Nonconcave Utility Maximization

Constantino Lagoa

The Pennsylvania State University, USA

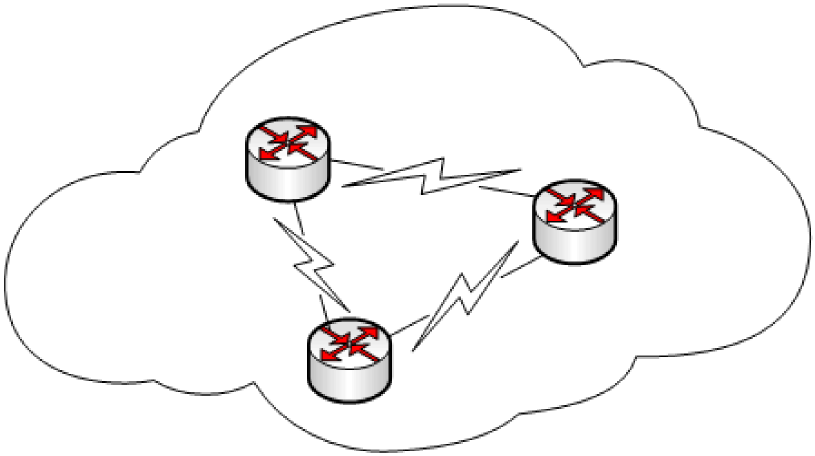
Collaborators

- Mahmoud Ashour
- Jingyao Wang
- N. Serhat Aybat
- Hao Che

Objectives

- Optimization problems motivated by practical applications
- Develop algorithms that can be "easily implemented"
- Decentralization is a must

Motivation



Problem statement:

$$\max \sum_{i \in \mathcal{S}} U_i(r_i) \text{ s.t. network resource constraints}$$

Motivation (cont.)

Challenges:

1. IP video traffic will be 82% of all consumer Internet traffic by 2020, up from 70% in 2015[1].
2. Large-scale networks.

---

[1] ***Cisco VNI Forecast and Methodology, 2015-2020.***

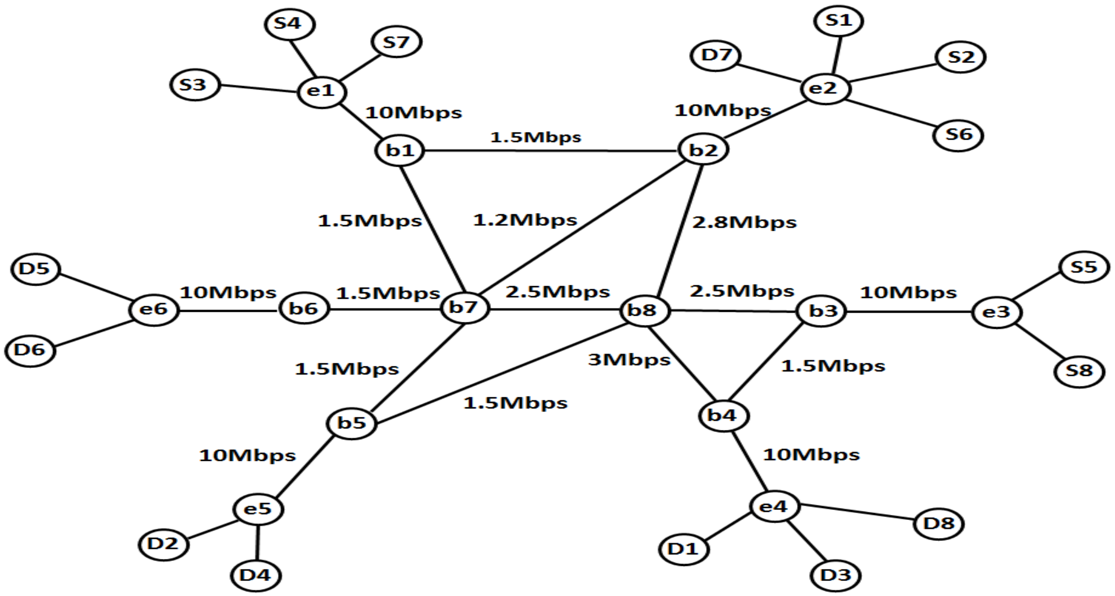
Assumptions:

1. Nonconcave network utility functions[2].
2. No global information available at any network entity.

---

[2] ***Yin, X et al, A control-Theoretic Approach for Dynamic Adaptive Video Streaming over http, SIGCOMM, 2015.***

Flow Control



Notation:

$(\mathcal{S}, \mathcal{L})$: set of sources and links, respectively.

x_i^p : data rate of source i on path p, $p \subset \mathcal{L}$.

