4th IFAC Workshop on Distributed Estimation
and Control in Networked Systems
September 25-26, 2013. Rhine-Moselle-Hall, Koblenz, Germany

Decomposition of Data Rate Allocation for Stabilization over Networks

Tingyue Gan * **Michael C. Rotkowitz** **

* *Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: tgan@umd.edu).*
** *Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: mcrotk@umd.edu).*

Abstract: We consider the problem of stabilizing a network consisting of linear time-invariant plants, sensors, controllers, and relays, where the links can be rate-limited. A previous result shows how to characterize such networks for which stabilizing controllers exist, and then shows how to synthesize coding and control laws to stabilize the network. A key component is finding pseudorates, which determine how much of a given network link is used to help stabilize a given unstable mode on the network. In this paper, we seek to determine these pseudorates using only local information, and show that this can be achieved using dual decomposition for most objectives of interest. For a 1-norm objective that often allows a sparse portion of the network to be used for stabilization, we instead develop a method using techniques similar to ADMM and show that all but one step of the algorithm can be decomposed.

Keywords: Networked control, convex optimization, dual decomposition, ADMM

1. INTRODUCTION

We consider a network of linear time-invariant plants, sensors, controllers, and relays. Each sensor gets measurements from a particular plant, and can then pass that on either to a relay, or directly to a controller, which would then give an input for a particular plant. Any of these links may be rate-limited. The framework for this paper mostly follows the development in (Nair and Evans (2007)), and that paper should be consulted for more detail on any of the setup, definitions, or assumptions which may be described in less detail here.

In that work, conditions were developed which characterized when such a network can be stabilized in terms of the available data rates, the eigenvalues of the plant modes, and the network topology. This was achieved using pseudorates, which corresponded to the amount of each channel being used to stabilize a particular mode, and then using these pseudorates to synthesize coding and control laws which stabilize the network. The pseudorates associated with every possible nontrivial irreducible cycle emanating from an unstable plant mode needed to satisfy certain inequalities.

In (Rotkowitz and Nair (2008)), it was shown how, by generalizing some definitions, the conditions for stabilizability could be written as a Linear Program (LP), without any need to work out the nontrivial irreducible cycles connected to each mode. This is important because an LP can be solved with standard software, and can either find the global optimum or determine the problem to be infeasible efficiently in polynomial time.

In this paper, we seek to decompose the computation of these optimal pseudorates, to address whether, or to what extent, we can compute them based only on information that is local to the respective link of the network. This could prove useful for very large networks, or possibly for time-varying networks.

The paper is organized as follows. In Section 2, we develop the framework and some preliminaries. In Section 3, we review some of the results on the sufficient conditions on the network for stabilization to be achievable, and state a general form of the optimization problem that we address here. In Section 4, we show how the optimization problem can indeed be decomposed using dual decomposition for most cost functions; namely, for p-norms of the pseudorate vector where $p > 1$. We work through a numerical example of this in Section 5 using the 2-norm. We then develop a method to address the 1-norm in Section 6, and show that we can decompose all but one step of the algorithm, and work through a numerical example of this in Section 7.

2. PRELIMINARIES

2.1 Problem Setup

We consider the problem of stabilizing discrete-time, linear time-invariant plants over a network of directed, point-to-point, error-less, finite bit-rate digital channels.

- The network topology can be represented via a ***directed graph*** $G = (V, E)$, where V is the set of nodes and E is the set of links. A ***link*** is an ordered pair (q, r) of nodes, meaning that this link leaves node q and enters node r (link $(q, r) : q \longrightarrow r$).

978-3-902823-55-7/2013 © IFAC 10.3182/20130925-2-DE-4044.00070

- $N = |V|$ is the number of nodes and $L = |E|$ is the number of links in G.
- $V_p \subseteq V$ is the subset of plant nodes. $P = |V_p|$ is the number of plant nodes.
- Label the nodes with consecutive integers from 1 to N, and then let $V = \{1, \dots, N\}$. Without loss of generality, we label the plant nodes from 1 to P, and the other nodes from $P+1$ to N, so we have $V_p = \{1, \dots, P\}$.
- Label the links with consecutive integers from 1 to L, and then let $E = \{1, \dots, L\}$.
- As in (Nair and Evans (2007)), assume that each plant has distinct eigenvalue, thus each plant node h represents a 1-dimensional plant with dynamics matrix $\eta_h \in \mathbb{R}$, $h = 1, \dots, P$.

2.2 Some Definitions

- In (Nair and Evans (2007)), for any plant node h, an ***irreducible*** x_h***-cycle*** c is defined as any finite sequence of nodes such that
  (1) The first and last element is h, but every other element in the sequence occurs only once.
  (2) Any other plant node h', $h' \neq h$, in the sequence must be followed by a sensor node which can observe it, and preceded by a controller node which can affect it.
- For each node $h \in V$, associate a value H_h which can be interpreted as the entropy generated at that node (Rotkowitz and Nair (2008)).

$$H_h = \begin{cases} \max\{\log|\eta_h|, 0\}, & \text{if } h \in V_p \\ 0, & \text{otherwise.} \end{cases}$$

- For each pair of nodes $q, r \in V$, associate a data rate $R_{q,r}$(Rotkowitz and Nair (2008)).

$$R_{q,r} = \begin{cases} \text{average channel data rate,} & \\ & \text{if } q \text{ is a relay communicating to } r \\ & \text{if } q \text{ is a sensor communicating to } r \\ \infty, & \text{if } q \text{ is a controller affecting } r \\ & \text{if } r \text{ is sensing from } q \\ 0, & \text{otherwise.} \end{cases}$$

3. STABILIZABILITY

In this section, we review and develop conditions that allow one to construct stabilizing encoders, decoders, and control laws. In Section 3.1, we further develop some notations allowing us to restate these stabilizablity conditions in a compact form that will be useful throughout, and then, in Section 3.2, we formulate our optimization problem that we will address in the remainder of the paper.

