Stability for Network Flow Control Against Disturbances and Time-Delays

W. Bian, M. French and A. Prügel-Bennett
School of Electronics and Computer Science,
University of Southampton, Southampton SO17 1BJ, UK
wb@ecs.soton.ac.uk, mcf@ecs.soton.ac.uk, apb@ecs.soton.ac.uk

Abstract

New stability results are presented for Kelly's algorithm for network traffic control. These are applicable to a broad class of algorithms including most of those previously discussed in the literature. Three types of results are derived; an asymptotic stability result for networks with external disturbances but without time delays, an asymptotic stability result for a class of algorithms with a linear cost with both external disturbances and time-delays, and a gain stability result for more general algorithms with both disturbances and time-delays. The time-delay results are obtained using a gap metric approach.

1 Introduction

Distributed communication networks, as exemplified by the Internet, are of huge social and economic importance. Controlling such networks to efficiently utilise the resources is unsurprisingly a very active area of research. The control mechanisms of the Internet (e.g. TCP) has evolved over time to ensure a stable flow of information. However, there are a number of ways it falls short of an ideal mechanism. It was not designed to explicitly optimise any desired properties and consequently it is clearly wasteful of the available bandwidth. A second limitation of the current protocol is that it does not allow packets of information to be treated differently depending on the urgency felt by the user (that is, it does not allow different quality-of-service). To remedy these deficiencies Kelly advanced a mathematical model for controlling the Internet or any similar communication systems in a series of seminal papers [12, 14, 15, 17] and [22]. In these model the users specify a utility function which encode their willingness to "pay" for an increased sending rate. Information about the congestion on the links used in transferring packets is collected by the packet and fedback to the sender. Kelly describes a totally decentralized mechanism for controlling the rate of flow of information based on feedback; the mechanism drives the system towards an equilibrium state which maximises a global utility function.

Kelly's solution, and its now many variants, have clearly desirable properties, but raise the question of how practical they are in real networks where some of the mathematical assumptions no longer hold. Ever since the approach was proposed there have been many attempts to characterise its robustness. One question of particular significance is that of the effect of delays in the network. Such delays arise from the time it takes for a packet to collect information

about the loading on the links of the network and feed it back to the sender. However, there are many other sources of error in the model, such as: errors in the feedback information due, for example, to rounding errors (which may be very significant if we want to minimise the space used in each packet for storing this information); discretisation errors due to the fact that discrete packets are being sent; transient effects caused by changing use over time; etc. Analysing the stability of the algorithms in the presence of time delays has attracted many researchers from both the control world and the communication network world. Under various conditions, algorithms are proved to be asymptotically stable, globally or locally. For example, linearised models are considered in [11, 20, 21, 24] and studies for nonlinear algorithms can be found in [4, 9, 16, 18] and references therein, each considers one specific dynamic controller either for single-path network or for multi-path network and most use a frequency domain approach.

This current paper also addresses these stability issues. The paper presents theorems applicable to a whole class of algorithms. This approach is particularly influenced by [4], however the analysis developed here encompasses a wide class of algorithms and develops results for systems with round trip delays by a distinct approach based on the gap metric. The results generalise many of the existing results for a variety of algorithms which have previously been considered separately. The main contributions of our paper are a number of results on the stability of algorithms to external disturbances (including in some cases time delays), where stability is interpreted in an input/output gain setting, namely that of the perturbation to the system caused by a disturbance is required to be bounded by a quantity proportional to that disturbance. This is clearly a desirable property of a controller, it is not however guaranteed. There are seeming sensible controllers which can fail catastrophically when an apparently innocuous disturbance is added to the system. The constants in the bounds we obtain can also suggest how to increase the stability of the system. By choosing a system to minimise the effect of an external disturbance we increase the theoretical guarantee on the robustness of the algorithm, which then in turn can be used to establish bounds on tolerable time delays. The only previous results of this nature that we are aware of is given in [4] where the effect of external disturbances is addressed for the case of a primal algorithm and for the corresponding dual algorithm in the case where there is no time delay. However, the conditions of [4] on the utility function is restrictive.

We derive two types of stability results; namely asymptotic stability results, and slightly weaker gain stability results. Asymptotic stability is obtained from a sensitivity analysis and has been carried out for a large class of controllers (both primal and dual algorithms) in the presence of external noise, but without delays. We also obtain an asymptotic stability result for a controller with a linear price function suffering both external disturbances and time delays. To obtain results on gain stability we use the recently developed gap metric techniques. We obtain a gain stability result for a much larger class of controllers suffering both external disturbances and time delays. The results on stability with time delays are obtained by treating the time delay as the gap uncertainty. The consideration is in an $L^\infty \times W^{1,\infty}$ setting and it is established that the gap distance between the systems with and without time delays is small when the time delay is small.

In section 5, we apply our theorems to a particular system and show how the size of the bounds obtained depend on the details of the algorithm. In particular, we consider how these depend on the structure of the network and illustrate how different network topologies of arbitrary size will influence stability, and where network topologies possess scalable stability guarantees.

2 Algorithms

2.1 The model

Consider a network consisting of a set J of n links and a set R of m routes, where each route is a collection of links and each link j has capacity c_j. Each user $r \in R$ is identified in terms of the route between its source and destination. The rate on route r at time t is denoted by $x_r(t)$. The user specifies a utility function denoted by $U_r(x_r)$ which expresses their willingness to pay for a rate x_r. Let $c = (c_j)_{j \in J}$ be the vector of link capacities; $x = (x_r)_{r \in R}$ be a vector of transmission rates and A be the incidence or routing matrix of the network with coefficients $a_{jr} = 1$ if link j is on the route r and $a_{jr} = 0$ otherwise. Using this notation, bandwidth constraints can be written $Ax \leq c$, i.e., so that for each link j, the aggregate transmission rate $(\sum_{r \in R} a_{jr} x_r)$ through any given link cannot exceed the link's capacity (c_j).

Suppose each link j is associated with a charging price (or penalty function) $p_j(y)$ per unit flow through resource j when the total flow through j is y. Kelly's approach (see [9, 15, 22]) is to seek decentralised schemes which maximise the total system utility:

$$\mathcal{U}(x) = \sum_{r \in R} U_r(x_r) - \sum_{j \in J} \int_0^{\sum_{s:j \in s} x_s} p_j(y) dy \qquad (2.1)$$

over the range

$$\{x = (x_r)_{r \in R} : 0 \leq x, Ax \leq c\}. \qquad (2.2)$$

We assume throughout that for each r, the utility function U_r is continuously differentiable, increasing and strictly concave over $[0, \infty)$, and the penalty function p_j is continuous, non-decreasing on $(-\infty, +\infty)$[1] and $\int_0^y p_j(y) dy \to \infty$ as $y \to \infty$. Consequently, the function $\mathcal{U}$ is continuous and strictly concave and admits a unique maximum. The rate allocation $x^* = (x_r^*)_{r \in R}$ at which the unconstrained maximum is reached satisfies

$$\left. \frac{\partial \mathcal{U}}{\partial x_r} \right|_{x_r = x_r^*} = U_r'(x_r^*) - \sum_{j:j \in r} p_j \left(\sum_{s:j \in s} x_s^* \right) = 0, \quad \text{for all } r \in R. \qquad (2.3)$$

Let $q_r^* = \sum_{j:j \in r} p_j \left(\sum_{s:j \in s} x_s^* \right)$. Then the equilibrium point of the system is such that

$$q_r^* = U_r'(x_r^*), \quad \text{for all } r \in R. \qquad (2.4)$$

2.2 Primal algorithms

Congestion control operates through sources or users controlling their rates x_r depending on the level of congestion in the network. The information about the network congestion is accumulated at each link and fed back to the sources through a handshake (i.e. a packet back to the source). This price information is denoted

$$q_r(t) = \sum_{j:j \in r} p_j \left(\sum_{s:j \in s} x_s(t) \right), \qquad (2.5)$$

---

[1]Usually p is defined on $[0, \infty)$ only. But when disturbances are considered, the input y can become negative.

where $p_j(y)$ is a 'price per unit flow' charged by link j when the total flow through the link is

$$y_j = \sum_{s:j \in s} x_s.$$

This information is supplied in a decentralised way by the links and fed back to the sources.

To solve the decentralised problem (2.3), we consider the following general gradient algorithm:

$$\dot{x}_r(t) = f_r(x_r(t), U'_r(x_r(t)), q_r(t)), \ r \in R, \tag{2.6}$$

where for each $r \in R$, $f_r : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a piecewise continuous function such that the equation has unique solution for any initial value $x_r(0) \geq 0$ and at the equilibrium as per (2.4):

$$f_r(x_r^*, U'_r(x_r^*), q_r^*) = f_r(x_r^*, U'_r(x_r^*), U'_r(x_r^*)) = 0. \tag{2.7}$$

Equation (2.6) describes a response by user r either to increase the transmission rate ($\dot{x}_r > 0$) or to decrease it ($\dot{x}_r < 0$) depending on user's current flow rate, utility and the price information received. The solution x_r of (2.6) is not necessarily the same as the equilibrium x_r^* but f_r will be chosen to ensure that x_r approximates x_r^* asymptotically or in some other sense.

Further to the flow rate vector $x = (x_r)_{r \in R}$ given in (2.2), we also let

$$q = (q_r)_{r \in R}, \ y = (y_j)_{j \in J}, \ \mu = (\mu_j)_{j \in J}, \ x^* = (x_r^*)_{r \in R}, \ q^* = (q_r^*)_{r \in R} \tag{2.8}$$

be the other rate vectors and let

$$U'(x) = (U'_r(x_r))_{r \in R}, \ f(x, U'(x), q) = (f_r(x_r, U'_r(x_r), q_r))_{r \in R}, \ p(y) = (p_j(y_j))_{j \in J} \tag{2.9}$$

be the function vectors of utility, algorithm and price, respectively. Then (2.6),(2.5) can be re-written in vector form:

$$\dot{x}(t) = f(x(t), U'(x(t)), q(t)), \tag{2.10}$$

and

$$q(t) = A^\top \mu(t), \ \mu(t) = p(y(t)), \ y(t) = Ax(t). \tag{2.11}$$

Together (2.10)-(2.11) form the following closed loop control system (Figure 1):
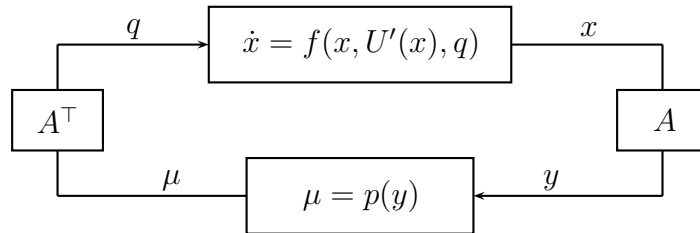

Figure 1: The closed-loop for the primal algorithm.

Equation (2.7) guarantees that x^* is an equilibrium point of equation (2.6). If, in addition,

$$\sum_{r \in R}(U_r'(x_r) - q_r)f_r(x_r, U_r'(x_r), q_r) > 0 \quad \text{for all } x \geq 0, q \geq 0, (x, q) \neq (x^*, q^*), \tag{2.12}$$

then,

$$\frac{d\mathcal{U}}{dt} = \sum_{r \in R}\frac{\partial \mathcal{U}}{\partial x_r}\dot{x}_r = \sum_{r \in R}\left(U_r'(x_r(t)) - q_r(t)\right)f_r(x_r(t), U_r'(x_r(t)), q_r(t)) \geq 0$$

and $V = \mathcal{U}(x^*) - \mathcal{U}(x)$ is a Lyapunov function for the system of equations (2.10)–(2.11). Hence, starting from any initial condition $\{x_r(0) \geq 0\}$, the unique value x^* maximising $\mathcal{U}$ is a stable point of system of equations (2.10)-(2.11), to which all trajectories converge.

A physical requirement of the system is the the flow rate $x_r(t)$ should be non-negative for all $t \geq 0$ provided it starts from a positive initial value $x_r(0) \geq 0$. This can be achieved by assuming[2]

$$f_r(0, U_r'(0), q_r) \geq 0, \quad \text{for all } r \in R, \; q_r \in \mathbb{R}. \tag{2.13}$$

Then, starting from $x_r(0) > 0$, once the flow decreases to 0, then $\dot{x}_r(t) \geq 0$ leading to $x_t(t)$ increasing or staying at 0, therefore $x_r(t) \geq 0$ for all t.

Several versions of this algorithm have been considered in the literature. All of these can be treated as special cases of (2.6)-(2.7). We now list them as examples.

Example 2.1. ([22])   Let $f_r(x_r, U_r'(x_r), q_r) = \kappa_r(x_r)(U_r'(x_r) - q_r)$ with $\kappa_r(x_r) > 0$ a differentiable function. Then we obtain the algorithm :

$$\dot{x}_r(t) = \kappa_r(x_r(t))(U_r'(x_r(t)) - q_r(t)). \tag{2.14}$$

This function f_r is locally Lipschitz, $f_r(x_r^*, U_r'(x_r^*), q_r^*) = \kappa_r(x_r^*)(U_r'(x_r^*) - q_r^*) = 0$ since $q_r^* = U_r'(x^*)$ and

$$(U_r'(x_r) - q_r)f_r(x_r, U_r'(x_r), q_r) = \kappa_r(x_r)(U_r'(x_r) - q_r)^2 \geq 0 \quad \text{for all } x_r, q_r,$$

the equality holds only at x^*, q_r^*. However, an extra requirement is needed for (2.13) to hold. A condition that is clearly sufficient and frequently assumed is

$$\lim_{x_r \to 0^+} U_r'(x_r) = +\infty.$$

Example 2.2. ([4])   We define $a = (b)_c^+$ to mean $a = b$ if $c > 0$ and $a = \max\{0, b\}$ if $c = 0$. In (2.6), let $f_r(x_r, U_r'(x_r), q_r) = \kappa_r\left(U_r'(x_r) - q_r\right)_{x_r}^+$, where $\kappa_r > 0$ is constant. Then we obtain the following algorithm:

$$\dot{x}_r(t) = \kappa_r\left(U_r'(x_r(t)) - q_r(t)\right)_{x_r(t)}^+. \tag{2.15}$$

This f_r is piecewise continuous; $f_r(x_r^*, U_r'(x_r^*), q_r^*) = 0$ and

$$(U_r'(x_r) - q_r)f_r(x_r, U_r'(x_r), q_r) \geq 0 \quad \text{for all } x_r, q_r.$$

By the definition of notation $(b)_c^+$, (2.13) holds, so $x_r(t) \geq 0$ for all t if it starts from $x_r(0) \geq 0$.

---

[2]Here we allow q_r to be negative since this is the case when disturbances are considered in the next section

Example 2.3. Consider the algorithm studied in [16]:

$$\dot{x}_r(t) = \kappa_r x_r(t) \left(1 - \frac{q_r(t)}{U_r'(x_r(t))} \right)^+_{x_r(t)}, \tag{2.16}$$

which is the case where

$$f_r(x_r, U_r'(x_r), q_r) = \kappa_r x_r(t) \left(1 - \frac{q_r(t)}{U_r'(x_r(t))} \right)^+_{x_r(t)} = \kappa_r x_r(t) \left(\frac{U_r'(x_r(t)) - q_r(t)}{U_r'(x_r(t))} \right)^+_{x_r(t)}.$$

Clearly $f_r(x_r^*, U_r'(x_r^*), q_r^*) = 0$ since $U_r'(x_r^*) = q_r^*$. If $U_r'(x) \geq 0$ for $x \neq 0$, we also have $(U_r'(x) - q_r)f_r(x, U_r'(0), q_r) \geq 0$.

Example 2.4. ([9, 13, 15]) We let $U_r(x_r) = w_r \log x_r$ and let $f_r(x_r, U_r'(x_r), q_r) = \kappa_r x_r(U_r'(x_r) - q_r)$, where $w_r > 0, \kappa_r > 0$ are constants. Then we have the algorithm

$$\dot{x}_r(t) = \kappa_r \left(w_r - x_r(t)q_r(t) \right). \tag{2.17}$$

This f_r satisfies (2.13), $f_r(x_r^*, U_r'(x_r^*), q_r^*) = 0$, and

$$(U_r'(x_r) - q_r)f_r(x_r, U_r'(x_r), q_r) = \kappa_r \frac{1}{x_r}(w_r - x_r q_r)^2 \geq 0 \quad \text{for all } x_r \geq 0, q_r \geq 0$$

with equality if and only if $w_r = x_r q_r$: that is $U'(x_r) = q_r$ which leads to $x_r = x_r^*, q_r = q_r^*$ since the equilibrium is unique.

In the above algorithms, the flow rate $x_r(t)$ could be very large at some t, and indeed may even exceed the link capacities, particularly when there are disturbances. For example, if $d_r(t)$ is a disturbance to the aggregate price q_r, algorithm (2.15) will become

$$\dot{x}_r(t) = \kappa_r \left(U_r'(x_r(t)) - q_r(t) + d_r(t) \right)^+_{x_r(t)}.$$

As $d_r(t)$ varies, $x_r(t)$ may diverge. Furthermore, observe that the primal algorithms do not even impose the requirement that the link capacities are respected at the equilibrium rate $x = x^*$, although the price p_j is typically chosen to ensure reasonable equilibrium behaviour. Hence, technically and practically, it is useful to have an upper bound for each $x_r(t)$. In many studies, it is assumed that the utility U_r should satisfy

$$U_r''(z) \leq -\lambda \quad \text{for all } z$$

for some $\lambda > 0$ ([4, 19]). The basic utility functions $U_r(x) = \log x$ and $U_r(x) = x^\alpha$ $(0 < \alpha < 1)$ do not satisfy such a requirement, however if each x_r is uniformly bounded for all $t \geq 0$, then $U_r''(x_r(t)) \leq -\lambda, \forall t \geq 0$ is achievable. Secondly, as pointed out in [19], at equilibrium, every user desires a strictly positive amount of variable bandwidth, which means x_r should lie in a finite interval. Finally, the finite capacity of each link also demands a finite flow rate for each user if the link capacities are respected.

