2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Krasovskii's Method in the Stability of Network Control

Diego Feijer and Fernando Paganini

Abstract— **We consider network resource allocation problems based on convex optimization, and their decentralized solutions by means of primal, dual, or primal-dual subgradient control. We show how Krasovskii's method, that seeks Lyapunov functions which are quadratic forms of the vector field, provides new global stability proofs for various problems of this kind. Applications include congestion control, cross-layer congestion and contention control, and other general network utility maximization problems. We show more generally how this proof method applies to concave-convex saddle point problems solved by subgradient methods.**

I. Introduction

Since the seminal work of Kelly in [4], it has become standard to study network resource allocation in the following terms: (i) a global, convex optimization objective, usually in terms of network utility maximization; (ii) decentralized, subgradient algorithms for its solution over a network, based on primal and/or dual decompositions; (iii) global convergence proofs of these algorithms to the optimum, usually based on Lyapunov techniques. For the congestion control problem, an excellent summary of many contributions with the above features is [14]. More recently, generalizations of this strategy to multiple layers of network protocols have been pursued, as discussed in the surveys [3], [9].

Finding a Lyapunov function to establish global stability is sometimes straightforward, when the dynamics are a gradient law for a single set of variables: this happens for primal congestion control laws [4], where rates follow the gradient of a modified objective function (utility minus a congestion cost), for dual laws that follow the gradient of the dual [10], and also for routing problems that follow the negative gradient of a congestion cost (see [13]). In these cases, the objective itself must be monotonic along trajectories. When multiple sets of variables are simultaneously controlled (primal-dual, or cross layer problems), finding a Lyapunov function is less obvious: for instance, primal-dual laws are seeking a saddle point of the Lagrangian, around which it is not sign definite; the Lagrangian cannot be a Lyapunov function. Proofs in these cases use quadratic Lyapunov functions (see [14]), motivated in some cases by a *passivity* decomposition [15]. However, such decompositions become less obvious for cross-layer problems with multiple optimization variables.

In this paper we pursue another source of Lyapunov proofs for such problems: Krasovskii's method [6], [5], that applies to the system $\dot{z} = F(z)$ a Lyapunov function

$$V(z) = \dot{z}^T Q \dot{z} = F(z)^T Q F(z).$$

Diego Feijer and Fernando Paganini are with the Department of Electrical Engineering, Universidad ORT Uruguay, Cuareim 1451, 11.100 Montevideo (e-mails: {feijer,paganini}@ort.edu.uy). Research supported by AFOSR-US, and by ANII-Uruguay, grant FCE 2007_265.

If $Q > 0$ is found such that $\left(\frac{\partial F}{\partial z}\right)^T Q + Q\left(\frac{\partial F}{\partial z}\right)$ is negative semidefinite at every point z, then $\dot{V} \leq 0$, which provides the essential step for a Lyapunov proof, possibly supplemented by the use of LaSalle's invariance principle (see [5]).

In Section II we apply this method to give new stability proofs in primal, dual, and primal-dual congestion control algorithms. It is in the last case where the method is most interesting, exploiting a particular symmetry in the equations to lead to a natural candidate for Q. This motivates in Section III the generalization to primal-dual methods for cross-layer problems. One such problem is joint congestion and contention control for wireless networks, introduced in [7], [8], where convexity is achieved by a logarithmic change of variables. [8] proposes primal algorithms for this problem, and [16] studied a dual version. We are mostly interested in the primal-dual approach, first studied in [17] with stochastic models. Here, using fluid models, we embed the primal-dual case in a class of problems for which Krasovskii's method provides a Lyapunov stability proof.

In Section IV, we see that the symmetry principle in our primal-dual construction appears more generally in any gradient method for a saddle point problem, as studied classically in [1]; thus, the Krasovskii method also covers this situation. Conclusions and future work are discussed in Section V. The Appendix contains some technical proofs.

II. Network Congestion Control

We work here with the standard setup for congestion control as in [14]. The network is made of a set of sources $\mathcal{S}$ and a set of links $\mathcal{L}$, of cardinality n and m, respectively. Each source s injects packets into the network at a rate x_s, and has an associated strictly concave utility function $U_s(x_s)$. These packets use a subset $\mathcal{L}(s) \subseteq \mathcal{L}$ of links. On the other hand, each link l transports traffic of a subset of sources $\mathcal{S}(l) \subseteq \mathcal{S}$, with total rate $y_l = \sum_{s \in \mathcal{S}(l)} x_s$. The link capacity is c_l.

We can summarize the relationship between the vectors of source and link rates through $y = Rx$, where the routing matrix R of dimension $m \times n$, is defined by:

$$R_{ls} = \begin{cases} 1 & \text{if } s \in \mathcal{S}(l), \\ 0 & \text{otherwise.} \end{cases}$$

The resource allocation proposed by Kelly [4] is the solution of the following network utility maximization.