Let $\mathcal{C}_h$ be the set of all irreducible x_h-cycles, and for any x_h-cycle $c \in \mathcal{C}_h$, let $\rho_{h,c} \geq 0$ be the pseudorate along this cycle being used to stabilize mode h. Consider the following inequalities from (Nair and Evans (2007)),

$$R_{q,r} \geq \sum_h \sum_{c \in \mathcal{C}_h : (q,r) \in c} \rho_{h,c} \qquad \forall\ (q,r) \in E \tag{1}$$

$$\sum_{c \in \mathcal{C}_h} \rho_{h,c} \geq \max\{\log|\eta_h|, 0\} \qquad \forall\ \ h \in V_p \tag{2}$$

Assume that each node performs only routing but no network coding, then this condition is necessary and almost sufficient for stabilizability.

In (Rotkowitz and Nair (2008)), a set of equivalent inequalities is developed. There they defined pseudorates $\rho_{h,q,r} \geq 0$, associated with a triple of nodes $h, q, r \in V$, to be the amount of data rate along link (q, r) being used to stabilize mode h. And with the generalized definitions of H_h and $R_{q,r}$ as in 2.2, it can be considered for any triple. Feasibility of (1) and (2) is then equivalent to the feasibility of the following linear inequalities,

$$R_{q,r} \geq \sum_h \rho_{h,q,r} \qquad \forall\ q, r \in V \tag{3}$$

$$\sum_q \rho_{h,q,r} \geq \sum_s \rho_{h,r,s} \qquad \forall\ h, r \in V \tag{4}$$

$$\sum_q \rho_{h,q,h} \geq H_h \qquad \forall\ h \in V \tag{5}$$

$$\rho_{h,i,j} \geq 0 \qquad \forall\ h, i, j \in V \tag{6}$$

3.1 Equivalent Compact Form

We have the following observations:

- In (3), if $(q, r) \notin E$, then $R_{q,r} = 0$, together with (6), we immediately have $\rho_{h,q,r} = 0$ for all $h \in V$. So we need only to solve for those $\rho_{h,q,r}$ such that $(q, r) \in E$.
- In (5), if $h \notin V_p$, then $H_h = 0$. Thus for any $(q, h) \in E$, as long as $\rho_{h,q,h} \geq 0$, i.e. (6) is satisfied, (5) is automatically satisfied. Because of (3), the data rate constraint on each link $(q, r) \in E$, and (5), the stabilizability constraint for each plant node $h \in V_p$, it would be desired to set $\rho_{h,q,r} = 0$ for any $h \notin V_p$, $(q, r) \in E$, so as to leave room for those $\rho_{h,q,r}$ such that $h \in V_p$, $(q, r) \in E$.
- Since the pseudorates travel in cycles, in order for (4) to hold for all nodes in any cycle, equality has to be enforced.

Based on the observations above, to solve the feasibility problem of (3)-(6), it suffices to solve the feasibility problem of the following set of linear equality and inequalities,

$$R_{q,r} \geq \sum_{h \in V_p} \rho_{h,q,r} \qquad \forall\ (q,r) \in E \tag{7}$$

$$\sum_{q:(q,r) \in E} \rho_{h,q,r} = \sum_{s:(r,s) \in E} \rho_{h,r,s} \quad \forall\ h \in V_p,\ \forall\ r \in V \tag{8}$$

$$\sum_{q:(q,h) \in E} \rho_{h,q,h} \geq H_h \qquad \forall\ h \in V_p \tag{9}$$

$$\rho_{h,q,r} \geq 0 \qquad \forall\ h \in V_p,\ \forall\ (q,r) \in E \tag{10}$$

With the labeling of nodes and links as in 2.1, we now let R_l be the data rate of link l, and $\rho_{h,l}$ be the pseudorate along link l being used to stabilize mode h. Equations (7)-(10) can then be rewritten as:

$$R_l \geq \sum_{h=1}^{P} \rho_{h,l} \qquad \forall \ l = 1, \ldots, L \tag{11}$$

$$\sum_{l \text{ enters } r} \rho_{h,l} = \sum_{l' \text{ leaves } r} \rho_{h,l'} \quad \forall \ h = 1, \ldots, P, \ \ r = 1, \ldots, N \tag{12}$$

$$\sum_{l \text{ enters } h} \rho_{h,l} \geq H_h \qquad \forall \ h = 1, \ldots, P \tag{13}$$

$$\rho_{h,l} \geq 0 \qquad \forall \ h = 1, \ldots, P, \ \ l = 1, \ldots, L \tag{14}$$

Next, we will further rewrite (11)-(14) in a more compact form.

For all $h = 1, \ldots, P$, let $\rho_h = [\rho_{h,1}, \ldots, \rho_{h,L}]^T \in \mathbb{R}^L$, further let $\rho = [\rho_1^T, \ldots, \rho_P^T]^T \in \mathbb{R}^{PL}$.

- The network topology can be represented by a ***node-link incidence matrix*** $\tilde{A} \in \mathbb{R}^{N \times L}$

$$\tilde{A}(n,l) = \begin{cases} -1, & \text{link } l \text{ leaves node } n \\ 1, & \text{link } l \text{ enters node } n \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

$$n = 1, \ldots, N, \ l = 1, \ldots, L$$

  Then (12) can be written as

$$\tilde{A}\rho_h = 0, \quad h = 1, \ldots, P \tag{16}$$

  Furthermore, let $A = \mathrm{diag}(\tilde{A}, \ldots, \tilde{A}) \in \mathbb{R}^{PN \times PL}$, then (12) can be compactly written as

$$A\rho = 0 \tag{17}$$

- Let $\lfloor \tilde{A} \rfloor_+$ be the projection of $\tilde{A}$ to $\mathbb{R}_+^{N \times L}$, i.e $\lfloor \tilde{A} \rfloor_+(n,l) = \max\{\tilde{A}(n,l), 0\}$, $n = 1, \ldots, N$, $l = 1, \ldots, L$. And then let $\tilde{b}_1^T, \ldots, \tilde{b}_P^T$ be the first P rows of $\lfloor \tilde{A} \rfloor_+$(recall that we label the plant nodes from 1 to P), then $\tilde{b}_h^T \rho_h$, $h = 1, \ldots, P$, is precisely the sum of pseudorates being used to stabilize mode h that enter the plant node h. Now let $B = \mathrm{diag}(\tilde{b}_1^T, \ldots, \tilde{b}_P^T) \in \mathbb{R}^{P \times PL}$, $H = [H_1, \ldots, H_P]^T$, then (13) can be compactly written as

$$B\rho \succeq H \tag{18}$$

- Let $I \in \mathbf{R}^{L \times L}$ be the identity matrix, and $C = [I, \ldots, I] \in \mathbb{R}^{L \times PL}$. Let $c_1^T, \ldots, c_L^T$ be the rows of C, then for all $l = 1, \ldots, L$, $c_l^T \rho$ is the total amount of pseudorates passing through link l. Let $R = [R_1, \ldots, R_L]^T \in \mathbb{R}^L$, then (11) can be compactly written as

$$C\rho \preceq R \tag{19}$$

Finally, (11)-(14) can be compactly written as

$$\begin{aligned} A\rho &= 0 \\ B\rho &\succeq H \\ C\rho &\preceq R \\ \rho &\succeq 0 \end{aligned} \tag{20}$$