To ensure $x_r(t)$ is bounded, we may assume $m_r > 0$ and

$$f_r(m_r, U_r'(m_r), q_r) \leq 0 \quad \text{for all } r \in R, q_r \in \mathbb{R}. \tag{2.18}$$

Hence once $x_r(t)$ reaches m_r, we will have $\dot{x}_r(t) \leq 0$ and $x_r(t)$ begin to decrease.

A concrete example satisfying the above is as follows:

6

Example 2.5. We let

$$\dot{x}_r = \kappa_r(x_r)(U'_r(x_r) - q_r)_{x_r}^{m_r}, \tag{2.19}$$

where $m_r > x_r^* \geq 0$ are a-priori determined constants and

$$(b)_x^{m_r} = \begin{cases} \min\{0, b\} & \text{if } x = m_r, \\ b & \text{if } x \in (0, m_r), \\ \max\{0, b\} & \text{if } x = 0. \end{cases}$$

Clearly, this algorithm maintains the same equilibrium point x^* together with positivity and convergence of x_r as in algorithm (2.15). Furthermore, $0 \leq x_r(t) \leq m_r$ for all $t \geq 0$.

In (2.19), the number m_r is understood as the maximal flow rate that user r can use. It is reasonable to choose $m_r = \min\{c_j : j \in r\}$. However as the model is an approximation, m_r could be allowed larger.

For our main results on primal algorithm in the next two sections, we will suppose that both (2.13) and (2.18) are satisfied.

2.3 Dual algorithms

The dual algorithm is obtained by maximising the sum of personal utilities $\sum_r U_r(x_r)$ subject to the bandwidth constraints imposed through Lagrange multipliers. The Lagrange multipliers can be interpreted as shadow prices μ_j which adjust dynamically depending on the flow rates through the links [15, 19, 22]. This leads to a dynamic equation for the shadow price

$$\dot{\mu}_j = p_j(y_j, \mu_j)$$

which is maintained by each link. Here each p_j is a function from $\mathbb{R}^2$ to $\mathbb{R}_+$. For the dual algorithm, we let $x_r = U'^{-1}_r(q_r)$ be the source controllers, (invertibility of U'_r follows from the assumption of strict concavity). Write $U'^{-1}(q) = (U'^{-1}_r(q_r))_{r \in R}$. Similar to (2.10) and (2.11), the system equations can also be written in vector form:

$$\dot{\mu}(t) = p(y(t), \mu(t)) =: (p_j(y_j(t), \mu_j(t)))_{j \in J}, \tag{2.20}$$

$$y(t) = Ax(t), \quad x(t) = U'^{-1}(q(t)), \quad q(t) = A^\top \mu(t). \tag{2.21}$$

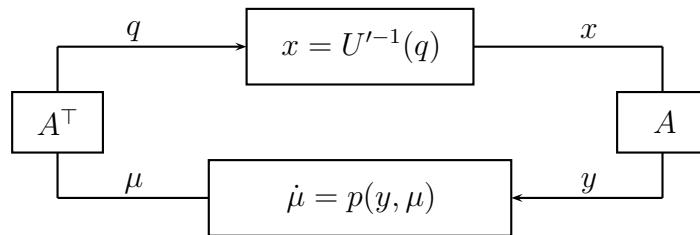Again this can be modelled as a plant and controller interconnection as described in Figure 2.

Figure 2: The closed-loop for dual algorithm.

Let x^* be the the equilibrium of (2.1) and let q^* from (2.4), $y^* = Ax^*$. Suppose that there exists a unique μ^* such that $q^* = A^\top \mu^*$, which can be guaranteed if the routing matrix A has full row rank (as assumed in [4, 22]). Also suppose

$$p(y^*, \mu^*) = 0.$$

Then μ^* is an equilibrium of (2.20)-(2.21).

Example 2.6. ([22]) A concrete dual link algorithm is given in [22] as

$$\dot{\mu}_j = h_j(\mu_j)(y_j - c_j)^+_{\mu_j} \tag{2.22}$$

where h_j is a continuous function and c_j the capacity of link j. For this algorithm, the equilibrium μ^* is such that $h_j(\mu_j^*)(y_j^* - c_j)^+_{\mu_j^*} = 0$, i.e. either $\mu_j^* = 0, y_j^* \leq c_j$ or $\mu_j^* > 0, y_j^* = c_j$.

As discussed in the last subsection on the boundedness of transmission rate x_r, we may also impose the following condition so that, under the action of the dynamic controller, the price μ_j lies in an interval $[0, \chi_j]$ with $\chi_j > 0$:

$$p_j(y_j, 0) \geq 0 \quad \text{and} \quad p_j(y_j, \chi_j) \leq 0 \quad \text{for all } y_j \in \mathbb{R}. \tag{2.23}$$

Example 2.7. Corresponding to (2.22), the algorithm

$$\dot{\mu}_j = h_j(\mu_j)(y_j - c_j)^{\chi_j}_{\mu_j}$$

satisfies condition (2.23). We will discuss this algorithm in Corollary 3.10.

Example 2.8. A further example is:

$$\dot{\mu}_j(t) = p_j(y_j(t), \mu_j(t)) =: \kappa_j(y_j(t) - h_j(\mu_j(t)))^{\chi_j}_{\mu_j(t)}, \tag{2.24}$$

where $\kappa_j > 0$ and $h_j(\mu_j)$ is the flow through link j which generates a price. The right hand side of (2.24) is described as the vector of excess demand at price μ and $\chi_j \geq 0$ is a number representing the possible maximal charge for each link. If there is no superscript 'χ_j' and subscript 'μ_j', it is the algorithm studied in [13, 15]. This closed-loop can be recognised as a tatonnement process by which prices adjust according to supply and demand ([15, 23]). Stability for this algorithm will be discussed in Corollary 3.8.

3   Stability in the presence of external disturbances

In this section, we suppose that external disturbances exist and consider the stability of both primal and dual algorithms against the disturbances. Subsequently these external disturbances will be used to model time delay effects via a graph metric/robust stability analysis.

All functions representing flow rate, price and disturbances in this section are supposed to be in the space $L^\infty(\mathbb{R}_+)$ with norm $\|x\|_{L^\infty} = \operatorname{esssup} \|x(t)\|$. Where the notation is unambiguous, we simply write $\|x\|_{L^\infty}$ as $\|x\|$. In the next section, the flow rate and the disturbance e will be assumed to be functions in $W^{1,\infty}(\mathbb{R}_+)$. For elements z_1, z_2 in any linear spaces, we define the line segment:

$$[z_1, z_2] = \{(1 - k)z_1 + kz_2 : 0 \leq k \leq 1\}.$$

3.1 Primal algorithms

A standard diagram for the closed-loop of the primal algorithm with external disturbances is as shown in Figure 3. With the vector notations given by (2.8)-(2.9), this corresponds to the equations,

$$\dot{x} = f(x, U'(x), q), \tag{3.1}$$

$$q = A^\top \mu + d, \ \mu = p(y), \ y = Ax + Ae \tag{3.2}$$

with d, e the external disturbances. We allow both d and e to be arbitrarily large in an $L^\infty(\mathbb{R}_+)$ sense with possibly negative values. The domains of functions f and p are $\mathbb{R}^3$ and $\mathbb{R}$, respectively, as stated in the last section.
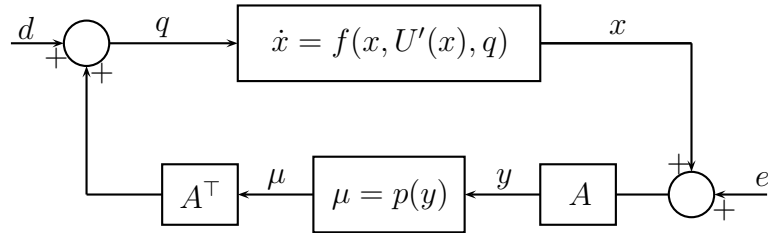

Figure 3: The closed-loop for the primal algorithm with external disturbances.

In the presence of disturbances, asymptotic stability may no longer persist. Hence we introduce and consider the following notions of stability. The definitions are given in a general setting for the primal algorithms, and stability for the dual algorithms can be defined accordingly.

Definition 3.1. *Consider the system described by Figure 3. Suppose that the disturbance e is from a normed space X with norm $\| \cdot \|_X$ and d is from a normed vector space Y with norm $\| \cdot \|_Y$. The system is said to be asymptotically stable on a subset $\mathcal{D} \subset X \times Y$ against the disturbances if there exists $k \geq 0$ such that the solution of (3.1)-(3.2) satisfies:*

$$\limsup_{t \to \infty} |x(t) - x^*| \leq k(\|e\|_X + \|d\|_Y) \ \text{ for all } (e, d) \in \mathcal{D}.$$

The system is said to be asymptotically stable with respect to $w_0 = (u_1, u_2)$ on $\mathcal{D} \subset X \times Y$ against the disturbances if there exist $k \geq 0$ such that the solution of (3.1)-(3.2) satisfies:

$$\limsup_{t \to \infty} |x(t) - \bar{x}(t)| \leq k(\|e - u_1\|_X + \|d - u_2\|_Y) \ \text{ for all } (e, d) \in \mathcal{D}$$

where $u_1 \in X, u_2 \in Y$ and $\bar{x}$ is the solution to the closed-loop when $d = u_1, e = u_2$.
If $\mathcal{D} = X \times Y$, we say the algorithm is asymptotically stable (with respect to $w_0 = (u_1, u_2)$).

It is noticed that without disturbances ($d = e = 0$), asymptotic stability implies that the trajectory converges to the equilibrium asymptotically, and with disturbances, it means the trajectory will tend to a neighbourhood of the equilibrium whereby the the size of neighbourhood depends on the size of the disturbances. If the system is asymptotically stable with respect to $w_0 = (u_1, u_2)$, then the solution converges to a neighbourhood of disturbance free trajectory. This will be the case when we consider the system with round trip delay in the next section.

Some systems which we consider are not asymptotically stable. In these cases we will consider a further relaxed notion of stability.

Definition 3.2. *Let $a \geq 0$. A system described by Figure 3 is said to be a-gain stable on $\mathcal{D} \subset X \times Y$ against the disturbances if there exist $b \geq 0$ such that the solution x to the closed-loop satisfies:*

$$\|x - x^*\| \leq a + b(\|e\|_X + \|d\|_Y) \quad \text{for all} \ (e, d) \in \mathcal{D}.$$

Note that if the system starts from a non-zero initial value, we cannot have classical gain stability (i.e. with $a = 0$). The above notion shows that small disturbances will keep the solution under a certain limit (e.g. link capacity) as the constant a usually depends on the initial value and the equilibrium. Since the trajectory x is continuous as a solution of a differential equation, asymptotic stability implies gain stability.

Throughout this section, we let $X = Y = L^\infty(\mathbb{R}_+)$. From the framework developed we are now in a position to state our first theorem concerning the asymptotic stability of a wide class of primal algorithms given by (3.1)-(3.2). The result shows that this algorithm is asymptotically stable against external disturbances under reasonable assumptions on the algorithm function f, price function p and utility function U. The bounds in the stability inequality (3.5) depend on various rate constants k_1, k_2, λ_1 and α, which have specific interpretations when applied to e.g. Examples 2.1-2.5 (see discussions in the Section 2 and that after Corollary 3.4), and on the network structure A which will be discussed in detail after the theorem and in Section 5.

Theorem 3.3. *Consider the equations (3.1)-(3.2) where x, f, U, q, μ, p as defined in (2.8)-(2.9), A is the routing matrix and $e \in X, d \in Y$. Suppose that*

1) inequality (2.18) holds with $m_r > x_r^* \geq 0$ for all $r \in R$, where x_r^* is the equilibrium given in (2.3);

2) inequality (2.13) holds;

3) for each $r \in R$, there exists a function $\tilde{f}_r : \mathbb{R}^3 \to \mathbb{R}$ such that $\tilde{f}_r(x_r, q_r^*, q_r^*) = 0$ and

$$(x_r - x_r^*)f_r(x_r, U_r'(x_r), q_r) \leq (x_r - x_r^*)\tilde{f}_r(x_r, U_r'(x_r), q_r) \quad \text{for all } x_r \in [0, m_r], q_r \in \mathbb{R}; \quad (3.3)$$

4) for each $r \in R$ and each $x_r \in [0, m_r]$, $\tilde{f}_r(x_r, \cdot, \cdot)$ is differentiable and there exist $k_1 > 0, k_2 \geq 0$ such that

$$\frac{\partial \tilde{f}_r}{\partial z_r}(x_r, z_r, q_r) \geq k_1, \quad -k_2 \leq \frac{\partial \tilde{f}_r}{\partial q_r}(x_r, z_r, q_r) \leq 0 \ \ \text{for all } x_r \in [0, m_r], z_r, q_r \in \mathbb{R}; \quad (3.4)$$

5) there exist $\alpha, \lambda_1 > 0$ such that

$$0 \leq p_j'(z_1) \leq \alpha \ \text{ and } \ U_r''(z_2) \leq -\lambda_1 \ \ \text{for all } z_1 \in \mathbb{R}_+, z_2 \in [0, m_r], j \in J, r \in R.$$

Then (3.1)-(3.2) is asymptotically stable on $\mathcal{D} = L^\infty(\mathbb{R}_+) \times L^\infty(\mathbb{R}_+)$ against external disturbances and

$$|x(t) - x^*| \leq |x(0) - x^*|e^{-2\lambda_1 k_1 t} + \frac{k_2(\|d\| + \alpha\|A^\top A\|\|e\|)}{\lambda_1 k_1}(1 - e^{-2\lambda_1 k_1 t}) \quad (3.5)$$

for all $x_r(0) \in [0, m_r]$.

Proof. Let $\tilde{x} = x - x^*$ and $V = \tilde{x}^\top \tilde{x}$. By our differentiability assumption on each $\tilde{f}_r$ and the Mean Value Theorem, we see that for all $r \in R$ there exist z_r, η_r such that

$$\tilde{f}_r(x_r, U_r'(x_r), q_r) - \tilde{f}_r(x_r, U_r'(x_r^*), q_r) = \left.\frac{\partial \tilde{f}_r(x_r, y_r, q_r)}{\partial y_r}\right|_{y_r = z_r}(U_r'(x_r) - U_r'(x_r^*)),$$

$$\tilde{f}_r(x_r, q_r^*, q_r) - \tilde{f}_r(x_r, q_r^*, q_r^*) = \left.\frac{\partial \tilde{f}_r(x_r, q_r^*, q_r)}{\partial q_r}\right|_{q_r = \eta}(q_r - q_r^*).$$

10

Let $\omega_r = \left.\frac{\partial \tilde{f}_r(x_r, y_r, q_r)}{\partial y_r}\right|_{y_r = z_r}$ and $\hat{\omega}_r = \left.\frac{\partial \tilde{f}_r(x_r, q_r^*, q_r)}{\partial q_r}\right|_{q_r = \eta}$, and let $K_1 = \mathrm{diag}(\omega_r)$ and $K_2 = \mathrm{diag}(\hat{\omega}_r)$ be corresponding diagonal matrices. Define $\tilde{f}(x, z, q) = (\tilde{f}_r(x_r, z_r, q_r))_{r \in R}$. Since x^* is constant, we have

$$
\begin{aligned}
\dot{V} &= 2\tilde{x}^\top \dot{x} = 2\tilde{x}^\top f(x, U'(x), q) \\
&\leq 2\tilde{x}^\top \tilde{f}(x, U'(x), q) \\
&= 2\tilde{x}^\top (\tilde{f}(x, U'(x), q) - \tilde{f}(x, U'(x^*), q)) + 2\tilde{x}^\top (\tilde{f}(x, q^*, q) - \tilde{f}(x, q^*, q^*)) \\
&= 2\tilde{x}^\top K_1(U'(x) - U'(x^*)) + 2\tilde{x}^\top K_2(q - q^*) \\
&= 2\tilde{x}^\top K_1(U'(x) - U'(x^*)) + 2\tilde{x}^\top K_2\big(d + A^\top p(Ax + Ae) - A^\top p(Ax^*)\big).
\end{aligned}
$$

By (2.13) and (2.18) it follows that $x_r \in [0, m_r]$. Since $x_r^* \in [0, m_r]$, by the Mean Value Theorem, it follows that there exist $\xi_r \in [0, m_r], \zeta_j \in \mathbb{R}$ such that

$$
\begin{aligned}
U'(x) - U'(x^*) &= \mathrm{diag}(U_r''(\xi_r))(x - x^*) = \mathrm{diag}(U_r''(\xi_r))\tilde{x}, \\
p(Ax + Ae) - p(Ax^*) &= \mathrm{diag}(p_j'(\zeta_j))(Ax + Ae - Ax^*) = \mathrm{diag}(p_j'(\zeta_j))A\tilde{x} + \mathrm{diag}(p_j'(\zeta_j))Ae.
\end{aligned}
$$

Hence

$$
\dot{V} \leq 2\tilde{x}^\top K_1 \mathrm{diag}(U_r''(\xi_r))\tilde{x} + 2\tilde{x}^\top K_2 A^\top \mathrm{diag}(p_j'(\zeta_j))A\tilde{x} + 2\tilde{x}^\top K_2\big(d + A^\top \mathrm{diag}(p_j'(\zeta_j))Ae\big).
$$

Since $0 \leq p_j'(\zeta_j) \leq \alpha$, $U_r''(\xi_r) \leq -\lambda_1$, $\omega_r \geq k_1 > 0$ and $0 \geq \hat{\omega} \geq -k_2$, we see that

$$
\tilde{x}^\top K_1 \, \mathrm{diag}(U_r''(\xi_r))\tilde{x} \leq -\lambda_1 k_1 \|\tilde{x}\|^2, \quad \tilde{x}^\top K_2 A^\top \mathrm{diag}(p_j'(\zeta_j))A\tilde{x} \leq 0
$$

and

$$
\|\tilde{x}^\top K_2\big(d + A^\top \mathrm{diag}(p_j'(\zeta_j))Ae\big)\| \leq k_2\big(\|d\| + \alpha\|A^\top A\|\|e\|\big)|\tilde{x}(t)|.
$$

This shows

$$
\frac{1}{2}\frac{d}{dt}|\tilde{x}(t)|^2 = \frac{1}{2}\dot{V} \leq -\lambda_1 k_1 |\tilde{x}(t)|^2 + k_2\big(\|d\| + \alpha\|A^\top A\|\|e\|\big)|\tilde{x}(t)| \tag{3.6}
$$

and

$$
|\tilde{x}(t)|^2 \leq -2\lambda_1 k_1 \int_0^t |\tilde{x}(s)|^2 ds + 2k_2\big(\|d\| + \alpha\|A^\top A\|\|e\|\big)\int_0^t |\tilde{x}(s)| ds.
$$

By Lemma 6.1, we have

$$
|\tilde{x}(t)| \leq |\tilde{x}(0)|e^{-2\lambda_1 k_1 t} + \frac{k_2(\|d\| + \alpha\|A^\top A\|\|e\|)}{\lambda_1 k_1}(1 - e^{-2\lambda_1 k_1 t}),
$$

as required.

