudo code:

- For each ACK on route r, $w_r \leftarrow w_r + I_r(\mathbf{w}_s)$.
- For each loss on route r, $w_r \leftarrow w_r - D_r(\mathbf{w}_s)$.

We now model the above pseudo codes by the fluid model (3). Let δw_r be the net change to window on route r in each round trip time. Then δw_r is roughly

$$\begin{aligned} \delta w_r &= (I_r(\mathbf{w}_s)(1 - q_r) - D_r(\mathbf{w}_s) q_r) w_r \\ &\approx (I_r(\mathbf{w}_s) - D_r(\mathbf{w}_s) q_r) w_r \end{aligned}$$

since the loss probability q_r is small. On the other hand

$$\delta w_r \approx \dot{w}_r \tau_r = \dot{x}_r \tau_r^2$$

Hence

$$\dot{x}_r = \frac{x_r}{\tau_r} \left(I_r(\mathbf{w}_s) - D_r(\mathbf{w}_s) q_r\right)$$

From (3) we have

$$\left\{ \begin{array}{lcl} k_r(\mathbf{x}_s) & = & \frac{x_r}{\tau_r} D_r(\mathbf{w}_s) \\ \phi_r(\mathbf{x}_s) & = & \frac{I_r(\mathbf{w}_s)}{D_r(\mathbf{w}_s)} \end{array} \right. \qquad (5)$$

We now apply this to the algorithms in the literature. We first summarize these algorithms in the form of a pseudo-code and then use (5) to derive parameters $k_r(\mathbf{x}_s)$ and $\phi_r(\mathbf{x}_s)$ of the fluid model (3).

Single-path TCP (TCP-NewReno): Single-path TCP is a special case of MP-TCP algorithm with $|s| = 1$. Hence x_s is a scalar and we identify each source with its route $r = s$. TCP-NewReno adjusts the window as follows:

- For each ACK on route r, $w_r \leftarrow w_r + 1/w_r$.
- For each loss on route r, $w_r \leftarrow w_r/2$.

From (5), this can be modeled by the fluid model (3) with

$$k_r(x_s) = \frac{1}{2}x_r^2, \quad \phi_r(x_s) = \frac{2}{\tau_r^2 x_r^2}$$

We now summarize some existing MP-TCP algorithms, all of which degenerate to TCP NewReno if there is only one route per source.

EWTCP [3]: EWTCP algorithm applies TCP-NewReno like algorithm on each route independently of other routes. It adjusts the window on multiple routes as follows:

- For each ACK on route r, $w_r \leftarrow w_r + a/w_r$.
- For each loss on route r, $w_r \leftarrow w_r/2$.

From (5), this can be modeled by the fluid model (3) with

$$k_r(\mathbf{x}_s) = \frac{1}{2}x_r^2, \quad \phi_r(\mathbf{x}_s) = \frac{2a}{\tau_r^2 x_r^2}$$

where $a > 0$ is a constant.

Coupled MPTCP [5], [6]: The Coupled MPTCP algorithm adjusts the window on multiple routes in a coordinated fashion as follows:

- For each ACK on route r, $w_r \leftarrow w_r + \frac{w_r}{(\sum_{k\in s} w_k)^2}$.
- For each loss on route r, $w_r \leftarrow w_r/2$.

From (5), this can be modeled by the fluid model (3) with

$$k_r(\mathbf{x}_s) = \frac{1}{2}x_r^2, \quad \phi_r(\mathbf{x}_s) = \frac{2}{(\sum_{k\in s} x_k \tau_k)^2}$$

Semicoupled MPTCP [7]: The Semi-coupled MPTCP algorithm adjusts the window on multiple routes as follows:

- For each ACK on route r, $w_r \leftarrow w_r + \frac{1}{\sum_{k\in s} w_k}$.
- For each loss on route r, $w_r \leftarrow w_r/2$.

From (5), this can be modeled by the fluid model (3) with

$$k_r(\mathbf{x}_s) = \frac{1}{2}x_r^2, \quad \phi_r(\mathbf{x}_s) = \frac{2}{x_r \tau_r (\sum_{k\in s} x_k \tau_k)}$$

Max MPTCP [7]: The Max MPTCP algorithm adjusts the window on multiple routes as follows:

- For each ACK on route r, $w_r \leftarrow w_r + \min\left\{\frac{\max\{w_k/\tau_k^2\}}{(\sum w_k/\tau_k)^2}, \frac{1}{w_r}\right\}$.
- For each loss on route r, $w_r \leftarrow w_r/2$.

From (5), this can be modeled by the fluid model (3) with

$$k_r(\mathbf{x}_s) = \frac{1}{2}x_r^2, \quad \phi_r(\mathbf{x}_s) = \frac{2\max\{x_k/\tau_k\}}{x_r \tau_r (\sum_{k\in s} x_k)^2}$$

where we have ignored taking the minimum with the $1/w_r$ term since the performance is mainly captured by $\frac{\max\{w_k/\tau_k^2\}}{(\sum w_k/\tau_k)^2}$.

TABLE I: MP-TCP algorithms

	C0	C1	C2, C3	C4	C5
EWTCP	Yes	Yes	Yes	Yes	Yes
Coupled	Yes	Yes	No	Yes	Yes
Semicoupled	No	Yes	Yes	Yes	Yes
Max	No	Yes	Yes	Yes	Yes
Generalized	No	Yes	Yes	Yes	Yes
Theorem	3.1	3.2, 3.3, 3.5		3.4	3.6

III. Structural properties

Throughout this paper we assume, for all $\mathbf{x}_s$, $r \in s$, $s \in S$, $k_r(\mathbf{x}_s) > 0$ and $\phi_r(\mathbf{x}_s) = 0$ only if $x_k = \infty$ for some $k \in s$. A point $(\mathbf{x}, \mathbf{p})$ is called an *equilibrium* of (3)–(4) if it satisfies, for all $r \in s, s \in S$ and $l \in L$,

$$k_r(\mathbf{x}_s)\left(\phi_r(\mathbf{x}_s) - q_r\right)^+_{x_r} = 0$$
$$\gamma_l\left(y_l - c_l\right)^+_{p_l} = 0$$

or equivalently,

$$x_r \geq 0,\ \phi_r(\mathbf{x}_s) \leq q_r \quad \text{and} \quad \phi_r(\mathbf{x}_s) = q_r \text{ if } x_r > 0 \qquad (6)$$
$$p_l \geq 0,\ y_l \leq c_l \quad \text{and} \quad y_l = c_l \text{ if } p_l > 0 \qquad (7)$$

We make two remarks. First an equilibrium $(\mathbf{x}, \mathbf{p})$ does not depend on K_s, but only on Φ_s. The design $(K_s, s \in S)$ however affects dynamic properties such as stability and responsiveness as we show below. Second, since $k_r(\mathbf{x}_s) > 0$ and $\phi_r(\mathbf{x}_s) = 0$ only if $x_k = \infty$ for some $k \in s$ by assumption, any finite equilibrium $(\mathbf{x}, \mathbf{p})$ must have $q_r > 0$ for all r. In the following we always restrict ourselves to finite equilibria.

In this section we denote an MP-TCP algorithm by $(K, \Phi) := (K_s, \Phi_s, s \in S)$. We characterize MP-TCP designs (K, Φ) that guarantee the existence, uniqueness, and stability of system equilibrium. We identify design criteria that determine TCP-friendliness, responsiveness and window oscillation and prove an inevitable tradeoff among these properties. We discuss in the next section the implications of these structural properties on existing algorithms. All proofs are relegated to the Appendices.

A. Summary

We first present some properties of an MP-TCP algorithm (K, Φ) that we have identified. We then interpret them and summarize their implications.

C0: For each $s \in S$ and each $\mathbf{x}_s$, the Jacobians of $\Phi_s(\mathbf{x}_s)$ is continuous and symmetric, i.e.,

$$\frac{\partial \Phi_s}{\partial \mathbf{x}_s}(\mathbf{x}_s) \quad = \quad \left[\frac{\partial \Phi_s}{\partial \mathbf{x}_s}(\mathbf{x}_s)\right]^T$$

C1: For each $s \in S$ there exists a nonnegative solution $\mathbf{x}_s := \mathbf{x}_s(\mathbf{p})$ to (6) for any finite $\mathbf{p} \geq 0$ such that $q_r > 0$ for all r. Moreover,

$$\frac{\partial y_l^s(\mathbf{p})}{\partial p_l} \leq 0, \qquad \lim_{p_l \to \infty} y_l^s(\mathbf{p}) = 0$$

where $y_l^s(\mathbf{p}) := \sum_{r\in s} H_{lr} x_r(\mathbf{p})$ is the aggregate traffic at link l from source s.

C2: For each $s \in S$ and each $\mathbf{x}_s$, $\Phi_s(\mathbf{x}_s)$ is continuously differentiable; moreover the symmetric part $[\partial\Phi_s(\mathbf{x}_s)/\partial\mathbf{x}_s]^+$ of the Jacobian is negative definite.
C3: For each $r \in R$, $\phi_r(\mathbf{x}_s) = \infty$ if and only if $x_r = 0$. The routing matrix H has full row rank.
C4: For each $r \in s$, $s \in S$, $\sum_{j\in s}[D_s]_{jr}(\mathbf{x}_s) \leq 0$ where $D_s(\mathbf{x}_s) := \left[\frac{\partial\Phi_s}{\partial\mathbf{x}_s}(\mathbf{x}_s)\right]^{-1}$.
C5: For each $r \in R$ and each $\mathbf{x}_{-r}$, $\lim_{x_r\to\infty}\phi_r(\mathbf{x}_s) = 0$.

These design criteria are intuitive and usually (but not always) satisfied; see Table I.

Condition C0 guarantees the existence of utility functions $U_s(\mathbf{x}_s)$ that an equilibrium $(\mathbf{x}, \mathbf{p})$ of a multipath TCP/AQM (3)–(4) implicitly maximizes (Theorem 3.1). It is always satisfied when there is only a single path ($|s| = 1$ for all s) but not when $|s| > 1$.

Conditions C1–C3 guarantee the existence, uniqueness, and global asymptotic stability of the equilibrium $(\mathbf{x}, \mathbf{p})$ (Theorems 3.2 and 3.3). C1 says that the aggregate traffic rate through a link l from source s decreases when the congestion price p_l on that link increases, and it decreases to 0 as p_l increases without bounds. C2 implies that at steady state, if $\mathbf{x}_s, \mathbf{q}_s$ are perturbed by $\delta\mathbf{x}_s, \delta\mathbf{q}_s$ respectively, then $(\delta\mathbf{x}_s)^T\delta\mathbf{q}_s < 0$. In the case of single-path TCP ($|s| = 1$ for all s), C2 is equivalent to the curvature of the utility function $U_s(x_s)$ being negative, i.e., $U_s(x_s)$ is strictly concave. C3 means that the rate on route r is zero if and only if it sees infinite price on that route.

Condition C4 is natural and satisfied by all the algorithms considered in this paper. It allows us to formally compare MP-TCP algorithms in terms of their TCP-friendliness (see formal definition below): under C1–C4, an MP-TCP algorithm (K, Φ) is more friendly if $\phi_r(\mathbf{x}_s)$ is smaller (Theorem 3.4). The existence of D_s in C4 is ensured by C2. To interpret C4, note that Lemma B.2 in Appendix B implies that $\Phi_s(\mathbf{x}_s^*) = \mathbf{q}_s^*$ at equilibrium. The implicit function theorem then implies $\mathbf{1}^T\frac{\partial\mathbf{x}_s}{\partial q_r} = \sum_{j\in s} D_{jr}$ at equilibrium for all $r \in s$. Hence C4 says that the aggregate throughput $\mathbf{1}^T\mathbf{x}_s$ at equilibrium over all routes $r \in s$ of an MP-TCP flow is a nonincreasing function of the price q_r.

