An Overview on F-Lipschitz Optimization with Wireless Networks Applications

University of Padova
Padova, July 21, 2010

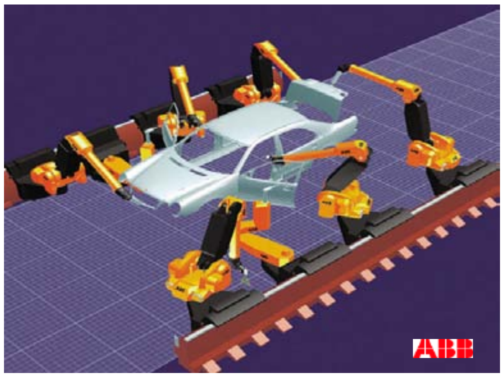
Carlo Fischione

carlofi@ee.kth.se

Electrical Engineering
KTH Royal Institute of Technology

Optimization in Networked Systems

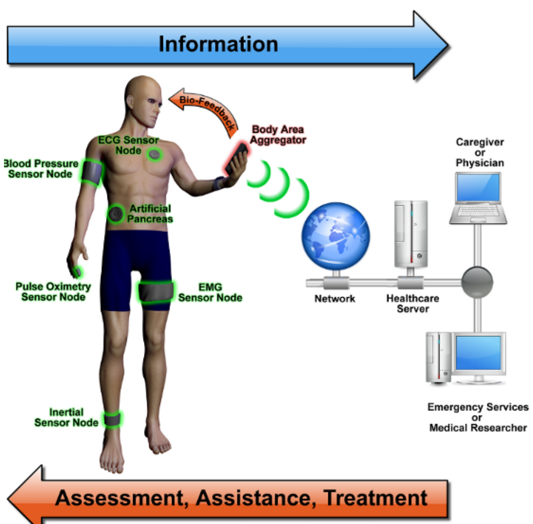
Industrial control



Environmental monitoring

Transportation

Health care



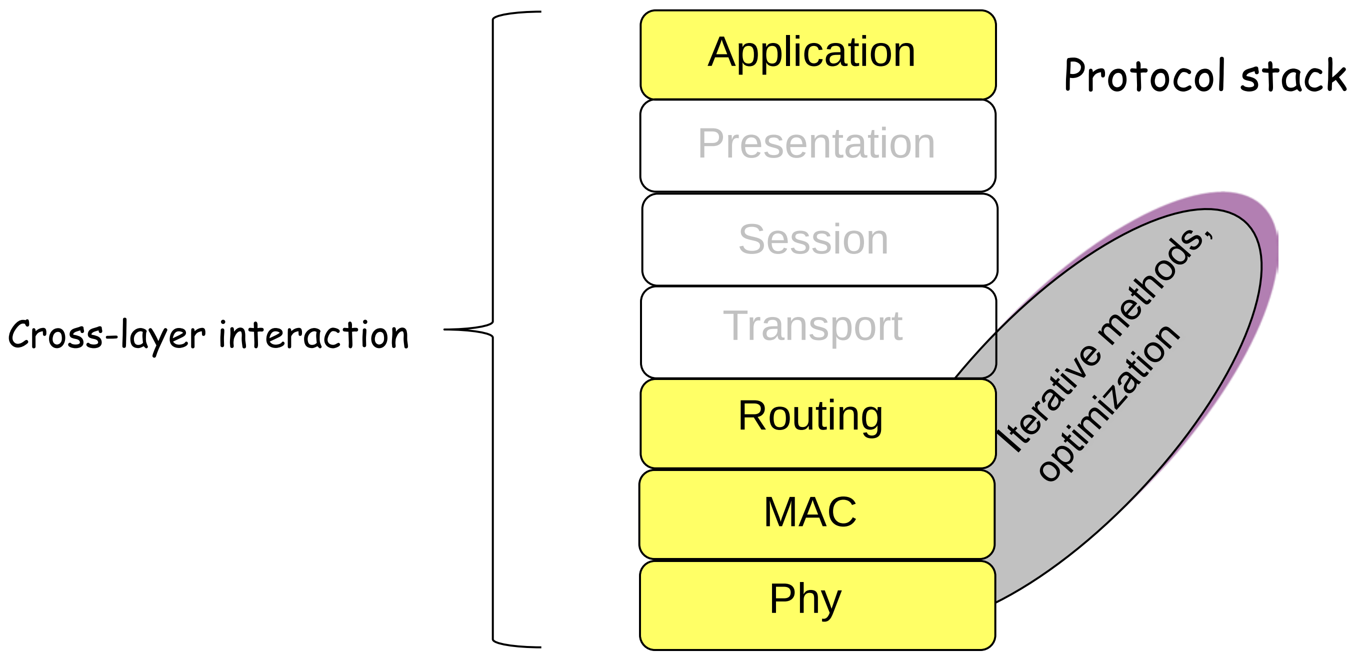
Smart grids

Outline

- **Motivating examples for fast optimization in Wireless Networks**
  - Physical layer
  - Medium access control and Routing
  - Peer-to-peer estimation
- F-Lipschitz optimization
  - Existence and uniqueness of the Pareto optimal solution
  - Centralized computation of the solutoin
  - Distributed computation of the solution
- Some F-Lipschitz applications
  - Interference function theory as a particular case of F-Lipschitz optimization
  - Problems in canonical form
  - Convex optimization and geometric programming
- Peer-to-peer estimation via F-Lipschitz optimization
- Conclusions & future work

Wireless Networks Protocols

- The operations of a node are specified by a set of protocols, or set of rules.



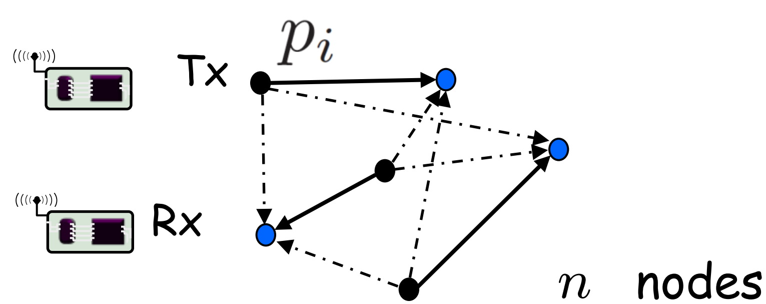
A. Bonivento, C. Fischione, L. Necchi, F. Pianegiani, A. Sangiovanni-Vincentelli, "*System Level Design for Clustered Wireless Sensor Networks*", IEEE Transactions on Industrial Informatics, 2007 (best paper award of TII 2007).

Radio power control

- Let $\mathbf{p} = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^n, \mathbf{p} \succeq 0,$ be a vector of radio powers
  - Each element is the power used by a node for transmission
- Let $I_j(\mathbf{p}) : \mathbb{R}^n \to \mathbb{R}$ be the interference that the radio power has to overcome so that the receiver can detect the transmitted signal

- Interference Function $I(\mathbf{p}) = (I_1(\mathbf{p}), I_2(p), \ldots, I_n(\mathbf{p}))$
- The radio powers of every sensor must be minimized subject to quality of communication constraints:

$$\min_{\mathbf{p}} \quad \mathbf{p}$$
$$\text{s.t.} \quad \mathbf{p} \geq I(\mathbf{p})$$



C. Fischione, K. H. Johansson, A. Sangiovanni-Vincentelli, B. Zurita Ares *"Energy Consumption of Minimum Energy Coding in CDMA Wireless Sensor Networks"*, IEEE Transactions on Wireless Communications, February 2009.