$\square$

The conclusion of the theorem tells us that in the L^∞ setting, that although the rate allocation may not converge to the existing equilibrium, nevertheless the deviation from the equilibrium is bounded by the size of disturbances. In the case when there is no disturbance, any solution of the algorithm, starting from an arbitrary initial value, will converge to the equilibrium asymptotically. Furthermore, by (3.6), we have

$$
|\tilde{x}(t)| \leq \max\left\{|\tilde{x}(0)|, \frac{k_2(\|d\| + \alpha\|A^\top A\|\|e\|)}{\lambda_1 k_1}\right\}.
$$

In the case when the disturbances are so small that $\frac{k_2(\|d\|+\alpha\|A^\top A\|\|e\|)}{\lambda_1 k_1} \le |\tilde{x}(0)|$, this shows $|\tilde{x}(t)| \le |\tilde{x}(0)|$ and it means that the transmission rate $x(t)$ will always remain within the initial distance of the equilibrium.

It also follows that the constants in inequality (3.5) depend not only on the constants k_1, k_2, λ_1 and α)which are related to users' algorithm and independent of network topology determined by A), but also on the $\|A^\top A\|$ which is determined by the network. Note that the system matrix A is independent of the constants k_1, k_2, λ_1 and α: it only depends on the number of users on each link and number of links used by each user. The more users using the same link, the larger the term $\|A^\top A\|$: this term affects the size of the transient and the amplification of the disturbance e. However, since $\|A^\top A\|$ is bounded by $\|A^\top\|\|A\|$ which is equal to the length of the longest route multiplied by the number of users on the most used link, if the network is designed such that each link has at most a fixed number of users and each user uses at most a fixed number of links, then $\|A^\top A\|$ remains bounded even when new users are added. One fact is that once the number of users increases, some links must be used by more users, which seems a contradiction to the above requirement. This problem can be solved by the following network design: for any link used by more users, its capacity must be very large, we can decompose this link into a group of sub-links, the more routes through the link, the more sub-links it decomposes, and each sub-links allow at most the given number of users. For example, in the network consist of two blocks of nodes (represented by circles) and links as shown in Figure 4, each node is linked to a TCP station (represented by a square frame) which choose the shortest route between any two nodes and control the numbers of users going to each link. Suppose that every link has super capacity and is treated as a group of sub-links
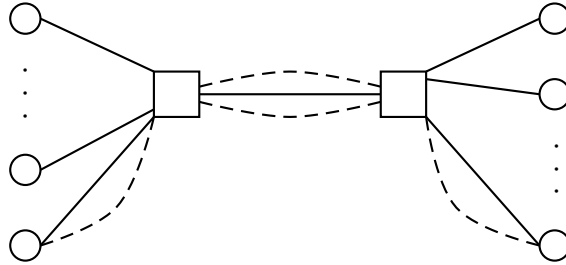

Figure 4: A network with many users.

(the dashed links), and each block has many nodes. Then at each joint of two links, there are enough sub-links available, and we are able to let at most a fixed number, say 3, of users use one sub-links. Since there are at most three links between any two nodes, we see that $\|A^\top A\| \le 9$, no matter how many users in the network. This will be further discussed in Section 5.

Applying the theorem to the concrete algorithm given in Example 2.5, we have

Corollary 3.4. *Consider the system given by equations* (2.19) *and* (3.2) *where* x, f, U, q, μ, p *as defined in* (2.8)-(2.9), A *is the routing matrix and* $e \in X, d \in Y$. *Suppose* $0 < k_1 \le \kappa_r(x_r) \le k_2$ *for all* $r \in R, x_r \in \mathbb{R}_+$. *Suppose there exist* $\alpha, \lambda_1 > 0$ *such that* $0 \le p'_j(z_1) \le \alpha$ *for all* $z_1 \in \mathbb{R}_+$ *and* $U''_r(z_2) \le -\lambda_1$ *for all* $z_2 \in [0, m_r]$, *then the system given by* (2.19) *and* (3.2) *is asymptotically stable on* $\mathcal{D} = L^\infty(\mathbb{R}_+) \times L^\infty(\mathbb{R}_+)$ *against the external disturbances and inequality* (3.5) *holds for all* $x_r(0) \in [0, m_r]$.

Proof. Let $f_r(x_r, z_r, q_r) = \kappa_r(x_r)(z_r - q_r)_{x_r}^{m_r}$, $\tilde{f}_r(x_r, z_r, q_r) = \kappa_r(x_r)(z_r - q_r)$. Then $\tilde{f}_r(x_r, q_r^*, q_r^*) = 0$, $\tilde{f}_r$ is differentiable with respect to both z_r and q_r and the derivatives satisfy inequality (3.4).

To establish inequality (3.3), we note that if $0 < x_r < m_r$, then $(x_r - x_r^*)f_r(x_r, z_r, q_r) = (x_r - x_r^*)\tilde{f}_r(x_r, z_r, q_r)$. If $x_r = 0$, since $\max\{0, z_r - q_r\} \geq z_r - q_r$ and $x_r^* \geq 0, \kappa_r(x_r) \geq 0$, we have

$$(x_r - x_r^*)f_r(x_r, z_r, q_r) = -x^*\kappa_r(0)\max\{0, z_r - q_r\} \leq -x_r^*\kappa_r(0)(z_r - q_r) = (x_r - x_r^*)\tilde{f}_r(x_r, z_r, q_r).$$

If $x_r = m_r$, since $x_r^* \in [0, m_r]$, we have

$$\begin{aligned}
(x_r - x_r^*)f_r(x_r, z_r, q_r) &= (m_r - x_r^*)\kappa_r(m_r)\min\{0, z_r - q_r\} \\
&\leq (m_r - x_r^*)\kappa_r(m_r)(z_r - q_r) = (x_r - x_r^*)\tilde{f}_r(x_r, z_r, q_r).
\end{aligned}$$

Since inequalities (2.13) and (2.18) hold for the function $f_r(x_r, z_r, q_r) = (\kappa_r(x_r)(z_r - q_r)_{x_r}^{m_r}$, all assumptions of Theorem 3.3 are satisfied and the conclusion follows. $\qquad\square$

This corollary shows that the conditions for Theorem 3.3 are easy to satisfy. For the algorithm described in Example 2.5, by Corollary 3.4 we have $0 \leq x_r(t) \leq m_r$. Then, e.g. for $U_r(x_r) = w_r \log(x_r + \varepsilon)$ with $\varepsilon > 0$ and domain(U_r) $= [0, m_r]$, we have

$$U_r''(x_r) = -\frac{w_r}{(x_r + \varepsilon)^2} \in \left[-\frac{w_r}{\varepsilon^2}, -\frac{w_r}{(m_r + \varepsilon)^2}\right],$$

and $\lambda_1 = -w_r/(m_r + \varepsilon)^2$. Although inequality (2.18) is not satisfied for some existing algorithms such as those given in Example 2.1 and Example 2.2, a similar conclusion holds provided $U_r''(x_r) \leq -\lambda_1$ for all $x_r \in [0, \infty)$ as in [4, 19]. However, no such $\lambda_1 > 0$ exists for the basic utility function $U_r(x_r) = \log x_r, \log(x_r + \varepsilon)$ or x_r^α since $U_r''(x_r) \to 0$ as $x_r \to \infty$. But it works when U_r is a quadratic function, for example, when $U_r(x_r) = -\frac{\lambda_1}{2}x_r^2 + ax + b$ with $a, b \geq 0$.

In Theorem 3.3, inequality (2.18) provides a sufficient condition that the rate x_r remains bounded by the constant m_r. In the following, we remove this condition and guarantee the boundedness of x_r by restricting on the sizes of disturbances and initial conditions.

Theorem 3.5. *Consider the system given by equations* (3.1)-(3.2) *where* x, f, U, q, μ, p *as defined in* (2.8)-(2.9), *A is the routing matrix and $e \in X, d \in Y$. Suppose that Assumptions 2),3), 4) and 5) of Theorem 3.3 are all satisfied with $m_r = 2L$, where $L \geq \|x^*\|$ is a given number. Then system* (3.1)-(3.2) *is asymptotically stable on*

$$\mathcal{D} = \left\{(e, d) \in L^\infty(\mathbb{R}_+) \times L^\infty(\mathbb{R}_+) : k_2(\|d\| + \alpha\|A^\top A\|\|e\|) \leq k_1 L\lambda_1\right\} \tag{3.7}$$

against the external disturbances and

$$|\tilde{x}(t)| \leq |\tilde{x}(0)|e^{-2\lambda_1 k_1 t} + \frac{k_2(\|d\| + \alpha\|A^\top A\|\|e\|)}{\lambda_L k_1}(1 - e^{-2\lambda_1 k_1 t}) \tag{3.8}$$

provided $\|x(0) - x^*\| \leq L$.

Proof. Let $(e, d) \in \mathcal{D}$. Let $V, \tilde{x}, K_1$ and K_2 be the same as in Theorem 3.3. Using the same method, we obtain

$$\begin{aligned}
\dot{V} =&\, 2\tilde{x}^\top K_1 \mathrm{diag}(U_r''(\xi_r))\tilde{x} + 2\tilde{x}^\top K_2 A^\top \mathrm{diag}(p_j'(\zeta_j))A\tilde{x} + 2\tilde{x}^\top K_2\big(d + A^\top \mathrm{diag}(p_j'(\zeta_j))Ae\big) \\
\leq&\, 2\tilde{x}^\top K_1 \mathrm{diag}(U_r''(\xi_r))\tilde{x} + 2\tilde{x}^\top K_2\big(d + A^\top \mathrm{diag}(p_j'(\zeta_j))Ae\big)
\end{aligned}$$

13

where $\xi_r \in [x_r^*, x_r(t)], \zeta_j \in \mathbb{R}$.

Let
$$\Omega = \{\tilde{x} = (\tilde{x}_r)_{r \in R} : \|\tilde{x}\| \leq L\}.$$

If $\tilde{x}(t) \in \partial\Omega$, i.e. $\|\tilde{x}(t)\| = L$, then

$$\|x(t)\| = \|\tilde{x}(t) + x^*\| \leq L + \|x^*\| \leq 2L$$

since $\|x^*\| \leq L$. By our assumption, we see

$$
\begin{aligned}
\dot{V} &\leq 2k_1 \max\{U_r''(\xi_r) : \xi_r \in [0, 2L], r \in R\}\|\tilde{x}\|^2 + 2k_2\big(\|d\| + \alpha\|A^\top A\|\|e\|\big)\|\tilde{x}\| \qquad (3.9) \\
&\leq 2k_1 L^2 \max\{U_r''(\xi_r) : \xi_r \in [0, 2L], r \in R\} + 2k_2 L\big(\|d\| + \alpha\|A^\top A\|\|e\|\big) \leq 0.
\end{aligned}
$$

Hence $V(\tilde{x}(t)) = \|\tilde{x}\|^2$ is decreasing at times $t \geq 0$ when $\tilde{x}(t) \in \partial\Omega$. This shows $\tilde{x}(t) \in \Omega$ and (3.9) remains valid provided $\tilde{x}(0) \in \Omega$. So by Lemma 6.1, we obtain (3.8). $\qquad\square$

Similar to Corollary 3.4, for the algorithm given by Example 2.2, we have:

Corollary 3.6. *Consider the system given by equations* (2.15) *and* (3.2) *where* x, f, U, q, μ, p *as defined in* (2.8)-(2.9), *A is the routing matrix and* $e \in X, d \in Y$. *Suppose* $0 < k_1 \leq \kappa_r(x_r) \leq k_2$ *for all* $r \in R, x_r \in \mathbb{R}_+$. *If there exist* $\alpha, \lambda_1 > 0$ *such that* $0 \leq p_j'(z_1) \leq \alpha$ *for all* $z_1 \in \mathbb{R}_+$ *and* $U_r''(z_2) \leq -\lambda_1$ *for all* $z_2 \in [0, 2L]$ *with* $L \geq \|x^*\|$, *then the system is locally asymptotically stable on* $\mathcal{D}$ *against the external disturbances and inequality* (3.8) *hold as in Theorem 3.5, provided* $|x(0) - x^*| \leq L$.

3.2   Dual algorithms

Now we consider stability in the presence of external disturbances for the dual algorithms (2.20)-(2.21). The diagram in Figure 5 shows the relationships between variables of the algorithm,
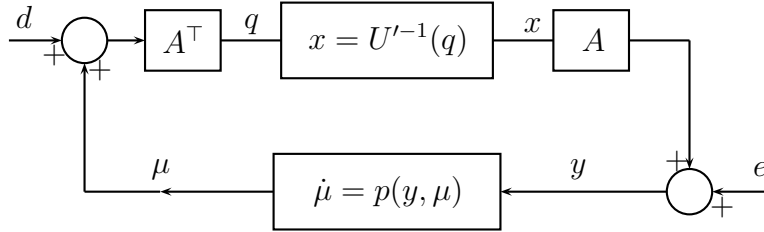

Figure 5: The closed-loop of dual algorithm with external disturbances.

and corresponds to the vector equations:

$$\dot{\mu} = p(y, \mu), \qquad (3.10)$$
$$y = e + Ax, \ x = U'^{-1}(q), \ q = A^\top(d + \mu). \qquad (3.11)$$

We suppose the penalty function p satisfies (2.23), hence the shadow price information μ_j lies in $[0, \chi_j]$. Since the utility function U_r is defined from $\mathbb{R}_+$ to itself, the domain of $U_r'^{-1}$ is $\mathbb{R}_+$. Thus to ensure the closed loop is well defined, we need to restrict the disturbance d_j such that $q = A^\top(\mu + d) \geq 0$.

Notice that the routing matrices A and $A^\top$ are lumped with the source controllers rather than the link controllers as in the primal algorithm case. This is not necessary, the routing matrices can be considered as part of the link controllers, giving an alternative mapping of the disturbances onto the network.

We now establish sufficient conditions for asymptotic stability of dual algorithm (3.10)-(3.11) against the external disturbances.

Theorem 3.7. *Consider the system given by equations (3.10)-(3.11) where p, U'^{-1} the vector functions as described in Section 2.3 and A is the routing matrix. Suppose that for all $j \in J$:*

1) there exist $\chi_j \geq 0$ such that $\mu_j^* \in [0, \chi_j], j \in J$;

2) each p_j satisfies (2.23) with the given $\chi_j \geq 0$;

3) there exist functions $\hat{p}_j : \mathbb{R}^2 \to \mathbb{R}$ such that $\hat{p}_j(y_j^*, \mu_j^*) = 0$ and

$$(\mu_j - \mu_j^*)p_j(y_j, \mu_j) \leq (\mu_j - \mu_j^*)\hat{p}_j(y_j, \mu_j) \quad \text{for all } y_j \in \mathbb{R}, \mu_j \in [0, \chi_j]; \tag{3.12}$$

4) for each j, $\hat{p}_j$ is differentiable and there exist $k_1 \geq 0, k_2 > 0$ such that

$$0 \leq \frac{\partial \hat{p}_j(y_j, \mu_j)}{\partial y_j} \leq k_1, \quad \frac{\partial \hat{p}_j(y_j, \mu_j)}{\partial \mu_j} \leq -k_2 \;\; \text{for all } y_j \in \mathbb{R}, \mu_j \in [0, \chi_j]; \tag{3.13}$$

5) there exists $\lambda_1 > 0$ such that

$$-\lambda_1 \leq (U_r'^{-1})'(z) \leq 0, \quad \text{for all } z \in \left[0, 2\sum_{j \in r} \chi_j\right], r \in R.$$

Then (3.10)-(3.11) is asymptotically stable against disturbances on

$$\mathcal{D} = \{(e, d) \in L^\infty(\mathbb{R}_+) \times L^\infty(\mathbb{R}_+) : 0 \leq d_j(t) \leq \chi_j \;\; \text{for all } \; j \in J, t \in \mathbb{R}_+\}.$$

Furthermore, we have

$$\|\mu(t) - \mu^*\| \leq \|\mu(0) - \mu^*\|e^{-2k_2 t} + k_2^{-1}k_1\big(\|e\| + \lambda_1\|AA^\top\|\|d\|\big)\big(1 - e^{-2k_2 t}\big) \tag{3.14}$$

for all $\mu(0)$ with $\mu_j(0) \in [0, \chi_j], j \in J$.

Proof. We omit the variable t unless it is necessary.