$\mathbf{x}_i = [x_i^p]_{p \in \mathcal{P}_i}$, $r_i = \sum_{p \in \mathcal{P}_i} x_i^p$.

$\mathbf{A}_i = [a_{l,p}^i]_{l \in \mathcal{L}, p \in \mathcal{P}_i}$.

Problem Formulation

$$
\begin{aligned}
\max_{\mathbf{x},\mathbf{r}} \sum_{i\in\mathcal{S}} U_i(r_i) \\
\text{s.t.} \ \sum_{i\in\mathcal{S}} \mathbf{A}_i\mathbf{x}_i \preceq \mathbf{c} \\
r_i = \sum_{p\in\mathcal{P}_i} x_i^p \\
0 \le b_i \le r_i \le B_i \\
x_i^p \ge 0
\end{aligned}
$$

Structure to be exploited:

- Utility function "decentralized"
- $a_{i,j} \ge 0$

Class of Utility Functions

We consider utility functions of the form

$$U_i(r_i) = \sum_{j=0}^{\ell} p_{i,j} \left(r_i^{1/\ell}\right)^j$$

Motivation:

- "General set" of utility functions
- Leads to a formulation solved efficiently by distributed algorithms.

Measures and Optimization

Theorem. *Let f be a real-valued function, $\mathcal{G}$ be a compact set, and μ be a probability measure with support* $\operatorname{supp}(\mu)$. *Then,*

$$\inf_x \{f(x) : x \in \mathcal{G}\} = \inf_\mu \left\{ \int f d\mu : \operatorname{supp}(\mu) \subset \mathcal{G} \right\}$$

Moments of Univariate Measures

Theorem. *Given* $\{m_j\}_{j=1}^{\ell}$, *there exists a Borel measure* $\mu(.)$ *with support contained in* $\mathcal{I} = [-\epsilon, \epsilon]$ *such that* $\mu(\mathcal{I}) = 1$ *and* $m_j = \int_{\mathcal{I}} y^j d\mu$ *if and only if* $\mathbf{M}(0, \ell) \succeq 0$, $\epsilon^2 \mathbf{M}(1, \ell - 1) \succeq \mathbf{M}(2, \ell)$.

$$\mathrm{M}(k, k+2h) = \begin{bmatrix} m_k & m_{k+1} & \cdots & m_{k+h} \\ m_{k+1} & \iddots & \iddots & m_{k+h+1} \\ \vdots & \iddots & \iddots & \vdots \\ m_{k+h} & \cdots & \cdots & m_{k+2h} \end{bmatrix}$$

Back to "our" Problem

$$
\begin{aligned}
\max_{\mathbf{x},\mathbf{r}} \; & \sum_{i\in\mathcal{S}} \sum_{j=0}^{\ell} p_{i,j} r_i^{j/\ell} \\
\text{s.t.} \; & \sum_{i\in\mathcal{S}} \mathbf{A}_i \mathbf{x}_i \preceq \mathbf{c} \\
& r_i = \sum_{p\in\mathcal{P}_i} x_i^p \\
& 0 \le b_i \le r_i \le B_i \\
& x_i^p \ge 0
\end{aligned}
\quad \Leftrightarrow \quad
\begin{aligned}
\max_{\mathbf{x},\mathbf{r}} \; & \sum_{i\in\mathcal{S}} \sum_{j=0}^{\ell} p_{i,j} y_i^j \\
\text{s.t.} \; & \sum_{i\in\mathcal{S}} \mathbf{A}_i \mathbf{x}_i \preceq \mathbf{c} \\
& 0 \le y_i \le r_i^{1/\ell} \\
& r_i = \sum_{p\in\mathcal{P}_i} x_i^p \\
& 0 \le b_i \le r_i \le B_i \\
& x_i^p \ge 0
\end{aligned}
$$