Power control with unreliable components

- Unreliable transceivers introduce intermodulation powers difficult to compensate

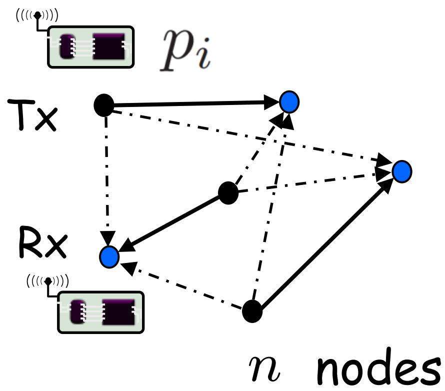
$$\mathrm{SINR}_i = \frac{G_{ii}p_i}{\sigma_i + \sum_{k\neq i} G_{ik}p_k + \sum_{k\neq i} M_{ik}p_i^2 p_k^2}$$

- How to minimize the radio power consumption?

$$\min_{\mathbf{p}} \quad \mathbf{p}$$

$$\text{s.t.} \quad \mathrm{SINR}_i \geq S_{\min}, \quad i = 1, \ldots, n\,,$$

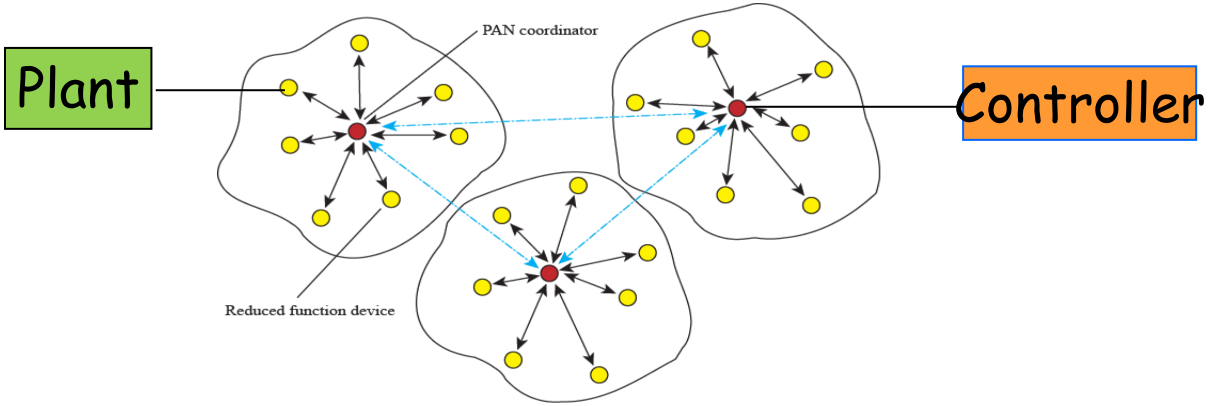
$$\mathbf{1}p_{\min} \leq \mathbf{p} \leq \mathbf{1}p_{\max}\,,$$

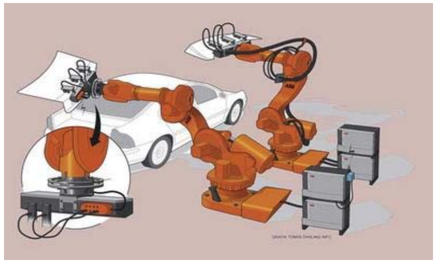


- There is a lack of theory on how to solve these optimization problems by simple and fast algorithms that run on resource constrained nodes.

MAC and Routing over IEEE 802.15.4 networks for Control

Application
Presentation
Session
Transport
Routing
MAC
Phy





- IEEE 802.15.4 wireless sensor networks
  - Nodes transmit their data directly to the cluster head
  - The controller is reached via cluster-head multi-hop routing.

IEEE 802.15.4 MAC and Routing for Control

- Energy, bounded delay and packet transmission requirements must be ensured by IEEE 802.15.4:
  - Control applications require a packet delivery within some deadline and with a guaranteed packet reception probability.

$$\min_{\mathbf{x}} \quad E(\mathbf{x})$$

$$\text{s.t.} \quad P_i(\mathbf{x}) \geq \Omega_i, \qquad i = 1, \ldots, n\,,$$

$$\Pr[D_i(\mathbf{x}) \leq \tau_i] \geq \Delta_i \qquad i = 1, \ldots, n$$

Energy Consumption

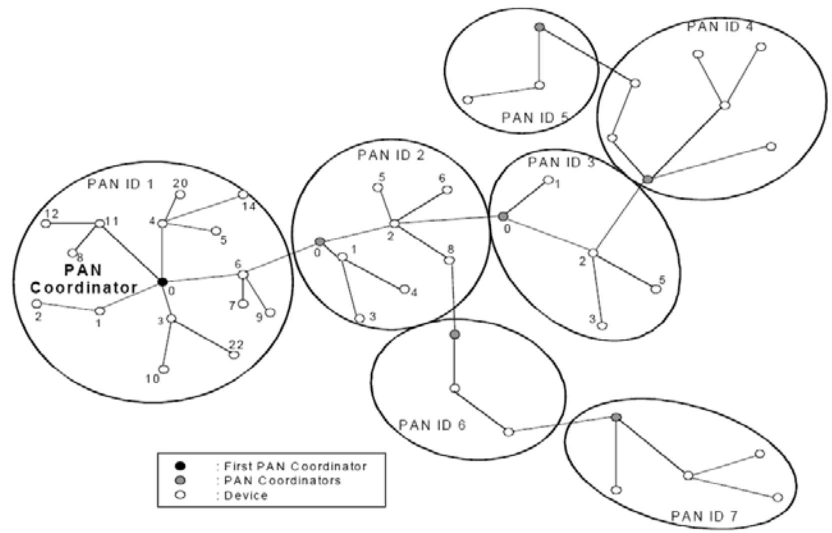
Reliability

Delay

- N clusters give N parallel and coupled optimization problems as the one above to solve without central coordination
  - How to do by nodes of reduced computational capability?

Göteborg: the IEEE 802.15.4 city

- October 2007: Ember & Göteborg Energi deployed 260.000 IEEE 802.15.4 smart meters for electricity monitoring and control
- http://www.ember.com



P. Park, P. Di Marco, P. Soldati, C. Fischione, K. H. Johansson, "*A Generalized Markov Model for an Effective Analysis of Slotted IEEE 802.1*5.4", IEEE Mobile Ad-hoc and Sensor Systems Conference, October 2009 (Best Paper Award).

Outline

- Motivating examples for fast optimization in WSNs
  - Physical layer
  - Medium access control
  - Routing
  - Peer-to-peer estimation
- **F-Lipschitz optimization**
  - Existence and uniqueness of the Pareto optimal solution
  - Centralized computation of the solutoin
  - Distributed computation of the solution
- Some F-Lipschitz applications
  - Interference function theory as a particular case of F-Lipschitz optimization
  - Problems in canonical form
  - Convex optimization and geometric programming
- Peer-to-peer estimation via F-Lipschitz optimization
- Conclusions & future work

The Fast-Lipschitz optimization

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\ \text{s.t.} \quad & x_i \leq f_i(\boldsymbol{x}) \,, \quad i = 1, \ldots, l \\ & x_i = h_i(\boldsymbol{x}) \,, \quad i = l+1, \ldots, n \\ & \boldsymbol{x} \in \mathscr{D} \,, \end{aligned}$$

$$f_0(\boldsymbol{x}) : \mathscr{D} \to \mathbb{R}^m \,, \qquad m \leq n$$

$$f_i(\boldsymbol{x}) : \mathscr{D} \to \mathbb{R} \,, \qquad i = 1, \ldots, l$$

$$h_i(\boldsymbol{x}) : \mathscr{D} \to \mathbb{R} \,, \qquad i = l+1, \ldots, n$$