Let $(e, d) \in \mathcal{D}$ and μ be the corresponding solution to (3.10)-(3.11). Let $\tilde{\mu} = \mu - \mu^*, V = \tilde{\mu}^\top \tilde{\mu}$. Then by the Mean Value Theorem, for all $j \in J$ there exist $\xi_j \in [\mu_j^*, \chi_j] \subset [0, \chi_j]$ and $\zeta_j \in [y_j^*, y_j]$ such that

$$\hat{p}_j(y_j, \mu_j^*) - \hat{p}_j(y_j^*, \mu_j^*) = \frac{\partial \hat{p}_j(y_j, \mu_j^*)}{\partial y_j}\bigg|_{y_j = \zeta_j} (y_j - y_j^*),$$

$$\hat{p}_j(y_j, \mu_j) - \hat{p}_j(y_j, \mu_j^*) = \frac{\partial \hat{p}_j(y_j, \mu_j)}{\partial \mu_j}\bigg|_{\mu_j = \xi_j} (\mu_j - \mu_j^*).$$

Let $K_1 = \text{diag}\big(\frac{\partial \hat{p}_j(y_j, \mu_j^*)}{\partial y_j}\big|_{y_j = \zeta_j}\big), K_2 = \text{diag}\big(\frac{\partial \hat{p}_j(y_j, \mu_j)}{\partial \mu_j}\big|_{\mu_j = \xi_j}\big)$ be the diagonal matrices. Since $d_j \geq 0$ for all $j \in J$, we see $x_j \geq 0$ and therefore, by (2.23), $0 \leq \mu_j \leq \chi_j$ for all $j \in J$. Then

$$\begin{aligned}
\dot{V} &= 2\tilde{\mu}^\top \mu' = 2\tilde{\mu}^\top p(y, \mu) \\
&\leq 2\tilde{\mu}^\top \hat{p}(y, \mu) \\
&= 2\tilde{\mu}^\top \big(\hat{p}(y, \mu) - \hat{p}(y, \mu^*) + \hat{p}(y, \mu^*) - \hat{p}(y^*, \mu^*)\big) \\
&= 2\tilde{\mu}^\top K_2(\mu - \mu^*) + 2\tilde{\mu}^\top K_1(y - y*) \\
&= 2\tilde{\mu}^\top K_2\tilde{\mu} + 2\tilde{\mu}^\top K_1\big(e + AU'^{-1}(A^\top(\mu + d)) - AU'^{-1}(A^\top \mu^*)\big) \\
&= 2\tilde{\mu}^\top K_2\tilde{\mu} + 2\tilde{\mu}^\top K_1\big(e + A(U'^{-1})'(\eta)A^\top(\tilde{\mu} + d)\big), \tag{3.15}
\end{aligned}$$

where by the Mean Value Theorem, $\eta = (\eta_r)_{r \in R}$ lies between $A^\top(\mu+d)$ and $A^\top\mu^*$, and $U'^{-1}(z) = \text{diag}(U_r'^{-1}(z_r))$ and $(U'^{-1})'(\eta) = \text{diag}((U_r'^{-1})'(\eta_r))$. Since d_j, μ_j and μ_j^* are all between 0 and χ_j for each $j \in J$, we see

$$0 \leq \eta_r \leq \sum_{j \in r} |\mu_j + d_j| \leq 2 \sum_{j \in r} \chi_j,$$

$$\tilde{\mu}^\top K_2 \tilde{\mu} \leq -k_2\|\tilde{\mu}(t)\|^2, \quad \tilde{\mu} K_1 A (U'^{-1})'(\eta) A^\top \tilde{\mu} \leq 0.$$

Hence by assumption 5),

$$\frac{1}{2}\frac{d}{dt}\|\tilde{\mu}(t)\|^2 = \tilde{\mu}^\top \dot{\tilde{\mu}} \leq -k_2\|\tilde{\mu}(t)\|^2 + k_1\big(\|e\| + \lambda_1\|AA^\top\|\|d\|\big)\|\tilde{\mu}(t)\|.$$

The conclusion follows from Lemma 6.1. $\qquad\square$

Note that, as in the primal case (Theorem 3.3), the constants k_1, k_2 and λ_1 in stability inequality (3.14) are independent of the network structure A, and $\|AA^\top\|$ scales well with network size if super-links are considered.

If $\chi_j = \infty$, then Assumption 5) can be replaced by the simpler global condition:

$$-\lambda_1 \leq (U_r'^{-1})'(z) \leq 0 \quad \text{for all } z \geq 0.$$

In this case, the disturbance d_j will be less restricted. However the most commonly used utility functions do not satisfy this condition. For example when $U_r(x) = \log(x)$, we have $(U_r'^{-1})'(z) = -1/z^2 \to -\infty$ as $z \to \infty$.

As an application, consider the dual algorithm given in Example 2.8, i.e.

$$\dot{\mu}_j(t) = p_j(y_j(t), \mu_j(t)) =: \kappa_j(y_j(t) - h_j(\mu_j(t)))_{\mu_j(t)}^{\chi_j}. \tag{3.16}$$

For this algorithm, Assumptions 1) and 2) of Theorem 3.7 hold by construction. Let $\hat{p}_j(y_j, \mu_j) = \kappa_j(y_j - h_j(\mu_j))$. Then

$$\frac{\partial \hat{p}_j}{\partial y_i} = \kappa_j, \frac{\partial \hat{p}_j}{\partial \mu_i} = -\kappa_j h_j'(\mu_j).$$

If there exists a $k_2 > 0$ such that

$$h_j \quad \text{is differentiable and } h_j'(\mu_j) \geq k_2/\kappa_j \quad \text{for all } j \in J, \mu_j \in [0, \chi_j], \tag{3.17}$$

then Assumptions 3) and 4) of Theorem 3.7 are also satisfied. Therefore, we have

Corollary 3.8. *Consider the system given by equations* (2.24) *in Example 2.8 and* (3.11) *where p, U'^{-1} the vector functions as described in Section 2.3 and A is the routing matrix. Let $\chi_j \geq 0, j \in J$ and suppose there exists a unique $\mu_j^* \in [0, \chi_j]$. Suppose* (3.17) *hold and there exists $\lambda_1 > 0$ such that $-\lambda_1 \leq (U_r'^{-1})'(z) \leq 0$ for all $0 \leq z \leq 2\sum_{j \in r} \chi_j$. Then the dual algorithm* (3.16) *with source controller* (3.11) *is asymptotically stable against disturbances on $\mathcal{D}$ as in Theorem 3.7. Furthermore, inequality* (3.14) *holds with $k_1 = \max\{\kappa_j : j \in J\}$.*

Similarly, we can study the stability of the dual algorithm given by Example 2.6 provided the function $h_j(\mu_j)$ in (2.22) is positive, uniformly bounded and $h_j'(\xi) \geq k_2 > 0$. Theorem 3.7 is no longer applicable when $h_j(x)$ is constant, such as in [4] where the considered dual algorithm is

$$\dot{\mu}_j = \gamma_j(y_j - c_j)_{\mu_j}^+. \tag{3.18}$$

So next we present a asymptotic stability result for dual algorithm (3.10)-(3.11), allowing $k_2 = 0$, provided the routing matrix A has full row rank.

Theorem 3.9. *Consider the system given by equations (3.10)-(3.11) where p, U'^{-1} the vector functions as described in Section 2 and A is the routing matrix. Let $\chi_j \geq 0, j \in J$. Suppose that for all $j \in J$:*

1) A has full row rank and $\mu_j^* \in [0, \chi_j]$;

2) each p_j satisfies (2.23) with the given $\chi_j \geq 0$;

3) there exist functions $\hat{p}_j : \mathbb{R}^2 \to \mathbb{R}$ such that $\hat{p}_j(y_j^*, \mu_j^*) = 0$ and

$$(\mu_j - \mu_j^*)p_j(y_j, \mu_j) \leq (\mu_j - \mu_j^*)\hat{p}_j(y_j, \mu_j) \quad \text{for all } y_j \in \mathbb{R}_+, \mu_j \in [0, \chi_j];$$

4) for each j, $\hat{p}_j$ is differentiable and there exist $k_1 \geq k_0 > 0, k_2 \geq 0$ such that

$$k_0 \leq \frac{\partial \hat{p}_j(y_j, \mu_j)}{\partial y_j} \leq k_1, \quad \frac{\partial \hat{p}_j(y_j, \mu_j)}{\partial \mu_j} \leq -k_2 \text{ for all } y_j \in \mathbb{R}, \mu_j \in [0, \chi_j]; . \tag{3.19}$$

5) there exist $\lambda_1, \lambda_2 > 0$ such that

$$-\lambda_2 \leq (U_r'^{-1})'(z) \leq -\lambda_1 \text{ for all } z \in \left[0, 2\sum_{j \in r} \chi_j\right], r \in R.$$

Then the algorithm (3.10)-(3.11) is asymptotically stable on the same $\mathcal{D}$ as in Theorem 3.7. Furthermore,

$$\|\mu(t) - \mu^*\| \leq \|\mu(0) - \mu^*\|e^{-2k_0 t} + \frac{k_1\left(\|e\| + \lambda_1^{-1}\|AA^\top\|\|d\|\right)\|(AA^\top)^{-1}\|}{\lambda_2 k_0}(1 - e^{-2k_0 t}) \tag{3.20}$$

for all $\mu(0)$ with $\mu_j(0) \in [0, \chi_j], j \in J$.

Proof. Let $(e, d) \in \mathcal{D}$ and μ be the corresponding solution to (3.10)-(3.11). Let $\tilde{\mu} = \mu - \mu^*, V = \tilde{\mu}^\top \tilde{\mu}$. Following the proof of Theorem 3.7 we obtain inequality (3.15), which implies

$$\frac{d\tilde{\mu}^\top \tilde{\mu}}{dt} \leq 2\tilde{\mu}^\top K_1 A(U'^{-1})'(\eta)A^\top \tilde{\mu} + 2\tilde{\mu}^\top K_1(e + A(U'^{-1})'(\eta)A^\top d) \tag{3.21}$$

with $\eta = (\eta_r)_{r \in R} \in [A^\top \mu^*, A^\top(\mu + d)]$. As assumed, $\mu_j \in [0, \chi_j], \|d_j\| \leq \chi_j$ and $\sum_{j \in r}(\mu_j + d_j) \geq 0$. So $\eta_r \in [0, 2\sum_{j \in r} \chi_j]$ and $(U_r'^{-1})'(\eta_r) \in [-\lambda_1, -\lambda_2]$ for all $r \in R$. Since $\tilde{\mu}^\top AA^\top \tilde{\mu} \geq 0$, we see

$$\tilde{\mu}^\top K_1 A(U'^{-1})'(\eta)A^\top \tilde{\mu} \leq -\lambda_2 k_0 \tilde{\mu}^\top AA^\top \tilde{\mu}.$$

Since A has full row rank and A is a $J \times R$ matrix, it follows that $AA^\top$ is invertible. So

$$\tilde{\mu}^\top K_1 A(U'^{-1})'(\eta)A^\top \tilde{\mu} \leq -\lambda_2 k_0 \|(AA^\top)^{-1}\|^{-1}\|\tilde{\mu}\|^2.$$

Since $\tilde{\mu}^\top K_2 \tilde{\mu} \leq 0$, by (3.21), we see

$$\frac{1}{2}\frac{d}{dt}\|\tilde{\mu}(t)\|^2 \leq -\lambda_2 k_0 \|(AA^\top)^{-1}\|^{-1}\|\tilde{\mu}\|^2 + k_1\left(\|e\| + \lambda_1\|AA^\top\|\|d\|\right)\|\tilde{\mu}(t)\|$$

and the conclusion follows. $\qquad \square$

By the definition of A, we see

$$\tilde{\mu}^\top A A^\top \tilde{\mu} = \sum_{r \in R} \left(\sum_{j \in r} \tilde{\mu}_j \right)^2.$$

The full row rank assumption indicates that each link must be used by at least one user, that is, each $\tilde{\mu}_j, j \in J$, must appear in at least one of the terms in $\sum_{j \in r} \tilde{\mu}_j$. Hence

$$\tilde{\mu}^\top A A^\top \tilde{\mu} \geq \sum_{j \in J} \tilde{\mu}_j^2.$$

In the case when the norm of $\hat{\mu}$ is defined as

$$\|\tilde{\mu}\| = \left(\sum_{j \in J} |\tilde{\mu}_j^2| \right)^{1/2} \quad \text{or} \quad \|\tilde{\mu}\| = \max_{j \in J} |\tilde{\mu}_j|,$$

we will have $\tilde{\mu}^\top A A^\top \tilde{\mu} \geq \|\tilde{\mu}\|$ and

$$\tilde{\mu}^\top K_1 A (U'^{-1})'(\eta) A^\top \tilde{\mu} \leq -\lambda_2 k_0 \|\tilde{\mu}\|^2,$$

therefore (3.20) becomes

$$\|\mu(t) - \mu^*\| \leq \|\mu(0) - \mu^*\| e^{-2k_0 t} + \frac{k_1 \left(\|e\| + \lambda_1^{-1} \|AA^\top\| \|d\| \right)}{\lambda_2 k_0} (1 - e^{-2k_0 t}).$$

This does not necessarily mean that this new estimation is better than (3.20) since $\|(AA^\top)^{-1}\|$ could be strictly smaller than 1.

Now let us consider the dual algorithm given by Example 2.7 with $h_j(\mu_j) \equiv \gamma_j \geq 0$, i.e.

$$\dot{\mu}_j = p_j(y_j, \mu_j) =: \gamma_j (y_j - c_j)_{\mu_j}^{\chi_j}, \tag{3.22}$$

Let $\hat{p}_j = \gamma_j (y_j - c_j)$. If $\mu_j \neq 0$ or $\mu_j = 0$ but $y_i - c_j \geq 0$, we have $(y_j - c_j)_{\mu_j}^{\chi_j} = y_j - c_j$. If $\mu_j = 0$ and $y_j - c_j < 0$, then

$$(\mu_j - \mu^*) p_j(y_j, \mu_j) = 0 < -\mu^* \gamma_j (y_j - c_j) = (\mu - \mu^*) \hat{p}_j(y_j, \mu_j) \quad \text{for all } \mu_j \in [0, \chi_j], y_j \in \mathbb{R}$$

and

$$\frac{\partial \hat{p}_j}{\partial y_j} = \gamma_j, \quad \frac{\partial \hat{p}_j}{\partial \mu_j} = 0 \quad \text{for all } \mu_j \in [0, \chi_j], y_j \in \mathbb{R}.$$

This shows that Assumptions 2), 3) and 4) of Theorem 3.9 are satisfied with $k_0 = \min\{\gamma_j : j \in J\}$, $k_1 = \max\{\gamma_j : j \in J\}$ and $k_2 = 0$. Hence, we have

Corollary 3.10. *Consider the system given by equations (3.22) and (3.11) with p, U'^{-1} the vector functions as described in Section 2.3 and A is the routing matrix. Let $\chi_j \geq 0, j \in J$ and suppose there exists a unique $\mu_j^* \in [0, \chi_j]$. If A has full row rank and $-\lambda_2 \leq (U_r'^{-1})'(z) \leq -\lambda_1$ for all $z \in [0, 2\sum_{j \in r} \chi_j]$, then the dual algorithm (3.22) with controller (3.11) is asymptotically stable against disturbances on the same $\mathcal{D}$ as in Theorem 3.9.*

This is a conclusion similar to Theorem 2 of [4], which considered algorithm (3.18) and assumed that $-\lambda_2 \leq U_r''(z) \leq -\lambda_1$ for all $z \in \mathbb{R}_+$, an assumption which is not satisfied by the commonly used utility functions.

4 Robustness with round trip delays

In this section, we consider the network with both external disturbances and round trip delay. Round trip delays are caused by propagation delays from links to users and from users to links. We will treat the round trip delay as an uncertainty and exploit the theory of gap metric to study the stability. That is, we consider the stable closed loop without round trip delay and calculate the gap between this closed loop and the one with round trip delay, both starting from the same starting initial value. If we can show that smaller round trip delays make the gap smaller, then we will be able to use the theory of robust stability to ensure stability for the network with small delays.

Throughout this section, we let $X = W^{1,\infty}(\mathbb{R}_+)$ and $Y = C(\mathbb{R}_+)$ as per Definitions 3.1 and 3.2, that is, the external disturbance e is such that $e, e' \in L^\infty(\mathbb{R}_+)$ and the disturbance $d \in L^\infty(\mathbb{R}_+)$ is continuous.

We first give some notation, definitions and results on the gap metric, based on [5].

4.1 A gap metric approach

The gap metric theory, developed over the past decade (see [2, 6, 7, 8, 10] and references therein), is an important approach for analyzing robust stability properties of control systems. It captures the idea that any sensible controller stabilizing a nominal plant also stabilizes a perturbed plant provided that an approximate distance between the two plants, measured by a gap metric, is small enough. Most existing gap metrics and related robustness results require an equilibrium of a system operator at 0 and therefore are not directly applicable to the study of network flow control. In this section, we present two results (one is from [8]), using biased norms of operators which overcome the equilibrium restriction.

Let $\mathcal{X}$ be a nonempty set. For $0 < \omega \le \infty$, let $\mathcal{S}_\omega$ denote the set of all locally integrable functions from $[0, \omega)$ to $\mathcal{X}$. For $\sigma \in (0, \omega)$, define a truncation operator T_σ and a restriction operator R_σ, respectively, as follows:

$$T_\sigma : \mathcal{S}_\omega \to \mathcal{S}_\infty, (T_\sigma x)(t) = \begin{cases} x(t), & \text{for } t \in [0, \sigma); \\ 0, & \text{otherwise.} \end{cases}$$

$$R_\sigma : \mathcal{S}_\infty \to \mathcal{S}_\omega, (R_\sigma x)(t) = x(t), \ t \in [0, \sigma).$$

Let $\mathcal{V} \subset \mathcal{S}_\infty$ be a normed vector space, and the norm $\|\cdot\| = \|\cdot\|_\mathcal{V}$ be defined for signals of the form $T_\sigma v, v \in \mathcal{V}, \sigma > 0$. We can define a norm $\|\cdot\|_\sigma$ on $\mathcal{S}_\sigma$ by $\|v\|_\sigma = \|T_\sigma v\|$. We associate spaces as follows:

$\mathcal{V}[0, \sigma) = \{v \in \mathcal{S}_\sigma : v = R_\sigma w, w \in \mathcal{V}, \|v\|_\sigma < \infty\}$, the interval space for $\sigma > 0$;
$\mathcal{V}_e = \{v \in \mathcal{S}_\infty : \forall \sigma > 0, R_\sigma v \in \mathcal{V}[0, \sigma)\}$, the extended space ;
$\mathcal{V}_\omega = \{v \in \mathcal{S}_\omega : \forall \sigma \in (0, \omega), R_\sigma v \in \mathcal{V}[0, \sigma)\}$, for $0 < \omega \le \infty$;
$\mathcal{V}_a = \cup_{\omega \in (0, \infty]} \mathcal{V}_\omega$, the ambient space.