3.2 Optimization problem formulation

Let $\phi(\rho) = \sum_{h=1}^{P} \sum_{l=1}^{L} \phi_{h,l}(\rho_{h,l})$ be a separable convex rate cost function, meant to capture in what manner we wish to keep our psuedorates small while satisfying the stabilization conditions. We can now state our ***stabilization over network optimzation problem*** as:

$$\begin{aligned} \text{minimize:} \quad & \phi(\rho) \\ \text{subject to:} \quad & A\rho = 0 && (21) \\ & B\rho \succeq H && (22) \\ & C\rho \preceq R && (23) \end{aligned}$$

with variable $\rho \in \mathbb{R}_+^{PL}$. This is a finite-dimensional convex optimization problem. We wish to see whether we can decompose the problem and solve it using only local information.

4. DUAL DECOMPOSITION METHOD

In this section, we first attempt to use dual decomposition.

4.1 Dual Problem

For (21), (22), (23), introduce dual variables $\nu \in \mathbb{R}^{PN}$, $\lambda \in \mathbb{R}_+^P$, $\gamma \in \mathbb{R}_+^L$, respectively. Furthermore, let $\nu = [{\nu_1}^T, \ldots, {\nu_P}^T]^T$, where $\nu_h = [\nu_{h,1}, \ldots, \nu_{h,N}]^T \in \mathbb{R}^N$, $h = 1, \ldots, P$. This results in the following Lagrangian,

$$\begin{aligned} & L(\rho, \nu, \lambda, \gamma) \\ &= \phi(\rho) + \nu^T(-A\rho) + \lambda^T(H - B\rho) + \gamma^T(C\rho - R) \\ &= \lambda^T H - \gamma^T R + \sum_{h=1}^{P} \sum_{l=1}^{L} [\phi_{h,l}(\rho_{h,l}) - (a_{(h-1)L+l}^T \nu)\rho_{h,l} \\ &\quad - (b_{(h-1)L+l}^T \lambda)\rho_{h,l} + (c_{(h-1)L+l}^T \gamma)\rho_{h,l}] \\ &= \lambda^T H - \gamma^T R + \sum_{h=1}^{P} \sum_{l=1}^{L} [\phi_{h,l}(\rho_{h,l}) - (\tilde{a}_l^T \nu_h)\rho_{h,l} \\ &\quad - (b_{h,l}\lambda_h)\rho_{h,l} + \gamma_l \rho_{h,l}] \end{aligned} \tag{24}$$

and the Lagrangian dual function,

$$\begin{aligned} g(\nu, \lambda, \gamma) &= \inf_{\rho \succeq 0} L(\rho, \nu, \lambda, \gamma) \\ &= \lambda^T H - \gamma^T R + \sum_{h=1}^{P} \sum_{l=1}^{L} \inf_{\rho_{h,l} \geq 0} [\phi_{h,l}(\rho_{h,l}) \\ &\quad - (\tilde{a}_l^T \nu_h)\rho_{h,l} - (b_{h,l}\lambda_h)\rho_{h,l} + \gamma_l \rho_{h,l}] \end{aligned} \tag{25}$$

Here $a_{(h-1)L+l}$, $b_{(h-1)L+l}$, $c_{(h-1)L+l}$ are the $[(h-1)L + l]$th column of matrices A, B, C respectively. $\tilde{a}_l$ is the lth column of matrix $\tilde{A}$. First, from the construction of matrix A, we have $a_{(h-1)L+l}^T \nu = \tilde{a}_l^T \nu_h$. Second, since $B = \mathrm{diag}(\tilde{b}_1^T, \ldots, \tilde{b}_P^T) \in \mathbb{R}^{P \times PL}$, further let $\tilde{b}_h^T = (b_{h,1}, \ldots, b_{h,L})$, we then have $b_{(h-1)L+l}^T \lambda = b_{h,l}\lambda_h$. Last, since $C = [I, \ldots, I] \in \mathbb{R}^{L \times PL}$, we have $c_{(h-1)L+l}^T \gamma = \gamma_l$.

4.2 Interpretation for Dual Variables

We now give some interpretations of the dual variables.

- Interpret $\nu_{h,n}$ as the potential at node n associated with mode h. Denote $\tilde{a}_l^T \nu_h$ as $\Delta\nu_{h,l}$, which can then be seen as the potential difference across link l associated with mode h.
- Denote $b_{h,l}\lambda_h$ as $\Delta\lambda_{h,l}$, which can be interpreted as the supply to plant node h via link l.
- Interpret γ_l as the price of using link l.

4.3 Solving the Dual Problem via Subgradient Method

Suppose we can uniquely determine

$$\rho^* = \arg \inf_{\rho \succeq 0} \sum_{h=1}^{P} \sum_{l=1}^{L} [\phi_{h,l}(\rho_{h,l}) - (\tilde{a}_l^T \nu_h)\rho_{h,l} - (b_{h,l}\lambda_h)\rho_{h,l} + \gamma_l \rho_{h,l}], \quad (26)$$

then the subgradient of negative dual is

$$A\rho^* \in \partial(-g)(\nu) \quad (27)$$
$$B\rho^* - H \in \partial(-g)(\lambda) \quad (28)$$
$$R - C\rho^* \in \partial(-g)(\gamma) \quad (29)$$

Using the subgradient method, we can have the following algorithm.