$\mathscr{D} \subset \mathbb{R}^n$ nonempty compact set

F-Lipschitz 3 qualifying properties

$$
\begin{aligned}
\max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\
\text{s.t.} \quad & x_i \le f_i(\boldsymbol{x}), \quad i = 1, \ldots, l \\
& x_i = h_i(\boldsymbol{x}), \quad i = l+1, \ldots, n \\
& \boldsymbol{x} \in \mathscr{D},
\end{aligned}
$$

1. $\forall \boldsymbol{x} \in \mathscr{D}, \nabla f_0(\boldsymbol{x})$ is a continuous stryctly increasing function;

2. $\forall \boldsymbol{x} \in \mathscr{D}$, either $\nabla_j f_i(\boldsymbol{x}) \le 0, \quad \nabla_j h_i(\boldsymbol{x}) \le 0, \quad \forall i \ne j$;

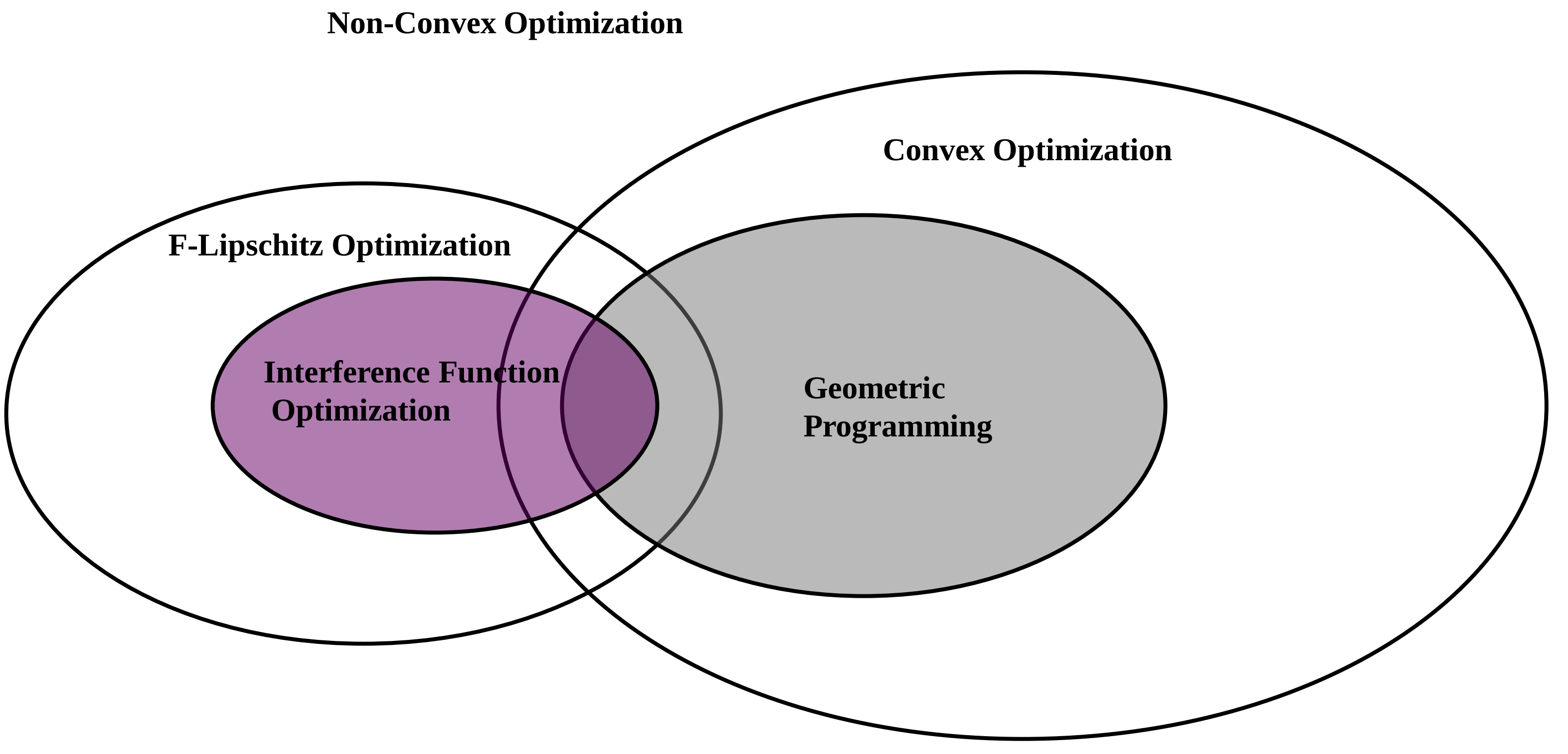
   or $\nabla_i f_0(\boldsymbol{x}) = \nabla_j f_0(\boldsymbol{x}) \quad \forall i \ne j$, and $\nabla_j f_i(\boldsymbol{x}) \ge 0, \quad \nabla_j h_i(\boldsymbol{x}) \ge 0, \quad \forall i \ne j$;

3. Lipschitz contractivity: $\forall \boldsymbol{x}, \mathbf{y} \in \mathscr{D}, \quad |f_i(\boldsymbol{x}) - f_i(\mathbf{y})| \le \alpha_i ||\boldsymbol{x} - \mathbf{y}||, \quad i = 1, \ldots, l$,

   and $|h_i(\boldsymbol{x}) - h_i(\mathbf{y})| \le \alpha_i ||\boldsymbol{x} - \mathbf{y}||, \quad i = l+1, \ldots, n$, with $\alpha_i \in [0, 1) \forall i$

$f_0(\boldsymbol{x})$, $f_i(\boldsymbol{x})$ and $h_i(\boldsymbol{x})$ can be non-convex

The F-Lipschitz optimization



Objective function

$$
\begin{aligned}
\max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\
\text{s.t.} \quad & x_i \leq f_i(\boldsymbol{x}), \quad i = 1, \ldots, l \\
& x_i = h_i(\boldsymbol{x}), \quad i = l+1, \ldots, n \\
& \boldsymbol{x} \in \mathscr{D},
\end{aligned}
$$

- It is allowed to be both a composable or decomposable function of the decision variables.
- It can be a scalar or a vector, for example

$$f_0(\boldsymbol{x}) = \boldsymbol{x}$$

$$f_0(\boldsymbol{x}) = \mathbf{b}^T \boldsymbol{x}, \quad \mathbf{b} \in \mathbb{R}^n, \quad \mathbf{b} \succ 0$$

$$f_0(\boldsymbol{x}) = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}, \quad \mathbf{A}\boldsymbol{x} \succ 0, \mathbf{A} \in \mathbb{R}^n$$

- An F-Lipschitz problem is in general a vector optimization problem with multi-objective function.