Let $\mathcal{U}, \mathcal{Y}$ be two normed signal spaces over $\mathbb{R}_+$ (such as $L^p(\mathbb{R}_+, \mathbb{R}^n), 1 \le p \le \infty$) with norm $\|\cdot\|_\mathcal{U}, \|\cdot\|_\mathcal{Y}$ respectively. If the notation is unambiguous, we will drop the subscripts. A mapping $Q : \mathcal{U}_a \to \mathcal{Y}_a$ is said to be causal if for any $x, y \in \mathcal{U}_a$ and any $\sigma \in \text{Dom}(x) \cap \text{Dom}(Qx)$, we have

$$R_\sigma x = R_\sigma y \quad \text{implies } R_\sigma(Qx) = R_\sigma(Qy).$$

Let $P : \mathcal{U}_a \to \mathcal{Y}_a$ and $C : \mathcal{Y}_a \to \mathcal{U}_a$ be two causal mappings representing the plant and the controller, respectively. We consider the system of equations[3]

$$u_0 = u_1 - u_2, \quad y_0 = y_1 - y_2, \quad y_1 = Pu_1, \quad u_2 = Cy_2 \tag{4.1}$$

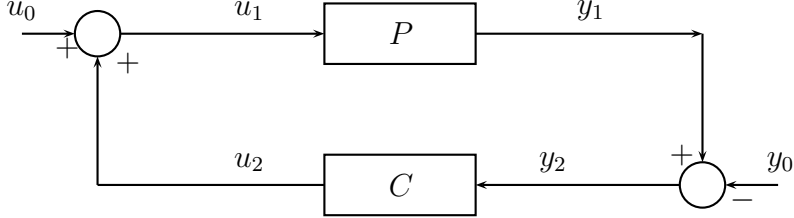corresponding to the closed-loop feedback configuration in Figure 5.

Figure 6: Standard feedback interconnection.

Let $\mathcal{W} = \mathcal{U} \times \mathcal{Y}$ with the product norm $\|(u, y)^\top\|_\mathcal{W} = \max\{\|u\|_\mathcal{U}, \|y\|_\mathcal{Y}\}$. Let

$$\text{Graph}(P) = \left\{ \begin{pmatrix} u \\ Pu \end{pmatrix} : u \in \mathcal{U}, Pu \in \mathcal{Y} \right\}, \quad \text{Graph}(C) = \left\{ \begin{pmatrix} Cy \\ y \end{pmatrix} : Cy \in \mathcal{U}, y \in \mathcal{Y} \right\}.$$

denote the graph of P and the graph of C, respectively. Suppose both P and C are stabilizable, i.e. for all $w = (u, y)^\top \in \mathcal{W}_a$ satisfying $Pu = y$ (resp. $Cy = u$) and for all $\sigma \in \text{Dom}(w)$, there exists $w' \in \text{Graph}(P)$ (resp. $\text{Graph}(C)$) such that $R_\sigma w = R_\sigma w'$.

For $w_0 = (u_0, y_0)^\top \in \mathcal{W}$, a pair $w_1 = (u_1, y_1)^\top, w_2 = (u_2, y_2)^\top \in \mathcal{W}_a$ is said to be a solution of the system if (4.1) holds on $\text{Dom}(w_1) \cap \text{Dom}(w_2)$ which is an interval $[0, \omega)$ with $\omega > 0$. Let $\mathcal{Z}_{w_0}$ be the set of all solutions to the system corresponding to the given w_0, which could be empty. Assume that $[P, C]$ has both the existence property, i.e. $\mathcal{Z}_{w_0} \neq \emptyset$ for each $w_0 \in \mathcal{W}$, and the uniqueness property, i.e.

$$(w_1, w_2), (\tilde{w}_1, \tilde{w}_2) \in \mathcal{Z}_{w_0} \text{ implies } (w_1, w_2) = (\tilde{w}_1, \tilde{w}_2) \text{ on } \text{Dom}(w_1, w_2) \cap \text{Dom}(\tilde{w}_1, \tilde{w}_2).$$

Then, for each $w_0 \in \mathcal{W}$, we define a number ω_{w_0} by

$$[0, \omega_{w_0}) = \cup_{(\tilde{w}_1, \tilde{w}_2) \in \mathcal{Z}_{w_0}} \text{Dom}(\tilde{w}_1) \cap \text{Dom}(\tilde{w}_2),$$

and define a pair $(w_1, w_2) \in \mathcal{W}_a \times \mathcal{W}_a$, with domain $\text{Dom}(w_1, w_2) = [0, \omega_{w_0})$, by the property $R_t(w_1, w_2) \in \mathcal{Z}_{w_0}$ for all $t < \omega_{w_0}$. This induces the system operator:

$$H_{P,C} : \mathcal{W} \to \mathcal{W}_a \times \mathcal{W}_a, \quad H_{P,C}w_0 = (w_1, w_2).$$

Let $\Pi_i : \mathcal{W}_a \times \mathcal{W}_a \to \mathcal{W}_a$ be the projection onto the i-th component of $\mathcal{W}_a \times \mathcal{W}_a$ for $i = 1, 2$. We define

$$\Pi_{P//C} = \Pi_1 H_{P,C}, \quad \text{and} \quad \Pi_{C//P} = \Pi_2 H_{P,C}.$$

Clearly, $H_{P,C} = (\Pi_{P//C}, \Pi_{C//P})$ and $\Pi_{P//C} - \Pi_{C//P} = I$.

---

[3]In literature, the equation is $u_0 = u_1 + u_2, y_0 = y_1 + y_2, y_1 = Pu_1$ and $u_2 = Cy_2$. The slight change here is to make the theorems easier to apply to the network and does not affect existing results except for a minor modification in the proofs.

Definition 4.1. *Let $\Omega \subset \mathcal{W}$. The closed-loop (4.1) is said to be:*

- *locally well posed on Ω if it has the existence and uniqueness properties and the operator $H_{P,C}|_\Omega : \Omega \to \mathcal{W}_a \times \mathcal{W}_a$ is causal;*

- *globally well posed on Ω if it is locally well posed and $H_{P,C}(\Omega) \subset \mathcal{W}_e \times \mathcal{W}_e$;*

- *regularly well posed if it is locally well posed and for all $w_0 \in \mathcal{W}$ with $\omega_{w_0} < \infty$, we have*

$$\|H_{P,C} w_0\|_\sigma \to \infty \text{ as } \sigma \to \omega_{w_0}.$$

Definition 4.2. *Let $\Omega \subset \mathcal{W}$. The closed-loop $[P,C]$ given by (4.1) is said to be:*

- *gain stable on Ω if it is globally well posed on Ω and*

$$\|H_{P,C}\| := \inf\{\alpha \geq 0 : \|R_\sigma H_{P,C} w\|_{\mathcal{W}[0,\sigma)} \leq \alpha \|R_\sigma w\|_{\mathcal{W}[0,\sigma)}, \sigma > 0, w \in \Omega\} < \infty;$$

- *gain stable w.r.t. $\tilde{w} \in \mathcal{W}$ on Ω if it is globally well posed on Ω and*

$$\|H_{P,C}\|_{(\tilde{w})} := \inf \left\{ \alpha \geq 0 : \begin{array}{c} \|R_\sigma H_{P,C} w - R_\sigma H_{P,C} \tilde{w}\|_{\mathcal{W}[0,\sigma)} \leq \alpha \|R_\sigma w - R_\sigma \tilde{w}\|_{\mathcal{W}[0,\sigma)} \\ \text{for all } \sigma > 0, w \in \Omega \end{array} \right\} < \infty;$$

- *β-gain stable on Ω for a $\beta > 0$ if it is globally well posed on Ω and*

$$\|H_{P,C}\|_\beta := \inf \left\{ \alpha \geq 0 : \|R_\sigma H_{P,C} w\|_{\mathcal{W}[0,\sigma)} \leq \alpha \|R_\sigma w\|_{\mathcal{W}[0,\sigma)} + \beta \text{ for all } \sigma > 0, w \in \Omega \right\} < \infty.$$

It is noticed that stability of $[P,C]$ is equivalent to the same stability of either $\Pi_{P//C}$ or $\Pi_{C//P}$. So the stability of a closed loop $[P,C]$ can be determined by the calculation of the induced norm, or either of the two bias norms, of the operator $\Pi_{P//C}$. For robustness, given P the nominal plant and P_1 the perturbed plant, we aim to bound $\|\Pi_{P_1//C}\|$ (resp. $\|\Pi_{P_1//C}\|_{(\tilde{w})}, \|\Pi_{P_1//C}\|_\beta$) in terms of $\|\Pi_{P//C}\|$ (resp. $\|\Pi_{P//C}\|_{(u_0)}, \|\Pi_{P//C}\|_\beta$). The gap metric framework provides a practical way of doing so. The nonlinear gap metric is presented in [7] by Georgiou and Smith using surjective mappings between graphs of the plant and controllers:

Definition 4.3. *The gap metric distance between causal operators $P, P_1 : \mathcal{U}_a \to \mathcal{Y}_a$ is defined to be*

$$\vec{\delta}(P, P_1) = \begin{cases} \inf_{\Phi \in \Theta} \|(I - \Phi)|_{\text{Graph}(P)}\| & \text{if } \Theta \neq \emptyset, \\ \infty & \text{if } \Theta = \emptyset \end{cases}$$

with

$$\Theta = \left\{ \Phi : \begin{array}{c} \Phi : \text{Dom}(\Phi) \subset \text{Graph}(P) \to \text{Graph}(P_1) \text{ is a} \\ \text{causal, gain stable and surjective mapping} \end{array} \right\}.$$

We remark that, if P and P_1 are the transfer functions of linear systems, this gap can be computed via a standard $\mathcal{H}^\infty$ optimisation [6]:

$$\vec{\delta}_0(P, P_1) = \inf_{Q \in \mathcal{H}_\infty} \left\| \begin{pmatrix} D \\ N \end{pmatrix} - \begin{pmatrix} D_1 \\ N_1 \end{pmatrix} Q \right\|,$$

where $(D, N), (D_1, N_1)$ are normalized right coprime factorizations of P and P_1 respectively. In the nonlinear setting, similar equivalent expressions can be found in [2] and [10].

The next lemma, which is a generalization of the main result of [7], shows that this gap metric can also be used to study the robustness of β-gain stability provided $\vec{\delta}(P, P_1)$ is sufficiently small. A similar result can be found in [3].

Lemma 4.4. *Let $[P, C]$ be β-gain stable and suppose $[P_1, C]$ is either globally or regulaly well-posed. If*

$$\vec{\delta}(P, P_1)\|\Pi_{P//C}\|_\beta < 1 \text{ with } \beta \geq 0,$$

then $[P_1, C]$ is globally well-posed and

$$\|\Pi_{P_1//C}w\|_{\beta_1} \leq \|\Pi_{P//C}\|_\beta \frac{1 + \vec{\delta}(P, P_1)}{1 - \vec{\delta}(P, P_1)\|\Pi_{P//C}\|_\beta} \tag{4.2}$$

where

$$\beta_1 = \frac{1 + \vec{\delta}(P, P_1)}{1 - \vec{\delta}(P, P_1)\|\Pi_{P//C}\|_\beta}\beta. \tag{4.3}$$

Proof. Let $\rho = \vec{\delta}(P, P_1)$ and $w \in \mathcal{W}$. By assumption, there exists $\varepsilon_0 > 0$ such that $(\rho + \varepsilon_0)\|\Pi_{P//C}\|_\beta < 1$. Let $\varepsilon \in (0, \varepsilon_0)$ be given. By the definition of $\vec{\delta}(P, P_1)$, there exist a surjective mapping $\Phi : \mathcal{D} \subset \mathrm{Graph}(P) \rightarrow \mathrm{Graph}(P_1)$ such that $\|\Phi - I\| \leq \rho + \varepsilon < \rho + \varepsilon_0$. Let $H_{P_1//C}w = (w_1, w_2)$. By the stabilizability assumption, for any $0 < \sigma < \omega_w$ there exist $w_1' \in \mathrm{Graph}(P_1), w_2' \in \mathrm{Graph}(C)$ such that $R_\sigma w_1 = R_\sigma w_1', R_\sigma w_2 = R_\sigma w_2'$. Since Φ is surjective, there exists $w_3 \in \mathrm{Graph}(P)$ such that $R_\sigma \Phi w_3 = R_\sigma w_1'$. Let $x = w_3 - w_2'$. Since $[P, C]$ is globally well-posed, it follows from the uniqueness property and the definitions of $\Pi_{C//P}$ and $\Pi_{P//C}$ that $\Pi_{P//C}x = w_3, R_\sigma \Pi_{C//P} = R_\sigma w_2'$. Since $R_\sigma w = R_\sigma w_1' - R_\sigma w_2'$, by the causality of Φ and $\Pi_{P//C}$, we have

$$R_\sigma w = R_\sigma \Phi w_3 - R_\sigma w_2' = R_\sigma \Phi \Pi_{P//C}x - R_\sigma \Pi_{C//P}x$$
$$= R_\sigma x + R_\sigma (\Phi - I)\Pi_{P//C}x$$

and

$$R_\sigma \Pi_{P_1//C}w = R_\sigma w_1 = R_\sigma w_1' = R_\sigma \Phi \Pi_{P//C}x.$$

Hence

$$\|R_\sigma x\| \leq \|R_\sigma w\| + \|R_\sigma (\Phi - I)\Pi_{P//C}x\| \leq \|R_\sigma w\| + \|\Phi - I\|(\|\Pi_{P//C}\|_\beta\|R_\sigma x\| + \beta).$$

Therefore

$$\|R_\sigma x\| \leq \frac{\|R_\sigma w\| + \|\Phi - I\|\beta}{1 - \|\Phi - I\|\|\Pi_{P//C}\|_\beta}$$

and

$$\|R_\sigma \Pi_{P_1//C}w\| \leq \|R_\sigma \Pi_{P//C}x\| + \|R_\sigma (\Phi - I)\Pi_{P//C}x\|$$
$$\leq \|R_\sigma \Pi_{P//C}\|_\beta\|R_\sigma x\| + \beta + \|\Phi - I\|\|\Pi_{P//C}\|_\beta\|R_\sigma x\| + \|\Phi - I\|\beta$$
$$\leq \|\Pi_{P//C}\|_\beta \frac{1 + \|\Phi - I\|}{1 - \|\Phi - I\|\|\Pi_{P//C}\|_\beta}(\|R_\sigma w\| + \|\Phi - I\|\beta) + \beta(1 + \|\Phi - I\|)$$
$$\leq \|\Pi_{P//C}\|_\beta \frac{1 + \rho + \varepsilon}{1 - (\rho + \varepsilon)\|\Pi_{P//C}\|_\beta}(\|R_\sigma w\| + (\rho + \varepsilon)\beta) + \beta(1 + \rho + \varepsilon).$$

Since this uniform bound holds for all $0 < \sigma < \omega_w$, it follows that if $[P_1, C]$ is regularly well posed, then $[P_1, C]$ is also globally well posed. Since $[P_1, C]$ is globally well posed, either by the above or by assumption, inequality (4.2) holds by letting $\varepsilon \rightarrow 0$. This completes the proof. $\square$

Finally we recall the following result of Georgiou and Smith [8] on the robustness of stability with respect to some point u_0. In [8], operators are mappings between normed spaces and the split of disturbances is $u_0 = u_1 + u_2$, $y_0 = y_1 + y_2$, here we present it within the extended/ambient space framework, which is less restrictive. Since the proof is similar, we only state the result.

Lemma 4.5. *Let $[P, C]$ be globally well-posed and $[P_1, C]$ be either globally or regularly well-posed. Suppose $\|\Pi_{P//C}\|_{(u_0)} < \infty$ for some $u_0 \in \mathcal{W}$. Let $\Phi : \mathrm{Graph}(P) \to \mathrm{Graph}(P_1)$ be a surjective mapping such that $\|\Phi - I\|_{(\Pi_{P//C}u_0)}\|\Pi_{P//C}\|_{(u_0)} < 1$. Define $w_0 = (I + (\Phi - I)\Pi_{P//C})u_0$. Then $[P_1, C]$ is globally well posed and*

$$\|\Pi_{P_1//C}\|_{(w_0)} \leq \frac{1 + \|\Phi - I\|_{(\Pi_{P//C}u_0)}}{1 - \|\Phi - I\|_{(\Pi_{P//C}u_0)}\|\Pi_{P//C}\|_{(u_0)}}\|\Pi_{P//C}\|_{(u_0)}.$$

4.2 Asymptotic stability

In this subsection, we consider the asymptotic stability of the following primal algorithm:

$$\dot{x}_r(t) = f_r(U'_r(x_r(t)), q_r(t)), \tag{4.4}$$

$$q_r(t) = \sum_{j \in r} \mu_j(t - \tau_{j,r}), \quad \mu_j(t) = p_j(y_j(t)), \quad y_j(t) = \sum_{s:j \in s} x_s(t - \tau_{s,j}), \tag{4.5}$$

$$p_j(y_j) = \rho_j + \alpha_j y_j \text{ with } \alpha_j \geq 0 \text{ for all } y_j, \tag{4.6}$$

where $\tau_{j,r}$ is the propagation delay from link j to user r, $\tau_{s,j}$ is the delay from route s to link j and $\rho_j \geq 0$ is a constant representing the minimum charge of link j. In comparison to the algorithms of the last section, here the function f_r has only two independent variables and the price function p_j is affine.