Condition C5 is also satisfied by all the algorithms considered in this paper. It means that the sending rate on a route r grows unbounded when the congestion price q_r is zero. Under C1–C3, an MP-TCP algorithm (K, Φ) is more responsive (see formal definition below) if the Jacobian of $\Phi_s(\mathbf{x}_s)$ is more negative definite (Theorem 3.5). C5 then implies an inevitable tradeoff: an MP-TCP algorithm that is more responsive is necessarily less TCP-friendly (Theorem 3.6).

We now elaborate on each of these properties.

B. *Utility maximization*

For single-path TCP (SP-TCP), one can associate a utility function $U_s(x_s) \in \mathbb{R}_+ \to \mathbb{R}$ with each flow s (x_s is a scalar and $|s| = 1$) and interpret (3)–(4) as a distributed algorithm to maximize aggregate users' utility, e.g. [8]–[11]. Indeed, for SP-TCP, an $(\mathbf{x}, \mathbf{p})$ is an equilibrium if and only if $\mathbf{x}$ is optimal for

$$\text{maximize} \sum_{s\in S} U_s(x_s) \quad \text{s.t. } y_l \leq c_l \quad l \in L \tag{8}$$

and $\mathbf{p}$ is optimal for the associated dual problem. Here $y_l \leq c_l$ means the aggregate traffic y_l at each link does not exceed its capacity c_l. In fact this holds for a much wider class of SP-TCP algorithms than those specified by (3)–(4) [11]. Furthermore all the main TCP algorithms proposed in the literature have strictly concave utility functions, implying a unique stable equilibrium.

The case of MP-TCP is much more delicate: whether an underlying utility function exists depends on the design choice of Φ_s and not all MP-TCP algorithms have one. Consider the multipath equivalent of (8):

$$\text{maximize} \sum_{s\in S} U_s(\mathbf{x}_s) \quad \text{s.t. } y_l \leq c_l \quad l \in L \tag{9}$$

where $\mathbf{x}_s := (x_r, r \in s)$ is the rate vector of flow s and $U_s : \mathbb{R}_+^{|s|} \to \mathbb{R}$ is a concave function.

Theorem 3.1 (utility maximization): There exists a twice continuously differentiable and concave $U_s(\mathbf{x}_s)$ such that an equilibrium $(\mathbf{x}, \mathbf{p})$ of (3)–(4) solves (9) and its dual problem if and only if condition C0 holds.

Condition C0 is satisfied trivially by SP-TCP when $|s| = 1$. For MP-TCP ($|s| > 1$), the models derived in Section II-B show that only EWTCP and Coupled algorithms satisfy C0 and have underlying utility functions. It therefore follows from the theory for SP-TCP that EWTCP has a unique stable equilibrium while Coupled algorithm may have multiple equilibria since its corresponding utility function is not strictly concave. The other MP-TCP algorithms all have asymmetric Jacobian $\frac{\partial\Phi_s}{\partial\mathbf{x}_s}$ and does not satisfy C0.

C. *Existence, uniqueness and stability of equilibrium*

Even though a multipath TCP algorithm (K, Φ) may not have a utility maximization interpretation, a unique equilibrium exists if conditions C1–C3 are satisfied.

Theorem 3.2 (existence and uniqueness):

1) Suppose C1 holds. Then (3)–(4) has at least one equilibrium.
2) Suppose C2 and C3 hold. Then (3)–(4) has at most one equilibrium

Thus (3)–(4) has a unique equilibrium $(\mathbf{x}^*, \mathbf{p}^*)$ under C1–C3.

Conditions C1-C3 not only guarantee the existence and uniqueness of the equilibrium, they also ensure that the equilibrium is globally asymptotically stable, when the gain $k_r(\mathbf{x}_s)$ is only a function of x_r itself, i.e., $k_r(\mathbf{x}_s) \equiv k_r(x_r)$ for all $r \in R$. This is satisfied by all the existing algorithms presented in Section II-B.

Theorem 3.3 (stability): Suppose C1-C3 hold and $k_r(\mathbf{x}_s) \equiv k_r(x_r)$ for all $r \in R$. Then the unique equilibrium $(\mathbf{x}^*, \mathbf{p}^*)$ is globally asymptotically stable. In particular, starting from any initial point $\mathbf{x}(0) \in \mathbb{R}_+^{|R|}$ and $\mathbf{p}(0) \in \mathbb{R}_+^{|L|}$, the trajectory $(\mathbf{x}(t), \mathbf{p}(t))$ generated by the MP-TCP algorithm (3)–(4) converges to the equilibrium $(\mathbf{x}^*, \mathbf{p}^*)$ as $t \to \infty$.

Our proposed algorithm does not satisfy $k_r(\mathbf{x}_s) \equiv k_r(x_r)$ even though it seems to be stable in our simulations. This

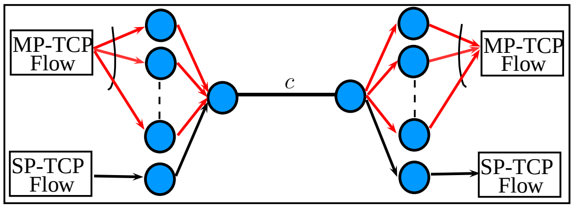

Fig. 1: Test network for the definition of TCP friendliness. The link in the middle is the only bottleneck link with capacity c.

condition is only sufficient and needed in our Lyapunov stability proof; see Appendix C. When $k_r(\mathbf{x}_s)$ depends on $\mathbf{x}_s$, one can replace $k_r(x_r)$ in the definition of the Lyapunov function V in (21) with $k_r(\mathbf{x}_s^*)$ evaluated at the equilibrium and the same argument there proves that $(\mathbf{x}^*, \mathbf{p}^*)$ is (locally) asymptotically stable. Also see Theorem 3.5 below for an alternative proof of local stability.

D. TCP friendliness

Informally, an MP-TCP flow is said to be 'TCP friendly' if it does not dominate the available bandwidth when it shares the same network with a SP-TCP flow [2]. To define this precisely we use the test network shared by a SP-TCP flow and a MP-TCP flow under test as shown in Fig. 1. All paths traverse a single bottleneck link with capacity c, with all other links with capacities strictly higher than c. The links have fixed but possibly different delays. To compare the friendliness of two MP-TCP algorithms $\hat{M} := (\hat{K}, \hat{\Phi})$ and $\tilde{M} := (\tilde{K}, \tilde{\Phi})$, suppose that when $\hat{M}$ shares the test network with a SP-TCP it achieves a throughput of $\|\hat{\mathbf{x}}\|_1$ in equilibrium aggregated over the available paths (the SP-TCP therefore attains a throughput of $c - \|\hat{\mathbf{x}}\|_1$). Suppose $\tilde{M}$ achieves a throughput of $\|\tilde{\mathbf{x}}\|_1$ in equilibrium when it shares the test network with the same SP-TCP. Then we say that $\hat{M}$ is *friendlier (or more TCP-friendly)* than $\tilde{M}$ if $\|\hat{\mathbf{x}}\|_1 \leq \|\tilde{\mathbf{x}}\|_1$, i.e., if $\hat{M}$ receives no more bandwidth than $\tilde{M}$ does when they *separately* share the test network in Fig. 1 with the same SP-TCP flow.

From the theory for single-path TCP ($|s| = 1$ for all $s \in S$), it is known that a design is more TCP-friendly if it has a smaller marginal utility $U_s'(x_s) = \Phi_s(x_s)$. The same intuition holds for MP-TCP algorithms even though the utility functions may not exist for MP-TCP algorithm.

Theorem 3.4 (friendliness): Consider two MP-TCP algorithms $\hat{M} := (\hat{K}, \hat{\Phi})$ and $\tilde{M} := (\tilde{K}, \tilde{\Phi})$. Suppose both satisfy C1–C4. Then $\hat{M}$ is friendlier than $\tilde{M}$ if $\hat{\Phi}_s(\mathbf{x}_s) \leq \tilde{\Phi}_s(\mathbf{x}_s)$ for all $s \in S$.

E. Responsiveness around equilibrium

Suppose conditions C1–C3 hold and there is a unique equilibrium $\mathbf{z}^* := (\mathbf{x}^*, \mathbf{p}^*)$. Assume all links in L are active with $p_l^* > 0$; otherwise remove from L all links with prices $p_l^* = 0$. Let $\delta\mathbf{z}(t) := \mathbf{z}(t) - \mathbf{z}^*$. The behavior of (3)–(4) around the equilibrium is defined by the linearized system:

$$\delta\dot{\mathbf{z}} \;=\; J^* \;\delta\mathbf{z}(t) \tag{10}$$

Here J^* is the Jacobian of (3)–(4) at the equilibrium $\mathbf{z}^*$:

$$J^* \;:=\; J(\mathbf{x}^*) \;:=\; \begin{bmatrix} \Lambda_k \frac{\partial \Phi}{\partial \mathbf{x}} & -\Lambda_k H^T \\ \Lambda_\gamma H & 0 \end{bmatrix}$$

where $\Lambda_k = diag\{k_r(\mathbf{x}_s^*), r \in R\}$, $\Lambda_\gamma = diag\{\gamma_l, l \in L\}$, and $\frac{\partial \Phi}{\partial \mathbf{x}}$ is evaluated at $\mathbf{x}^*$.

The stability and responsiveness of the linearized system (10) (how fast does the system converges to the equilibrium locally) is determined by the real parts of the eigenvalues of J^*. Specifically the linearized system is stable if the real parts of all eigenvalues of J^* are negative; moreover the more negative the real parts are the faster the linearized system converges to the equilibrium. We now show that the linearized system (10) is stable (i.e., converges exponentially fast to $\mathbf{z}^*$ locally) and characterize its responsiveness in terms of the design choices (K, Φ).

Let $Z = \{\mathbf{z} := (\mathbf{x}, \mathbf{p}) \in \mathbb{C}^{|R|+|L|} \mid \|\mathbf{z}\|_2 = 1\}$.

Theorem 3.5 (responsiveness): Suppose C1–C3 hold. Then

1) The linearized system (10) is stable, i.e., $\mathbf{Re}(\lambda) < 0$ for any eigenvalue λ of J^*. Moreover $\mathbf{Re}(\lambda) \leq \overline{\lambda}(J^*)$ where

$$\overline{\lambda}(J^*) \;:=\; \max_{\mathbf{z} \in Z} \left\{ \frac{\mathbf{x}^H \left[\frac{\partial \Phi}{\partial \mathbf{x}}\right]^+ \mathbf{x}}{\mathbf{x}^H \Lambda_k^{-1} \mathbf{x} + \mathbf{p}^H \Lambda_\gamma^{-1} \mathbf{p}} \right\} \;\leq\; 0$$

where Λ_k and $\frac{\partial \Phi_s}{\partial \mathbf{x}_s}$ are evaluated at the equilibrium point z^*.

2) For two MP-TCP algorithms $(\hat{K}, \hat{\Phi})$ and $(\tilde{K}, \tilde{\Phi})$, $\overline{\lambda}(\hat{J}^*) \leq \overline{\lambda}(\tilde{J}^*)$ provided

$$\hat{K}_s \;\geq\; \tilde{K}_s \quad \text{and} \quad \frac{\partial \hat{\Phi}_s}{\partial \mathbf{x}_s} \preceq \frac{\partial \tilde{\Phi}_s}{\partial \mathbf{x}_s} \qquad \text{for all } s \in S$$

Theorem 3.5 motivates the following definition of responsiveness. Given two MP-TCP $\hat{M}$ and $\tilde{M}$, we say that $\hat{M}$ is *more responsive than* $\tilde{M}$ if $\overline{\lambda}(\hat{J}^*) \leq \overline{\lambda}(\tilde{J}^*)$. Theorem 3.5(2) implies that an MP-TCP algorithm with a larger $K_s(\mathbf{x}_s^*)$ or more negative definite $\left[\frac{\partial \Phi_s}{\partial \mathbf{x}_s}(\mathbf{x}_s^*)\right]^+$ is more responsive, in the sense that the real parts of the eigenvalues of the Jacobian J^* have a smaller more negative upper bound.