Problem 1: Maximize $\sum_{s \in \mathcal{S}} U_s(x_s)$, over $x_s \geq 0$, subject to the capacity constraints

$$y_l \leq c_l, \qquad l \in \mathcal{L}. \tag{1}$$

A. Primal Congestion Control

A *primal* algorithm for the resource allocation involves a dynamic law for the source rates x_s. In order to enforce (approximately) the capacity constraints, a *price* signal $\lambda_l = \varphi_l(y_l)$ is introduced, through an increasing penalty function $\varphi_l(\cdot)$ for the constraint $y_l \leq c_l$. The total price per source

$$q_s = \sum_{l \in \mathcal{L}(s)} \lambda_l \tag{2}$$

acts as a penalty term for the control of the source rates,

$$\dot{x_s} = \kappa_s [U_s'(x_s) - q_s]^+_{x_s}. \tag{3}$$

Here $\kappa_s > 0$, and the positive projection is defined by

$$[w]^+_z := \begin{cases} w, & \text{if } w > 0 \text{ or } z > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The projection is said to be active if the second case applies. The above algorithm does not solve Problem 1 exactly, but rather the barrier approximation

$$\max_{x \geq 0} \left[\sum_{s \in \mathcal{S}} U_s(x_s) - \sum_{l \in \mathcal{L}} \int_0^{y_l} \varphi_l(y) dy \right]. \tag{5}$$

In fact, $\dot{x}$ in (3) is a subgradient of the modified objective in (5), which makes it a natural Lyapunov function for proving convergence, as indeed was done in [4], [14]. Here, as a first simple example of the use of Krasovskii's method, we wish to consider the alternate function

$$V(x) = \frac{1}{2} \dot{x}^T K^{-1} \dot{x} = \sum_{s \in \mathcal{S}} \frac{\dot{x}_s^2}{2\kappa_s}, \tag{6}$$

where $K = \text{diag}(\kappa_s)$. This function can be discontinuous at switches of the projection operation; we will show later how these can be treated. For now, we avoid this issue, differentiating V at a point with $x_s > 0$ to get

$$\begin{aligned} \dot{V}(x) &= \sum_{s \in \mathcal{S}} \frac{\dot{x}_s \ddot{x}_s}{\kappa_s} \\ &= \sum_{s \in \mathcal{S}} \dot{x}_s \left(U_s''(x_s)\dot{x}_s - \sum_{l \in \mathcal{L}(s)} \varphi_l'(y_l)\dot{y}_l \right) \\ &= \dot{x}^T \text{diag}\{U_s''(x_s)\}\dot{x} - \dot{x}^T R^T \text{diag}\{\varphi_l'(y_l)\} R\dot{x}. \end{aligned}$$

Since U_s is strictly concave for all $s \in \mathcal{S}$, and φ_l is non decreasing for all $l \in \mathcal{L}$, we see that $\dot{V}(x) \leq 0$. A LaSalle argument can complete the asymptotic stability proof, analogous to others covered later in this paper.

B. Dual Congestion Control

Problem 1 can be solved exactly by using duality: let $\lambda = (\lambda_l)_{l \in \mathcal{L}}$ be the Lagrange multipliers associated with the capacity constraints (1). The Lagrangian is

$$L(x, \lambda) = \sum_{s \in \mathcal{S}} U_s(x_s) - \sum_{l \in \mathcal{L}} \lambda_l (y_l - c_l).$$

Strong duality implies that Problem 1 is equivalent to the dual optimization $\min_{\lambda \geq 0} D(\lambda)$, where the Lagrange dual function $D : \mathbb{R}^m_+ \to \mathbb{R}$ is

$$D(\lambda) = \max_{x \geq 0} L(x, \lambda). \tag{7}$$

The condition for x to be the maximizer of (7) is

$$x_s = U_s'^{-1}(q_s), \quad \text{or } x_s = 0 \text{ and } U_s'(0) < q_s, \tag{8}$$

where q_s is defined as in (2), for these new prices λ. The above condition defines a *decreasing* "demand curve" $x_s = f_s(q_s)$; we make the standing assumption

$$\lim_{q_s \to \infty} f_s(q_s) = 0. \tag{9}$$

We denote by $\hat{x}$ and $\hat{\lambda}$ the primal-dual optimal points, that satisfy the Karush-Kuhn-Tucker conditions for this problem: (8) and the complementary slackness condition

$$\hat{\lambda}_l \begin{cases} = 0 & \hat{y}_l < c_l, \\ \geq 0 & \hat{y}_l = c_l. \end{cases} \tag{10}$$