Algorithm 1

- **given** initial potential vector ν^0, plant supply vector $\lambda^0 \succeq 0$, and link price vector $\gamma^0 \succeq 0$
- **repeat**
  (1) Determine
  $$\rho_{h,l}^{k+1} = \arg \inf_{\rho_{h,l} \geq 0} [\phi_{h,l}(\rho_{h,l}) - (\tilde{a}_l^T \nu_h^k)\rho_{h,l} - (b_{h,l}\lambda_h^k)\rho_{h,l} + \gamma_l^k \rho_{h,l}]$$
  $$= \arg \inf_{\rho_{h,l} \geq 0} [\phi_{h,l}(\rho_{h,l}) - \Delta\nu_{h,l}^k \rho_{h,l} - \Delta\lambda_{h,l}^k \rho_{h,l} + \gamma_l^k \rho_{h,l}]$$
  (2) Compute rate surplus
  $$S_{h,n}^k = a_{(h-1)N+n}^T \rho^{k+1} = \tilde{a}_n^T \rho_h^{k+1}$$
  Here $a_{(h-1)N+n}^T$ is the $[(h-1)N+n]$ th row of matix A, and $\tilde{a}_n^T$ is the n th row of $\tilde{A}$.
  (3) Update node potentials
  $$\nu_{h,n}^{k+1} = \nu_{h,n}^k - \alpha_k S_{h,n}^k$$
  (4) Compute demand for each plant node
  $$D_h^k = (H_h - \tilde{b}_h^T \rho_h^{k+1})$$
  (5) Update plant supply
  $$\lambda_h^{k+1} = \lfloor \lambda_h^k + \alpha_k D_h^k \rfloor_+$$
  (6) Compute link capacity margins
  $$M_l^k = R_l - \sum_{h=1}^{P} \rho_{h,l}^{k+1}$$
  (7) Update link price
  $$\gamma_l^{k+1} = \lfloor \gamma_l^k - \alpha_k M_l^k \rfloor_+$$
  (8) Update dual objective
  $$d^{k+1} = g(\nu^{k+1}, \lambda^{k+1}, \gamma^{k+1})$$

Where α_k is an appropriate positive scalar step-size.

This algorithm is decentralized, since

- Rate $\rho_{h,l}$ is calculated from $\Delta\nu_{h,l} = \tilde{a}_l^T \nu_h$, the potential difference across link l associated with mode h, $\Delta\lambda_{h,l} = b_{h,l}\lambda_h$, the supply to plant node h via link l, and γ_l, the price of using link l.
- Node potential $\nu_{h,n}$ is updated from its own rate surplus.
- Supply λ_h is updated from its own demand.
- Link price γ_l is updated from its own capacity margin.

Node potential $\nu_{h,n}^k$, supply λ_h^k, and link price γ_l^k converge to optimal, so does pseudorate $\rho_{h,l}^k$. Iterates can be (and often are) infeasible, i.e. $A\rho^k \neq 0$, $B\rho^k \not\succeq H$, $C\rho^k \not\preceq R$, but we do have $A\rho^\infty = 0$, $B\rho^\infty \succeq H$, $C\rho^\infty \preceq R$, in the limit. $g(\nu, \lambda, \gamma)$ gives a lower bound on primal optimal.

5. NUMERICAL EXAMPLE OF SUBGRADIENT METHOD

We now demonstrate the algorithm developed in the previous section on a numerical example.

5.1 Parameters of the Networked system

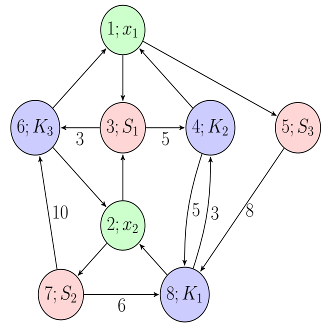


Fig. 1. Networked System Example. Nodes are given an overall enumeration, as well as indicated as being a dynamical node with state x_i, a sensor node S_i, or a controller node K_i. Edges are labelled with their average channel data rate where it is finite.

- Networked system with $N = 8$ nodes and $L = 15$ links (Figure 1).
  · Two plant nodes: $1 - x_1$, $2 - x_2$
  · Three sensor nodes: $3 - S_1$, $5 - S_3$, $7 - S_2$
  · Three controller nodes: $4 - K_3$, $6 - K_3$, $8 - K_1$
- Label the links as:
  $1 - (1,3)$, $2 - (1,5)$, $3 - (2,3)$, $4 - (2,7)$, $5 - (3,4)$, $6 - (3,6)$, $7 - (4,1)$, $8 - (4,8)$, $9 - (5,8)$, $10 - (6,1)$, $11 - (6,2)$, $12 - (7,6)$, $13 - (7,8)$, $14 - (8,2)$, $15 - (8,4)$
- $R = [\infty; \infty; \infty; \infty; 5; 3; \infty; 5; 8; \infty; \infty; 10; 6; \infty; 3]$
- $H = [H_1; H_2] = [10; 7]$

5.2 Objective $\phi(\rho) = 1/2\|\rho\|_2^2$

First we solve the primal problem using the `cvx` (Grant and Boyd (2012)) package for Matlab, allowing the computation to be centralized, then we solve the dual problem using Algorithm 1, our decentralized algorithm, and then compare the results.

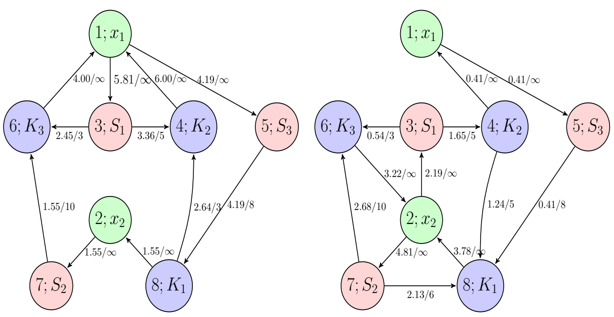
Solving the Primal The solution given by `cvx` is:

$$[\rho_1^* \ \rho_2^*] = \begin{bmatrix} 5.809 & 0 \\ 4.191 & 0.412 \\ 0 & 2.191 \\ 1.546 & 4.809 \\ 3.353 & 1.647 \\ 2.456 & 0.544 \\ 5.997 & 0.412 \\ 0 & 1.235 \\ 4.191 & 0.412 \\ 4.003 & 0 \\ 0 & 3.224 \\ 1.546 & 2.68 \\ 0 & 2.129 \\ 1.546 & 3.776 \\ 2.645 & 0 \end{bmatrix}$$

and the optimal primal objective is $p^* = 110.822$.