Pareto optimal solution

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\ \text{s.t.} \quad & x_i \le f_i(\boldsymbol{x}), \quad i = 1, \dots, l \\ & x_i = h_i(\boldsymbol{x}), \quad i = l+1, \dots, n \\ & \boldsymbol{x} \in \mathscr{D}, \end{aligned}$$

Definition : Consider the following set

$$\begin{aligned} \mathscr{A} = \{ \boldsymbol{x} \in \mathscr{D} : x_i \le f_i(\boldsymbol{x}), i = 1, \dots, l, \\ x_i = h_i(\boldsymbol{x}), i = l+1, \dots, n \}, \end{aligned}$$

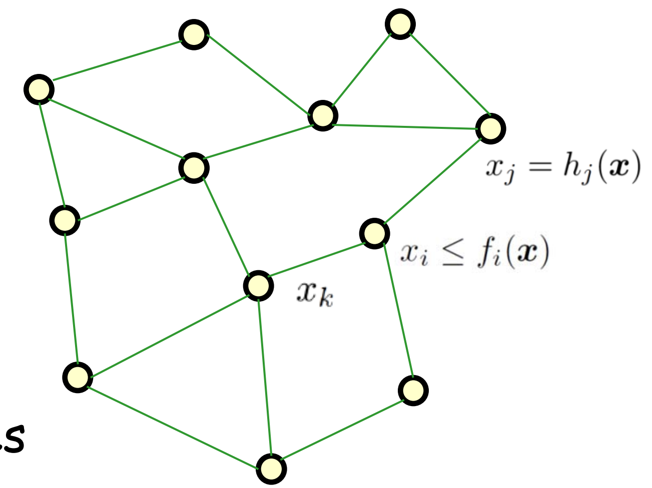
and let $\mathscr{B} \in \mathbb{R}^l$ be the image set of $f_0(\boldsymbol{x})$, namely $f_0(\boldsymbol{x}) : \mathscr{A} \to \mathscr{B}$. Then, we make the natural assumption that the set $\mathscr{B}$ is partially ordered in a natural way, namely if $\boldsymbol{x}, \mathbf{y} \in \mathscr{B}$ then $\boldsymbol{x} \succeq \mathbf{y}$ if $x_i \ge y_i \; \forall i$ (e.g., $\mathbb{R}^l_+$ is the ordering cone).

Definition (Pareto Optimal): A vector $\boldsymbol{x}^*$ is called a Pareto optimal (or an Edgeworth-Pareto optimal) point if there is no $\boldsymbol{x} \in \mathscr{A}$ such that $f_0(\boldsymbol{x}) \succeq f_0(\boldsymbol{x}^*)$ (i.e., if $f_0(\boldsymbol{x}^*)$ is the maximal element of the set $\mathscr{B}$ with respect to the natural partial ordering defined by the cone $\mathbb{R}^l_+$).

Computation of the solution

- Centralized optimization
  - Problem solved by a central processor
- Distributed optimization
  - Decision variables and constraints are associated to distributed nodes that compute the solution

$$
\begin{aligned}
\max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\
\text{s.t.} \quad & x_i \leq f_i(\boldsymbol{x}), \quad i = 1, \ldots, l \\
& x_i = h_i(\boldsymbol{x}), \quad i = l+1, \ldots, n \\
& \boldsymbol{x} \in \mathscr{D},
\end{aligned}
$$



Network of n nodes

Optimal Solution

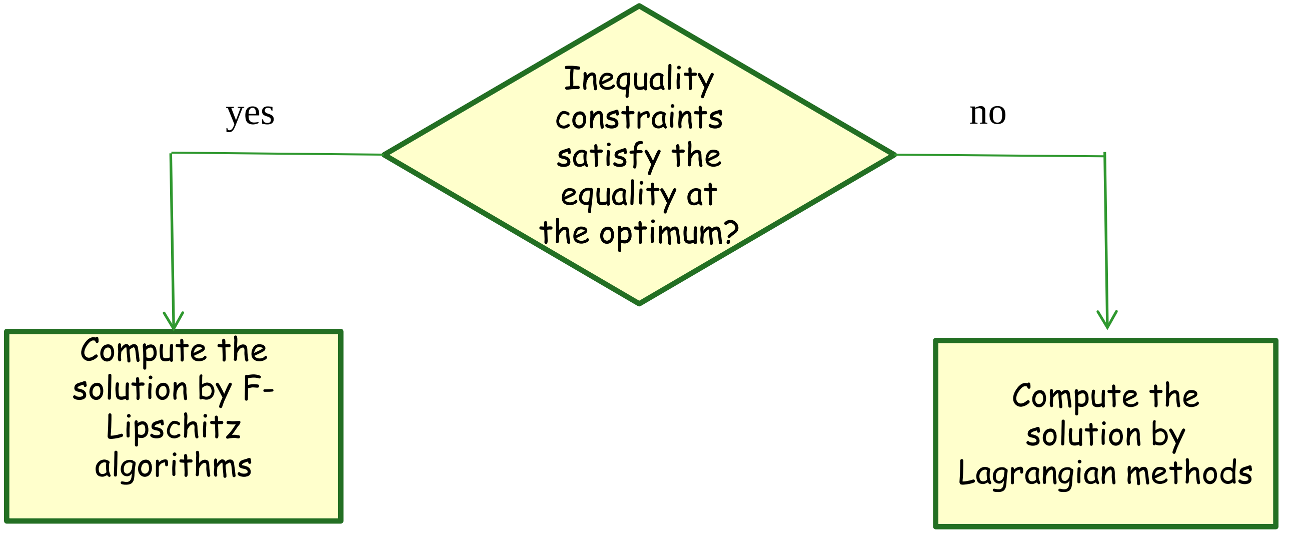
$$\max_{\boldsymbol{x}} \quad f_0(\boldsymbol{x})$$
$$\text{s.t.} \quad x_i \leq f_i(\boldsymbol{x}), \quad i = 1, \ldots, l$$
$$x_i = h_i(\boldsymbol{x}), \quad i = l+1, \ldots, n$$
$$\boldsymbol{x} \in \mathscr{D},$$

Theorem : Let an F-Lipschitz optimization problem be feasible. Then, the problem admits a unique Pareto optimum $\boldsymbol{x}^* \in \mathscr{D}$ given by the solutions of the following set of equations:

$$x_i^* = [f_i(\boldsymbol{x}^*)]^{\mathscr{D}} \quad i = 1, \ldots, l$$
$$x_i^* = h_i(\boldsymbol{x}^*) \quad i = l+1, \ldots, m.$$

- The Pareto optimal solution is just given by a set of (in general non-linear) equations.
- Solving a set of equations is much easier than solving an optimization problem by traditional Lagrangian methods.

F-Lipschitz Optimization



- The F-Lipschitz optimization defines a class of problems for which all the constraints are active at the optimal solution.
- The solution to the set of equations given by the projected constraints is the optimal solution.
- This avoids using Lagrangian methods, which are computationally expensive.

Centralized optimization

$$\max_{\boldsymbol{x}} \quad f_0(\boldsymbol{x})$$
$$\text{s.t.} \quad x_i \leq f_i(\boldsymbol{x}), \quad i = 1, \ldots, l$$
$$x_i = h_i(\boldsymbol{x}), \quad i = l+1, \ldots, n$$
$$\boldsymbol{x} \in \mathscr{D},$$

- The optimal solution is given by iterative methods to solve systems of non-linear equations, such as the Newton's method

$$\boldsymbol{x}(k+1) = \left[\boldsymbol{x}(k) - \beta \left(I - \nabla \mathbf{F}(\boldsymbol{x}(k))\right)^{-1} \left(\boldsymbol{x}(k) - \mathbf{F}(\boldsymbol{x}(k))\right)\right]^{\mathscr{D}}$$

$$\mathbf{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_l(\boldsymbol{x})]^T$$

$$\mathbf{h}(\boldsymbol{x}) = [h_{l+1}(\boldsymbol{x}), h_{l+2}(\boldsymbol{x}), \ldots, h_n(\boldsymbol{x})]^T$$

$$\mathbf{F}(\boldsymbol{x}) = [\mathbf{f}(\boldsymbol{x})^T \mathbf{h}(\boldsymbol{x})^T]^T$$

β is a positive scalar to choose so that convergence speed is maximized.

- Many other methods are available, e.g., heavy balls.

Traditional Lagrangian methods

$$
\begin{aligned}
\max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\
\text{s.t.} \quad & x_i \le f_i(\boldsymbol{x})\,, \quad i = 1, \ldots, l \\
& x_i = h_i(\boldsymbol{x})\,, \quad i = l+1, \ldots, n \\
& \boldsymbol{x} \in \mathscr{D}\,,
\end{aligned}
$$

An F-Lipschitz optimization can be solved by Lagrangian methods.