To apply the gap metric theory of Section 4.1, we re-draw the closed loop diagram and introduce two new variables, $\bar{\mu}$ and $\bar{y}$, as shown in Figure 7, where A, y, μ, p, q, f, U' and x are as described in Section 2. We suppose that the disturbance vector e is continuously differentiable and the disturbance vector d is continuous. We consider the algorithm as a closed loop in the functional spaces: $\mathcal{U} = W^{1,\infty}(\mathbb{R}_+)$ with norm $\|u\|_{\mathcal{U}} = \sup_{t \geq 0}\|u(t)\| + \sup_{t \geq 0}\|\dot{u}(t)\|$ and $\mathcal{Y} = C(\mathbb{R}_+)$ with norm $\|y\|_{\mathcal{Y}} = \sup_{t \geq 0}\|y(t)\|$.
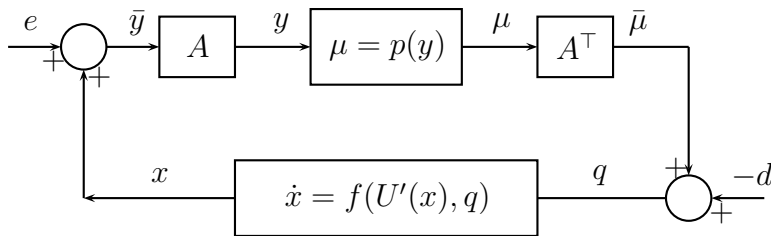

Figure 7: The closed-loop of primal algorithm with external disturbances.

Define operators $P, P_1 : \mathcal{U}_a \to \mathcal{Y}_a$ and $C : \mathcal{Y}_a \to \mathcal{U}_a$ as:

$$P : \bar{y} \mapsto \bar{\mu}, \quad \bar{\mu}(t) = \left(\sum_{j \in r} \left(\rho_j + \alpha_j \left(\sum_{s:j \in s} \bar{y}_s(t) \right) \right) \right)_{r \in R}, \tag{4.7}$$

$$P_1 : \bar{y} \mapsto \bar{\mu}, \quad \bar{\mu}(t) = \left(\sum_{j \in r} \left[\rho_j + \alpha_j \left(\sum_{s:j \in s} \bar{y}_s(t - \tau_{j,r} - \tau_{s,j}) \right) \right] \right)_{r \in R}, \tag{4.8}$$

$$C : q \mapsto x, \quad x = (x_r)_{r \in R}, \quad \dot{x}_r(t) = f_r(U_r'(x_r(t)), q_r(t)), \quad x_r(0) = x_{r,0}, \tag{4.9}$$

where $y(-\tau) = y(0)$ for any $\tau > 0$ and $y \in \mathcal{U}_a$. Then, both $[P, C]$ and $[P_1, C]$ form closed loop control systems, representing the networks without or with round trip delays, respectively.

We now apply Lemma 4.5 to prove that, with reasonable assumptions, there exists $w_0 \in \mathcal{U} \times \mathcal{Y}$ such that (4.4)-(4.6) is asymptotically stable with respect to w_0 on $\mathcal{U} \times \mathcal{Y}$.

Theorem 4.6. *Consider the system $[P_1, C]$ given by (4.8)-(4.9) with f_r and U_r the functions as described in Section 2. Suppose*
(i) there exist $k_1 > 0, k_2 \geq 0$ and a differentiable function $g_r : \mathbb{R}^2 \to \mathbb{R}$ with $g_r(q_r^*, q_r^*) = 0$ such that, for all $x_r, y_r \in \mathbb{R}_+, q_x, q_y \in \mathbb{R}$, we have

$$(x_r - y_r)\big(f_r(U_r'(x_r), q_x) - f_r(U_r'(y_r) - q_y)\big) \leq (x_r - y_r)\big(g_r(U_r'(x_r), q_x) - g_r(U_r'(y_r), q_y)\big), \tag{4.10}$$

$$\frac{\partial g_r}{\partial z_r}(x_r, z_r, q_r) \geq k_1, \quad -k_2 \leq \frac{\partial g_r}{\partial q_r}(x_r, z_r, q_r) \leq 0, \tag{4.11}$$

(ii) for all $z_r, q_r \in \mathbb{R}$, the functions $z \mapsto f_r(z_r, q_r)$ and $q_r \mapsto f_r(z_r, q_r)$ are both left and right differentiable everywhere, and for all $z_r, q_r \in \mathbb{R}$ there exists $k_3 \geq 0$ such that

$$\max \left\{ \left| \frac{\partial^+ f_r}{\partial z_r}(z_r, q_r) \right|, \left| \frac{\partial^- f_r}{\partial z_r}(z_r, q_r) \right|, \left| \frac{\partial^+ f_r}{\partial q_r}(z_r, q_r) \right|, \left| \frac{\partial^- f_r}{\partial q_r}(z_r, q_r) \right| \right\} \leq k_3, \tag{4.12}$$

where $\frac{\partial^+ f_r}{\partial z_r}, \frac{\partial^- f_r}{\partial z_r}$ denote the left and right partial derivative of f_r w.r.t. z_r respectively,
(iii) each function $f_r : \mathbb{R}^2 \to \mathbb{R}$ satisfies inequalities (2.13), (2.18) with the given $m_r \geq 0$ and further there exist $\lambda_1, \lambda_2 > 0$ such that

$$-\lambda_2 \leq U_r''(z) \leq -\lambda_1 \quad \text{for all } z \in [0, m_r], r \in R. \tag{4.13}$$

Let A be the routing matrix and let

$$\tau = \max\{\tau_{j,r} + \tau_{s_j} : j \in J, r, s \in R\},$$
$$\alpha = \max\{\alpha_j : j \in J\},$$
$$\delta_1 = 1 + k_3 + \frac{k_2}{\lambda_1 k_1}\big(1 + \alpha \|A^\top A\| + k_3(\lambda_2 + \alpha \|A^\top A\|)\big), \tag{4.14}$$
$$\delta_2 = 1 + \alpha \|A^\top A\| \left(1 + k_3 + \frac{1 + \alpha \|A^\top A\| + k_3(\lambda_2 + \alpha \|A^\top A\|)}{\lambda_1 k_1} k_2 \right). \tag{4.15}$$

If $\tau \alpha \|A^\top A\| \max\{\delta_1, \delta_2\} < 1$, then (4.4)-(4.6) is asymptotically stable with respect to $w_0 = (0, P_1 \tilde{x} - \tilde{q})^\top$ on $\mathcal{D} = C(\mathbb{R}_+) \times W^{1,\infty}(\mathbb{R}_+)$, where $\tilde{x}, \tilde{q}$ are the solution to the closed loop $[P, C]$ when external disturbances are 0.

Proof. As in the last section, we sometimes adopt the vector forms for equations (4.4)-(4.6), using the notation given in (2.8)-(2.9), or the operators C, P and P_1 for convenience.

Let $\tilde{x}, x$ be the solutions to the closed-loop $[P, C]$ without and with external disturbances respectively, starting from the same initial value x_0, that is

$$\dot{x} = f(U'(x), q), \ x(0) = x_0, \ q = -d + A^\top p(Ax + Ae),$$
$$\tilde{x}' = f(U'(\tilde{x}), \tilde{q}), \ \tilde{x}(0) = x_0, \ \tilde{q} = A^\top p(A\tilde{x})$$

As assumed in (iii), $x_r, \tilde{x}_r \in [0, m_r]$ for each $r \in R$.

We first estimate $\|\Pi_{P//C}\|_{(0)}$ in the space $\mathcal{W} =: \mathcal{U} \times \mathcal{Y}$. Let $\sigma \geq 0$. By definition of $\Pi_{P//C}$ and the causality of $\Pi_{P//C}$, we have

$$\left\| R_\sigma \Pi_{P//C} \begin{pmatrix} e \\ d \end{pmatrix} - R_\sigma \Pi_{P//C} 0 \right\|_{\mathcal{W}[0,\sigma)} = \left\| R_\sigma \begin{pmatrix} \bar{y} \\ \bar{\mu} \end{pmatrix} - R_\sigma \begin{pmatrix} \tilde{x} \\ \tilde{q} \end{pmatrix} \right\|_{\mathcal{W}[0,\sigma)}$$
$$= \left\| R_\sigma \begin{pmatrix} x + e - \tilde{x} \\ q + d - \tilde{q} \end{pmatrix} \right\|_{\mathcal{W}[0,\sigma)}. \tag{4.16}$$

By our differentiability assumption and the mean value theorem (Lemma 6.2), we see there exist $z_r, \eta_r \in \mathbb{R}_+, \xi_r \in [0, m_r]$ and $\omega_{r,1}, \omega_{r,2} \in \mathbb{R}$ such that

$$\omega_{r,1} \in \left[\left. \frac{\partial^- f_r(\phi_r, q_r)}{\partial \phi_r} \right|_{\phi_r = z_r}, \left. \frac{\partial^+ f_r(\phi_r, q_r)}{\partial \phi_r} \right|_{\phi_r = z_r} \right],$$

$$\omega_{r,2} \in \left[\left. \frac{\partial^- f_r(U'_r(\tilde{x}), \psi_r)}{\partial \psi_r} \right|_{\psi_r = \eta_r}, \left. \frac{\partial^+ f_r(U'_r(\tilde{x}), \psi_r)}{\partial \psi_r} \right|_{\psi_r = \eta_r} \right],$$

$$f_r(U'_r(x_r), q_r) - f_r(U'_r(\tilde{x}_r), q_r) = \omega_{r,1}(U'_r(x_r) - U'_r(\tilde{x}_r)) = \omega_{r,1} U''_r(\xi_r)(x_r - \tilde{x}_r),$$

$$f_r(U'_r(\tilde{x}_r), q_r) - f_r(U'_r(\tilde{x}_r), \tilde{q}_r) = \omega_{r,2}(q_r - \tilde{q}_r))$$

and

$$q - \tilde{q} = -d + A^\top \text{diag}(\alpha_j) A(e + x - \tilde{x}). \tag{4.17}$$

Let $K_1 = \text{diag}(\omega_{r,1} U''_r(\xi_r))$, $K_2 = \text{diag}(\omega_{r,2})$ be diagonal matrices. Then

$$\dot{x} - \tilde{x}' = f(U'(x), q) - f(U'(\tilde{x}), \tilde{q})$$
$$= K_1(x - \tilde{x}) + K_2 A^\top \text{diag}(\alpha_j) A(x - \tilde{x}) + K_2(-d + A^\top \text{diag}(\alpha_j) Ae). \tag{4.18}$$

Applying the above procedure to function g_r and using our assumption (ii), we see that there exist $\bar{z}_r, \bar{\xi}_r, \bar{\eta}_r$ such that the matrices

$$\tilde{K}_1 = \text{diag} \left(\left. \frac{\partial g_r(\phi_r, q_r)}{\partial \phi_r} \right|_{\phi_r = \bar{z}_r} U''_r(\bar{\xi}_r) \right), \tilde{K}_2 = \text{diag} \left(\left. \frac{\partial g_r(U'_r(\tilde{x}_r), \psi_r)}{\partial \psi_r} \right|_{\psi_r = \bar{\eta}_r} \right)$$

satisfy

$$(x - \tilde{x})^\top (\dot{x} - \tilde{x}') \leq (x - \tilde{x})^\top \tilde{K}_1 (x - \tilde{x}) + (x - \tilde{x})^\top \tilde{K}_2 A^\top \text{diag}(\alpha_j) A(x - \tilde{x})$$
$$+ (x - \tilde{x})^\top K_2(-d + A^\top \text{diag}(\alpha_j) Ae).$$

From the assumptions on the derivatives of U_r and g_r, it follows

$$(x - \tilde{x})^\top \tilde{K}_1 (x - \tilde{x}) \leq -\lambda_1 k_1 \|x - \tilde{x}\|^2, \quad (x - \tilde{x})^\top \tilde{K}_2 A^\top \mathrm{diag}(\alpha_j) A (x - \tilde{x}) \leq 0.$$

All the above inequalities and equations hold for all $t \in \mathbb{R}_+$, but we restrict our consideration on $[0, \sigma]$. Thus,

$$\frac{1}{2} \frac{d}{dt} \|x(t) - \tilde{x}(t)\|^2 = (x - \tilde{x})^\top (\dot{x} - \tilde{x}')$$
$$\leq -\lambda_1 k_1 \|x(t) - \tilde{x}(t)\|^2 + k_2 (\|R_\sigma d\| + \alpha \|A^\top A\| \|R_\sigma e\|) \|x(t) - \tilde{x}(t)\|$$

and, since $x(0) = \tilde{x}(0) = x_0$,

$$\|x(t) - \tilde{x}(t)\|^2 \leq -\lambda_1 k_1 \int_0^t \|x(s) - \tilde{x}(s)\|^2 ds + k_2 (\|R_\sigma d\| + \alpha \|A^\top A\| \|R_\sigma e\|) \int_0^t \|x(s) - \tilde{x}(s)\| ds$$

for all $t \in [0, \sigma]$. By Lemma 6.1, we have

$$\|x(t) - \tilde{x}(t)\| \leq \frac{k_2 (\|R_\sigma d\| + \alpha \|A^\top A\| \|R_\sigma e\|)}{\lambda_1 k_1}, \quad \text{for } t \in [0, \sigma]. \tag{4.19}$$

Inequality (4.19) remains valid when $\tau = \sigma$. Since $\tilde{x}(t) \to x^*$ as proved in the last section, we see the algorithm with external disturbances is asymptotically stable. By (4.17) and (4.18)

$$\|q(t) - \tilde{q}(t)\| \leq \|R_\sigma d\| + \alpha \|A^\top A\| (\|R_\sigma e\| + \|x(t) - \tilde{x}(t)\|)\|, \tag{4.20}$$
$$\|\dot{x}(t) - \tilde{x}'(t)\| \leq k_3 (\lambda_2 + \alpha \|A^\top A\|) \|x(t) - \tilde{x}(t)\| + k_3 (\|R_\sigma d\| + \alpha \|A^\top A\| \|R_\sigma e\|) \tag{4.21}$$

for all $t \in [0, \sigma]$. Substituting (4.19),(4.20) and (4.21) into (4.16) and simplifying to obtain

$$\left\| R_\sigma \Pi_{P//C} \begin{pmatrix} e \\ d \end{pmatrix} - R_\sigma \Pi_{P//C} 0 \right\|_{\mathcal{W}[0,\sigma)} \leq \max\{\delta_1, \delta_2\} \left\| R_\sigma \begin{pmatrix} e \\ d \end{pmatrix} \right\|_{\mathcal{W}[0,\sigma)}$$

which shows $\|\Pi_{P//C}\|_{(0)} \leq \max\{\delta_1, \delta_2\}$.

Consider the mapping $\Phi : \mathrm{Graph}(P) \to \mathrm{Graph}(P_1)$ defined by

$$\Phi \begin{pmatrix} \bar{y}(t) \\ P\bar{y}(t) \end{pmatrix} = \begin{pmatrix} \bar{y}(t) \\ P_1\bar{y}(t) \end{pmatrix}.$$

Clearly, Φ is surjective and causal since both P and P_1 are causal. Also we have

$$\left\| R_\sigma (\Phi - I) \begin{pmatrix} \bar{y} \\ P\bar{y} \end{pmatrix} - R_\sigma (\Phi - I) \begin{pmatrix} \tilde{x} \\ \tilde{q} \end{pmatrix} \right\|_{\mathcal{W}[0,\sigma)} = \sup_{t \in [0,\sigma)} \|R_\sigma (P\bar{y} - P_1\bar{y} - P\tilde{x} + P_1\tilde{x})(t)\|$$

$$= \sup_{t \in [0,\sigma)} \left\| R_\sigma \left(\sum_{j \in r} \alpha_j \left(\sum_{s : j \in s} [(\bar{y}_s - \tilde{x}_s)(t) - (\bar{y}_s - \tilde{x}_s)(t - \tau_{j,r} - \tau_{s,j})] \right) \right)_{r \in R} \right\|$$

$$= \sup_{t \in [0,\sigma)} \left\| R_\sigma \left(\sum_{j \in r} \alpha_j \left(\sum_{s : j \in s} [(\bar{y}_s' - \tilde{x}_s')(t_j)] (\tau_{j,r} + \tau_{s,j}) \right) \right)_{r \in R} \right\|$$

$$\leq \alpha \tau \|A^\top A\| \|R_\sigma (\bar{y} - \tilde{x})\|_{\mathcal{U}[0,\sigma)} \leq \alpha \tau \|A^\top A\| \left\| R_\sigma \begin{pmatrix} \bar{y} \\ P\bar{y} \end{pmatrix} - T_\sigma \begin{pmatrix} \tilde{x} \\ \tilde{q} \end{pmatrix} \right\|_{\mathcal{W}[0,\sigma)},$$

26

where $t_j \in [t - \tau_{j,r} - \tau_{s,j}, t]$. From this it follows $\|\Phi - I\|_{(v_0)} \leq \alpha \|A^\top A\| \tau$.

By our assumptions, function f is local Lipschitz. So for any $(e, d) \in W^{1,\infty}(\mathbb{R}_+) \times C(\mathbb{R}_+)$ and any initial value, the system of differential equations (4.4)-(4.5) has a locally unique solution, and $[P_1, C]$ is regularly well posed. By Lemma 4.5, $[P_1, C]$ is globally well posed, and

$$\|\Pi_{P_1//C}\|_{(w_0)} \leq \frac{1 + \alpha \|A^\top A\| \tau}{1 - \alpha \|A^\top A\| \tau \max\{\delta_1, \delta_2\}} \max\{\delta_1, \delta_2\}. \tag{4.22}$$

Let

$$w_0 = (\Phi - I)\Pi_{P//C} 0 = \begin{pmatrix} 0 \\ P_1 \tilde{x} - \tilde{q} \end{pmatrix}, \quad \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \Pi_{P_1//C} w_0, \quad \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \Pi_{P_1//C} \begin{pmatrix} e \\ d \end{pmatrix}.$$

Then from (4.22), it follows

$$\left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\|_{\mathcal{W}} \leq \frac{1 + \alpha \tau \|A^\top A\|}{1 - \alpha \|A^\top A\| \tau \max\{\delta_1, \delta_2\}} \max\{\delta_1, \delta_2\} \left\| \begin{pmatrix} e \\ d \end{pmatrix} - w_0 \right\|_{\mathcal{W}}.$$

Let $\hat{x} = z_1 - e$. Then $\hat{x}$ is the solution of closed-loop (4.4)-(4.5) and for all $t \geq 0$:

$$\|\hat{x}(t) - w_1(t)\| \leq \|\hat{x} - w_1\|_{\mathcal{U}} \leq \|z_1 - w_1\|_{\mathcal{U}} + \|e\|_{\mathcal{U}}$$
$$\leq \frac{1 + \alpha \tau \|A^\top A\|}{1 - \alpha \|A^\top A\| \tau \max\{\delta_1, \delta_2\}} \max\{\delta_1, \delta_2\} \left\| \begin{pmatrix} e \\ d \end{pmatrix} - w_0 \right\|_{\mathcal{U} \times \mathcal{Y}} + \|e\|_{\mathcal{U}}.$$

Hence the system with round trip delay is asymptotically stable with respect to $w_0 = (0, P_1 \tilde{x} - \tilde{q})$ according to Definition 3.1. $\qquad \square$

This result shows that the solution of the algorithms stay in a neighbourhood of w_2, but w_2 is not necessarily the equilibrium of the system optimisation problem. The size of the neighbourhood depends on both the disturbances and w_0, and the centre of the neighbourhood is actually the solution of the algorithm when $e = 0, d = P_1 \tilde{x} - \tilde{q}$. That is, the disturbances and round trip delay together cause the 'equilibrium' to move. We notice that both w_2 and w_0 are independent of the external disturbances, but they depend on the time delay. If $\tau = 0$, we may let $\Phi = I$ to see $w_0 = 0$ and, in this case, the solution of the system converges to the equilibrium x^* since w_1 converges to x^* as shown in the last section.