Moment Formulation

$$\begin{aligned} \max_{\mathbf{m},\mathbf{x},\mathbf{r}} & \sum_{i\in\mathcal{S}} \mathbf{p}_i^\top \mathbf{m}_i \\ \text{s.t.} \;\; & \mathbf{x} \in \mathcal{C} \\ & (\mathbf{m}_i, \mathbf{x}_i, r_i) \in \mathcal{K}_i, \;\; i \in \mathcal{S} \end{aligned}$$

$$\mathcal{C} = \left\{ \mathbf{x} : \sum_{i\in\mathcal{S}} \mathbf{A}_i \mathbf{x}_i \preceq \mathbf{c} \right\}$$

$$\mathcal{X}_i = \left\{ (\mathbf{x}_i, r_i) \in \mathbb{R}_+^{|\mathcal{P}_i|} \times \mathbb{R}_+ : r_i = \sum_{p\in\mathcal{P}_i} x_i^p, \;\; b_i \leq r_i \leq B_i \right\}$$

$$\begin{aligned} \mathcal{K}_i = \Big\{ & (\mathbf{m}_i, \mathbf{x}_i, r_i) : \mathbf{M}_i(0,\alpha) \succeq 0, \;\; B_i^2 \mathbf{M}_i(1, \alpha-1) \succeq \mathbf{M}_i(2,\alpha), \\ & m_{i,j} \leq r_i^{j/\ell}, j = 1, \ldots, \alpha, (\mathbf{x}_i, r_i) \in \mathcal{X}_i \Big\} \end{aligned}$$

Comments on Moment Formulation

- Converges to the desired optimum as $\alpha \to \infty$
- "Nice" separation between local and global constraints
- Convex if $\alpha \leq \ell$

Decentralized Implementation

Algorithm 1: Distributed Traffic Allocation Algorithm (DTTA) $(\rho, \{\alpha^n\}_{n\in\mathbb{Z}_+}, \{\tau^k\}_{k\in\mathbb{N}}, \{\lambda^k\}_{k\in\mathbb{N}}, \mathbf{z}^0 = [\mathbf{z}_i^0]_{i\in\mathcal{S}}, \mathbf{u}^0 = [\mathbf{u}_i^0]_{i\in\mathcal{S}})$

Initialize $\mathbf{z}^0$, $\mathbf{u}^0$

for $k = 0, 1, \ldots$ **do**

$$(\mathbf{m}_i, \mathbf{x}_i, r_i)^{k+1} \leftarrow \operatorname{argmax}\left\{\mathbf{p}_i^\top \mathbf{m}_i - \frac{\rho}{2}\left\|\mathbf{x}_i - \mathbf{z}_i^k + \mathbf{u}_i^k\right\|^2 : (\mathbf{m}_i, \mathbf{x}_i, r_i) \in \mathcal{K}_i\right\}, i \in \mathcal{S}.$$

Initialize $\mathbf{z}^{k,1} \leftarrow \mathbf{z}^k$

for $n = 1, \ldots, \tau^k - 1$ **do**

Each source $i \in \mathcal{S}$ transmits data at rate $(z_i^p)^{k,n}$ along the path $p \in \mathcal{P}_i$.

For each link $l \in \mathcal{L}$: $b_l^{k,n} \leftarrow 1$ if link l is congested; $b_l^{k,n} \leftarrow 0$ otherwise.

Each link $l \in \mathcal{L}$ sends one bit $b_l^{k,n}$ to all $i \in \mathcal{S}_l$.

$$\mathbf{g}_i^{k,n} = \mathbf{z}_i^{k,n} - \mathbf{x}_i^{k+1} - \mathbf{u}_i^k + \lambda^k \sum_{l\in\mathcal{L}} b_l^{k,n} \mathbf{a}_l^i, \; i \in \mathcal{S}.$$

$$\mathbf{z}_i^{k,n+1} = \left(\mathbf{z}_i^{k,n} - \alpha^n \mathbf{g}_i^{k,n}\right)^+, \; i \in \mathcal{S}.$$

$\mathbf{z}^{k+1} \leftarrow \mathbf{z}^{k,\tau^k}$

$\mathbf{u}_i^{k+1} \leftarrow \mathbf{u}_i^k + \mathbf{x}_i^{k+1} - \mathbf{z}_i^{k+1}, \; i \in \mathcal{S}.$

Why does it work ?