Optimal cycles can be identified as below:

- Mode 1 (Figure 2a)
  $x_1 \longrightarrow S_1 \longrightarrow K_3 \longrightarrow x_1$ rate $= 2.45$
  $x_1 \longrightarrow S_1 \longrightarrow K_2 \longrightarrow x_1$ rate $= 3.36$
  $x_1 \longrightarrow S_3 \longrightarrow K_1 \longrightarrow K_2 \longrightarrow x_1$ rate $= 2.64$
  $x_1 \longrightarrow S_3 \longrightarrow K_1 \longrightarrow x_2 \longrightarrow S_2 \longrightarrow K_3 \longrightarrow x_1$ rate $= 1.55$
- Mode 2 (Figure 2b)
  $x_2 \longrightarrow S_1 \longrightarrow K_2 \longrightarrow x_1 \longrightarrow S_3 \longrightarrow K_1 \longrightarrow x_2$ rate $= 0.41$
  $x_2 \longrightarrow S_1 \longrightarrow K_2 \longrightarrow K_1 \longrightarrow x_2$ rate $= 1.24$
  $x_2 \longrightarrow S_1 \longrightarrow K_3 \longrightarrow x_2$ rate $= 0.54$
  $x_2 \longrightarrow S_2 \longrightarrow K_3 \longrightarrow x_2$ rate $= 2.68$
  $x_2 \longrightarrow S_2 \longrightarrow K_1 \longrightarrow x_2$ rate $= 2.13$



(a) Optimal cycles for mode 1 (b) Optimal cycles for mode 2

Fig. 2. Optimal cycles for $\phi(\rho) = 1/2\|\rho\|_2^2$. Edges are labelled with the pseudorate used for the given mode, and the total allowable average channel data rate.

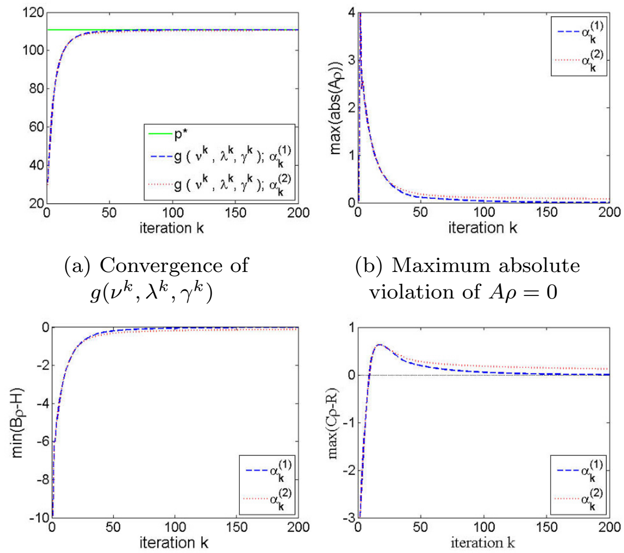
Solving the Dual In each iteration, for all $h = 1, \ldots, P$, $l = 1, \ldots, L$, optimal pseudorate can be uniquely determined by

$$\begin{aligned} &\rho_{h,l}^{k+1}(\Delta\nu_{h,l}^k,\ \Delta\lambda_{h,l}^k,\ \gamma_l^k) \\ &= \arg\inf_{\rho_{h,l}\geq 0} \ [\phi_{h,l}(\rho_{h,l}) - \Delta\nu_{h,l}^k\rho_{h,l} - \Delta\lambda_{h,l}^k\rho_{h,l} + \gamma_l^k\rho_{h,l}] \\ &= \arg\inf_{\rho_{h,l}\geq 0} \ [\frac{1}{2}\rho_{h,l}^2 - (\Delta\nu_{h,l}^k + \Delta\lambda_{h,l}^k - \gamma_l^k)\rho_{h,l}] \\ &= \max\{\Delta\nu_{h,l}^k + \Delta\lambda_{h,l}^k - \gamma_l^k,\ 0\} \end{aligned} \quad (30)$$

We set the initial node potential $\nu^0 = 0$, plant supply $\lambda^0 = 0$, link price $\gamma^0 = 0$, and we would like to try two different step-size rules:

(1) Constant step-size rule: $\alpha_k = 0.2$. Since when $\phi(\rho) = 1/2\|\rho\|_2^2$, the Lagrangian dual function $g(\nu, \lambda, \gamma)$ is differentiable, so subgradient method with constant step-size yields convergence to the optimal value, provided the constant is small enough (Boyd et al. (2003); Bertsekas and Tsitsiklis (1997)).
(2) Mixed step-size rule: for the first 20 iterations, use step-size $\alpha_k = 0.2$, for the rest iterations, use diminishing step-size $\alpha_k = 4/k$.

Figure 3a, 3b, 3c, 3d show the convergence of the dual objective and feasibility violations of this simulation. The constant step-size rule (blue, dashed) and the mixed step-size rule (red, dotted) converge similarly, except with respect to the capacity violation, where the constant step-size rule converges much more quickly.



(a) Convergence of $g(\nu^k, \lambda^k, \gamma^k)$ (b) Maximum absolute violation of $A\rho = 0$ (c) Minimum violation of $B\rho \succeq H$ (d) Maximum capacity violation

Fig. 3. Simulation results for $\phi(\rho) = 1/2\|\rho\|_2^2$.