We now apply Theorem 4.6 to the algorithm given in Example 2.5, i.e. when the user controller is

$$\dot{x}_r(t) = \kappa_r(U_r'(x_r(t)) - q_r(t))_{x_r(t)}^{m_r}. \tag{4.23}$$

In this case the operator C given by (4.9) becomes

$$C' : q \mapsto x, \quad x = (x_r)_{r \in R}, \ \dot{x}_r(t) = \kappa_r(U_r'(x_r(t)) - q_r(t))_{x_r(t)}^{m_r}, \ x_r(0) = x_{r,0}. \tag{4.24}$$

We first need a lemma.

Lemma 4.7. *The function* $f_r(z_r, q) = \kappa_r(z_r - q_r)_{x_r}^{m_r}$ *is left and right differentiable with respect to either* z_r *or* q_r *and*

$$\frac{\partial^- f_r(z_r, q_r)}{\partial z_r}, \frac{\partial^+ f_r(z_r, q_r)}{\partial z_r} \in [0, \kappa_r], \quad \frac{\partial^- f_r(z_r, q_r)}{\partial z_r}, \frac{\partial^+ f_r(z_r, q_r)}{\partial z_r} \in [-\kappa_r, 0].$$

27

Proof. We only show the left and right differentiability with respect to z_r.

If $x_r \in (0, m_r)$, then $f_r(z_r, q_r) = \kappa_r(z_r - q_r)$ which is clearly differentiable and $\partial f_r / \partial z_r = \kappa_r$ for all z_r; If $x_r = 0$, then $f_r(z_r, q_r) = \kappa_r \max\{0, z_r - q_r\}$, which tells us that f_r is differentiable at any $z_r \neq q_r$, left and right differentiable when $z_r = q_r$, and

$$\text{for } z_r > q_r, \frac{\partial f_r}{\partial z_r} = \kappa_r, \quad \text{for } z_r < q_r, \frac{\partial f_r}{\partial z_r} = 0 \text{ and for } z_r = q_r, \frac{\partial^+ f_r}{\partial z_r} = \kappa_r, \frac{\partial^- f_r}{\partial z_r} = 0.$$

Similarly, if $x_r = m_r$, then $f_r(z_r, q_r) = \kappa_r \min\{0, z_r - q_r\}$ and

$$\text{for } z_r < q_r, \frac{\partial f_r}{\partial z_r} = \kappa_r, \quad \text{for } z_r > q_r, \frac{\partial f_r}{\partial z_r} = 0 \text{ and for } z_r = q_r, \frac{\partial^- f_r}{\partial z_r} = \kappa_r, \frac{\partial^+ f_r}{\partial z_r} = 0.$$

This proves the lemma. $\qquad \square$

Corollary 4.8. *Consider the system $[P_1, C']$ given by (4.8) and (4.24) with U_r as described in section 2. Let $m_r > 0$ and let A be the routing matrix. Suppose $0 < k_1 \leq \kappa_r \leq k_2$ for all $r \in R$. If there exist $\lambda_1, \lambda_2 \geq 0$ such that $-\lambda_2 \leq U_r''(z) \leq -\lambda_1$ for all $z \in [0, m_r]$ and if $\alpha \|A^\top A\| \tau \max\{\delta_1, \delta_2\} < 1$ where δ_1, δ_2 are as in Theorem 4.6, then the algorithm (4.23) with controller (4.5)-(4.6) is asymptotically stable with respect to w_0.*

Proof. Let $f_r(z_r, q_r) = \kappa_r(z_r - q_r)_{x_r}^{m_r}$, $g_r(z_r, q_r) = \kappa_r(z_r - q_r)$. Then $g_r(q_r^*, q_r^*) = 0$, g_r is differentiable with respect to both z_r and q_r and the derivatives satisfy (4.11). By the assumptions and Lemma 4.7, (4.13) and (4.12) are also satisfied. So, only (4.10) is left to be verified.

Let x_r, y_r, q_x, q_y be given. Let $LHS = (x_r - y_r)\big(f_r(U_r'(x_r), q_x) - f_r(U_r'(y_r), q_y)\big)$ and $RHS = (x_r - y_r)\big(g_r(U_r'(x_r), q_x) - g_r(U_r'(y_r), q_y)\big)$. When $x_r = y_r$, the inequality is obvious. We may suppose $x_r > y_r$, if necessary, by interchanging the roles of x_r, y_r. Then, there are four cases to consider:

$$m_r > x_r > y_r = 0; \ m_r > x_r > y_r > 0; \ m_r = x_r > y_r > 0 \ \text{ and } m_r = x_r > y_r = 0.$$

In the case where $m_r > x_r > y_r = 0$, we have

$$LHS = (x_r - y_r)\kappa_r\big(U_r'(x_r) - q_x - \max\{0, U_r'(y_r) - q_y\}\big)$$
$$\leq (x_r - y_r)\kappa_r\big(U_r'(x_r) - q_x - (U_r'(y_r) - q_y)\big) = RHS.$$

In the case where $m_r > x_r > y_r > 0$, we have

$$LHS = (x_r - y_r)\kappa_r\big(U_r'(x_r) - q_x - U_r'(y_r) + q_y\big) = RHS.$$

In the case where $m_r = x_r > y_r > 0$, we have

$$LHS = (x_r - y_r)\kappa_r\big(\min\{0, U_r'(x_r) - q_x\} - (U_r'(y_r) - q_y)\big)$$
$$\leq (x_r - y_r)\kappa_r[U_r'(x_r) - q_x - (U_r'(y_r) - q_y)] = RHS.$$

Finally, if $m_r = x_r > y_r = 0$ then

$$LHS = (x_r - y_r)\kappa_r\big(\min\{0, U_r'(x_r) - q_x\} - \max\{0, U_r'(y_r) - q_y\}\big)$$
$$\leq (x_r - y_r)\kappa_r\big(U_r'(x_r) - q_x - (U_r'(y_r) - q_y)\big) = RHS.$$

Hence (4.10) holds for all x_r, y_r.

By Theorem 4.6, the conclusion follows. $\qquad \square$

28

4.3 Gain stability

In the last subsection, asymptotic stability of algorithms required the price function to be affine when round trip time delay was considered. If the price function is not affine, we consider weaker notions of stability. So in this subsection, we consider gain stability as defined in Definition 3.2 for the general system:

$$\dot{x}_r(t) = f_r(x_r(t), U_r'(x_r(t)), q_r(t)), \tag{4.25}$$

$$q_r(t) = \sum_{j \in r} \mu_j(t - \tau_{j,r}), \ \mu_j(t) = p_j(y_i(t)), \ y_j(t) = \sum_{s:j \in s} x_s(t - \tau_{s,j}), \tag{4.26}$$

where, in contrast with the last subsection, p_j is not necessarily affine.

As in the last subsection, we let $\mathcal{Y} = C(\mathbb{R}_+)$ and $\mathcal{U} = W^{1,\infty}(\mathbb{R}_+)$ and let

$$P : \bar{y} \mapsto \bar{\mu}, \ \ \bar{\mu}(t) = A^\top p(A(\bar{y}(t))) = \left(\sum_{j \in r} p_j \left(\sum_{s:j \in s} y_j(t) \right) \right)_{r \in R}, \tag{4.27}$$

$$P_1 : \bar{y} \mapsto \bar{\mu}, \ \ \bar{\mu}(t) = \left(\sum_{j \in r} p_j \left(\sum_{s:j \in s} y_j(t - \tau_{j,r} - \tau_{s,j}) \right) \right)_{r \in R}, \tag{4.28}$$

$$C : q \mapsto x, \ \ \dot{x} = f(x, U, (x), q), \ x(0) = x_0. \tag{4.29}$$

Here x, f, U', q are vectors and A is the routing matrix, as described in Section 2.

Theorem 4.9. *Consider the system $[P_1, C]$ given by (4.28)-(4.29) with the functions f_r, U_r and p_j and A is the routing matrix as described in Section 2. Suppose that all assumptions of Theorem 3.3 are all satisfied. Suppose, in addition, $|f_r(x_r, z_r, q_r)| \leq k_3|z_r - q_r|$ for all $x_r \geq 0, z_r, q_r \in \mathbb{R}$ and $-\lambda_2 \leq U_r''(z_r)$ for all $z_r \in [0, m_r]$. Let*

$$\tau = \max\{\tau_{j,r} + \tau_{s,j} : r, s \in R, j \in J\}$$

$$\delta_1 = k_3 + \left(1 + k_3\lambda_2 + (1 + k_3)\alpha\|A^\top A\|\right)\frac{\alpha k_2\|A^\top A\|}{\lambda_1 k_1},$$

$$\delta_2 = 1 + (1 + k_3)\alpha\|A^\top A\| + \left(1 + k_3\lambda_2 + (1 + k_3)\alpha\|A^\top A\|\right)\frac{\alpha k_2\|A^\top A\|}{\lambda_1 k_1}.$$

If

$$\tau\alpha\|A^\top A\| \max\{\delta_1, \delta_2\} < 1, \tag{4.30}$$

then the system (4.25),(4.26) is β_1-gain stable, where

$$\beta_1 = \frac{1 + \tau\alpha\|A^\top A\|}{1 - \tau\alpha\|A^\top A\| \max\{\delta_1, \delta_2\}}\beta,$$

$$\beta = \|x^*\| + \|A^\top p(Ax^*)\| + \left(1 + k_3\lambda_2 + (1 + k_2)\alpha\|A^\top A\|\right)\|x(0) - x^*\|.$$

Proof. We first estimate $\|\Pi_{P//C}\|_\beta$ for some $\beta \geq 0$. Let $\sigma \geq 0$. Applying Theorem 3.3 to the time interval $[0, \sigma]$ we see that any solution x to $[P, C]$ satisfies:

$$\|x(t) - x^*\| \leq \|x(0) - x^*\| + \frac{k_2(\|R_\sigma d\| + \alpha\|A^\top A\|\|R_\sigma e\|)}{\lambda_1 k_1} \ \text{ for } t \in [0, \sigma]. \tag{4.31}$$

Let $\bar{\mu}^* = A^\top p(Ax^*)$. Since $p'_j(z) \leq \alpha$ for all z, by the Mean Value Theorem,

$$\|\bar{\mu}(t) - \bar{\mu}^*\| = \|A^\top p(Ae + Ax) - A^\top p(Ax^*)\| = \|A^\top p'(z)A(e + x(t) - x^*)\|$$
$$\leq \alpha \|A^\top A\|(\|R_\sigma e\| + \|x(t) - x^*\|). \tag{4.32}$$

By the additional assumption,

$$\|\dot{x}(t)\| = \|f(x, U'(x), q)\| \leq k_3 \|U'(x^*) - q\| = k_3 \|(U'(x) - U'(x^*)) - (q - q^*)\|$$
$$= k_3 \|U''_r(\xi_r))(x - x^*) - (\bar{\mu} - \bar{\mu}^*) - d\| \leq k_3 \lambda_2 \|x - x^*\| + k_3 \|\bar{\mu} - \bar{\mu}^*\| + k_3 \|R_\sigma d\|$$
$$\leq k_3(\lambda_2 + \alpha)\|A^\top A\|)\|x - x^*\| + k_3\|R_\sigma d\| + k_3 \alpha \|A^\top A\|\|R_\sigma e\| \tag{4.33}$$

for all $t \in [0, \sigma]$. Since $\bar{y} = x + e$, we have

$$\|R_\sigma \bar{y}\|_{\mathbf{X}} = \|R_\sigma \bar{y}\| + \|R_\sigma \bar{y}'\| \leq \|R_\sigma x\| + \|R_\sigma \dot{x}\| + \|R_\sigma e\| + \|R_\sigma e'\| \leq \|R_\sigma x\| + \|x R'_\sigma\| + \|R_\sigma e\|.$$

Together with (4.31), (4.32) and (4.33), we have

$$\|\Pi_{P//C}\|_\beta < \max\{\delta_1, \delta_2\}.$$

To calculate the gap between P and P_1, for a given $\bar{y} \in \mathcal{U}$, we let $\bar{\mu}(t) = P\bar{y}(t))$ and $\hat{\mu}(t) = P_1 \bar{y}(t)$. Let $\Phi : \text{Graph}(P) \to \text{Graph}(P_1)$ be the mapping defined by

$$\Phi \begin{pmatrix} \bar{y}(t) \\ \bar{\mu}(t) \end{pmatrix} = \begin{pmatrix} \bar{y}(t) \\ \hat{\mu}(t) \end{pmatrix}.$$

Then Φ is a surjective mapping from $\text{Graph}(C)$ to $\text{Graph}(C_1)$ and

$$\|R_\sigma(\Phi - I)(\bar{y}, \bar{\mu})^\top\| = \sup_{t \in [0,\sigma)} \|R_\sigma \bar{\mu}(t) - R_\sigma \hat{\mu}(t)\|$$

$$= \sup_{t \in [0,\sigma)} \left\| R_\sigma \left(\sum_{j \in r} p'_j(\xi_j) \left(\sum_{s:j \in s} (y_s(t) - y_s(t - \tau_{j,r} - \tau_{s,j})) \right) \right)_{r \in R} \right\|$$

$$= \sup_{t \in [0,\sigma)} \left\| R_\sigma \left(\sum_{j \in r} p'_j(\xi_j) \left(\sum_{s:j \in s} (y'_s(t_j)(\tau_{j,r} + \tau_{s,j})) \right) \right)_{r \in R} \right\|$$

$$\leq \alpha \tau \|A^\top A\| \|T_\sigma y\|_{\mathcal{U}},$$

from which it follows $\vec{\delta}(P, P_1) \leq \alpha \tau \|A^\top A\|$. Therefore, by Lemma 4.4, if (4.30) holds, then $[P_1, C]$ is β_1-gain stable and

$$\|R_\sigma \Pi_{P_1//C}(e, d)^\top\| \leq \frac{1 + \tau \alpha \|A^\top A\|}{1 - \tau \alpha \|A^\top A\| \max\{\delta_1, \delta_2\}} \max\{\delta_1, \delta_2\} \|R_\sigma(e, d)^\top\| + \beta_1$$

$$\leq \frac{1 + \tau \alpha \|A^\top A\|}{1 - \tau \alpha \|A^\top A\| \max\{\delta_1, \delta_2\}} \max\{\delta_1, \delta_2\} \|(e, d)^\top\| + \beta_1.$$

Since σ is arbitrary, we have

$$\|x\|_{\mathcal{U}} = \|\bar{y} - e\|_{\mathcal{U}} \leq \frac{1 + \tau \alpha \|A^\top A\|}{1 - \tau \alpha \|A^\top A\| \max\{\delta_1, \delta_2\}} \max\{\delta_1, \delta_2\} \|(e, d)^\top\| + \|e\| + \beta_1.$$

$\square$

5   Examples

Our first example considers a network with a small number of nodes in order to explicitly demonstrate results in a simple setting. The second example considers a more realistic scenario of arbitrary size with a fixed super-link architecture.

Example 5.1. The network shown in Figure 8 has 4 nodes or ISPs $(B_1, \cdots, B_4)$ and 5 links $(l_1, \cdots, l_5)$,
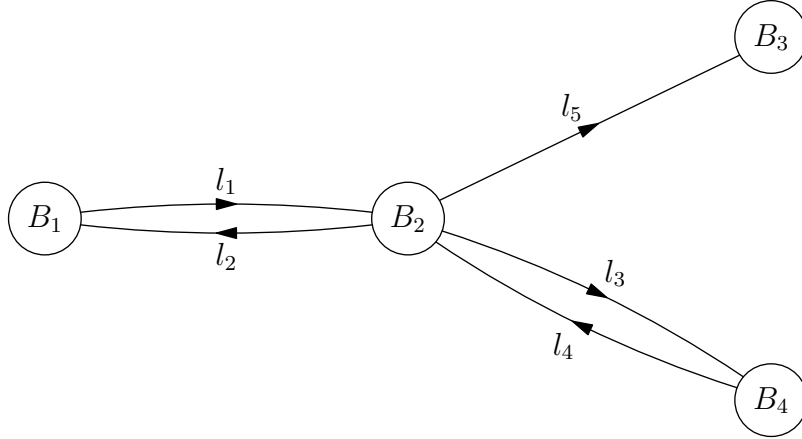

Figure 8: A simple network.

where the link between B_2 and B_4 is one way only. All possible routes are

$$R = \{B_1B_2, B_1B_2B_3, B_1B_2B_4, B_2B_1, B_2B_3, B_2B_4, B_3B_2, B_3B_2B_1, B_3B_2B_4\}.$$

Here B_iB_k means the route from B_i to B_k without going through other ISPs and $B_iB_kB_s$ represent the route from B_i to B_s via B_k. The routing matrix is

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

and $\|A\| = 2, \|A^\top\| = 3, \|AA^\top\| = \|A^\top A\| = 5$.