Then the next result suggests an inevitable tradeoff between responsiveness and friendliness.

Theorem 3.6 (tradeoff): Consider two MP-TCP algorithms $(K, \hat{\Phi})$ and $(K, \tilde{\Phi})$ with the same gain K. Suppose both satisfy C1-C3 and C5. Then for all $s \in S$

$$\frac{\partial \hat{\Phi}_s(\mathbf{x}_s)}{\partial \mathbf{x}_s} \;\preceq\; \frac{\partial \tilde{\Phi}_s(\mathbf{x}_s)}{\partial \mathbf{x}_s} \quad \Rightarrow \quad \hat{\Phi}_s(\mathbf{x}_s) \;\geq\; \tilde{\Phi}_s(\mathbf{x}_s)$$

In light of Theorems 3.4 and 3.5, Theorem 3.6 says that a more responsive MP-TCP design is inevitably less friendly if they have the same K.

The theorem is easier to understand in the case of SP-TCP, i.e., when $|s| = 1$ for all $s \in S$ and $\Phi_s(x_s) = U_s'(x_s)$. Then it implies that a more concave utility function $U_s(x_s)$ has a larger marginal utility, and hence less friendly.

F. Window oscillation

Window oscillations are inherent in loss-based additive increase multiplicative decrease (AIMD) TCP algorithms. We close this section by discussing informally why a larger design $K_s(\mathbf{x}_s)$ generally creates more severe window oscillations. This implies a tradeoff between responsiveness (which is enhanced by a large $K_s(\mathbf{x}_s)$) and oscillation (which is reduced with a small $K_s(\mathbf{x}_s)$).

The effect of $K_s(\mathbf{x}_s)$ on window fluctuations can be understood by studying how it affects the decrease $D_r(\mathbf{w}_s)$ per packet loss in the following packet level model:

- For each ACK on route r, $w_r \leftarrow w_r + I_r(\mathbf{w}_s)$.
- For each loss on route r, $w_r \leftarrow w_r - D_r(\mathbf{w}_s)$.

Let $Z_r \in \{0,1\}$ be an indicator variable of whether a packet loss is observed on route r at an arbitrary time in steady state. Then

$$D_s(\mathbf{x}_s) \quad := \quad \frac{1}{\|\mathbf{x}_s\|_1}\, \mathbb{E}\left(\sum_{r\in s} \frac{D_r(\mathbf{w}_s)}{\tau_r} Z_r \,\middle|\, \sum_{k\in s} Z_k \geq 1 \right)$$

represents the expected relative reduction in *aggregate* throughput $\sum_{r\in s} D_r(\mathbf{w}_s)/\tau_r$, given that there is at least one packet loss on some route $r \in s$. It is a measure of throughput fluctuation for each packet loss that an application experiences. For TCP-NewReno (for which $s = \{r\}$ and w_s is a scalar), the window size is halved on each packet loss, $D_r(w_s) = w_r/2$, and hence $D_s(x_s) = 1/2$.

To understand $D_s(\mathbf{x}_s)$ for MP-TCP algorithms, we need the following result.

Lemma 3.1: Let $A_i := \{a_{i1}, a_{i2}, \ldots\}$ with $|A_i|$ elements. Each element a_{ij} is an independent binary random variable with $\mathbb{P}(a_{ij} = 1) = 1 - \mathbb{P}(a_{ij} = 0) = q_i$. Define $D_i(A_i) := d_i 1_{(\sum_j a_{ij} \geq 1)}$. Then

$$\mathbb{E}\left(\sum_k D_k(A_k) \,\middle|\, \sum_{i,j} a_{ij} \geq 1 \right) \;=\; \frac{\sum_k d_k q_k |A_k|}{\sum_k q_k |A_k|} + o\left(\sum_k q_k \right)$$

Suppose each route has a fixed loss probability q_r. Then within each RTT, Lemma 3.1 implies

$$D_s(\mathbf{x}_s) \;=\; \frac{1}{\|\mathbf{x}_s\|_1} \left(\frac{\sum_{r\in s} w_r q_r D_r(\mathbf{w}_s)/\tau_r}{\sum_{r\in s} q_r w_r} + o\left(\sum_{r\in s} q_r \right) \right)$$

Substituting $w_r = x_r \tau_r$ and $x_r D_r(\mathbf{w}_s) = \tau_r k_r(\mathbf{x}_s)$ from (5), we get, ignoring the high-order terms,

$$D_s(\mathbf{x}_s) \quad = \quad \frac{1}{\|\mathbf{x}_s\|_1} \left(\frac{\sum_{r\in s} \tau_r q_r k_r(\mathbf{x}_s)}{\sum_{r\in s} \tau_r q_r x_r} \right) \qquad (11)$$

to the first order. Note that $k_r(\mathbf{x}_s)$ does not affect the *equilibrium* rates $\mathbf{x}_s$. Hence, with the assumption that τ_r are constants, $D_s(\mathbf{x}_s)$ is determined by the functions $k_r(\mathbf{x}_s)$ in steady state.

Specifically an MP-TCP algorithm with a larger $K_s(\mathbf{x}_s)$ tends to have a larger $D_s(\mathbf{x}_s)$ and hence more severe window oscillations. Theorem 3.5 however suggests that a larger $K_s(\mathbf{x}_s)$ also leads to better responsiveness, suggesting an inevitable tradeoff between responsiveness and window oscillation.

IV. Implications and a new algorithm

In this section we discuss the implications of these structural properties on the behavior of existing MP-TCP algorithms. They are further illustrated in simulation results in Section V. The discussion motivates a new design that generalizes the existing MP-TCP algorithm.

A. Implications on existing algorithms

Recall Table I that summarizes the conditions satisfied by the various algorithms. Only EWTCP and Coupled algorithm satisfy C0. Their equilibrium properties can be studied in the standard utility maximization model as done for single-path TCP. Semicoupled and Max algorithms do not satisfy C0 and therefore analysis through utility maximization is not applicable. However Theorem 4.1 below implies that, both Semicoupled and Max algorithms satisfy C1–C3 provided they enable no more than 8 routes. Theorem 3.2 and 3.3 then implies that they have a unique and globally stable equilibrium. It is also easy to show that EWTCP satisfies C1-C3. The Coupled algorithm does not satisfy C2 and is found to have multiple equilibria in [5].

Next we discuss friendliness of existing MP-TCP algorithms. It can be shown that the $\phi_r(\mathbf{x}_s)$ corresponding to these algorithms satisfy:

$$\phi_r^{ewtcp}(\mathbf{x}_s) \geq \phi_r^{semicoupled}(\mathbf{x}_s) \geq \phi_r^{max}(\mathbf{x}_s) \geq \phi_r^{coupled}(\mathbf{x}_s)$$

for all $\mathbf{x}_s \geq 0$ if all routes $r \in s$ have the same round trip time. Since all of them satisfy C4, Theorem 3.4 implies that their friendliness will be in the same order, i.e., their throughputs in the test network of Fig. 1 are ordered as follows:

EWTCP($a \geq 1$)[1] $\geq$Semicoupled$\geq$Max$\geq$Coupled

This is confirmed by the ns2 simulation.

Third we will discuss responsiveness of existing MP-TCP algorithms. These algorithms have the same gain function $k_r(\mathbf{x}_s) = 0.5x_r^2$ and

$$(\frac{\partial \Phi_s}{\partial \mathbf{x}_s})^{ewtcp} \preceq (\frac{\partial \Phi_s}{\partial \mathbf{x}_s})^{semicoupled} \preceq (\frac{\partial \Phi_s}{\partial \mathbf{x}_s})^{max} \preceq (\frac{\partial \Phi_s}{\partial \mathbf{x}_s})^{coupled}$$

Theorem 3.5 then implies that their responsiveness should be in the same order, as confirmed by our simulations in section V.

Finally we discuss window oscillation of existing MP-TCP algorithms using $D_s(\mathbf{x}_s)$ as the metric. As mentioned in Section III-F, $D_s(\mathbf{x}_s) = 0.5$ for TCP NewReno, a benchmark single-path TCP algorithm. According to (11), if $k_r(\mathbf{x}_s) \leq 0.5x_r\|\mathbf{x}_s\|_1$, we have, to the first order

$$D_s(\mathbf{x}_s) \quad \leq \quad \frac{1}{2} \frac{\sum_{r\in s} \tau_r q_r x_r \|\mathbf{x}_s\|_1}{\|\mathbf{x}_s\|_1 \sum_{r\in s} \tau_r q_r x_r} \;=\; \frac{1}{2}$$

All existing MP-TCP algorithms have the same $k_r(\mathbf{x}_s) = 0.5x_r^2 \leq 0.5x_r\|\mathbf{x}_s\|_1$, with strict inequality if $|s| > 1$ and $x_r > 0$ for at least two $r \in s$. Thus enabling MP-TCP always tend to reduce window oscillation for existing algorithms compared to TCP NewReno. Moreover, the window

[1] When $a < 1$, the MP-TCP source can obtain even smaller throughput than the competing single-path TCP source.

oscillation is always reduced compared to TCP NewReno when $k_r(\mathbf{x}_s) \le 0.5 x_r \|\mathbf{x}_s\|_1$.

B. A generalized algorithm

Consider the class of algorithms parametrized by (β, n, η) as follows:

$$\begin{cases} k_r(\mathbf{x}_s) &= \frac{1}{2} x_r (x_r + \eta(\|\mathbf{x}_s\|_\infty - x_r)), \quad \eta \ge 0 \\ \phi_r(\mathbf{x}_s) &= \frac{2((1-\beta)x_r + \beta \|\mathbf{x}_s\|_n)}{\tau_r^2 x_r \|\mathbf{x}_s\|_1^2}, \quad n \in \mathbb{N}_+, \beta \ge 0 \end{cases} \tag{12}$$

This class includes the Max ($\beta = 1, \eta = 0, n = \infty$), Coupled ($\beta = 0, \eta = 0$), and Semicoupled ($\beta = 1, \eta = 0, n = 1$) algorithms as special cases when all RTTs on different paths of the same source are the same, i.e., $\tau_r = \tau_s$, $r \in s$.

The next result characterizes a subclass that have a unique and locally stable equilibrium point.

Theorem 4.1: Fix any $\eta \ge 0$ and $n \in \mathbb{N}_+$. For any $s \in S$, the $\phi_r(\mathbf{x}_s)$ in (12) satisfies

1) C1 if $\beta \ge 0$.
2) C2–C3 if $0 < \beta \le 1$, $|s| \le 8$ and τ_r are the same for all $r \in s$ (assuming H has full row rank).

The requirement that $|s| \le 8$ is not restrictive since in practice a mobile device may typically enable no more than 3 paths. The requirement that τ_r are the same for all $r \in s$ is used in proving the negative definiteness of the (symmetric part of the) Jacobian of $\Phi_s(\mathbf{x}_s)$. Since a negative definite matrix remains negative definite after small enough perturbations of its entries, Theorem 4.1 holds if the RTTs of the subpaths do not differ much. This (sufficient) condition seems reasonable as two paths between the same source-destination pair often have similar RTTs if both are wireline paths or if any wireless links in the path are either high-speed WiFi or new generation (e.g., 4G) cellular links, as shown in [12].