- Strong duality always applies to F-Lipschitz problems
- The Pareto optimal solution is given be the Karush-Kuhn-Tucker (KKT) conditions:

$$
\begin{aligned}
& x_i - f_i(\boldsymbol{x}^*) \le 0 \quad i = 1, \ldots, l \\
& x_i - h_i(\boldsymbol{x}^*) = 0 \quad i = l+1, \ldots, n \\
& \lambda_i^* \ge 0 \qquad\qquad i = 1, \ldots, n \\
& \lambda_i^* f_i(\boldsymbol{x}^*) = 0 \quad i = 1, \ldots, n \\
& \nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = 0\,,
\end{aligned}
$$

$$
L(\boldsymbol{x}, \boldsymbol{\lambda}) = -\,\boldsymbol{\rho}^T f_0(\boldsymbol{x}) + \sum_{i=1}^{l} \lambda_i (x_i - f_i(\boldsymbol{x})) + \sum_{i=l+1}^{n} \lambda_i (x_i - h_i(\boldsymbol{x}))
$$
Lagrangian

$$
\begin{aligned}
\boldsymbol{x}(k+1) &= \boldsymbol{x}(k) - \beta \nabla_{\boldsymbol{x}} L(\boldsymbol{x}(k), \boldsymbol{\lambda}(k)) \\
\boldsymbol{\lambda}(k+1) &= \boldsymbol{\lambda}(k) - \beta \nabla_{\boldsymbol{\lambda}} L(\boldsymbol{x}(k), \boldsymbol{\lambda}(k))
\end{aligned}
$$

Distributed optimization

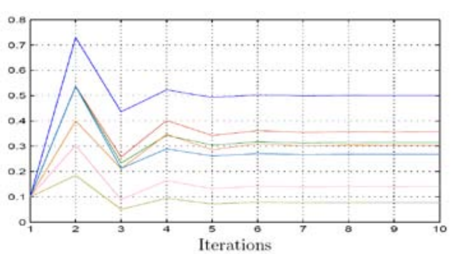
$$\begin{aligned} \max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\ \text{s.t.} \quad & x_i \leq f_i(\boldsymbol{x}), \quad i = 1, \ldots, l \\ & x_i = h_i(\boldsymbol{x}), \quad i = l+1, \ldots, n \\ & \boldsymbol{x} \in \mathscr{D}, \end{aligned}$$

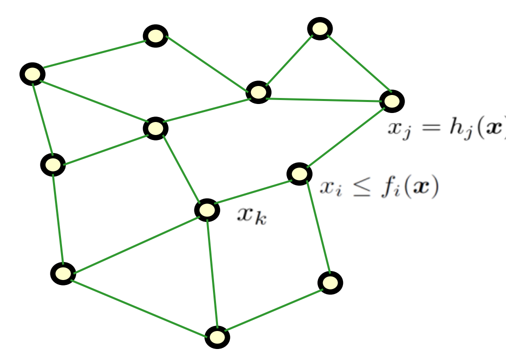
Proposition 3.7: Let $\boldsymbol{x}(0) \in$ be an initial guess of the optimal solution to a feasible F-Lipschitz problem. Let $\boldsymbol{x}^i(k) = [x_1(\tau_1^i(k)), x_2(\tau_2^i(k)), \ldots, x_n^i(\tau_n(k))]$ the vector of decision variables available at node i at time $k \in \mathbb{N}_+$, where $\tau_j^i(k)$ is the delay with which the decision variable of node j is communicated to node i. Then, the following iterative algorithm converges to the optimal solution:

$$x_i(k+1) = [f_i(\boldsymbol{x}^i(k))]^{\mathscr{D}} \quad i = 1, \ldots, l$$

$$x_i(k+1) = h_i(\boldsymbol{x}^i(k)) \quad i = l+1, \ldots, n$$

where $k \in \mathbb{N}_+$ is an integer associated to the iterations.





Outline

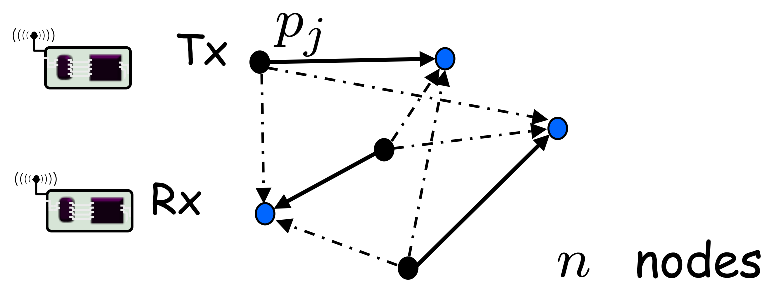
- Motivating examples for fast optimization in WSNs
  - Physical layer
  - Medium access control
  - Routing
  - Peer-to-peer estimation
- F-Lipschitz optimization
  - Existence and uniqueness of the Pareto optimal solution
  - Centralized computation of the solutoin
  - Distributed computation of the solution
- **Some F-Lipschitz applications**
  - Interference function theory as a particular case of F-Lipschitz optimization
  - Problems in canonical form
  - Convex optimization and geometric programming
- Peer-to-peer estimation via F-Lipschitz optimization
- Conclusions & future work

Interference function theory

- $\mathbf{p} = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^n, \mathbf{p} \succeq 0,$ vector of radio powers
- $I_j(\mathbf{p}) : \mathbb{R}^n \to \mathbb{R}$ interference that the radio power has to overcome

$$I(\mathbf{p}) = (I_1(\mathbf{p}), I_2(p), \ldots, I_n(\mathbf{p}))$$

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{p} \\ \text{s.t.} \quad & \mathbf{p} \geq I(\mathbf{p}) \end{aligned}$$



- Properties of the (Type-I) interference function

1. $\mathbf{I}(\mathbf{p}) \succ 0$
2. $\mathbf{p} \succeq \mathbf{q} \Longrightarrow \mathbf{I}(\mathbf{p}) \succeq \mathbf{I}(\mathbf{q})$
3. $c \in \mathbb{R}, \quad c > 1 \Longrightarrow c\mathbf{I}(\mathbf{p}) > \mathbf{I}(c\mathbf{p})$

Zander, ”*Distributed cochannel control in cellular radio systems interference*”, IEEE Trans. Veh. Technol., 1992.
Foschini, Miljanic, “*A simple distributed autonomous power control algorithm and its convergence*,” IEEE Trans. Veh. Technol., 1993.

Power control as an F-Lipschitz problem

- $x = -\mathrm{p} \qquad f_i(x) = -I_i(-x)$

1. $\mathrm{f}(x) \prec 0$
2. $x \succeq \mathrm{y} \Longrightarrow \mathrm{f}(x) \succeq \mathrm{f}(\mathrm{y})$
3. $c \in \mathbb{R}, \quad c > 1 \Longrightarrow \mathrm{f}(cx) > c\mathrm{f}(x)$

$$\max_{x} \quad x$$
$$\text{s.t.} \quad x \leq \mathbf{f}(x)$$

Theorem : Suppose that a function $f(x)$ satisfies the type-I properties, then it satisfies the F-Lipschitz qualifying properties.

- F-Lipschitz qualifying properties are much more general than the interference function properties.