Equivalent formulations:

$$\max_{\mathbf{m},\mathbf{x},\mathbf{r}} \left\{ \sum_{i \in \mathcal{S}} \mathbf{p}_i^\top \mathbf{m}_i : \mathbf{x} \in \mathcal{C},\ (\mathbf{m}_i, \mathbf{x}_i, r_i) \in \mathcal{K}_i,\ i \in \mathcal{S} \right\}$$

$$\max_{\mathbf{m},\mathbf{x},\mathbf{r},\mathbf{z}} \left\{ \sum_{i \in \mathcal{S}} \mathbf{p}_i^\top \mathbf{m}_i \ :\ \mathbf{z} \in \mathcal{C},\ (\mathbf{m}_i, \mathbf{x}_i, r_i) \in \mathcal{K}_i,\ \mathbf{z}_i = \mathbf{x}_i,\ i \in \mathcal{S} \right\}$$

Augmented Lagrangian:

$$\mathrm{L}(\mathbf{m}, \mathbf{x}, \mathbf{r}, \mathbf{z}, \nu) = \sum_{i \in \mathcal{S}} \left[\mathbf{p}_i^\top \mathbf{m}_i - \nu_i^\top (\mathbf{x}_i - \mathbf{z}_i) - (\rho/2) \|\mathbf{x}_i - \mathbf{z}_i\|^2 \right]$$

$$g(\nu) = \max_{\mathbf{m},\mathbf{x},\mathbf{r},\mathbf{z}} \{ \mathrm{L}(\mathbf{m}, \mathbf{x}, \mathbf{r}, \mathbf{z}, \nu) : (\mathbf{m}_i, \mathbf{x}_i, r_i) \in \mathcal{K}_i,\ i \in \mathcal{S}, \mathbf{z} \in \mathcal{C} \}$$

Why does it work? (cont.)

ADMM:

$$(\mathbf{m},\mathbf{x},\mathbf{r})^{k+1} = \operatorname*{argmax}_{\mathbf{m},\mathbf{x},\mathbf{r}} \left\{ \mathrm{L}(\mathbf{m},\mathbf{x},\mathbf{r},\mathbf{z}^k,\nu^k) : (\mathbf{m}_i,\mathbf{x}_i,r_i) \in \mathcal{K}_i, i \in \mathcal{S} \right\}$$

$$\mathbf{z}^{k+1} = \operatorname*{argmax}_{\mathbf{z}} \left\{ \mathrm{L}(\mathbf{m}^{k+1},\mathbf{x}^{k+1},\mathbf{r}^{k+1},\mathbf{z},\nu^k) : \mathbf{z} \in \mathcal{C} \right\}$$