For the constant step-size rule, after 200 iterations, we have dual objective $d^{200} = 110.821$, and the primal and dual iterates are

$$[\lambda_1^{200} \lambda_2^{200}] = [\,16.99 \quad 10.69\,]$$

$$[\rho_1^{200} \ \ \rho_2^{200}] = \begin{bmatrix} 5.814 & 0 \\ 4.179 & 0.413 \\ 0 & 2.202 \\ 1.543 & 4.793 \\ 3.355 & 1.649 \\ 2.456 & 0.549 \\ 5.994 & 0.412 \\ 0 & 1.239 \\ 4.178 & 0.414 \\ 4.002 & 0 \\ 0 & 3.222 \\ 1.544 & 2.673 \\ 0 & 2.121 \\ 1.542 & 3.775 \\ 2.637 & 0 \end{bmatrix} \qquad \gamma^{200} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1.825 \\ 4.717 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$[\nu_1^{200}\ \nu_2^{200}] = \begin{bmatrix} -7.959 & 1.133 \\ 1.940 & -4.955 \\ -2.145 & -2.754 \\ 3.035 & 0.721 \\ -3.780 & 1.545 \\ 5.027 & 2.512 \\ 3.483 & -0.161 \\ 0.398 & 1.960 \end{bmatrix}$$

The primal iterate ρ^{200} is very close to the solution given by `cvx`. And from the complementary slackness condition, we know from the dual variable γ^{200} that link 5 and 6 are the only two links that have no capacity margin left, and know from dual variable λ^{200} that the supply just meet the demand for each plant, i.e., we have $B\rho^{200} = H$.

6. MODIFIED ADMM METHOD FOR $\phi(\rho) = \|\rho\|_1$

When $\phi(\rho) = \|\rho\|_1$, we have

$$\begin{aligned} &\rho^*_{h,l}(\Delta\nu_{h,l},\ \Delta\lambda_{h,l},\ \gamma_l) \\ &\quad = \arg \inf_{\rho_{h,l}\geq 0} [\phi_{h,l}(\rho_{h,l}) - \Delta\nu_{h,l}\rho_{h,l} - \Delta\lambda_{h,l}\rho_{h,l} + \gamma_l\rho_{h,l}] \\ &\quad = \arg \inf_{\rho_{h,l}\geq 0} (1 - \Delta\nu_{h,l} - \Delta\lambda_{h,l} + \gamma_l)\ \rho_{h,l} \\ &\quad = \begin{cases} \{0\}, & \text{if } \ 1 - \ \Delta\nu_{h,l} - \Delta\lambda_{h,l} + \gamma_l > 0 \\ [0,\infty), & \text{if } \ 1 - \ \Delta\nu_{h,l} - \Delta\lambda_{h,l} + \gamma_l = 0 \\ \{\infty\}, & \text{if } \ 1 - \ \Delta\nu_{h,l} - \Delta\lambda_{h,l} + \gamma_l < 0 \end{cases} \end{aligned}$$

Here ρ^* cannot be uniquely determined, thus we cannot use Algorithm 1. We use some ideas from the Alternating Direction Method of Multipliers (ADMM) algorithm to handle this case. This involves splitting our objective into a function of two variables, which are then constrained to be equal. When formulating the Lagrangian, we then augment it with a term involving the squared norm of the mismatch of these variables; this maintains the key saddle-point property of the Lagrangian, while ensuring that it has a unique minimizer for any value of the dual variable. While adding this term would ruin the separability of the Lagrangian, convergence guarantees still exist if the primal variables are minimized separately, and this allows us to preserve most of our decomposition (Boyd et al. (2011)).

Rewrite the problem as below:

$$\begin{aligned} \text{minimize:} \quad & \|\rho\|_1 + f(z) \\ \text{subject to:} \quad & \rho - z = 0 \\ & B\rho \succeq H \\ & C\rho \preceq R \\ \text{with variables} \quad & \rho \in \mathbb{R}^{PL}_+,\ z \in \mathbb{R}^{PL} \end{aligned} \tag{31}$$

where f is the indicator function of $\{z \in \mathbb{R}^{PL} \mid Az = 0\}$

$$f(z) = \begin{cases} 0 & \text{if } Az = 0 \\ \infty & \text{otherwise} \end{cases}$$

Since $\tilde{A} \in \mathbb{R}^{N\times L}$ is a node-link incidence matrix of a directed graph, every column of $\tilde{A}$ represents a link and contains exactly a 1 and a -1. The rank of $\tilde{A}$ is precisely one less than the number of the rows of $\tilde{A}$. Take away the last row of $\tilde{A}$, and denote the resulting matrix as $\bar{\tilde{A}}$, then $\bar{\tilde{A}} \in \mathbb{R}^{(N-1)\times L}$ is of full row rank, and contains all the information about the topology of the directed graph. Since we construct A as $A = \operatorname{diag}(\tilde{A}, \dots, \tilde{A}) \in \mathbb{R}^{PN\times PL}$, let $\bar{A} = \operatorname{diag}(\bar{\tilde{A}}, \dots, \bar{\tilde{A}}) \in \mathbb{R}^{P(N-1)\times PL}$, then we have $\{z \in \mathbb{R}^{PL} \mid Az = 0\} = \{z \in \mathbb{R}^{PL} \mid \bar{A}z = 0\}$, and $\bar{A}$ is of full row rank.

The augmented Lagrangian is

$$\begin{aligned} &L_\beta(\rho, z, \eta, \lambda, \gamma) \\ &\quad = \|\rho\|_1 + f(z) + \eta^T(\rho - z) + \lambda^T(H - B\rho) \\ &\qquad + \gamma^T(C\rho - R) + \frac{\beta}{2}\|z - \rho\|_2^2 \\ &\quad = \|\rho\|_1 + f(z) + \lambda^T(H - B\rho) + \gamma^T(C\rho - R) \\ &\qquad + \frac{\beta}{2}\|z - \rho - \frac{1}{\beta}\eta\|_2^2 - \frac{1}{2\beta}\eta^T\eta \end{aligned}$$

where $\eta \in \mathbb{R}^{PL}, \lambda \in \mathbb{R}^P_+, \gamma \in \mathbb{R}^L_+$ are the dual variables, and β is an appropriate positive scalar constant.

We are now ready to state the Modified ADMM Algorithm for this problem.