For this network, we consider the primal algorithm

$$\dot{x}_r = (U'_r(x_r) - q_r)^{9.5}_{x_r}, \quad \mu_j = p_j\left(\sum_{s:j\in s} x_s\right) \tag{5.1}$$

with

$$U_r(x) = \log(x + 0.5), \quad p_j(z) = \rho_j + 0.1z.$$

Then $x_r \in [0, 9.5], p'_j(z) = 0.1$ and $U''_r(x) = -1/(x+0.5)^2 \in [-4, -0.01]$. By letting $g_r(x, z, q) = z - q$, we see $\partial g/\partial z = -1, \partial g/\partial q = -1$ and all assumptions of Corollaries 3.4 and 4.8 are satisfied with $k_1 = k_2 = k_3 = 1, \alpha = 0.1, \lambda_1 = 0.01, \lambda_2 = 4$. So by Corollary 3.4,

$$\|x(t) - x^*\| \leq \|x(0) - x^*\|e^{-0.02t} + 100(\|d\| + 0.5\|e\|).$$

Substituting into (4.14) and (4.15), we have $\delta_1 = 602, \delta_2 = 302$. By Corollary 4.8, if

$$\tau < \frac{1}{\alpha \|A^\top A\| \max\{\delta_1, \delta_2\}} = \frac{1}{301},$$

the algorithm

$$\dot{x}_r = (U_r'(x_r) - q_r)_{x_r}^{9.5}, \quad q_r(t) = \sum_{j \in r} \mu_j(t - \tau_{j,r}), \ \mu_j(t) = \rho_j + 0.1 \sum_{s:j \in s} x_s(t - \tau_{s,j}) \qquad (5.2)$$

will be asymptotically stable in the sense that

$$|x_r(t) - w_{1,r}(t)| \leq \frac{602(1 + 0.5\tau)}{1 - 300.5\tau} \left\| \begin{pmatrix} e \\ d \end{pmatrix} - w_0 \right\|$$

where $w_{1,r}$ is the solution of system (5.2) when disturbances $e = 0, d = (\tilde{d}_r - \tilde{q}_r)_{r \in R}$ with

$$\tilde{d}_r = \sum_{j \in r} \left[\rho_j + 0.1 \left(\sum_{s:j \in s} \tilde{x}(t - \tau_{j,r} - \tau_{s,j}) \right) \right]$$

and $(\tilde{x}_r, \tilde{q}_r)$ is the solution to system (5.1) when $e = d = 0$.
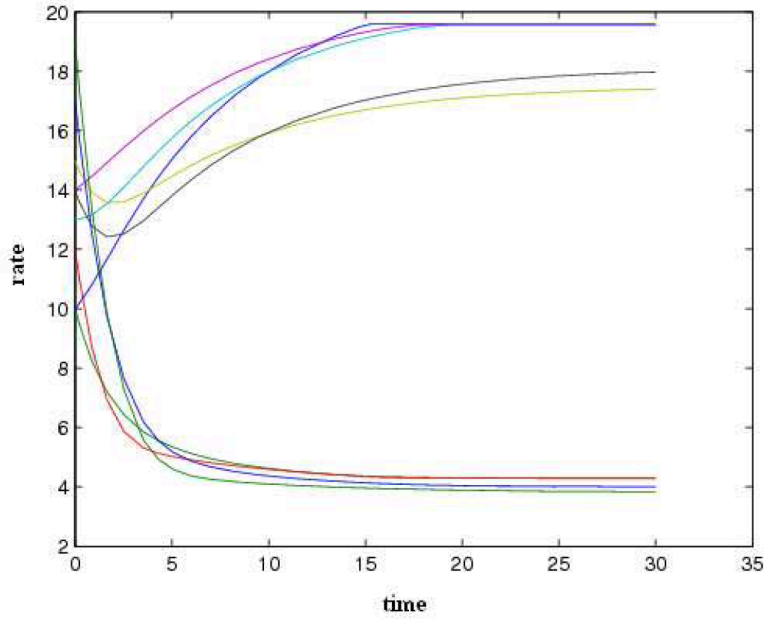

Figure 9: Simulation of network with time delay $\tau = 0.5$ and $m_r = 19.5$. The different lines show the rates of the different users.

It is noted that the upper bound for the tolerable time delay τ is not optimal, Figure 9 shows the convergence of the x_r, starting from different initials, when $m_r = 19.5, \tau = 0.5$, $d = [0.1, 0.2, 0.04, 0.03, 0.06, 0.1, 0.1, 0.06, 0.09]$ and $e = [-0.1, -0.07, 0.02, 0.5, 6, -3, -9, 6, 9]$.
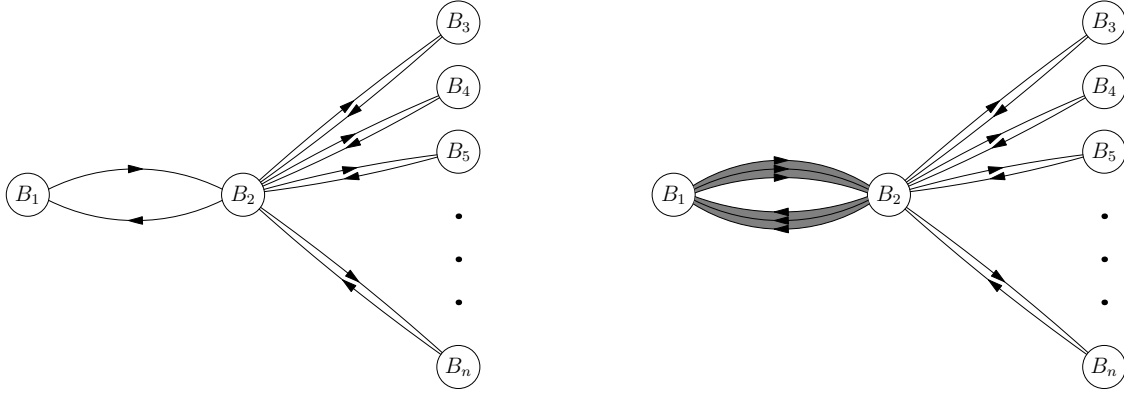
32

Figure 10: A network with more nodes.

Example 5.2. The network in Figure 8 has only four nodes. But in the real internet, the number of nodes is very large, hence the numbers of users through some links and/or the numbers of links used by some users will be large. Consequently $\|A\|, \|A^\top\|$ and $\|AA^\top\|$ could be very large and the tolerable time delay τ could be unrealistically small. As remarked after Theorem 3.3, if each extensively used link is divided into a group of sub-links, the more users use the link, the more sub-links it decomposes, then we can control the number of links used by any user and the number of users using the same links without considering the number of nodes. For example, suppose there are more nodes added to the network of Example 5.1 as shown left in Figure 10: a totality of n nodes. Suppose each link has super capacity and can be treated as a group of sub-links and each sub-link only allows at most 3 routes (this can be achieved via the TCP at each node) (for example we have shown multiple sub links for $B_1 B_2$ in Figure 10 (right). Then, by re-counting the number of links in terms of sub-links, we see each user uses at most 2 links and each link is used by at each 3 users. For this network, the system matrix A has at most $(n^2 + n - 2)/2$ columns and $(n-1)^2/3$ rows, but each column has at most 2 and each row has at most 3 non-zero entries. So $\|AA^\top\| \le 6$ and, therefore, the tolerable time delay τ is not affected by the number of nodes. The convergence of each flow rate is almost the same as in Example 5.1.

We remark that the technique of Example 5.2 is the key to networks of arbitrary size with so-called small world features. Here we define a 'small world' network to be one in which there is a uniform (small) bound l on the number of links along each route, and we let λ denote the number of routes per link (which can be kept small by the technique of Example 5.2). It is straightforward to then observe, as above:

$$\|AA^\top\|_2 \le \|A\|_2 \|A^\top\|_2 \le \|A\|_1 \|A^\top\|_1 \le \max_i \sum_j |a_{ij}| \cdot \max_j \sum_i |a_{ij}| \le l\lambda.$$

We finally remark that sensible routing algorithms themselves include an objective to keep the number of links per route small.

6 Conclusions

In this paper, we studied the stability of network congestion control algorithms with the presence of both external disturbances and round trip delay, which are important in the implemen-

tation of mathematical model. We considered both primal and dual algorithms under different conditions. For both algorithms, the controllers are taken to be of general form satisfying, mainly, certain smoothness assumptions. In the nominal case when there is no round trip delay involved, we considered both asymptotic stability and gain stability. The results shows that although the solution may not necessarily converge to the equilibrium, nevertheless it tends to a neighbourhood of the equilibrium whose size depends on the magnitude of the disturbances. This is a reflection of the fact that the disturbances change the equilibrium of the network and small disturbances will eventually keep these changes small. In the case when the round trip delay is considered, we investigated the stability of the solution against the disturbance. To our knowledge, there is no previous research addressing this question. We have used the gap metric approach for our study. When the price feedback function is linear, we obtained an asymptotic stability result for a primal algorithm. In general, the feedback function is assumed to nonlinear but having linear growth, and gain stability results are obtained for both primal and dual algorithms.

It is important to recognise that a treatment based on a perturbation analysis such as that given here (or in [4]) will necessarily be linked to the case of relatively small round trip delays. The delays present in a real network may be larger than can be handled by such an analysis and may demand a more detailed analysis of this structure effect. Nevertheless we contend that this perturbative analysis does reveal important qualitative effects and trade-offs and leads to significant considerations for systematic design.

Although our consideration is for single-path networking, the results could be generalized to multi-path network problems provided the rate distribution matrix H (see [9, 16]) satisfies certain conditions. We notice that our results depend on the norm of routing matrix A. This limits the scalability of the network. However, if the network is designed such that every link has at most a fixed number of users, the results are scale-invariant: for example the tolerable trip delays are then invariant to the network size.

References

[1] Bian W., Existence results for second order nonlinear evolution inclusions, Indian Journal on Pure and Applied Mathematics, 29 (1998) 1177-1193

[2] Bian W. and French M., Graph topologies, gap metrics and robust stability for nonlinear systems. SIAM Journal of Control and Optimization, 44 (2005) 418-443

[3] French M. and Bian W., A biased approach to nonlinear robust stability with applications, Proceeding of 48th IEEE CDC & 28th CCC, Shanghai, 2009, 1393-1398

[4] Fan, X. Arcak M. and Wen J. T., Robustness of network flow control against disturbance and time-delay. Systems & Control Letters 153 (2004) 13-29

[5] French M., Ilchmann A. and Mueller M., Robust stabilization by linear output delay feedback. SIAM J. Control Optim., 48(4): 2533-2561, 2009

[6] Georgiou T., On the computation of the gap metric, Systems & Control Letters, 11(1988), 253–257

[7] Georgiou T. and Smith M. C., Robustness analysis of nonlinear feedback systems: an input-output approach, IEEE Transaction on Automatic Control, 42(1997) 1200–1221

[8] Georgiou T. and Smith M. C., Biased norms and robustness analysis for nonlinear feedback systems, Proceeding of 36th conf. on Decision and Control, San Diego, 1997, 642-643

[9] Han H., Shakkottai S., Hollot C., Srikant R. and Towsley D., Overlay TCP for multipath routing and congestion control. In ENS-INRIA ARC-TCP Workshop, Paris, France, 2003.

[10] James M.R., Smith M.C. and Vinnicombe G., Gap metrics, representations, and nonlinear robust stability. SIAM Journal of Control and Optimization, 43 (2005) 1535-1582

[11] Johari R. and Tan D., End-to-end congestion control for the internet: delays and stability, IEEE/ACM, Transaction on Networking, 9(2001) 818-832

[12] Kelly F. P., Charging and rate control for elastic traffic. European Transactions on Telecommunications, 8 (1997) 33-37

[13] Kelly F. P., Fairness and stability of end-to-end congestion control. European Journal of Control, 9 (2003) 159-176.

[14] Kelly F. P., Models for a self-managed internet. Philosophical Transactions of the Royal Society, A358 (2000) 2335-2348

[15] Kelly F. P., Maulloo A. and Tan D., Rate control in communication networks: shadow prices, proportional fairness and stability. Journal of the Operational Research Society, 49 (1998) 237-252

[16] Kelly F. P. and Voice T., Stability of end-to-end algorithms for joint routing and rate control. Computer Communication Review, 35:2 (2005) 5-12.

[17] Kelly F. P., Mathematical modelling of the Internet. In: Engquist B, Schmid W (eds). Mathematics Unlimited 2001 and Beyond. Springer- Verlag, Berlin. 2001, pp 685-702.

[18] Lestas I. and Vinnicombe G., Scalable decentralized robust stability certificates for networks of interconnected heterogeneous dynamical systems. IEEE Transactions on Automatic Control, 51 (2006) 1613- 1625

[19] Low S. H. and Lapsley D. E., Optimization flow control - I: basic algorithm and convergence. IEEE Transactions on Networking, 7 (1999) 861- 874

[20] Massoulié L., Stability of distributed congesti on control with heterogeneous feedback delays, IEEE Transaction on Automatic Control, 46(2002) 895-902

[21] Paganini F., Doyle J. and Low S., Scalability for stable network congestion control, Proceeding of 40th conf. on Decision and Control, Orlando, 2001, 185-190

[22] Srikant R., The Mathematics of Internet Congestion Control. Birkhauser, 2003.

[23] Varian H. R., Microenonomic Analysis. Third edition, Norton: New York. 1992

[24] Vinnicombe G., On the stability of end-to-end congestion control for the internet, University of Cambridge, Tech. Rep. CUED/F-INFENG/TR. 398, 2000.

6.1 Appendix

The following is a generalized Gronwall inequality.

Lemma 6.1. *[1] Let h_1, h_2 be continuous function, $\phi \in L^\infty(\mathbb{R}_+, \mathbb{R}_+)$, $k \geq 0$ and suppose:*

$$\phi^2(t) \leq k^2 + 2 \int_0^t \big(h_1(s)\phi(s) + h_2(s)\phi^2(s)\big)ds.$$

Then

$$\phi(t) \leq \int_0^t h_1(s) \exp\left(\int_s^t h_2(\nu)d\nu\right) ds + k \exp\left(\int_0^t h_2(s)ds\right).$$

A generalized mean value theorem is given below which will be used for the stability of congestion control algorithms involving non-differentiable function such as those in (2.15), (2.16) and (2.19). It is possibly known but we present it with proof.

Lemma 6.2 (Mean Value Theorem). *Suppose that $h : [a, b] \to \mathbb{R}$ is continuous and both left and right differentiable on (a, b). Then there exists $x_0 \in (a, b)$ such that*

$$\frac{h(b) - h(a)}{b - a} \in [h'_-(x_0), h'_+(x_0)]. \tag{6.1}$$

Here and after, $[h'_-(x_0), h'_+(x_0)] = \{(1 - t)h'_-(x_0) + th'_+(x_0) : t \in [0, 1]\}$.

Proof. Let $\tilde{h}(x) = h(x) - \frac{h(b)-h(a)}{b-a}x$. Then $\tilde{h}(a) = h_1(b)$, $\tilde{h}$ is continuous on $[a, b]$, both left and right differentiable on (a, b), and

$$\tilde{h}'_-(x) = h'_-(x) - \frac{h(b) - h(a)}{b - a}, \quad \tilde{h}'_+(x) = h'_+(x) - \frac{h(b) - h(a)}{b - a}. \tag{6.2}$$

If $\tilde{h}$ is constant, then both $\tilde{h}$ and h are differentiable and, for any $x_0 \in (a, b)$,

$$0 = \tilde{h}'(x_0) = h'(x_0) - \frac{h(b) - h(a)}{b - a}$$

which proves the claim.

If $\tilde{h}$ is not constant, then $\tilde{h}$ must has local maximum or minimum at some $x_0 \in (a, b)$. Without loss of generality, suppose it is a minimum. Then h_1 decreases on the left of x_0 and increases on the right. This shows $\tilde{h}'_-(x_0) \leq 0 \leq \tilde{h}'_+(x_0)$. Together with (6.2), we obtain (6.1). $\square$

For differentiable functions, this lemma is the classical mean value theorem. The lemma also shows that if a function h is both left and right differentiable and all derivatives are finite, then h is locally Lipschitz.

Review commentary

Associate Editor (Remarks to Author):

In this second round of review, some additional comments are raised, including the complexity of the application, and static networks. Hence the authors are encouraged to revise their paper to satisfy the comments of the Reviewer 2. See below.

Referee 1 (Remarks to the Author):

My previous concerns are fully addressed and I do not have further questions.

Referee 2 (Remarks to the Author):

This paper presents stability results for algorithms for network traffic control. Robustness is investigated for externa disturbances and time delays. The comments on this paper are as follows:

1) The paper, in general, is well-written and appears to be devoid of any technical issues.

2) However, since the application sought is the control of information flow on the internet, the example given does not justify the complexity of the application. The authors discuss the case of higher nodes but a representative internet network will also have far higher number of branches. Overall, the example provided gives a sense that the application of the results could be limited. However, that does not undermine the theoretical contribution of the paper, which appears to be sound.

We have added further discussion on the applicability to larger scale networks at the end of Example 5.2.

3) The internet network is also dynamic. That is new nodes and branches are continuously plugging in or out of the network. This paper appears to consider static networks. A dynamic network could lead to jumps in the dimensions of the system. It would be curious to know what inplication that would have on network stability. Can the existing methods be used or extended to the dynamic case?

The reviewer raises an important issue concerning the direct applicability of the underlying Kelly framework to the case of dynamic networks. Although well outside the scope of this contribution we believe it is reasonable to speculate that the approach can be extended to classes of dynamic networks whereby essentially the sum of the variations of the network topology are suitably constrained, for example if the matrix A is time varying and denoted by A_t and $\|\sum_t A_t(\sum_t A_t^\top)\| \leq M < \infty$. More fundamentally however, where such conditions are violated, but the time variation is suitably slow, one might also reasonably expect this analysis to form part of a more general consideration. A combination of the two approaches may very well be sufficient to model realistic dynamic variations (that is fast changes which do not violate connectivity assumptions, and slow variations which do). Note that these are entirely speculative comments, but by which we indicate that there are strong possibilities for moving towards the challenging reality whilst building from this base.