Dual flow control consists of a static update of source rates following (8), and a dynamic control of the prices,

$$\dot{\lambda}_l = \gamma_l [y_l - c_l]^+_{\lambda_l}. \tag{11}$$

Here $\gamma_l > 0$, and the positive projection is defined as in (4). The above price dynamics amounts to a subgradient algorithm for the dual function, indeed

$$y - c = -\frac{\partial D}{\partial \lambda}.$$

This observation makes $D(\lambda)$ a natural Lyapunov function for these dynamics, see e.g. [13]. Again, we would like to look at the alternative of Krasovskii's method, so consider the Lyapunov candidate (here $\Gamma = \text{diag}(\gamma_l)$)

$$V(\lambda) = \frac{1}{2} \dot{\lambda}^T \Gamma^{-1} \dot{\lambda} = \sum_{l \in \mathcal{L}} \frac{\dot{\lambda}_l^2}{2\gamma_l}. \tag{12}$$

We again ignore initially the projections and compute the derivative of V at a point $\lambda > 0$:

$$\begin{aligned} \dot{V} &= \sum_{l \in \mathcal{L}} \frac{\dot{\lambda}_l \ddot{\lambda}_l}{\gamma_l} = \dot{\lambda}^T R \dot{x} = \dot{\lambda}^T R \,\text{diag}\left\{ f_s'(q_s) \right\} \dot{q} \\ &= \dot{q}^T \text{diag}\left\{ f_s'(q_s) \right\} \dot{q} \leq 0, \end{aligned} \tag{13}$$

since the demand curve f_s is decreasing.

We now treat more carefully the projections, that make the dynamics (11) a hybrid system of the type considered in [2], [11]. Specifically, let the discrete state σ denote the subset of links for which the positive projection is active; σ has a finite set of alternatives. (11) is equivalent to

$$\dot{\lambda}_l = \begin{cases} \gamma_l (y_l - c_l), & l \notin \sigma(t); \\ 0 & l \in \sigma(t). \end{cases} \tag{14}$$

The Lyapunov function (12) takes the form

$$V(\sigma, \lambda) = \sum_{l \notin \sigma} \frac{\gamma_l}{2} (y_l - c_l)^2. \tag{15}$$

For an interval of time in which $\sigma(t) \equiv \sigma_0$ constant, the calculations in (13) are still valid: the saturated links do not appear since $\dot{\lambda}_l \equiv 0$ for $l \in \sigma_0$. Therefore V will still decrease in this interval. It remains to study the behavior of V at times when the set σ of active projections changes; here, V may be discontinuous. Notice, however:

- The set σ will be *enlarged* if a link l with $y_l(t^-) \leq c_l$ reaches $\lambda_l(t) = 0$. In that case, the sum in (15) loses a term, and so $V(t^-) \geq V(t^+)$. V is discontinuous but in the decreasing direction.
- Consider now a link l which had an active projection at t^-, and no longer has it at t^+. This happens only if $y_l - c_l$ went through zero, from negative to positive, at t. An extra term is added to (15), but this term is initially at zero. Hence there is no discontinuity of V.

Therefore, V satisfies the monotonicity conditions of a Lyapunov function. To prove asymptotic stability, a LaSalle argument is given in the Appendix.

C. Primal-Dual Congestion Control

The primal-dual control scheme uses subgradient dynamics in both primal and dual variables; i.e. , rates are updated as in (3), where q satisfies (2) with link prices generated by (11). In other words, we have

$$\dot{x} = K\left[\frac{\partial L}{\partial x}\right]_x^+ = K[U'(x) - R^T\lambda]_x^+, \tag{16}$$

$$\dot{\lambda} = \Gamma\left[-\frac{\partial L}{\partial \lambda}\right]_\lambda^+ = \Gamma[Rx - c_l]_\lambda^+. \tag{17}$$

Here $K = \mathrm{diag}(\kappa_s)$, $\Gamma = \mathrm{diag}(\gamma_l)$ and $U'(x)$ is the column vector of $U_s'(x_s)$. In this case the Lyapunov choice is less obvious, since the dynamics seek a *saddle* point, not an extremum, of the Lagrangian. The first global stability proof for this law in the network control literature was the passivity argument in [15], which results in a quadratic Lyapunov function not directly related to the primal or dual objectives[1].

However, when using Krasovskii's method, the primal-dual case is a natural combination of the previous cases. Superimposing the functions in (6) and (12), introduce the following Lyapunov function in the state $z = (x, \lambda)^T$:

$$V(z) = \dot{z}^T Q \dot{z}, \quad \text{with } Q = \frac{1}{2}\begin{bmatrix} K^{-1} & 0 \\ 0 & \Gamma^{-1} \end{bmatrix}. \tag{18}$$

Consider a point with $x > 0$ and $\lambda > 0$, so projections are inactive. The dynamics (16)-(17) have the form $\dot{z} = F(z)$, with the derivative of the vector field $F(z)$ satisfying

$$Q\left(\frac{\partial F}{\partial z}\right) = \frac{1}{2}\begin{bmatrix} \mathrm{diag}\{U_s''(x_s)\} & -R^T \\ R & 0 \end{bmatrix}.$$

The antisymmetric structure of the above matrix has its roots in the fact that (16)-(17) are derived from gradients of the same Lagrangian: as we will see later in more generality, we are exploiting the identity

$$\frac{\partial^2 L}{\partial x_s \partial \lambda_l} = \frac{\partial^2 L}{\partial \lambda_l \partial x_s}.$$

[1]In Section IV we see that the quadratic Lyapunov proof also follows as a special case of the classical study [1].