For the class of algorithms specified by (12), Theorem 4.1 motivates a design space defined by $\beta \in (0, 1], \eta \ge 0, n \in \mathbb{N}_+$. The algorithm given at the end of Section I corresponds to the choice $(\beta, n, \eta) = (0.2, 0.5, \infty)$. Theorem 4.1 guarantees that the algorithm satisfies C1–C3 when each source has no more than 8 paths and their round-trip times are equal. Moreover it can be shown that the algorithm also satisfies C4, as summarized in Table I.

Compared with existing MP-TCP algorithms, our $\eta = 0.5$ is larger than the Max, Copuled, and Semicoupled algorithms. Hence our design tends to have a higher gain K_s than these algorithms. Our $\beta = 0.2$ is smaller than the Max and Semcoupled algorithms (both have $\beta = 1$) and bigger than Coupled ($\beta = 0$). Hence our Φ_s is smaller than that of the Max and Semcoupled algorithms but bigger than that of the Coupled algorithm (for the same n). Note that responsiveness is mainly affected by subpaths with small throughput while window oscillation is mainly affected by subpaths with large throughput. Our design with nonzero η scales $k_r(\mathbf{x}_s)$ in the right way: a path r that has a large x_r has $k_r(\mathbf{x}_s) \approx 0.5 x_r^2$ and hence a similar degree of window oscillation as existing algorithms, while a path r with a small x_r has larger $k_r(\mathbf{x}_s)$ than that under a design with zero η and therefore is more responsive.

TABLE II: How design choices affect MP-TCP performance.

Performance	Parameter
TCP friendliness	$\phi_r(\mathbf{x}_s) \downarrow$
Responsiveness	$k_r(\mathbf{x}_s) \uparrow$, $-\partial\Phi_s/\partial\mathbf{x}_s \uparrow$
Window oscillation	$k_r(\mathbf{x}_s) \downarrow$

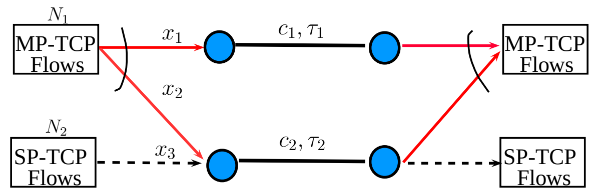

Fig. 2: Network for simulations on TCP friendliness and responsiveness, with N_1 MP-TCP flows and N_2 single-path TCP flows sharing 2 links of capacity c_1, c_2 and delay τ_1, τ_2. MP-TCP flows maintain two routes with rate x_1, x_2. Single-path TCP flows maintain one route with rate x_3.

We now explain the rationale underlying our design choice. We choose $n = \infty$ because then the computation does not require exponentiation and is much simpler. As discussed above the design of MP-TCP algorithms involves inevitable tradeoffs among responsiveness, friendliness, and the severity of window oscillation. Specifically a design is more responsive if it has a higher gain K_s or a more negative definite Jacobian $[\partial\Phi_s/\partial\mathbf{x}_s]^+$ (Theorem 3.5). However a larger K_s usually creates a bigger window oscillation; a more negative definite $[\partial\Phi_s/\partial\mathbf{x}_s]^+$ implies a larger Φ_s, usually hurting friendliness (Theorems 3.6 and 3.4). This is summarized in Table II. Since enabling multiple paths already reduces window oscillation compared to single-path TCP (section IV-A), MP-TCP can afford to use a relatively large gain K_s for responsiveness. This does not compromise too much on window oscillation, but allows us to use a less negative definite Jacobian $[\partial\Phi_s/\partial\mathbf{x}_s]^+$ with a smaller Φ_s to maintain sufficient TCP friendliness. This strikes a good balance among responsiveness, friendliness, and window oscillation, as illustrated by the simulation results in Section V.

V. Simulation

In this section we summarize our ns2 simulation results that illustrate the above analysis. Only the congestion avoidance phase on each path is modified to follow the MP-TCP algorithm. Every path follows the same slow start and loss recovery mechanism as TCP-NewReno, except that the minimum *ssthresh* is set to 1 instead of 2 when more than 1 path is available. We assume the advertised window *awnd* is set to infinity.

Our simulations are divided into three parts. First we compare TCP friendliness of our algorithm and prior algorithms. The result confirms that the Couple algorithm is the friendliest, ours is close to the Coupled algorithm and friendlier than the other algorithms. Second we compare the responsiveness of each algorithm in a dynamic environment where flows come and go. The result shows that the Coupled algorithm is unresponsive (illustrating the tradeoff between responsiveness and friendliness). EWTCP is the most responsive; ours is

TABLE III: TCP friendliness (same RTTs): throughputs of MPTCP and single-path TCP users.

	ewtcp	semi.	max	ours	coupled
MP-TCP Mbps	2.98	2.64	2.58	2.25	2.22
SP-TCP Mbps	1.01	1.32	1.35	1.61	1.67

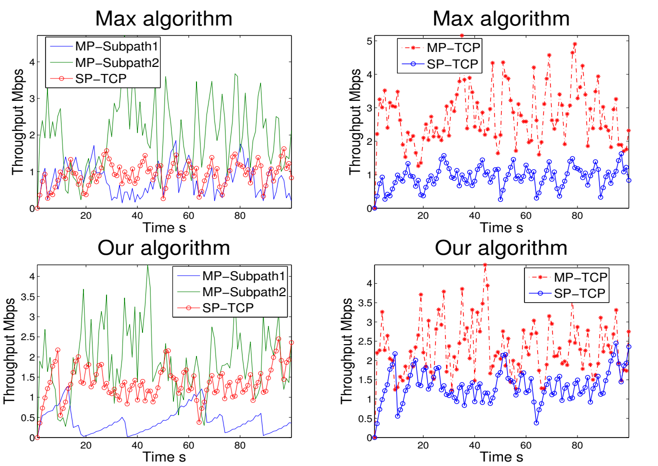

Fig. 3: TCP friendliness (different RTTs): the throughputs of MP-TCP and single-path TCP users. Left panel: throughput of each path. Right panel: aggregate throughput of each source.

similar in responsiveness but friendlier to single-path TCP flows. Finally we show that all MP-TCP algorithms have smaller window oscillations than SP-TCP.

These simulations confirm our analytical results and suggest our design choice strike a good balance among friendliness, responsiveness, and window oscillation.

A. *TCP friendliness*

We study TCP friendliness of each algorithm using the network topology in Fig. 2, first with equal RTTs and then with different RTTs. We assume all the flows are long lived and focus on the steady state throughput.

In the first set of simulations, we let $\tau_1 = \tau_2 = 20$ms, $c_1 = c_2 = 10$Mbps and $N_1 = N_2 = 5$. The average aggregate throughput of MP-TCP and single-path TCP users are shown in Table III. The EWTCP, Semicoupled and Max algorithms are significantly less friendly to SP-TCP. The Coupled algorithm is the friendliest and our algorithm is very close.

In the second set of simulations the link capacities and the number of flows are unchanged but the propagation delays are changed to $\tau_1 = 10$ms and $\tau_2 = 50$ms. Due to space limitation we only show the throughput trajectories of the Max algorithm and ours, in Fig. 3. Compared with the simulation results obtained where each path has similar round trip time, Max algorithm become more unfriendly towards single-path TCP. However, our algorithm is quite moderate in comparison with the Max algorithm.

B. *Responsiveness*

A good algorithm should react fast in a dynamic environment. We use the network in Fig. 2 with $c_1 = c_2 = 2$Mbps, $\tau_1 = \tau_2 = 20$ms and $N_1 = N_2 = 1$. The MP-TCP flow is long lived while the SP-TCP flow starts at 40s and end at 80s. We record the average throughput of the SP-TCP flow from 40-80s, which measures the friendliness of MP-TCP. We also measure the time for the throughput on the second path to recover[2] for MP-TCP users. It measures the responsiveness of MP-TCP. These measurements are shown in Table IV and the throughput trajectories of all algorithms are shown in Fig. 5. EWTCP is the most responsive among all the algorithms. Ours is as responsive as the Max algorithm, yet significantly friendlier than EWTCP. The Coupled algorithm took an excessively long time to recover. The excessively slow recovery of the throughput on the second path (see Fig. 5) is due to the design of Coupled algorithm that increases the window by $w_r/(\sum_{k\in s} w_k)^2$ on each ACK. After the single-path TCP flow has left, w_2 is small while w_1 is large, so that $w_2/(w_1+w_2)^2$ is very small. It therefore takes a long time for w_2 to increase to its steady state value. In general, under the Coupled algorithm, a route with a large throughput can greatly suppress the throughput on another route even though the other route is underutilized.

TABLE IV: Responsiveness: convergence time of MP-TCP and throughput of single-path TCP

	ewtcp	semi.	max	ours	coupled
Convergence time (s)	1.0	2.5	4.5	4.5	54.0
SP-TCP througput (Mbps)	1.02	1.17	1.30	1.57	1.72

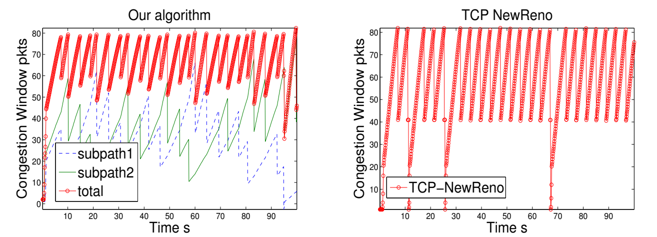

Fig. 4: Window oscillation: the red trajectories represent throughput fluctuations experience by the application that run on MP-TCP/SP-TCP.

C. *Window oscillation*

We use a single-link network model to compare window oscillation under MP-TCP and that under SP-TCP. First a MP-TCP flow initiates two subpaths through that link, and we measure the window size of each subpath and their aggregate window size. Then a TCP-NewReno flow traverse the same link and we measure its window size. The results are shown in Fig. 4 for our algorithm in comparison with SP-TCP (other MP-TCP algorithms have a similar behavior). They confirm that enabling multiple paths reduces window oscillation.

VI. Conclusion

We have presented a model for MP-TCP and identified designs that guarantee the existence, uniqueness and stability

[2]Defined as the first time the throughput on the second path is within 90% of the average throughput after the single-path user has left.

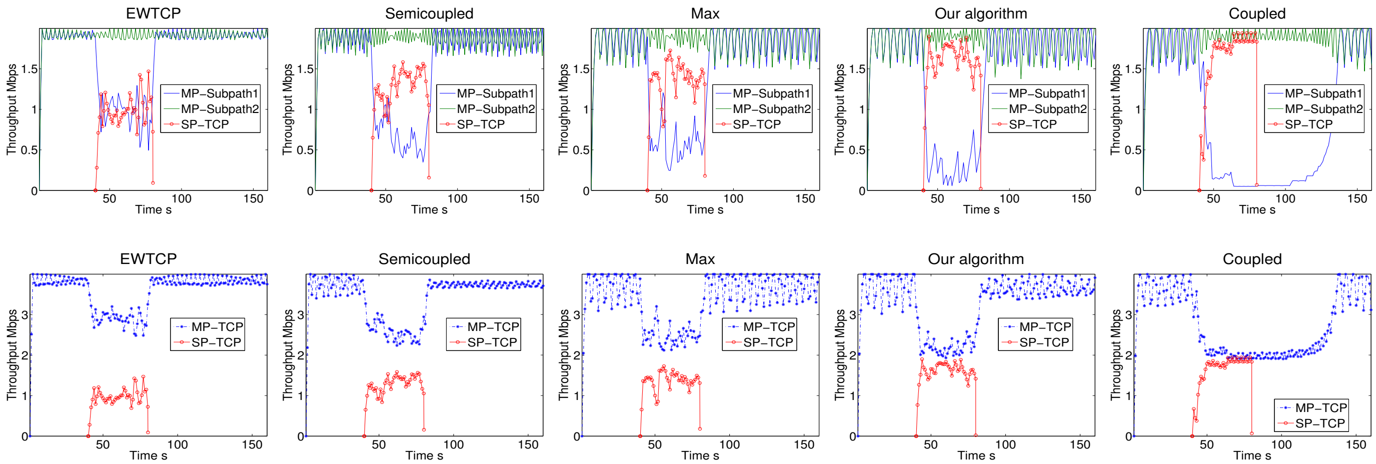

Fig. 5: Responsiveness: dynamic behavior of MP-TCP. Upper figures: throughput trajectory for each path. Lower figures: total throughput for each user.

of the network equilibrium. We have characterized the design space and study the tradeoff among TCP friendliness, responsiveness, and window oscillation. We have proposed a new MP-TCP algorithm that generalizes existing algorithms and strikes a good balance among these properties. We have presented preliminary ns2 simulations to illustrate our analysis.