Algorithm 2

- **given** initial dual variables $\eta^0, \lambda^0 \succeq 0, \gamma^0 \succeq 0$
- **repeat**
  (1) Update z, $z^{k+1} = \Pi(\rho^k + \frac{1}{\beta}\eta^k)$, where Π is the projection onto $\{z \in \mathbb{R}^{PL} \mid \bar{A}z = 0\}$, which involves solving a linearly constrained minimum Euclidean norm problem, can be written explicitly as
  $$z^{k+1} = (I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A})(\rho^k + \frac{1}{\beta}\eta^k)$$
  Denote $z = [z_1^T \dots z_p^T]^T$, where $z_h = [z_{h,1} \dots z_{h,L}]^T, h = 1, \dots, P$.
  (2) Update ρ
  $$\begin{aligned} \rho^{k+1} &= \arg \inf_{\rho\succeq 0} L_\beta(\rho, z^{k+1}, \eta^k, \lambda^k, \gamma^k) \\ &= \max\{z^{k+1} - \frac{1}{\beta}(\mathbf{1} + \eta^k - B^T\lambda^k + C^T\gamma^k), 0\} \end{aligned}$$
  Once we have z^{k+1}, the update of ρ splits into PL parallel updates of $\rho_{h,l}$.
  $$\rho^{k+1}_{h,l} = \max\{z^{k+1}_{h,l} - \frac{1}{\beta}(1 + \eta^k_{h,l} - \Delta\lambda^k_{h,l} + \gamma^k_l), 0\}$$
  (3) Update the dual variable η
  $$\eta^{k+1}_{h,l} = \eta^k_{h,l} - \beta(z^{k+1}_{h,l} - \rho^{k+1}_{h,l})$$
  (4) Compute demand for each plant node
  $$D^k_h = (H_h - \tilde{b}^T_h\rho^{k+1}_h)$$
  (5) Update plant supply
  $$\lambda^{k+1}_h = \lfloor\lambda^k_h + \alpha_k D^k_h\rfloor_+$$
  (6) Compute link capacity margins
  $$M^k_l = R_l - \sum_{h=1}^P \rho^{k+1}_{h,l}$$
  .
  (7) Update link price
  $$\gamma^{k+1}_l = \lfloor\gamma^k_l - \alpha_k M^k_l\rfloor_+$$
  (8) Update objective
  $$p^{k+1} = \phi(\rho^{k+1})$$

Where α_k is an appropriate positive step-size.

In this algorithm, except for the update of z, all of the other update steps of the primal and dual iterates are decentralized.

7. NUMERICAL EXAMPLE FOR THE MODIFIED ADMM

We use the same networked system example as in 5.1.

7.1 Solving the Primal

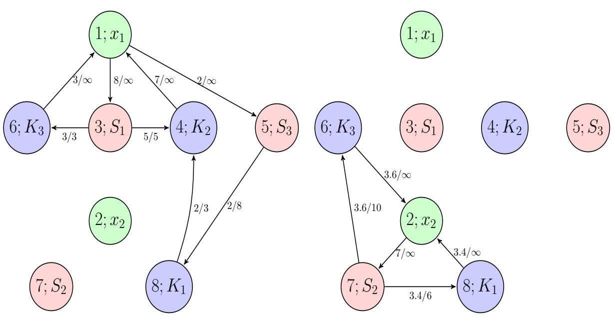
The solution given by `cvx` is:

$$[\rho_1^* \ \rho_2^*] = \begin{bmatrix} 8.000 & 0 \\ 2.000 & 0 \\ 0 & 0 \\ 0 & 7.000 \\ 5.000 & 0 \\ 3.000 & 0 \\ 7.000 & 0 \\ 0 & 0 \\ 2.000 & 0 \\ 3.000 & 0 \\ 0 & 3.628 \\ 0 & 3.628 \\ 0 & 3.372 \\ 0 & 3.372 \\ 2.000 & 0 \end{bmatrix}$$

and the optimal primal objective is $p^* = 53.000$.

Optimal cycles can be identified as below:

- Mode 1(Figure 4a)
  $x_1 \longrightarrow S_1 \longrightarrow K_3 \longrightarrow x_1$ rate $= 3$
  $x_1 \longrightarrow S_1 \longrightarrow K_2 \longrightarrow x_1$ rate $= 5$
  $x_1 \longrightarrow S_3 \longrightarrow K_1 \longrightarrow K_2 \longrightarrow x_1$ rate $= 2$
- Mode 2 (Figure 4b)
  $x_2 \longrightarrow S_2 \longrightarrow K_3 \longrightarrow x_2$ rate $= 3.6$
  $x_2 \longrightarrow S_2 \longrightarrow K_1 \longrightarrow x_2$ rate $= 3.4$



(a) Optimal cycles for mode 1 (b) Optimal cycles for mode 2

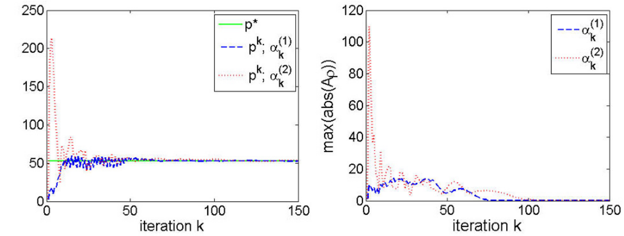
Fig. 4. Optimal cycles for $\phi(\rho) = \|\rho\|_1$. Edges are labelled with the pseudorate used for the given mode, and the total allowable average channel data rate.

7.2 Solving the dual via Modified ADMM

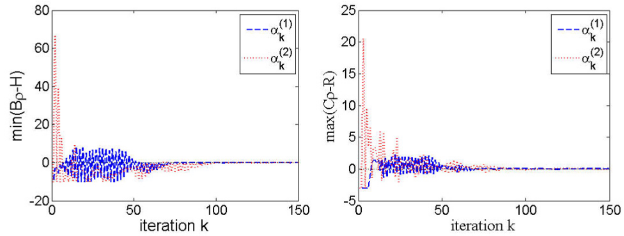
We set the initial primal and dual variables $\rho^0 = 0,\ z^0 = 0,\ \eta^0 = 0,\ \lambda^0 = 0,\ \gamma^0 = 0$ and the parameter $\beta = 0.2$. We compare performance for two different step-size rules:

(1) $\alpha_k = 1.2/\sqrt{k}$.
(2) $\alpha_k = 0.2$ for $1 \leq k \leq 40$, and $\alpha_k = 8/k$ for $k > 40$.