As a consequence of this symmetry we have

$$\left(\frac{\partial F}{\partial z}\right)^T Q + Q\left(\frac{\partial F}{\partial z}\right) = \begin{bmatrix} \mathrm{diag}\{U_s''(x_s)\} & 0 \\ 0 & 0 \end{bmatrix},$$

negative semidefinite; so again we have $\dot{V} \leq 0$, the basis for a Lyapunov argument. The details of the proof are relayed to Theorem 1 below, which generalizes primal-dual laws to a larger class of problems.

III. Generalizations and Cross-Layer Optimization

As a natural continuation of the congestion control studies, recent research has incorporated into the optimization framework other resources to be allocated in networks, involving other layers of protocols. In particular, many researchers have tackled the control of routing, medium access control (MAC), or physical layer control, most prominently for the case of wireless networks. Recent surveys which cover part of the substantial literature are [9], [3]. In particular, [3] advocates the network utility maximization problem (NUM), and its decompositions, as the unifying paradigm over which network architectures can be designed and controlled.

We now formulate one such NUM problem, which generalizes Problem 1.

Problem 2: Maximize $\sum_{s\in\mathcal{S}} U_s(x_s)$, subject to

$$h_l(x) \leq c_l(p), \quad \forall l \in \mathcal{L},$$

where $\{h_l\}$ and $\{c_l\}$ are twice continuously differentiable convex and concave functions, respectively.

Through the variable $p \in \mathcal{P} \subseteq \mathbb{R}^k$ above, we can represent the dependence of link capacity on parameters of the lower layers, e.g. the MAC protocols or physical layer parameters. Letting x enter the constraint in a possibly nonlinear (convex) way, provides more generality for the choice of primal variables, other than link rates themselves. An example is described in Section III-A.

We introduce the Lagrangian for this problem,

$$L(x, p, \lambda) = \sum_{s\in\mathcal{S}} U_s(x_s) - \sum_{l\in\mathcal{L}} \lambda_l(h_l(x) - c_l(p)). \tag{19}$$

The dual problem will be $\min_{\lambda\geq 0} D(\lambda)$, where the Lagrange dual function can be written as

$$D(\lambda) = \max_x \left\{\sum_{s\in\mathcal{S}} U_s(x_s) - \lambda^T h(x)\right\} + \max_p \lambda^T c(p). \tag{20}$$

Note the "dual decomposition" (between variables x and p) in the solution of (20). Below, we will study two control algorithms to solve the dual problem, with stability proofs using the Krasovskii method.

A. Cross-Layer Congestion and Contention Control

In this section we motivate the formulation in Problem 2 through the following application, which originates in [7], [8]. Consider an ad-hoc wireless network made up of a set of nodes, which use a random MAC. Each node n accesses the medium with probability P^n, and when transmitting it chooses one of its outgoing links with probability p_l. Thus, we have the following convex constraints (which define the set $\mathcal{P}$) for every node:

$$p_l \geq 0, \qquad \sum_{l \in \mathcal{L}_{out}(n)} p_l = P^n, \quad P^n \leq 1. \tag{21}$$

The transmission of a link is interfered by another link, if the receiver of the former is within range of the transmitter of the latter; when such a collision occurs, we assume no useful information is transmitted. So the capacity of a link l depends on the probability of accessing the channel without presence of the interfering nodes; we write

$$C_l(p) := c_l p_l \prod_{k \in \mathcal{N}_I(l)} (1 - P^k),$$

where $\mathcal{N}_I(l)$ is the set of nodes that interfere with link l.

Under this additional degree of freedom, we wish to allocate network resources (rates and transmission probabilities) to optimize an overall network utility. The optimal congestion and contention control problem is thus

$$\begin{aligned} \Xi: \quad & \text{maximize} \quad \sum_{s \in \mathcal{S}} U_s(x_s) \\ & \text{subject to (21) and} \quad \sum_{s \in \mathcal{S}(l)} x_s \leq C_l(p), \quad \forall l. \end{aligned} \tag{22}$$

The difficulty with the above problem is that the last constraint is non-convex; nevertheless, a change of variables was proposed in [7] that helps tackle this problem. Take the logarithm on both sides of inequality (22), and define $\tilde{x}_s := \log x_s$, $\tilde{U}_s(\tilde{x}_s) := U_s\left(e^{\tilde{x}_s}\right)$. Then problem Ξ is equivalent to

$$\begin{aligned} \tilde{\Xi}: \quad & \text{maximize} \quad \sum_{s \in \mathcal{S}} \tilde{U}_s(\tilde{x}_s) \\ & \text{subject to (21) and} \quad \log\left(\sum_{s \in \mathcal{S}(l)} e^{\tilde{x}_s}\right) \leq \tilde{C}_l(p), \quad \forall l \end{aligned} \tag{23}$$

where $\tilde{C}_l(p) := \log c_l + \log p_l + \sum_{k \in \mathcal{N}_I(l)} \log(1 - P^k)$.