REFERENCES

[1] Q. Peng, A. Walid, and S. H. Low, "Multipath TCP algorithms: theory and design," in *Proceedings of the ACM SIGMETRICS/international conference on Measurement and modeling of computer systems*. ACM, 2013, pp. 305–316.

[2] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure, "TCP extensions for multipath operation with multiple addresses," *IETF MPTCP proposal*, 2009.

[3] M. Honda, Y. Nishida, L. Eggert, P. Sarolahti, and H. Tokuda, "Multipath congestion control for shared bottleneck," in *Proc. PFLDNeT workshop*, 2009.

[4] J. R. Iyengar, P. D. Amer, and R. Stewart, "Concurrent multipath transfer using SCTP multihoming over independent end-to-end paths," *Networking, IEEE/ACM Transactions on*, vol. 14, no. 5, pp. 951–964, 2006.

[5] F. Kelly and T. Voice, "Stability of end-to-end algorithms for joint routing and rate control," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 2, pp. 5–12, 2005.

[6] H. Han, S. Shakkottai, C. Hollot, R. Srikant, and D. Towsley, "Overlay TCP for multi-path routing and congestion control," in *IMA Workshop on Measurements and Modeling of the Internet*, 2004.

[7] D. Wischik, C. Raiciu, A. Greenhalgh, and M. Handley, "Design, implementation and evaluation of congestion control for multipath TCP," in *Proceedings of the 8th USENIX conference on Networked systems design and implementation*. USENIX Association, 2011, pp. 8–8.

[8] S. Shakkottai and R. Srikant, "Network optimization and control," *Foundations and Trends® in Networking*, vol. 2, no. 3, pp. 271–379, 2007.

[9] F. P. Kelly, A. K. Maulloo, and D. K. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research society*, vol. 49, no. 3, pp. 237–252, 1998.

[10] S. H. Low and D. E. Lapsley, "Optimization flow control, I: basic algorithm and convergence," *IEEE/ACM Transactions on Networking (TON)*, vol. 7, no. 6, pp. 861–874, 1999.

[11] S. H. Low, "A duality model of tcp and queue management algorithms," *Networking, IEEE/ACM Transactions on*, vol. 11, no. 4, pp. 525–536, 2003.

[12] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "A close examination of performance and power characteristics of 4G LTE networks," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 225–238.

[13] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Prentice-Hall, Inc., 1996.

ACKNOWLEDGMENTS

This work was supported by ARO MURI through grant W911NF-08-1-0233, NSF NetSE through grant CNS 0911041 and Bell Labs, Lucent-Alcatel.

APPENDIX A
PROOF OF THEOREM 3.1 (UTILITY MAXIMIZATION)

The Lagrangian of (9) is:

$$
\begin{aligned}
L(\mathbf{x}, \mathbf{p}) &= \sum_{s \in S} U_s(\mathbf{x}_s) - \sum_{l \in L} p_l (y_l - c_l) \\
&= \sum_{s \in S} U_s(\mathbf{x}_s) - \sum_{l \in L} p_l \Big(\sum_{r \in R} H_{lr} x_r - c_l \Big) \\
&= \sum_{s \in S} \Big(U_s(\mathbf{x}_s) - \sum_{r \in s} x_r q_r \Big) + \sum_{l \in L} p_l c_l
\end{aligned}
$$

where $\mathbf{p} \geq \mathbf{0}$ are the dual variables and $q_r := \sum_{r \in R} H_{lr} p_l$. Then the dual problem is

$$
D(\mathbf{p}) = \sum_{s \in S} \max_{\mathbf{x}_s \geq 0} \{ B_s(\mathbf{x}_s, \mathbf{p}) \} + \sum_{l \in L} p_l c_l \quad \mathbf{p} \geq \mathbf{0}
$$

where $B_s(\mathbf{x}_s, \mathbf{p}) = U_s(\mathbf{x}_s) - \sum_{r \in s} x_r q_r$. The KKT condition implies that, at optimality, we have

$$
\frac{\partial U_s(\mathbf{x}_s)}{\partial x_r} < q_r \Rightarrow x_r = 0 \text{ and } x_r > 0 \Rightarrow \frac{\partial U_s(\mathbf{x}_s)}{\partial x_r} = q_r \tag{13}
$$

$$
y_l < c_l \Rightarrow p_l = 0 \text{ and } p_l > 0 \Rightarrow y_l = c_l \tag{14}
$$

Comparing with (6)–(7) we conclude that, if a MP-TCP algorithm defined by (3)–(4) has an underlying utility function U_s, then we must have

$$
\frac{\partial U_s(\mathbf{x}_s)}{\partial x_r} = \phi_r(\mathbf{x}_s) \quad r \in s, \ x_r > 0 \tag{15}
$$

Given $\phi_r(\mathbf{x}_s)$, (15) has a continuously differentiable solutions $U_s(\mathbf{x}_s)$ if and only if the Jacobian of $\Phi_s(\mathbf{x}_s)$ is symmetric, i.e., if and only if

$$\frac{\partial \Phi(\mathbf{x}_s)}{\partial \mathbf{x}_s} = \left[\frac{\partial \Phi(\mathbf{x}_s)}{\partial \mathbf{x}_s}\right]^T$$

Appendix B
Proof of Theorem 3.2 (existence and uniqueness)

A. Proof of part 1

For any link $l \in L$, let

$$\mathbf{p}_{-l} = \{p_1, \ldots, p_{l-1}, p_{l+1}, \ldots, p_{|L|}\},$$

whose component composes of all the elements in $\mathbf{p}$ except p_l. For $l \in L$, let

$$g_l(\mathbf{p}) := c_l - \sum_{r:l\in r} x_r = c_l - \sum_{s:r\in s, l\in r} y_l^s(p_l, \mathbf{p}_{-l})$$

and $h_l(\mathbf{p}) := -g_l^2(\mathbf{p})$. According to C1, we have the following two facts, which will be used in the proof.

- $g_l(\mathbf{p})$ is a nondecreasing function of p_l on $\mathbb{R}_+$ since $y_l^s(\mathbf{p})$ is a nonincreasing function of p_l.
- $\lim_{p_l\to\infty} g_l(p_l, \mathbf{p}_{-l}) = c_l$ since $\lim_{p_l\to\infty} y_l^s(\mathbf{p}) = 0$.

Next, we will show that $h_l(\mathbf{p})$ is a quasi-concave function of p_l. In other words, for any fixed $\mathbf{p}_{-l}$, the set $S_a := \{p_l \mid h_l(\mathbf{p}) \geq a\}$ is a convex set. If $g_l(0, \mathbf{p}_{-l}) \geq 0$, then

$$g_l(p_l, \mathbf{p}_{-l}) \geq g_l(0, \mathbf{p}_{-l}) \geq 0 \quad \forall p_l \geq 0,$$

which means $h_l(p_l, \mathbf{p}_{-l})$ is a nonincreasing function of p_l, hence is a quasi-concave function of p_l and

$$\arg\max_{p_l} h_l(p_l, \mathbf{p}_{-l}) = 0. \tag{16}$$

On the other hand, if $g_l(0, \mathbf{p}_{-l}) < 0$, then there exists a $p_l^* > 0$ such that $g_l(p_l^*, \mathbf{p}_{-l}) = 0$ since $g_l(\cdot)$ is continuous and $\lim_{p_l\to\infty} g_l(p_l, \mathbf{p}_{-l}) = c_l > 0$. Note that $g_l(\mathbf{p})$ is a nondecreasing function of p_l, then $h_l(p_l, \mathbf{p}_{-l})$ is nondecreasing for $p_l \in [0, p_l^*]$ and nonincreasing for $p_l \in [p_l^*, \infty)$. Hence, $h_l(p_l, \mathbf{p}_{-l})$ is also a quasi-concave function of p_l in this case and

$$\max_{p_l} h_l(p_l, \mathbf{p}_{-l}) = 0. \tag{17}$$

By Nash theorem, if $h_l(p_l, \mathbf{p}_{-l})$ is a quasi-concave function of p_l for all $l \in L$ and $\mathbf{p}$ is in a bounded set, then there exists a $\mathbf{p}^\star \in \mathbb{R}_+^{|L|}$ such that

$$p_l^\star = \arg\max_{p_l\in\mathbb{R}_+} h_l(p_l, \mathbf{p}_{-l}^*).$$

According to (16) and (17), for any $l \in L$, either $p_l^* > 0$ or $g_l^*(\mathbf{p}^*) > 0$ but not both holds at any time. Therefore $\mathbf{p}^*$ satisfies Eqn. (7). Since $\mathbf{q} = R^T\mathbf{p}$, there exists an $\mathbf{x}^*$ to (6). Hence there exists at least one solution $(\mathbf{x}, \mathbf{p})$ that satisfies (6) and (7).

B. Proof of part 2

Lemma B.1: Assume a function $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and $\left[\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x})\right]^+$ is negative definite for all $\mathbf{x}$. Then for any $\mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n$,

$$(\mathbf{x}_1 - \mathbf{x}_2)^T(F(\mathbf{x}_1) - F(\mathbf{x}_2)) < 0.$$

Proof: Fix any $\mathbf{x}_1 \neq \mathbf{x}_2 \in \mathbb{R}^n$. Define $A(t) := F(t\mathbf{x}_1 + (1-t)\mathbf{x}_2)$. Since $\partial F/\partial \mathbf{x}$ is continuous, there exists a $\lambda < 0$ such that the eigenvalues of $[\partial F/\partial \mathbf{x}]^+ \leq \lambda$ over the compact set $\{t\mathbf{x}_1 + (1-t)\mathbf{x}_2 \mid 0 \leq t \leq 1\}$. Then

$$\begin{aligned}
& (\mathbf{x}_1 - \mathbf{x}_2)^T(F(\mathbf{x}_1) - F(\mathbf{x}_2)) \\
= & \int_0^1 (\mathbf{x}_1 - \mathbf{x}_2)^T \frac{dA}{dt}(\tau)\, d\tau \\
= & \int_0^1 (\mathbf{x}_1 - \mathbf{x}_2)^T \frac{\partial F}{\partial \mathbf{x}} (\tau\mathbf{x}_1 + (1-\tau)\mathbf{x}_2)\, (\mathbf{x}_1 - \mathbf{x}_2)\, d\tau \\
\leq & \lambda \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 < 0
\end{aligned}$$

■

Lemma B.2: Suppose C3 holds. Then $x_r^* > 0$ at equilibrium for all $r \in R$.