$$\nu^{k+1} = \nu^k + \rho(\mathbf{x}^{k+1} - \mathbf{z}^{k+1})$$

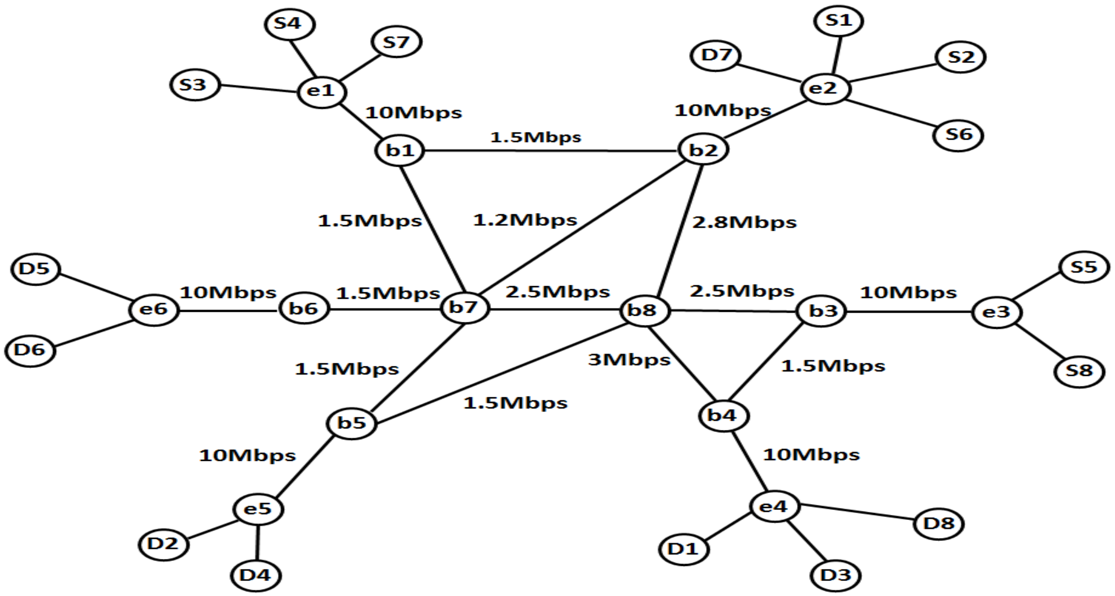
Decentralization:

$$(\mathbf{m}_i,\mathbf{x}_i,r_i)^{k+1} = \operatorname*{argmax}_{(\mathbf{m}_i,\mathbf{x}_i,r_i)\in\mathcal{K}_i} \left\{ \mathbf{p}_i^\top \mathbf{m}_i - (\rho/2) \left\| \mathbf{x}_i - \mathbf{z}_i^k + \mathbf{u}_i^k \right\|^2 \right\}$$

$$\mathbf{z}^{k+1} = \Pi_{\mathcal{C}} \left\{ \mathbf{x}^{k+1} + \mathbf{u}^k \right\}$$

$$\mathbf{u}_i^{k+1} = \mathbf{u}_i^k + \mathbf{x}_i^{k+1} - \mathbf{z}_i^{k+1}$$

Simulations



Parameter values:

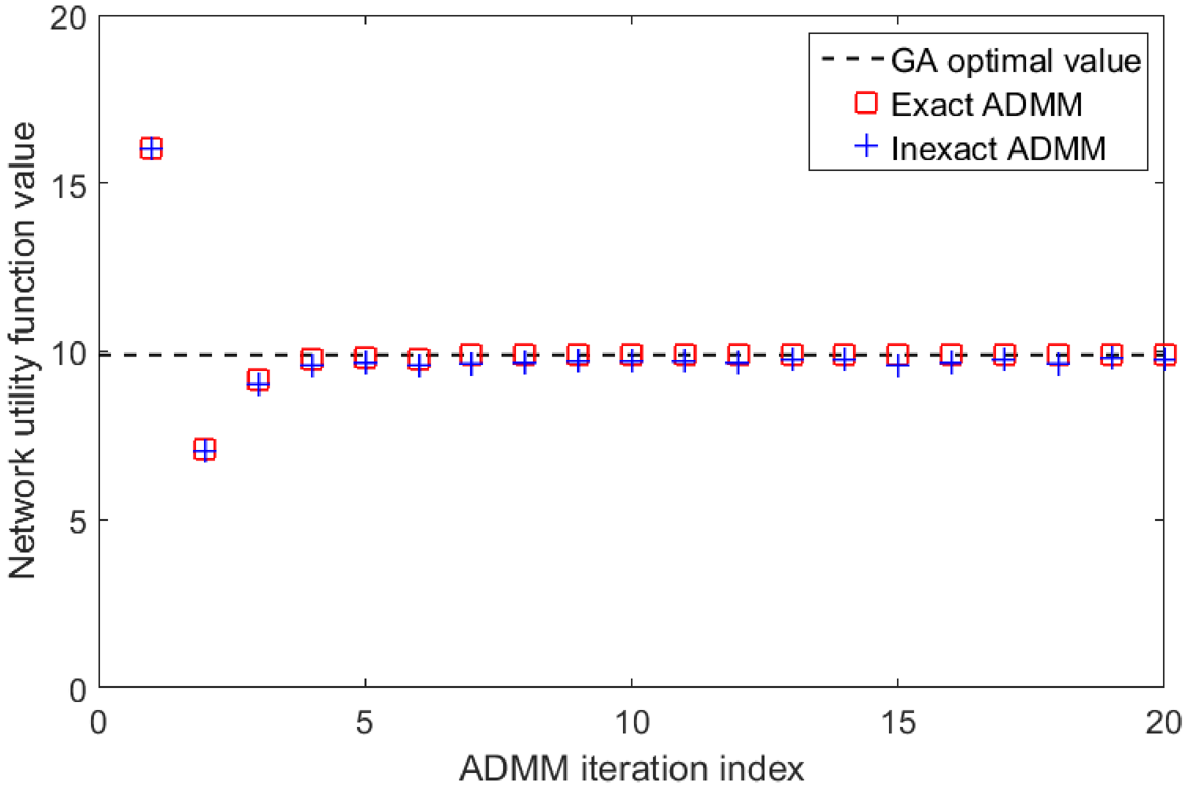
$$\ell = 6$$

$$b_i = 0, B_i = 10$$

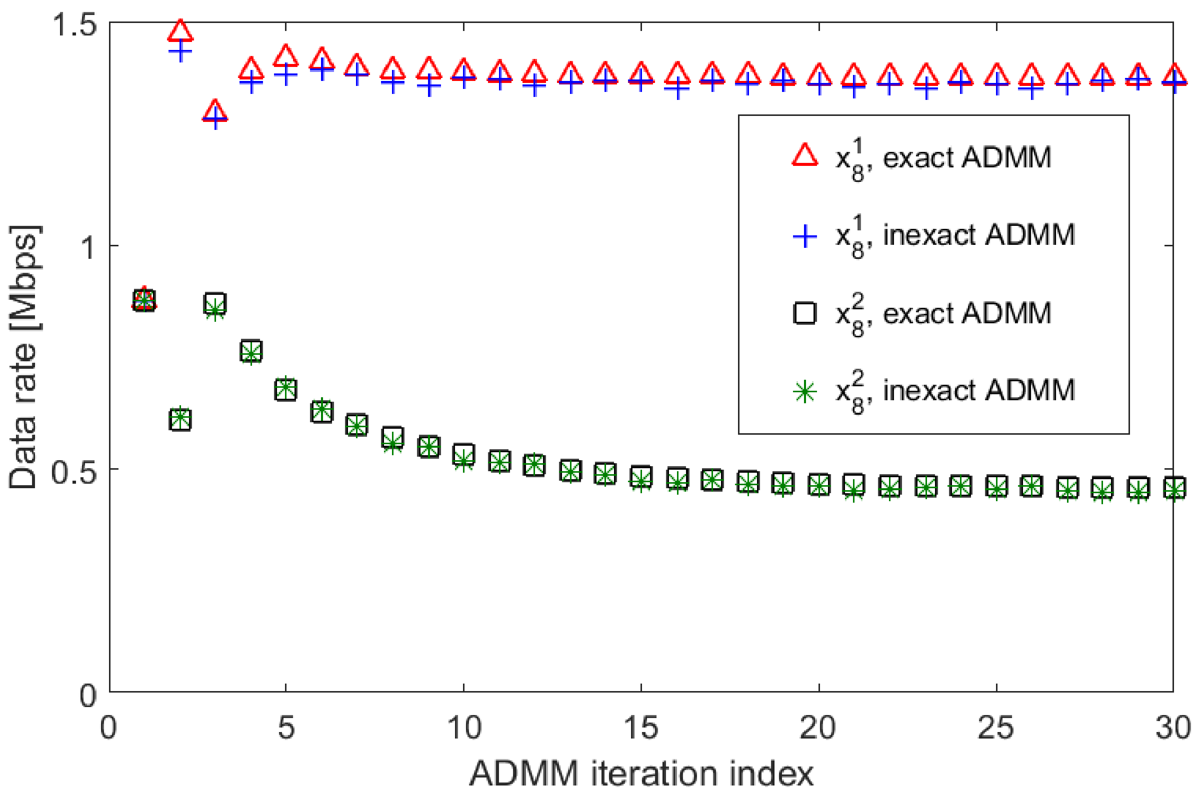
$$\rho = 1$$

$$\lambda^k = 10 \ \forall k \in \mathbb{N}, \tau^k = 10^3 \ \forall k \in \mathbb{N}, \alpha^n = 1/n$$