Note that $\tilde{C}_l$ is concave in the variables p_l, P^k, and the left hand side of (23) is convex in $\tilde{x}_s$. Therefore, the new formulation falls in the class of Problem 2, provided that the new objective function is concave in $\tilde{x}$: this happens (see [7]) for utility functions satisfying

$$\frac{d^2 U_s(x_s)}{dx_s^2} x_s + \frac{dU_s(x_s)}{dx_s} \leq 0. \tag{24}$$

The remaining challenge for formulation $\tilde{\Xi}$ is to find a distributed network solution. In this regard, note that the left hand side of (23) is not separable between primal variables. In a dual approach that would solve the problem exactly, the optimization over x in (20) is not easy to decentralize. [8] tried for this purpose a primal-based approach, where the gradient direction of x can be found in a distributed way, but is forced then to approximate $\tilde{\Xi}$ by a barrier function. Recently, [16] formulated a dual approach which achieves decentralization through additional variables per source and link, not a very scalable proposition. The separability concerns disappear if one uses a primal-dual approach; this was first studied in [17], where the focus is on stochastic issues associated with estimation of subgradients; a stability result was given involving quadratic Lyapunov functions in discrete time.

In this paper we will obtain, as a special case of Theorem 1 below, a global stability result for primal-dual laws for $\tilde{\Xi}$, based on deterministic fluid models in continuous time, and Krasovksii's method.

B. Primal-Dual control for Problem 2 and its stability: version with dual update for p.

We discuss here a solution of Problem 2 based on a primal-dual approach. Lagrange multipliers λ will be dynamically updated to follow a subgradient of the dual function, and the primal variables x will also use gradient control laws. The question arises as to how to deal with the primal variables p, for which in principle we could use either choice: solve directly for the optimal $p(\lambda)$ in (20), or a gradient approach. In this section we study the former option, attractive for problems where this calculation can be done in a *separable* way, like the one in Section III-A.

Specifically, define $\psi(\lambda) := \max_{p \in \mathcal{P}} \lambda^T c(p)$, and denote by $\bar{p}(\lambda)$ the maximizing p. The function $\psi(\lambda)$ is *convex* in λ (maximum of linear functions).

In the particular case of Section III-A, we have

$$\psi(\lambda) = \sum_{n \in \mathcal{N}} \left[\sum_{l \in \mathcal{L}(n)} \lambda_l \log \bar{p}_l + \sum_{k \in \mathcal{L}_I(n)} \lambda_k \log(1 - \bar{P}^n) \right],$$

$$\text{where} \quad \bar{p}_l(\lambda) := \frac{\lambda_l}{\sum_{l \in \mathcal{L}_{out}(n)} \lambda_l + \sum_{l \in \mathcal{L}_I(n)} \lambda_l}.$$

Note that in this case the optimal $\bar{p}_l(\lambda)$ depends only Lagrange multipliers of nodes neighboring link l; for this reason the calculation is amenable for distributed computation with message passing between nodes.

Returning now to the general case, define

$$\begin{aligned} \bar{L}(x, \lambda) &= \max_p L(x, p, \lambda) \\ &= \sum_{s \in \mathcal{S}} U_s(x_s) - \sum_{l \in \mathcal{L}} \lambda_l h_l(x) + \psi(\lambda). \end{aligned}$$

Our primal-dual computation for Problem 2 will be based on gradients of this function:

$$\dot{x} = K \left[\frac{\partial \bar{L}}{\partial x}\right]_x^+ = K \left[U'(x) - \left(\frac{\partial h}{\partial x}\right)^T \lambda\right]_x^+, \tag{25}$$

$$\begin{aligned} \dot{\lambda} &= \Gamma \left[-\frac{\partial \bar{L}}{\partial \lambda}\right]_\lambda^+ = \Gamma \left[h(x) - \nabla \psi^T(\lambda)\right]_\lambda^+ \\ &= \Gamma \left[h(x) - c(\bar{p}(\lambda))\right]_\lambda^+, \end{aligned} \tag{26}$$

where the second equality in (26) follows from the Envelope Theorem with $\bar{p}(\lambda) = \arg\max_p \lambda^T c(p)$. We state the following result.