Proof: Suppose $x_r^* = 0$. Then $q_r^* \geq \phi_r(\mathbf{x}_r^*) = \infty$ by C3 and hence there is a link $l \in r$ with $p_l^* = \infty$. But then, for all paths $r' \ni l$, $q_{r'}^* = \infty$ and hence $x_{r'}^* = 0$ by C3. This implies $y_l^* = 0 < c_l$, and hence $p_l^* = 0$ by (7), contradicting $p_l^* = \infty$. ■

Recall the vector notations that $\mathbf{x} := (\mathbf{x}_s, s \in S) := (x_r, r \in s, s \in S)$ and $\Phi(\mathbf{x}) := (\Phi_s(\mathbf{x}_s), s \in S) := (\Phi_r(\mathbf{x}_s), r \in s, s \in S)$. To prove uniqueness of the equilibrium, suppose for the sake of contradiction that there are two distinct equilibrium points $(\mathbf{x}, \mathbf{p})$ and $(\hat{\mathbf{x}}, \hat{\mathbf{p}})$. By Lemma B.2 we have $\mathbf{x} > 0$ and $\hat{\mathbf{x}} > 0$. Hence (6) implies $\Phi(\mathbf{x}) = \mathbf{q} = H^T\mathbf{p}$ and $\Phi(\hat{\mathbf{x}}) = \hat{\mathbf{q}} = H^T\hat{\mathbf{p}}$. By Lemma B.1 and assumption C2 we then have

$$\begin{aligned}
0 &> (\mathbf{x} - \hat{\mathbf{x}})^T(\Phi(\mathbf{x}) - \Phi(\hat{\mathbf{x}})) \\
&= (\mathbf{x} - \hat{\mathbf{x}})^T H^T(\mathbf{p} - \hat{\mathbf{p}}) \\
&= (\mathbf{p} - \hat{\mathbf{p}})^T(\mathbf{y} - \hat{\mathbf{y}})
\end{aligned}$$

Hence

$$\mathbf{p}^T\mathbf{y} + \hat{\mathbf{p}}^T\hat{\mathbf{y}} \;<\; \mathbf{p}^T\hat{\mathbf{y}} + \hat{\mathbf{p}}^T\mathbf{y} \tag{18}$$

Equilibrium condition (7) implies

$$\mathbf{p}^T(\mathbf{c} - \mathbf{y}) \;=\; 0 \quad \text{and} \quad \hat{\mathbf{p}}^T(\mathbf{c} - \hat{\mathbf{y}}) \;=\; 0 \tag{19}$$

$$\mathbf{y} \;\leq\; \mathbf{c} \quad \text{and} \quad \hat{\mathbf{y}} \;\leq\; \mathbf{c} \tag{20}$$

Substituting (19) into (18) yields

$$\begin{aligned}
\mathbf{p}^T\mathbf{c} + \hat{\mathbf{p}}^T\mathbf{c} &< \mathbf{p}^T\hat{\mathbf{y}} + \hat{\mathbf{p}}^T\mathbf{y} \\
\mathbf{p}^T(\mathbf{c} - \hat{\mathbf{y}}) + \hat{\mathbf{p}}^T(\mathbf{c} - \mathbf{y}) &< 0
\end{aligned}$$

But (20) implies that the left-hand side of the last inequality is nonnegative (since $\mathbf{p} \geq 0$, $\hat{\mathbf{p}} \geq 0$), a contradiction. Hence the equilibrium is unique.

Appendix C
Proof of Theorem 3.3 (stability)

We will construct a Lyapunov function and use LaSalle's invariance principle [13] to prove global asymptotic stability of the unique equilibrium point $(\mathbf{x}^*, \mathbf{p}^*)$. Define $\delta\mathbf{x} := \mathbf{x} - \mathbf{x}^\star$, $\delta\mathbf{p} := \mathbf{p} - \mathbf{p}^\star$. Consider the candidate Lyapunov function:

$$V(\mathbf{x}, \mathbf{p}) = \sum_{r\in R} \int_{x_r^*}^{x_r} \frac{z - x_r^*}{k_r(z)} dz + \frac{1}{2} \sum_{l\in L} \frac{\delta p_l^2}{\gamma_l} \qquad (21)$$

By definition, $V(\mathbf{x}, \mathbf{p}) > 0$ for all $(\mathbf{x}, \mathbf{p}) \neq (\mathbf{x}^*, \mathbf{p}^*)$ and $V(\mathbf{x}, \mathbf{p}) = 0$ if $(\mathbf{x}, \mathbf{p}) = (\mathbf{x}^*, \mathbf{p}^*)$. Furthermore V is radially unbounded, i.e., $V(\mathbf{x}, \mathbf{p}) \to \infty$ as $\|(\mathbf{x}, \mathbf{p})\|_2 \to \infty$. Finally

$$\dot{V}(\mathbf{x}, \mathbf{p}) = \sum_{r\in R} \frac{1}{k_r(x_r)} \delta x_r \dot{x}_r + \sum_{l\in L} \frac{1}{\gamma_l} \delta p_l \dot{p}_l$$

If $\delta x_r \neq 0$ then we have (since $k_r(\mathbf{x}_s) = k_r(x_r)$)

$$\begin{aligned} \frac{1}{k_r(x_r)} \delta x_r \dot{x}_r &= \delta x_r \left(\phi_r(\mathbf{x}_s) - q_r\right)_{x_r}^+ \\ &\leq \delta x_r \left(\phi_r(\mathbf{x}_s) - q_r\right) \\ &= \delta x_r \left(\phi_r(\mathbf{x}_s) - \phi_r(\mathbf{x}_s^*) - \delta q_r\right) \end{aligned}$$

The first inequality holds since $(\phi_r(\mathbf{x}_s) - q_r)_{x_r}^+ = \phi_r(\mathbf{x}_s) - q_r$ if $x_r > 0$ and $\phi_r(\mathbf{x}_s) - q_r \leq 0$, $\delta x_r = -x_r^*$ if $x_r = 0$. The last equality holds since $\phi_r(\mathbf{x}_s^*) = q_r^*$ by Lemma B.2 and (6). Hence

$$\begin{aligned} \sum_{r\in R} \frac{1}{k_r(x_r)} \delta x_r \dot{x}_r &\leq \delta\mathbf{x}^T (\Phi(\mathbf{x}) - \Phi(\mathbf{x}^*)) - \delta\mathbf{x}^T \delta\mathbf{q} \\ &< -\delta\mathbf{x}^T H^T \delta\mathbf{p} \end{aligned}$$

where the last inequality holds since $\delta\mathbf{x}^T \left(\phi(\mathbf{x}) - \phi(\mathbf{x}^*)\right) < 0$ by Lemma B.1 and assumption C2. Similarly

$$\frac{1}{\gamma_l} \delta p_l \dot{p}_l \;=\; \delta p_l (y_l - c_l)_{p_l}^+ \;\leq\; \delta p_l (y_l - c_l) \;\leq\; \delta p_l \delta y_l$$

where the last inequality holds since $\delta p_l c_l \geq \delta p_l y_l^*$ by the equilibrium condition (7). Hence

$$\sum_{l\in L} \frac{1}{\gamma_l} \delta p_l \dot{p}_l \quad \leq \quad \delta\mathbf{p}^T H \delta\mathbf{x}$$

Therefore if $\delta\mathbf{x} \neq 0$ then

$$\dot{V}(\mathbf{x}, \mathbf{p}) \quad < \quad -\delta\mathbf{x}^T H^T \delta\mathbf{p} + \delta\mathbf{p}^T H \delta\mathbf{x} \;=\; 0$$

and if $\delta\mathbf{x} = 0$ then $\dot{V}(\mathbf{x}, \mathbf{p}) = 0$. This means $\dot{V}(\mathbf{x}, \mathbf{p}) \leq 0$ and V is indeed a Lyapunov function.

Consider the set

$$Z \quad := \quad \{\ (\mathbf{x}(t), \mathbf{p}(t)) \,|\, \dot{V}(\mathbf{x}(t), \mathbf{p}(t)) = 0 \text{ for all } t \geq 0\ \}$$

of trajectories on which $\dot{V} \equiv 0$. If the only trajectory in Z is the trivial trajectory $(\mathbf{x}, \mathbf{p}) \equiv (\mathbf{x}^*, \mathbf{p}^*)$ then LaSalle's invariance principle implies that $(\mathbf{x}^*, \mathbf{p}^*)$ is globally asymptotically stable. We now show that this is indeed the case.

As shown above $\dot{V} \equiv 0$ implies $\delta\mathbf{x} \equiv 0$, i.e., any trajectory $(\mathbf{x}(t), \mathbf{p}(t))$ in Z must have $\mathbf{x}(t) = \mathbf{x}^*$ for all $t \geq 0$. This means $\dot{\mathbf{x}} \equiv 0$ and hence, for all $t \geq 0$, $\mathbf{q}(t) = \Phi(\mathbf{x}(t))$ since $\mathbf{x}(t) = \mathbf{x}^* > 0$ by Lemma B.2. That is, for all $t \geq 0$, $H^T \mathbf{p}(t) = \Phi(\mathbf{x}^*)$ and hence $\mathbf{p}(t) = \mathbf{p}^*$ since H has full row rank by C3. Therefore $(\mathbf{x}, \mathbf{p}) \equiv (\mathbf{x}^*, \mathbf{p}^*)$ is indeed the only trajectory in Z. This completes the proof of Theorem 3.3.

Appendix D
Proof of Theorem 3.4 (friendliness)

Let the MP-TCP source be defined by

$$\phi_r(\mathbf{x}_s; \mu) = \mu \tilde{\phi}_r(\mathbf{x}_s) + (1 - \mu) \hat{\phi}_r(\mathbf{x}_s), \quad \mu \in [0, 1]$$

Algorithm $\hat{M}$ and $\tilde{M}$ corresponds to $\mu = 0$ and $\mu = 1$ respectively. Let x_g and τ_g be the throughput and RTT of the TCP NewReno source in Fig. 1. The equilibrium is defined by $F(\mathbf{x}, \mu) = 0$ where $\mathbf{x} := (\mathbf{x}_s, x_g)$ and F is given by:

$$\begin{aligned} \Phi_s(\mathbf{x}_s; \mu) - \frac{1}{\tau_g^2 x_g^2} \mathbf{1} &= 0 \\ \mathbf{1}^T \mathbf{x}_s + x_g &= c \end{aligned}$$

where the first equation follows from

$$p^* = \frac{1}{\tau_g^2 x_g^2} = \phi_r(\mathbf{x}_s; \mu), \quad r \in s$$

and p^* is the congestion price at the bottleneck link. Applying the implicit function theorem, we get

$$\begin{aligned} \frac{d\mathbf{x}}{d\mu} &= -\left(\frac{\partial F}{\partial \mathbf{x}}\right)^{-1} \frac{\partial F}{\partial \mu} \\ &= -\begin{bmatrix} \frac{\partial \Phi_s}{\partial \mathbf{x}_s} & \frac{2}{x_g^3} \mathbf{1} \\ \mathbf{1}^T & 1 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\Phi}_s(\mathbf{x}_s) - \hat{\Phi}_s(\mathbf{x}_s) \\ 0 \end{bmatrix} \end{aligned}$$

where the inverse exists by condition C2. C2 also guarantees the inverse of $\frac{\partial \Phi_s}{\partial \mathbf{x}_s}(\mathbf{x}_s; \mu)$, denoted by $D(\mu)$; C4 ensures $\sum_{i\in s} D_{ij}(\mu) \leq 0$. Let

$$A \;:=\; \frac{\partial \Phi_s}{\partial \mathbf{x}_s} - \frac{2}{x_g^3} \mathbf{1}\mathbf{1}^T \quad \text{and} \quad d \;:=\; 1 - \frac{2}{x_g^3} \sum_{i,j} D_{ij}(\mu)$$

Then

$$\begin{bmatrix} \frac{\partial \Phi_s}{\partial \mathbf{x}_s} & \frac{2}{x_p^3} \mathbf{1} \\ \mathbf{1}^T & 1 \end{bmatrix}^{-1} \quad = \quad \begin{bmatrix} A^{-1} & -D\mathbf{1}d \\ -d\mathbf{1}^T A^{-1} & d^{-1} \end{bmatrix}$$