Problems in canonical form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & g_0(\boldsymbol{x}) \\ \text{s.t.} \quad & g_i(\boldsymbol{x}) \leq 0\,, \quad i = 1, \ldots, l \\ & p_i(\boldsymbol{x}) = 0\,, \quad i = l+1, \ldots, n \\ & \boldsymbol{x} \in \mathscr{D}\,, \end{aligned}$$

Canonical form

Bertsekas, *Non Linear Programming*, 2004.

$$\begin{aligned} f_0(\boldsymbol{x}) &= -g_0(\boldsymbol{x})\,, \\ f_i(\boldsymbol{x}) &= x_i - \gamma_i g_i(\boldsymbol{x})\,, \quad \gamma_i > 0 \\ h_i(\boldsymbol{x}) &= x_i - \mu_i p_i(\boldsymbol{x})\,, \quad \mu_i \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & f_0(\boldsymbol{x}) \\ \text{s.t.} \quad & x_i \leq f_i(\boldsymbol{x})\,, \quad i = 1, \ldots, l\,, \\ & x_i = h_i(\boldsymbol{x}) \quad i = l+1, \ldots, n\,, \\ & \boldsymbol{x} \in \mathscr{D}\,, \end{aligned}$$

F-Lipschitz form

Problems in canonical form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & g_0(\boldsymbol{x}) \\ \text{s.t.} \quad & g_i(\boldsymbol{x}) \le 0\,, \quad i = 1, \dots, l \\ & p_i(\boldsymbol{x}) = 0\,, \quad i = l+1, \dots, n \\ & \boldsymbol{x} \in \mathscr{D}\,, \end{aligned}$$

Theorem Consider an optimization problem in canonical form. Let

$$\mathscr{A} = \{\boldsymbol{x} \in \mathscr{D} | g_i(\boldsymbol{x}) \le 0, i = 1, \dots, l,$$
$$p_i(\boldsymbol{x}) = 0, \quad i = l+1, \dots, n\}\ .$$

For $\boldsymbol{x} \in \mathscr{A}$, suppose that

1. $\nabla g_0(\boldsymbol{x}) \prec 0$
2. $\nabla_i g_i(\boldsymbol{x}) > 0$ and $\nabla_i p_i(\boldsymbol{x}) > 0 \quad \forall i.$
3. Either $\nabla_j g_i(\boldsymbol{x}) \ge 0$ and $\nabla_j p_i(\boldsymbol{x}) \ge 0 \quad \forall j \ne i$ or $\nabla_i g_0(\boldsymbol{x}) = \nabla_j g_0(\boldsymbol{x})\,, \quad \nabla_j g_i(\boldsymbol{x}) \le 0$ and $\nabla_j p_i(\boldsymbol{x}) \le 0 \quad \forall j \ne i\,.$

Problems in canonical form

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & g_0(\boldsymbol{x}) \\ \text{s.t.} \quad & g_i(\boldsymbol{x}) \le 0\,, \quad i = 1, \ldots, l \\ & p_i(\boldsymbol{x}) = 0\,, \quad i = l+1, \ldots, n \\ & \boldsymbol{x} \in \mathscr{D}\,, \end{aligned}$$

Then, the problem is F-Lipschitz if for every $i = 1, \ldots, l$

either $\nabla_i g_i(\boldsymbol{x}) > \sum_{j \neq i} |\nabla_j g_i(\boldsymbol{x})|$

or $|\nabla_i g_i(\boldsymbol{x})| + \sum_{j \neq i} |\nabla_j g_i(\boldsymbol{x})| < L_{g_i}$ and $\nabla \mathbf{F}(\boldsymbol{x})$ has full rank

and for every $i = l + 1 \ldots, n$

either $\nabla_i p_i(\boldsymbol{x}) > \sum_{j \neq i} |\nabla_j p_i(\boldsymbol{x})|$

or $|\nabla_i p_i(\boldsymbol{x})| + \sum_{j \neq i} |\nabla_j p_i(\boldsymbol{x})| < L_{h_i}$ and $\nabla \mathbf{F}(\boldsymbol{x})$ has full rank .

Example 1: from a convex problem to an F-Lipschitz one

$$
\begin{aligned}
\min_{x,y} \quad & (ax^2 + cy^2)^{-1} \\
\text{s.t.} \quad & x - 0.5y - 1 \leq 0 \\
& -x + 2y \leq 0 \\
& x \geq 0\,, \quad y \geq 0\,,
\end{aligned}
$$

$$
a > 0,\ b > 0 \qquad x, y \in \mathbb{R}
$$

- The problem is convex: KKT conditions could be used to compute the optimal solution, but the problem is F-Lipschitz:

$$
\begin{aligned}
&\nabla_x(x - 0.5y - 1) = 1 > 0 \\
&\nabla_y(x - 0.5y - 1) = -0.5 < 0 \\
&\nabla_y(-x + 2y) = 2 > 0 \\
&\nabla_x(-x + 2y) = -1 < 0
\end{aligned}
$$

$$
\begin{aligned}
&\nabla_x(x - 0.5y - 1) = 1 > |\nabla_y(x - 0.5y - 1)| = 0.5 \\
&\nabla_y(-x + 2y) = 2 > |\nabla_x(-x + 2y)| = 1\,,
\end{aligned}
$$

- The solution is given by the constraints at the equality, trivially

$$
\begin{aligned}
x - 0.5y - 1 &= 0 \qquad & x^* &= 4/3 \\
-x + 2y &= 0\,, \qquad & y^* &= 2/3
\end{aligned}
$$

Example 2: a hidden F-Lipschitz problem

$$
\begin{aligned}
\min_{x,y,z} \quad & (ax^2y^2 + bz^{-1})^{-1} \\
\text{s.t.} \quad & x - 0.5y + z + 3 \le 0 \\
& -x + 2y - z^{-1} + 1 \le 0 \\
& -3x - y + z^{-2} + 2 \le 0 \\
& x_{\min} \le x \le x_{\max}, \quad y_{\min} \le y \le y_{\max}, \quad z_{\min} \le z \le z_{\max},
\end{aligned}
$$

A convex non-F-Lipschitz problem

$$\nabla_z(x - 0.5y + z + 3) > 0$$

$t = z^{-1}$ ⇓

$$
\begin{aligned}
\min_{x,y,t} \quad & (ax^2y^2 + bt)^{-1} \\
\text{s.t.} \quad & x - 0.5y + t^{-1} + 3 \le 0 \\
& -x + 2y - t + 1 \le 0 \\
& -3x - y + t^2 + 2 \le 0 \\
& x_{\min} \le x \le x_{\max}, \quad y_{\min} \le y \le y_{\max}, \quad 1/z_{\max} \le t \le 1/z_{\min}
\end{aligned}
$$

An F-Lipschitz problem

Geometric programming

$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & g_0(\boldsymbol{x}) \\
\text{s.t.} \quad & g_i(\boldsymbol{x}) \le 1 \quad i = 1, \ldots, l \\
& p_i(\boldsymbol{x}) = 1 \quad i = l+1, \ldots, m \\
& \boldsymbol{x} \in \mathscr{D} \quad \boldsymbol{x} \succ 0
\end{aligned}
$$

$$g_i(\boldsymbol{x}) = \sum_{k=1}^{K} c_{ik} x_1^{a_{i1k}} x_2^{a_{i2k}} \cdots x_m^{a_{imk}} \quad i = 0, \ldots, l$$ posynomial

$$p_i(\boldsymbol{x}) = c_i x_1^{b_{i1}} x_2^{b_{i2}} \cdots x_m^{b_{im}} \quad i = l+1, \ldots, m$$ monomial

$$c_{ik} > 0,\ a_{ijk} \in \mathbb{R},\ b_{ij} \in \mathbb{R},\ \forall i, j, k$$

- Geometric problems are convex (via a simple mechanical conversion) and are solved by Lagrangian methods (interior point methods).
- Geometric problems play an essential role in electrical circuit design.

S. Boyd, S. J. Kim, L. Vandenberghe, A. Hassibi, "A tutorial on geometric programming," *Optimization and Engineering, vol. 1, no. 1*, p. 1, 1 2006.