Theorem 1: The optimum of Problem 2 is a globally asymptotically stable equilibrium of the dynamics (25)-(26).

Proof: Consider the Lyapunov function defined in (18). Operating again initially at a point $z > 0$ where all projections are inactive, we have now

$$Q\left(\frac{\partial F}{\partial z}\right) = \frac{1}{2}\begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & -\left(\frac{\partial h}{\partial x}\right)^T \\ \left(\frac{\partial h}{\partial x}\right) & -\nabla^2\psi(\lambda) \end{bmatrix},$$

where

$$\frac{\partial^2 L}{\partial x^2} = \mathrm{diag}\{U_s''(x_s)\} - \sum_{l\in\mathcal{L}} \lambda_l \nabla^2 h_l(x). \tag{27}$$

Note from the strict concavity of $U_s(x_s)$ and the convexity of $h_l(x)$, that $\frac{\partial^2 L}{\partial x^2}$ is negative definite. Also, $-\nabla^2\psi(\lambda)$ is negative semidefinite because of the convexity of $\psi(\lambda)$. Therefore

$$\left(\frac{\partial F}{\partial z}\right)^T Q + Q\left(\frac{\partial F}{\partial z}\right) = \begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & 0 \\ 0 & -\nabla^2\psi(\lambda) \end{bmatrix} \leq 0, \tag{28}$$

hence $\dot{V} \leq 0$ along trajectories, as desired. The remainder of the argument, which considers the projections and LaSalle invariance, is covered in the Appendix. ■

C. Primal-Dual control for Problem 2 and its stability: version with primal update for p.

An alternative primal-dual law for Problem 2 involves a gradient update of the primal variable p: this could be suitable for a situation where explicit formulas for $\bar{p}(\lambda)$ are not available. Returning to the original Lagrangian (19), the control law would be

$$\dot{x} = K\left[\frac{\partial L}{\partial x}\right]_x^+ = K\left[U'(x) - \left(\frac{\partial h}{\partial x}\right)^T \lambda\right]_x^+, \tag{29}$$

$$\dot{p} = \Upsilon\left[\frac{\partial L}{\partial p}\right]_{\mathcal{P}} = \Upsilon\left[\lambda^T \frac{\partial c}{\partial p}\right]_{\mathcal{P}}, \tag{30}$$

$$\dot{\lambda} = \Gamma\left[-\frac{\partial L}{\partial \lambda}\right]_\lambda^+ = \Gamma\left[h(x) - c(p)\right]_\lambda^+, \tag{31}$$

where $\Upsilon = \mathrm{diag}\{\epsilon_k\}$, $\epsilon_k > 0$, and where $[\cdot]_{\mathcal{P}}$ denotes the projection on to $\mathcal{P}$.

Theorem 2: Suppose that $c(p)$ is a strictly concave function. Then the optimum of Problem 2 is a globally asymptotically stable equilibrium of the dynamics (29)-(31).

The proof is omitted due to space limitations; we will only indicate the relevant Lyapunov function based on Krasovskii's method,

$$V(z) = \dot{z}^T Q \dot{z}, \quad Q = \frac{1}{2}\begin{bmatrix} K^{-1} & 0 & 0 \\ 0 & \Upsilon^{-1} & 0 \\ 0 & 0 & \Gamma^{-1} \end{bmatrix}.$$

IV. Krasovskii's method and gradient laws for saddle points

All primal-dual laws considered have the general form

$$\dot{x} = K\left[\frac{\partial L}{\partial x}\right], \qquad \dot{\lambda} = \Gamma\left[-\frac{\partial L}{\partial \lambda}\right]; \tag{32}$$

they are thus gradient laws which seek a saddle point (maximum in x, minimum in λ), of a certain function $L(x,\lambda)$. Here for simplicity we have removed projections.

Such algorithms were studied in the classical work of Arrow, Hurwicz and Uzawa [1]; here it was proved, using a quadratic Lyapunov function and in essence an invariance argument (before LaSalle!), that the saddle is achieved for $L(x,\lambda)$ strictly concave in x and convex in λ. It is noteworthy that in the congestion control literature, such proofs were essentially rediscovered based on passivity arguments [15].

In this paper we have found that the Krasovskii method provides an alternate proof for these problems. Indeed, the Lyapunov function $\dot{z}^T Q \dot{z}$ in (18) with $z = (x,\lambda)^T$, satisfies for the dynamics (32) the condition

$$Q\left(\frac{\partial F}{\partial z}\right) = \frac{1}{2}\begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial \lambda} \\ -\frac{\partial^2 L}{\partial \lambda \partial x} & -\frac{\partial^2 L}{\partial \lambda^2} \end{bmatrix}.$$

As before, when computing the Lyapunov derivative, the terms off the block diagonal will cancel, rendering

$$\left[\left(\frac{\partial F}{\partial z}\right)^T Q + Q\left(\frac{\partial F}{\partial z}\right)\right] = \begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & 0 \\ 0 & -\frac{\partial^2 L}{\partial \lambda^2} \end{bmatrix} \leq 0,$$

for $L(x,\lambda)$ concave in x and convex in λ. Thus $\dot{V} \leq 0$, and a Lasalle argument follows if there is strict concavity in x.