Thus

$$\begin{aligned} \mathbf{1}^T \frac{\partial \mathbf{x}_s}{\partial \mu} &= -[\mathbf{1}^T 0] \left(\frac{\partial F}{\partial \mathbf{x}}\right)^{-1} \frac{\partial F}{\partial \mu} \\ &= -\mathbf{1}^T A^{-1} (\tilde{\Phi}_s(\mathbf{x}_s) - \hat{\Phi}_s(\mathbf{x}_s)) \end{aligned} \qquad (22)$$

By matrix inverse formula,

$$\begin{aligned} A^{-1} &= \left(\frac{\partial \Phi_s}{\partial \mathbf{x}_s} - \frac{2}{x_g^3} \mathbf{1}\mathbf{1}^T\right)^{-1} \\ &= D(\mu) + \frac{1}{\frac{x_g^3}{2} - \mathbf{1}^T D(\mu) \mathbf{1}} D(\mu) \mathbf{1}\mathbf{1}^T D(\mu) \end{aligned}$$

Substitute it into (22), we have

$$\begin{aligned} &\mathbf{1}^T A^{-1} (\hat{\Phi}_s(\mathbf{x}_s) - \tilde{\Phi}_s(\mathbf{x}_s)) \\ &= \left(1 + \frac{\mathbf{1}^T D(\mu) \mathbf{1}}{\frac{x_g^3}{2} - \mathbf{1}^T D(\mu) \mathbf{1}}\right) \mathbf{1}^T D(\mu) (\tilde{\Phi}_s(\mathbf{x}_s) - \hat{\Phi}_s(\mathbf{x}_s)) \\ &= \frac{x_g^3}{x_g^3 - 2\mathbf{1}^T D(\mu) \mathbf{1}} \sum_{r\in s} \left(\sum_{i\in s} D_{ir}(\mu)\right) (\tilde{\phi}_r(\mathbf{x}_s) - \hat{\phi}_r(\mathbf{x}_s)) \\ &\leq 0 \end{aligned}$$

where the inequality follows because $D(\mu)$ is negative definite, $\sum_{i\in s} D_{ir}(\mu) < 0$ and $\tilde{\phi}_r(\mathbf{x}_s) - \hat{\phi}_r(\mathbf{x}_s) \geq 0$. Thus we have $\mathbf{1}^T \frac{\partial \mathbf{x}_s}{\partial \mu} \geq 0$ for $\mu \in [0,1]$, i.e., the aggregate throughput of the MP-TCP over its available paths is increasing in μ. This means $\tilde{M}$ (corresponding to $\mu = 1$) will attain a higher throughput than $\hat{M}$ (corresponding to $\mu = 0$) when separately sharing the test network in Fig. 1 with the same SP-TCP.

APPENDIX E
PROOF OF THEOREM 3.5 (RESPONSIVENESS)

A. Proof of part 1

Fix any eigenvalue λ of J^*. Let $\mathbf{z} := (\mathbf{x}, \mathbf{p}) \in Z$ be the corresponding eigenvector with $\|\mathbf{z}\|_2 = 1$. Then we have

$$\lambda \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \Lambda_k & 0 \\ & \Lambda_\gamma \end{bmatrix} \begin{bmatrix} \frac{\partial \Phi}{\partial \mathbf{x}} & -H^T \\ H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}$$

Hence

$$\lambda \begin{bmatrix} \Lambda_k^{-1} & 0 \\ & \Lambda_\gamma^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \frac{\partial \Phi}{\partial \mathbf{x}} & -H^T \\ H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}$$

Premultiplying $\mathbf{z}^H$ on both sides, we have

$$\lambda \;=\; \frac{\mathbf{x}^H \frac{\partial \Phi}{\partial \mathbf{x}} \mathbf{x} \;+\; (\mathbf{p}^H H \mathbf{x} - \mathbf{x}^H H^T \mathbf{p})}{\mathbf{x}^H \Lambda_k^{-1} \mathbf{x} + \mathbf{p}^H \Lambda_\gamma^{-1} \mathbf{p}}$$

The denominator is real and positive, and $(\mathbf{p}^H H \mathbf{x} - \mathbf{x}^H H^T \mathbf{p})$ in the numerator is imaginary. Hence

$$\begin{aligned} \mathbf{Re}(\lambda) &= \frac{\mathbf{Re}\left(\mathbf{x}^H \frac{\partial \Phi}{\partial \mathbf{x}} \mathbf{x}\right)}{\mathbf{x}^H \Lambda_k^{-1} \mathbf{x} + \mathbf{p}^H \Lambda_\gamma^{-1} \mathbf{p}} \\ &= \frac{\mathbf{x}^H \left[\frac{\partial \Phi}{\partial \mathbf{x}}\right]^+ \mathbf{x}}{\mathbf{x}^H \Lambda_k^{-1} \mathbf{x} + \mathbf{p}^H \Lambda_\gamma^{-1} \mathbf{p}} \;<\; 0 \end{aligned}$$

where the last inequality holds because the numerator is negative by condition C2 and the denominator is positive. Since this holds for all eigenvalues λ of J^*, the linearized system (10) is stable. Moreover $\mathbf{Re}(\lambda) \leq \overline{\lambda}(J^*) \leq 0$ as desired.

B. Proof of part 2

Consider two MP-TCP algorithms $(\hat{K}, \hat{\Phi})$ and $(\tilde{K}, \tilde{\Phi})$ such that

$$\hat{K}_s \;\geq\; \tilde{K}_s \quad \text{and} \quad \frac{\partial \hat{\Phi}_s}{\partial \mathbf{x}_s} \preceq \frac{\partial \tilde{\Phi}_s}{\partial \mathbf{x}_s} \qquad \text{for all } s \in S$$

For any (nonzero) $\mathbf{z} = (\mathbf{x}, \mathbf{p}) \in Z$ we have

$$0 \;\leq\; \mathbf{x}^H \hat{\Lambda}_k^{-1} \mathbf{x} \;\;\leq\;\; \mathbf{x}^H \tilde{\Lambda}_k^{-1} \mathbf{x} \tag{23}$$

$$\mathbf{x}^H \left[\frac{\partial \hat{\Phi}}{\partial \mathbf{x}}\right]^+ \mathbf{x} \;\;\leq\;\; \mathbf{x}^H \left[\frac{\partial \tilde{\Phi}}{\partial \mathbf{x}}\right]^+ \mathbf{x} \;<\; 0 \tag{24}$$

Hence $\overline{\lambda}(\hat{J}^*) \leq \overline{\lambda}(\tilde{J}^*)$.

APPENDIX F
PROOF OF THEOREM 3.6 (TRADEOFF)

Fix an s. Let $f_r(\mathbf{x}_s) := \hat{\phi}_r(\mathbf{x}_s) - \tilde{\phi}_r(\mathbf{x}_s)$ and $F(\mathbf{x}_s) := (f_r(\mathbf{x}_s), r \in s) = \hat{\Phi}_s(\mathbf{x}_s) - \tilde{\Phi}_s(\mathbf{x}_s)$. Suppose for the sake of contradiction that $\partial \hat{\Phi}_s(\mathbf{x}_s)/\partial \mathbf{x}_s \;\preceq\; \partial \tilde{\Phi}_s(\mathbf{x}_s)/\partial \mathbf{x}_s$ but $\hat{\Phi}_s(\mathbf{x}_s) \geq \tilde{\Phi}_s(\mathbf{x}_s)$ does not hold, i.e., there exists a finite $\mathbf{x}_s^0$ and a $r \in s$ such that

$$f_r(\mathbf{x}_s^0) \;=\; \hat{\phi}_r(\mathbf{x}_s^0) - \tilde{\phi}_r(\mathbf{x}_s^0) \quad<\quad 0 \tag{25}$$

Since $[\partial F/\partial \mathbf{x}_s]^+ \;\preceq\; 0$ by assumption, a trivial modification of Lemma B.1 shows that, for all $\mathbf{x}_s \neq \mathbf{x}_s^0$, $(\mathbf{x}_s - \mathbf{x}_s^0)^T (F(\mathbf{x}_s) - F(\mathbf{x}_s^0)) \leq 0$, i.e.,

$$0 \;\;\geq\;\; \sum_{r' \in s} (x_{r'} - x_{r'}^0)\,(f_{r'}(\mathbf{x}_s) - f_{r'}(\mathbf{x}_s^0)) \tag{26}$$

Choose an $\mathbf{x}_s$ as follows: for all $r' \neq r$, choose $x_{r'} = x_{r'}^0$ and then use condition C5 to choose an $x_r < \infty$ large enough so that $x_r > x_r^0$ and $f_r(\mathbf{x}_s) > f_r(\mathbf{x}_s^0)/2$. With this $\mathbf{x}_s$, (26) becomes

$$\begin{aligned} 0 \;\;&\geq\;\; (x_r - x_r^0)\,(f_r(\mathbf{x}_s) - f_r(\mathbf{x}_s^0)) \\ &>\;\; (x_r - x_r^0) \left(-\frac{f_r(\mathbf{x}_s^0)}{2}\right) \;>\; 0 \end{aligned}$$

where the last inequality follows from (25). This is a contradiction and hence $\hat{\Phi}_s(\mathbf{x}_s) \geq \tilde{\Phi}_s(\mathbf{x}_s)$.

APPENDIX G
PROOF OF THEOREM 4.1

We will show the results hold for any $n \in \mathbb{N}_+$. Since $\lim_{n\to\infty} \|\mathbf{x}_s\|_n = \|\mathbf{x}_s\|_\infty$, the results also hold for $n = \infty$. When $\beta = 0$, it is easy to show that ϕ_r satisfies C1 and $\left[\frac{\partial \Phi_s}{\partial \mathbf{x}_s}\right]^+$ is negative semidefinite under the conditions of the theorem. We hence prove the theorem for $\beta > 0$.

A. Proof of part 1

Fix any $n \in \mathbb{N}_+$ and $\beta > 0$. Fix any finite $\mathbf{p} \geq 0$ such that $q_r > 0$ for all r. Fix any $s \in S$. We now show that there exists an $\mathbf{x}_s > 0$ that satisfies (6), in particular $\phi_r(\mathbf{x}_s) = q_r$, in two steps.

First, there exists an $\mathbf{x}_s$ that satisfies $\phi_r(\mathbf{x}_s) = q_r$ if and only if

$$\phi_r(\mathbf{x}_s) = \frac{2}{\tau_r^2\,\|\mathbf{x}_s\|_1^2} \left(1 \;+\; \beta\left(\frac{\|\mathbf{x}_s\|_n}{x_r} - 1\right)\right) = q_r, \tag{27}$$

which is equivalent to

$$\frac{x_r}{\|\mathbf{x}_s\|_n} \;\;=\;\; \frac{2\beta}{2\beta + q_r \tau_r^2 \|\mathbf{x}_s\|_1^2 - 2} \tag{28}$$

Since this holds for all $r \in s$, we have

$$\begin{aligned} 1 \;\;&=\;\; \sum_{r\in s} \left(\frac{x_r}{\|\mathbf{x}_s\|_n}\right)^n \\ &=\;\; \sum_{r\in s} \left(\frac{2\beta}{2\beta + q_r \tau_r^2 \|\mathbf{x}_s\|_1^2 - 2}\right)^n \;\;=:\;\; \psi\left(\|\mathbf{x}_s\|_1^2\right) \end{aligned} \tag{29}$$

Clearly $\psi(C) \to 0$ as $C \to \infty$. Let

$$\underline{C} \;:=\; \frac{2}{\min_{r\in s}\, q_r\tau_r^2} \tag{30}$$

Then $\underline{C} < \infty$ since $q_r > 0$ for all r by assumption. Moreover $q_r\tau_r^2\underline{C} \geq 2$ for all $r \in s$ and hence

$$\psi(\underline{C}) \;=\; 1 + \sum_{r\neq \underline{r}} \left(\frac{2\beta}{2\beta + q_r\tau_r^2\underline{C} - 2}\right)^n \;>\; 1$$

where $\underline{r}$ is a minimizing $r \in s$ in (30). Since $\psi(C)$ is continuous, there exists an $\tilde{C} \in [\underline{C}, \infty)$ with $\psi(\tilde{C}) = 1$. Moreover such a $\tilde{C}$ is unique since $\psi(C)$ is strictly decreasing.