Network Utility

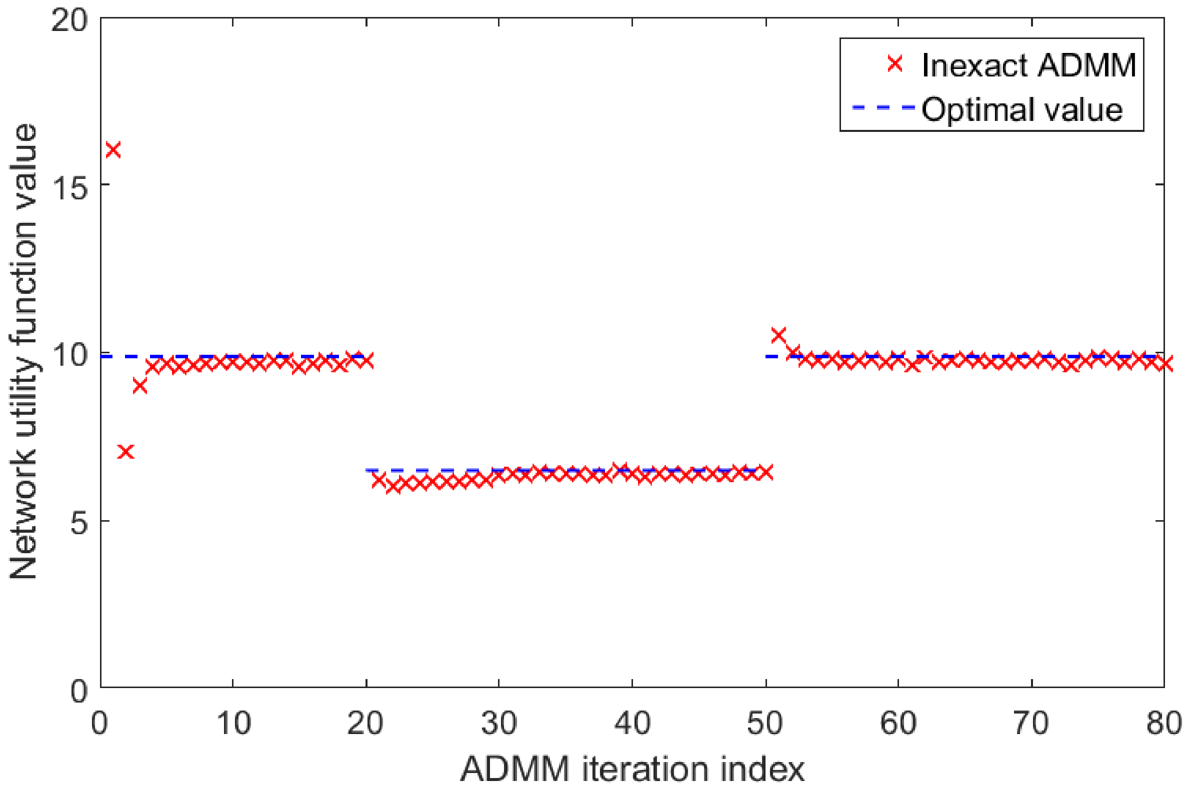


Data Rate Allocation



PennState College of Engineering | ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

Robustness to Link Failures



Concluding Remarks

- General approach to utility maximization
- Efficiently solution of a class of polynomial optimization problems
- Proposed algorithm is decentralized and can solve large problems

Further work:

- What other classes of polynomials optimization problems can be efficiently solved?

Acknowledgements

- National Science Foundation

Thank you

Merci beaucoup

Questions?