V. Conclusions

We studied network optimization problems from congestion control and extensions to cross-layer optimization. We found a new technique for proving global stability results in a variety of such problems, invoking Krasovskii's method for the choice of Lyapunov functions. This covers primal and dual algorithms, but is particularly suited for the primal-dual case, in which a certain symmetry of the gradient equations allows for a simplification in the Lyapunov derivative. Furthermore, we have seen that the method extends to gradient dynamics for saddle-point problems as studied classically in [1]. A natural future question is the applicability of this proof method to situations where strict concavity is relaxed.

Appendix: stability proofs based on LaSalle Invariance in switched Systems

Dual congestion control

We have shown that the function $V(\sigma,\lambda)$ in (15) is decreasing along trajectories of the system. We wish to invoke Theorem IV.1 in [11], which establishes a LaSalle invariance principle for switched systems of this kind.

Invoking [11], the dynamics must converge to an invariant set inside the set of trajectories that at all times satisfy either (i): σ fixed and $\frac{d}{dt}V(\sigma,\lambda(t)) \equiv 0$; or (ii): σ switches at time t between σ^- and σ^+, but $V(\sigma^-,\lambda(t)) = V(\sigma^+,\lambda(t))$. This implies $V(t)$ is constant for all time.

Focusing on case (i), we see from (13) that $\dot{q}_s \equiv 0$ for all s where $f_s'(q_s) < 0$; this always happens unless the projection in (8) occurs, in which case $x_s \equiv 0$. In any case the rates x_s are constant while (i) holds, and so are the link rates y_l. If there is a switch as in case (ii), referring to expression (15) and the discussion after it, we see that the rate y_l cannot change either during this switch. Therefore our invariant trajectory satisfies $y_l \equiv y_{l0}$, constant for all time.

If $y_{l0} < c_l$ for some l, the corresponding λ_l must be at zero with the projection active; otherwise, it would fall linearly and lead to a discontinuous switch in V, violating (ii) above. If $y_{l0} > c_l$, λ_l would tend to infinity, but this means through assumption (9) that all sources using this link would have rate going to zero, contradicting the constancy of y_l. Therefore, the only other option is $y_{l0} = c_l$ for some link with $\lambda_l \geq 0$. We see that these are precisely the KKT conditions for the problem, thus the invariant trajectory is at the optimum.

Proof of Theorem 1

The primal-dual dynamics (25)-(26) have projections in x and in λ; let these be represented by the discrete state $\sigma = (\sigma_x, \sigma_\lambda)$. The Lyapunov function from (18) is

$$V(\sigma, z) = \sum_{s \notin \sigma_x} \frac{\kappa_s}{2} \left[U_s'(x_s) - \sum_{l \in \mathcal{L}} \lambda_l \frac{\partial h_l}{\partial x_s} \right]^2 + \sum_{l \notin \sigma_\lambda} \frac{\gamma_l}{2} \left[h_l(x) - c_l(\bar{p}(\lambda)) \right]^2$$

A key observation is that the expression

$$\dot{V} = \dot{x}^T \left(\mathrm{diag}\{U_s''(x_s)\} - \sum_{l \in \mathcal{L}} \lambda_l \nabla^2 h_l(x) \right) \dot{x} - \dot{\lambda}^T \nabla^2 \psi \dot{\lambda},$$

implicit in (28), still applies on any interval of constant σ; the applicable projections in x or λ simply provide zero terms. Therefore, V is decreasing in this situation.

In the case of a projection switch, an analogous argument as the one in Section II-B applies: the only discontinuities that can arise in V are in the *decreasing* direction, when a new projection gets activated. Therefore we are still under the conditions of the LaSalle theory in [11].

It remains to characterize an invariant set within the conditions (i) or (ii) as discussed in the preceding proof for the dual. For an interval in case (i) ($\dot{V} \equiv 0$ without switching), the strict concavity assumed in $U_s(x_s)$, implies that again we will have that $\dot{x_s} \equiv 0$ for all $s \notin \sigma_x$, and trivially for the rest. Since the state variable x has no discontinuities at switching, x is constant over all time for the invariant trajectory, $x \equiv \hat{x} = \arg\max_x L(x, p, \lambda)$.