Finally consider the set of $\mathbf{x}_s$ with $\|\mathbf{x}_s\|_1^2 = \tilde{C}$. All such $\mathbf{x}_s$ satisfy (28) with

$$x_r \;=\; \frac{2\beta}{2\beta + q_r\tau_r^2\tilde{C} - 2}\,\|\mathbf{x}_s\|_n \;=:\; a_r\,\|\mathbf{x}_s\|_n \tag{31}$$

But $\tilde{C} = \|\mathbf{x}_s\|_1^2 = \left(\sum_{r\in s} a_r\,\|\mathbf{x}_s\|_n\right)^2$, implying

$$\|\mathbf{x}_s\|_n \;=\; \frac{\sqrt{\tilde{C}}}{\sum_{r\in s} a_r}$$

In summary, given any finite $\mathbf{p} \geq 0$ such that $q_r > 0$ for all r, a solution $\mathbf{x}_s > 0$ to (28) is *uniquely* given by

$$x_r \;=\; \frac{a_r}{\sum_{k\in s} a_k}\sqrt{\tilde{C}}, \qquad r \in s \tag{32}$$

where

$$a_r \;:=\; \frac{2\beta}{2\beta + q_r\tau_r^2\tilde{C} - 2}$$

and $\tilde{C} = \|\mathbf{x}_s\|_1^2$ is the unique value at which $\psi(\tilde{C}) = 1$.

We now prove the other conditions in C1:

$$\frac{\partial y_l^s(\mathbf{p})}{\partial p_l} \leq 0, \qquad \lim_{p_l\to\infty} y_l^s(\mathbf{p}) = 0$$

According to (29), we can show that $\tilde{C}$ is a decreasing function of q_r and $q_r\tau_r^2\tilde{C}$ is an increasing function of q_r for $r \in s$. Thus, $\tilde{C}$ is a decreasing function of p_l and $q_r\tau_r^2\tilde{C}$ is an increasing of p_l if $l \in r$ because $q_r = \sum_{l\in L} H_{lr}p_l$. For each $l \in L$, let $s_l := \{r \mid l \in r, r \in s\}$, then by definition and (32), we have

$$y_l^s(\mathbf{p}) = \frac{\sum_{r\in s_l} a_r}{\sum_{r\in s} a_r}\sqrt{\tilde{C}} = \frac{\sum_{r\in s_l} a_r}{\sum_{r\in s_l} a_r + \sum_{r\notin s_l} a_r}\sqrt{\tilde{C}}.$$

Since a_r is a decreasing function of $q_r\tau_r^2\tilde{C}$, it is also a decreasing function of p_l if $l \in r$. Recall that $\sqrt{\tilde{C}}$ is also a decreasing function of p_l, $y_l^s(\mathbf{p})$ is thus a decreasing function of p_l, in other words, $\frac{\partial y_l^s(\mathbf{p})}{\partial p_l} \leq 0$.

On the other hand, as $p_l \to \infty$, $q_r \to \infty$ for all paths r traversing l. Then $x_r \to 0$ by (27) for $l \in r$, which shows $\lim_{p_l\to\infty} y_l^s(\mathbf{p}) = 0$.

B. Proof of part 2

To prove $\phi_r(\mathbf{x}_s)$ satisfies C2 and C3 for $\beta > 0$, we will show that the Jacobian $\partial\Phi_s(\mathbf{x}_s)/\partial\mathbf{x}_s$ is negative definite if $0 < \beta \leq 1$, $|s| \leq 8$ and τ_r are the same for $r \in s$. Other properties of C2 and C3 are easy to prove and we omit the proof. Fix an s and let $\tau_r = \tau$, the common round-trip time for all $r \in s$.

Let $\Lambda_s := \text{diag}\{\mathbf{x}_s\}$ and

$$\mathbf{a}_s \;:=\; \left(\frac{2x_r}{\|\mathbf{x}_s\|_1} - \frac{x_r^n}{\|\mathbf{x}_s\|_n^n},\; r \in s\right)$$

Then the Jacobian of Φ_s at $\mathbf{x}_s$ is

$$\frac{\partial\Phi_s}{\partial\mathbf{x}_s} = -\frac{4(1-\beta)}{\tau^2\|\mathbf{x}_s\|_1^3}\mathbf{1}\mathbf{1}^T - 2\beta\frac{\|\mathbf{x}_s\|_n}{\tau^2\|\mathbf{x}_s\|_1^2}\Lambda_s^{-1}\left(I_{|s|} + \mathbf{1}\mathbf{a}_s^T\right)\Lambda_s^{-1}$$

and it is negative definite for $\beta > 0$ if $\left[I_{|s|} + \mathbf{1}\mathbf{a}_s^T\right]^+$ is positive definite. We now show that this is indeed the case when $|s| \leq 8$, i.e., for any $\mathbf{z}_s \in \mathbb{R}^{|s|}$,

$$\mathbf{z}_s^T(I_{|s|} + \mathbf{1}\mathbf{a}_s^T)\mathbf{z}_s = \|\mathbf{z}_s\|_2^2 + \sum_{r\in s} z_r \sum_{r\in s} a_r z_r \;>\; 0 \tag{33}$$

By Lemma G.1 below, $\mathbf{1}^T\mathbf{a}_s = 1$ and $\|\mathbf{a}_s\|_2^2 \leq 1$. Then (33) follows from Lemma G.2 below provided $|s| \leq 8$. Hence the Jacobian is negative definite.[3] The proof of Theorem 4.1 is complete after Lemmas G.1 and G.2 are proved.

To show that it satisfies C3, it follows directly from (27) that if $x_r = 0$ then $\phi_r(\mathbf{x}_s) = \infty$. It is also clear from (27) that the converse holds. This proves C3.

Lemma G.1: Fix any integer $p \geq 1$. Given any $\mathbf{x} \in \mathbb{R}_+^m$, define a vector $\mathbf{a}$ in $\mathbb{R}^m$ as follows:

$$a_i \;=\; \frac{2x_i}{\sum_{j=1}^m x_j} - \frac{x_i^p}{\sum_{j=1}^m x_j^p}, \quad 1 \leq i \leq m$$

Then $\sum_{i=1}^m a_i = 1$ and $\sum_{i=1}^m a_i^2 \leq 1$.

Proof: It is obvious that $\sum_{i=1}^m a_i = 1$. To show $\sum_{i=1}^m a_i^2 \leq 1$, we have

$$\begin{aligned}
\sum_{i=1}^m a_i^2 &= \frac{\sum_i x_i^{2p}}{\left(\sum_j x_j^p\right)^2} + \frac{4\sum_i x_i^2}{\left(\sum_j x_j\right)^2} - \frac{4\sum_i x_i^{p+1}}{\left(\sum_j x_j^p\right)\left(\sum_j x_j\right)} \\
&\leq 1 + \frac{4\sum_i x_i^2}{\left(\sum_j x_j\right)^2} - \frac{4\sum_i x_i^{p+1}}{\left(\sum_j x_j^p\right)\left(\sum_j x_j\right)} \\
&= 1 - 4\,\frac{\sum_{1\leq i<j\leq m} x_i x_j (x_i - x_j)\left(x_i^{p-1} - x_j^{p-1}\right)}{\left(\sum_j x_j\right)^2\left(\sum_j x_j^p\right)} \\
&\leq 1
\end{aligned}$$

■

[3] If $\beta = 0$ the Jacobian degenerates to

$$\frac{\partial\Phi_s}{\partial\mathbf{x}_s} = -\frac{4}{\tau^2\|\mathbf{x}_s\|_1^3}\mathbf{1}\mathbf{1}^T, \tag{34}$$

which is merely negative semidefinite.

Lemma G.2: Let $\mathbf{a} \in \mathbb{R}^m$ that satisfies $\sum_{i=1}^m a_i = 1$ and $\sum_{i=1}^m a_i^2 \leq 1$. Then for any nonzero $\mathbf{z} \in \mathbb{R}^m$ we have

$$f(\mathbf{z}) := \sum_{i=1}^m z_i^2 + \sum_{i=1}^m z_i \sum_{i=1}^m a_i z_i \quad > \quad 0$$

provided $m \leq 8$.

Proof: Given any M let $Z_M := \{z \mid \sum_{i=1}^m z_i = M\}$. It then suffices to show that, for every $M \in \mathbb{R}$, $f(z) > 0$ for $z \in Z_M$. Given any M, consider

$$\min_{\mathbf{z} \in Z_M} f(\mathbf{z}) \quad = \quad \min_{\mathbf{z} \in Z_M} \sum_{i=1}^m z_i^2 + M \sum_{i=1}^m a_i z_i \qquad (35)$$

Its Lagrangian is

$$L(\mathbf{z}, \mu) = \sum_{i=1}^m z_i^2 + M \sum_{i=1}^m a_i z_i + \mu \left(\sum_{i=1}^m z_i - M \right)$$

where μ is the Lagrange multiplier. Setting $\partial L / \partial z_i = 0$ for all $1 \leq i \leq m$ and substitute it into $\sum_{i=1}^m z_i = M$, we obtain the unique minimizer given by $\mu = -3M/m$ and $z_i = \frac{M}{2}(\frac{3}{m} - a_i)$. Then

$$\min_{\mathbf{z} \in Z_M} f(\mathbf{z}) \quad = \quad \frac{M^2}{4} \left(\frac{9}{m} - \sum_{i=1}^m a_i^2 \right) \;\geq\; \frac{M^2}{4} \left(\frac{9}{m} - 1 \right)$$

Hence, when $M \neq 0$, $\min_{\mathbf{z} \in Z_M} f(\mathbf{z}) > 0$ if $n < 9$. When $\mathbf{z}$ is nonzero but $M = 0$, then $f(\mathbf{z}) > 0$ from (35). ■

Appendix H
Proof of Lemma 3.1

By the definition of $D_k(A_k)$, we have

$$\begin{aligned}
\mathbb{E}\left[D_k(A_k) \mid \sum_{i,j} a_{ij} \geq 1 \right] &= d_k \mathbb{P}\left(\sum_j a_{kj} \geq 1 \mid \sum_{i,j} a_{ij} \geq 1 \right) \\
&= d_k \frac{\mathbb{P}(\sum_j a_{kj} \geq 1)}{\mathbb{P}(\sum_{i,j} a_{ij} \geq 1))} \\
&= d_k \frac{q_k |A_k|}{\sum_i q_i |A_i|} + o\left(\sum_i q_i \right),
\end{aligned}$$

where the last equality follows from the independence of a_{ij} and $\mathbb{P}(\sum_j a_{kj} \geq 1) = 1 - (1 - q_k)^{|A_k|} = |A_k| q_k + o(q_k)$, $\mathbb{P}(\sum_{ij} a_{ij} \geq 1) = 1 - \prod_i (1 - q_i)^{|A_i|} = \sum_i |A_i| q_i + \sum_i o(q_i)$. Thus,

$$\mathbb{E}\left[\sum_i D_k(A_k) \mid \sum_{i,j} a_{ij} \geq 1 \right] = \frac{\sum_k d_k q_k |A_k|}{\sum_k q_k |A_k|} + o\left(\sum_k q_k \right).$$