When geometric problems are F-Lipschitz

$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & g_0(\boldsymbol{x}) \\
\text{s.t.} \quad & g_i(\boldsymbol{x}) \leq 1 \quad i = 1, \ldots, l \\
& p_i(\boldsymbol{x}) = 1 \quad i = l+1, \ldots, m \\
& \boldsymbol{x} \in \mathscr{D} \quad \boldsymbol{x} \succ 0
\end{aligned}
$$

Corollary Consider a geometric optimization problem. Let

$$\mathscr{A} = \{\boldsymbol{x} \in \mathscr{D} | g_i(\boldsymbol{x}) \leq 1, i = 1, \ldots, l, p_i(\boldsymbol{x}) = 1\} .$$

The problem is an F-Lipschitz one if the following conditions simultaneously hold:

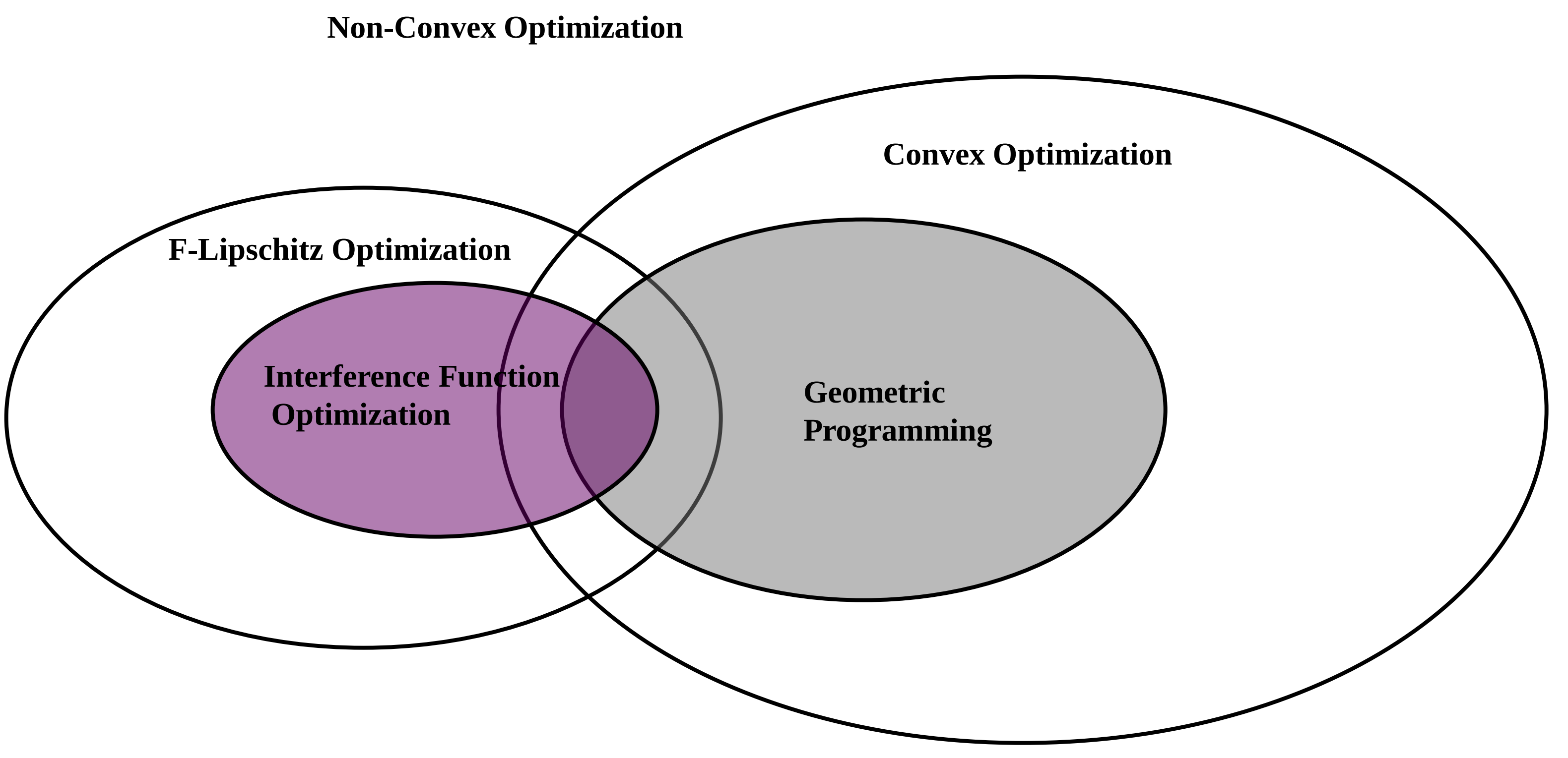
1) $\nabla g_0(\boldsymbol{x}) \prec 0 \ \forall \boldsymbol{x} \in \mathscr{A}$;
2) $a_{iik} > 0$ and $b_{ii} > 0 \ \forall i$;
3) either $a_{ijk} \geq 0 \ b_{ij} \geq 0$ or $\nabla_i g_0(\boldsymbol{x}) = \nabla_j g_0(\boldsymbol{x})$, $a_{ijk} \leq 0$, and $b_{ij} \leq 0$, $\quad \forall i$ and $\forall j \neq i$;
4) $\mathscr{D} = [x_{1,\min}, x_{1,\max}] \times [x_{2,\min}, x_{2,\max}] \times \ldots \times [x_{n,\min}, x_{n,\max}]$, with $0 < x_{i,\min} < x_{i,\max} < \infty \ \forall i$.

Example: a geometric problem is easily recognized as F-Lipschitz

$$
\begin{aligned}
\min_{x,y,z} \quad & x^{-2} + 9y^{-1} + z^{-3} \\
\text{s.t.} \quad & 3x^{3.1}y^{-1} + 4y^{-2} + z^{-1} \le 12 \\
& 5x^{-2} + 6y^{2}x^{-1} + z^{-1} \le 10 \\
& x^{-1}y^{-1}z^{2} = 10 \\
& \mathscr{D} = \{10^{-10} \le x \le 1, \quad 10^{-10} \le y \le 1, \quad 10^{-10} \le z \le 1\}
\end{aligned}
$$

The exponent of the i-th decicion variable of the i-th constraint is always positive, whereas the other exponents are negative...