Figure 5a, 5b, 5c, 5d show the convergence of the objective p^k and feasibility violations of this simulation. The convergence for both step-size rules are fast. Since for step-size rule (2), we use a constant step-size for the first 40 iterations, and then switch to a diminishing step-size, its curves behave more radical in the early iterations than those for the step-size rule (1), which uses a diminishing step-size from the very beginning. Also, we observe that the convergence for the step-size rule (1) is relatively faster than that for step-size rule (2).



(a) Convergence of the objective p^k (b) Maximum absolute violation of $A\rho = 0$



(c) Minimum violation of $B\rho \succeq H$ (d) Maximum capacity violation

Fig. 5. Simulation results for $\phi(\rho) = \|\rho\|_1$.

For step-size rule (1), after 150 iterations, we have the objective $p^{150} = 52.823$, and the primal and dual iterates are

$$[\lambda_1^{150} \ \lambda_2^{150}] = [\,4.008 \quad 3.001\,]$$

$$[\rho_1^{150} \ \rho_2^{150}] = \begin{bmatrix} 8.043 & 0 \\ 1.936 & 0 \\ 0 & 0 \\ 0 & 6.983 \\ 4.993 & 0 \\ 3.038 & 0 \\ 6.930 & 0 \\ 0 & 0 \\ 1.936 & 0 \\ 3.059 & 0 \\ 0 & 3.426 \\ 0 & 3.422 \\ 0 & 3.561 \\ 0 & 3.565 \\ 1.921 & 0 \end{bmatrix}$$

$$[z_1^{150}\ z_2^{150}] = \begin{bmatrix} 8.043 & 0 \\ 1.936 & 0 \\ -0.003 & 0.002 \\ 0.001 & 6.983 \\ 4.994 & 0 \\ 3.047 & 0.001 \\ 6.926 & 0 \\ -0.001 & 0 \\ 1.936 & 0 \\ 3.054 & 0 \\ -0.004 & 3.423 \\ 0.003 & 3.422 \\ -0.003 & 3.561 \\ 0.002 & 3.562 \\ 1.931 & 0 \end{bmatrix}$$

$$[\eta_1^{150}\ \eta_2^{150}] = \begin{bmatrix} -1.000 & 0.969 \\ -1.000 & -0.303 \\ 1.100 & -0.425 \\ -0.287 & -1.000 \\ -2.011 & -1.030 \\ -2.006 & -1.575 \\ 3.008 & 0.061 \\ 1.003 & -0.545 \\ -1.000 & -0.303 \\ 3.008 & 0.606 \\ 0.907 & 2.000 \\ -0.621 & -1.000 \\ 0.380 & -1.000 \\ -0.094 & 2.000 \\ -1.000 & 0.545 \end{bmatrix} \quad \gamma^{150} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1.011 \\ 1.008 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

We see that most entries of ρ^{150} are very close to the entries of ρ^* solved by `cvx`. The differences appear in the $11, 12, 13, 14$th entries of ρ_2. As we can see from the cycles for mode 2 in Figure 4b, when the rate cost function is $\phi(\rho) = \|\rho\|_1$, the solution of the primal variable ρ is not unique, since we can assign any positive pseduorate on link $11, 12, 13, 14$, as long as $\rho_{2,11} = \rho_{2,12} \leq 10$, $\rho_{2,13} = \rho_{2,14} \leq 6$, and $\rho_{2,11}+\rho_{2,14} = 7$ are satisfied. Similarly, from the complementary slackness condition, we know from the dual variable γ^{150} that link 5 and 6 are the only two links that have no capacity margin left, and know from dual variable λ^{150} that the supply just meet the demand for each plant, i.e., $B\rho^{150} = H$.

Our simulations converged using small constant β and diminishing α_k as described. The main ADMM convergence results do not directly apply to our problem however, and so classifying constants for which convergence can be guaranteed is the subject of future work.

The objective addressed in this section is intended to produce a sparse portion of the network that can be used for stabilization, and often does, as illustrated with our numerical example. There are cases where the optimization problem may indeed have a sparse solution, but may be indifferent between it and other solutions. When finding a sparse solution is desired in such a case, one possibility is to use the solution to this problem as an initial point for the nonconvex problem with objective norm $1 - \epsilon$, for a small $\epsilon > 0$, which would no longer be indifferent. Other ideas for dealing with situations where the 1-norm does not distinguish between sparse and non-sparse solutions are discussed in (Pilanci et al. (2012)). The decomposability of these algorithms could be another subject of future work.

8. CONCLUSION

We considered the problem of stabilizing plant nodes over a network that includes data rate-limited links. In particular we sought to decompose a key part of that stabilization problem into local optimization problems. We showed how the optimization problem can indeed be decomposed using dual decomposition for most cost functions; namely, for p-norms of the pseudorate vector where $p > 1$. We then developed a method to address the 1-norm which used ideas from ADMM, and showed that we could decompose all but one step of the algorithm.

ACKNOWLEDGEMENTS

The second author would like to thank Randy Cogill for a conversation where some of these ideas arose, as well as Girish Nair and Rob Evans for conversations about the main reference.

REFERENCES

Bertsekas, D. and Tsitsiklis, J. (1997). *Parallel and distributed computation: Numerical methods.* Athena Scientific.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1), 1–122.

Boyd, S., Xiao, L., and Mutapcic, A. (2003). Subgradient methods. *Lecture notes of EE392o, Stanford University, Autumn Quarter*, 2004.

Grant, M. and Boyd, S. (2012). CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`.

Nair, G.N. and Evans, R.J. (2007). Cooperative networked stabilisability of linear systems with measurement noise. In *Proc. 15th Mediterranean Conference on Control & Automation.*

Pilanci, M., El Ghaoui, L., and Chandrasekaran, V. (2012). Recovery of sparse probability measures via convex programming. In *Proc. Advances in Neural Information Processing Systems (NIPS).*

Rotkowitz, M. and Nair, G. (2008). An LP for stabilization over networks with rate constraints. *Proc. 18th Int. Symp. Mathematical Theory of Networks and Systems.*