Also, from the second term in $\dot{V}$ we see that $c_l(\bar{p}(\lambda)) = \frac{\partial \psi}{\partial \lambda_l}$ must remain constant as well. So the right-hand side of

$$\dot{\lambda}_l = \gamma_l [h_l(\hat{x}) - c_l(\bar{p}(\lambda))]$$

is constant in time for an invariant trajectory. Again, if it were negative it would lead to λ_l saturating at zero after a finite time, and thus a discontinuous switch in V. So we can assume $h_l(\hat{x}) - c_l(\bar{p}(\lambda)) \geq 0$ for all l; it remains to rule out the case where it is strictly positive. For this, we note that the trajectory is moving within the optimum of (20),

$$D(\lambda(t)) = \sum_{s \in \mathcal{S}} U_s(\hat{x}_s) - \sum_{l \in \mathcal{L}} \lambda_l(t)[h_l(\hat{x}) - c_l(\bar{p}(\lambda))]. \quad (33)$$

Hence, if $h_l(\hat{x}) - c_l(\bar{p}(\lambda)) > 0$ for some l, the corresponding λ_l grows linearly in time, and $D(\lambda(t))$ is strictly decreasing. To continue we invoke the following lemma proved in [12].

Lemma: Let $\bar{\lambda}$ be a vector such that the set $\mathcal{M}_{\bar{\lambda}} = \{\lambda \geq 0 : D(\lambda) \leq D(\bar{\lambda})\}$ is nonempty. Then $\mathcal{M}_{\bar{\lambda}}$ is bounded.

From the above Lemma we observe that $\lambda(t) \in \mathcal{M}_{\lambda(0)}$ for all $t \geq 0$. But since $\mathcal{M}_{\lambda(0)}$ is a bounded set, this contradicts the linear growth of $\lambda_l(t)$.

Therefore, the invariant set has λ also at an equilibrium $\hat{\lambda}$, which by the previous argument must satisfy the complementary slackness conditions:

$$\text{either } \lambda_l = 0, \text{ or } h_l(\hat{x}) - c_l(\bar{p}(\lambda)) = 0.$$

Note also that once λ converges to $\hat{\lambda}$, p also converges to $\hat{p}(\hat{\lambda}) = \arg\max_p \hat{\lambda}^T c(p)$. Therefore, the invariant trajectory is an optimal point of Problem 2. ■

REFERENCES

[1] K. J. Arrow, L. Hurwicz and H. Uzawa, *Studies in Linear and Non-Linear Programming*, Stanford University Press, Stanford, 1958.
[2] M.S. Branicky, "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems", *IEEE Transactions on Automatic Control*, April 1998, pp. 475–482.
[3] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle. "Layering as optimization decomposition: A mathematical theory of network architectures". *Proc. of the IEEE*, vol. 95, pages 255–312, Jan 2007.
[4] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability", *Jour. Oper. Res. Society*, vol. 49(3), pp 237-252, 1998.
[5] H. K. Khalil, *Nonlinear Systems*, Prentice-Hall, 1996.
[6] N. N. Krasovskii, *Problems of the Theory of Stability of Motion*, (Russian), 1959. English translation: Stanford University Press, 1963.
[7] J-W. Lee, M. Chiang and A. R. Calderbank, "Utility-optimal random-access control", *IEEE Transactions on Wireless Communication*, vol. 6, no. 7, pp. 2741–2751, July 2007.
[8] J-W. Lee, M. Chiang and A. R. Calderbank, "Jointly optimal congestion and contention control based on network utility maximization", *IEEE Communications Letters*, vol. 10, no. 3, pp. 216-218, 2006.
[9] X. Lin, N. Shroff and R. Srikant, "A Tutorial on Cross-Layer Optimization in Wireless Networks," *IEEE Journal on Selected Areas in Communications,* August 2006, pp. 1452-1463.
[10] S. H. Low and D. E. Lapsley, "Optimization flow control, I: basic algorithm and convergence", *IEEE/ACM Transactions on Networking*, vol.7, no.6,pp861-874, December 1999.
[11] J. Lygeros, K. H. Johansson, S. N. Simic, J. Zhang and S. Sastry, "Dynamical properties of hybrid automata", *IEEE Transactions on Automatic Control*, 2003, vol. 48, pp. 2–17.
[12] A. Nedic and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods", *LIDS Technical Report 2753*, MIT, March 2007.
[13] F. Paganini, "Control and economics in networking, via optimization: perspectives from an emerging discipline", European Control Conference, Kos, Greece, July 2007.
[14] R. Srikant, *The Mathematics of Internet Congestion Control,* Birkhauser, 2004.
[15] J. T. Wen and M. Arcak. "A unifying passivity framework for network flow control", *IEEE INFOCOM*, August 2003.
[16] Y. Yu and G. B. Giannakis, "Cross-layer congestion and contention control for wireless ad hoc networks", *IEEE Transactions on Wireless Communication*, vol. 7, no. 1, January 2008.
[17] J, Zhang, D. Zheng, and M. Chiang, "The impact of stochastic noisy feedback on distributed network utility maximization", *IEEE INFOCOM 2007.*