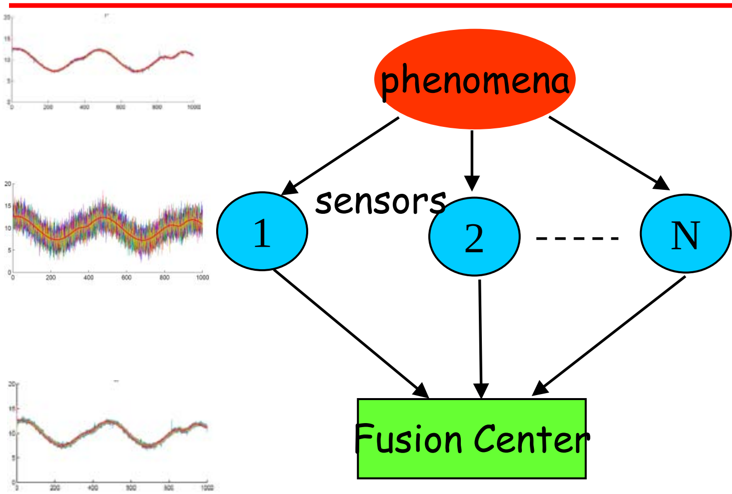
The F-Lipschitz optimization



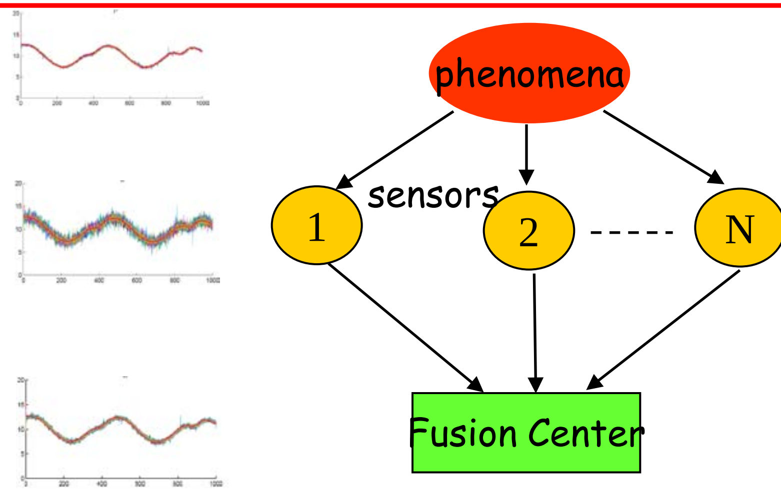
Outline

- Motivating examples for fast optimization in WSNs
  - Physical layer
  - Medium access control
  - Routing
  - Peer-to-peer estimation
- F-Lipschitz optimization
  - Existence and uniqueness of the Pareto optimal solution
  - Centralized computation of the solutoin
  - Distributed computation of the solution
- Some F-Lipschitz applications
  - Interference function theory as a particular case of F-Lipschitz optimization
  - Problems in canonical form
  - Convex optimization and geometric programming
- **Peer-to-peer estimation via F-Lipschitz optimization**
- Conclusions & future work

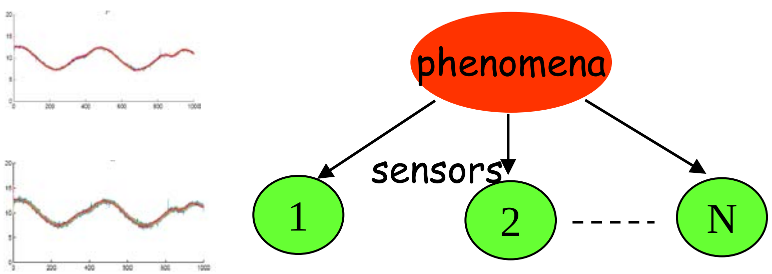
Estimation



Centralized Estimation:
no intelligence on sensors



Distributed Estimation:
some processing on sensors

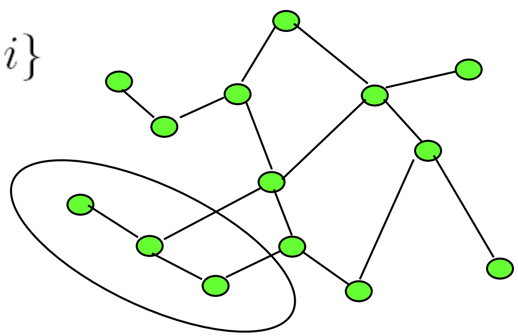


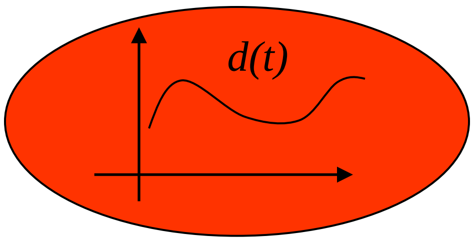
Peer-to-Peer Estimation:
no central coordination

Peer-to-Peer Estimation

A. Speranzon, C. Fischione, K. H. Johansson, A. Sangiovanni-Vincentelli, “*A Distributed Minimum Variance Estimator for Sensor Networks*”, IEEE Journal on Selected Areas in Communications, special issue on Control and Communications, May 2008.

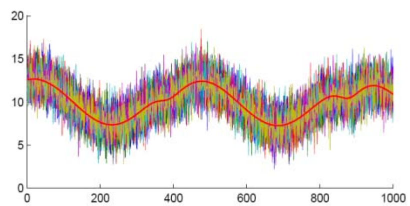
$$\mathcal{N}_i = \{j \in \mathcal{V} : (j,i) \in \mathcal{E}\} \cup \{i\}$$





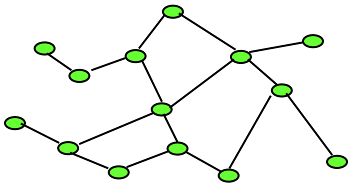
- Nodes perform a noisy measurement of a common time-varying signal

$$u_i(t) = d(t) + v_i(t)$$



- Communication subject to space-time varying packet losses

Peer-to-Peer Estimator



- Nodes exchange local measurements and estimates

$$z_i(t) = \sum_{j \in \mathcal{N}_i(t)} k_{ij}(t)\phi_{ij}(t)z_j(t-1) + \sum_{j \in \mathcal{N}_i(t)} h_{ij}(t)\phi_{ij}(t)u_j(t)$$

Local estimate at node *i*

$$\mathbf{z}(t) = (\mathbf{K}(t) \circ \mathbf{\Phi}(t))\, \mathbf{z}(t-1) + (\mathbf{H}(t) \circ \mathbf{\Phi}(t))\, \mathbf{u}(t)$$

Global vector of the estimates

$$\mathbf{K}(t) = [\mathbf{k}_i(t)] \in \mathbb{R}^{N \times N}$$
$$\mathbf{H}(t) = [\mathbf{h}_i(t)] \in \mathbb{R}^{N \times N}$$
$$\mathbf{\Phi}(t) = [\boldsymbol{\phi}_i(t)] \in \mathbb{R}^{N \times N}$$

- **Goal**: find locally the estimation coefficients $\mathbf{k}_i(t)$ and $\mathbf{h}_i(t)$ that minimize the variance of the estimation error.

Estimation Coefficients

$$\mathbf{e}(t) = \mathbf{z}(t) - d(t)\mathbf{1}$$ Estimation Error

- Estimation Coefficients given by minimizing the average estimation error, under stability constraints

$$\begin{aligned} \min_{\mathbf{K}(t),\mathbf{H}(t)} \quad & \mathbb{E}\,\mathbf{e}(t)^T\mathbf{e}(t) \\ \text{s.t.} \quad & ((\mathbf{K}(t)+\mathbf{H}(t)-\mathbf{I})\circ\mathbf{\Phi}(t))\,\mathbf{1}=0 \\ & \|\mathbf{K}(t)\circ\mathbf{\Phi}(t)\| \le \gamma_{\max} < 1 \end{aligned}$$

Small Bias

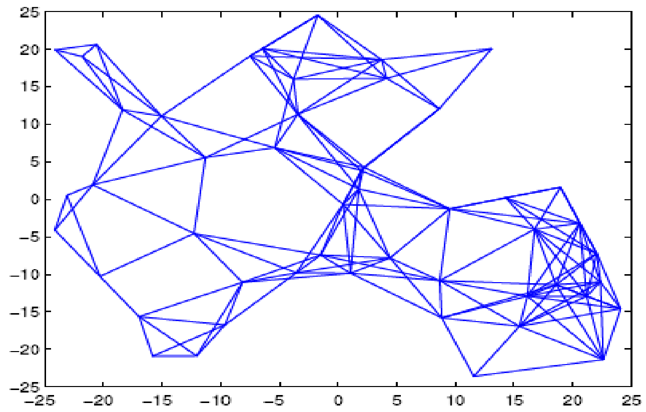
Stable Estimation Error

- A centralized optimization problem

- How to distribute the computation of the optimal solution?
  - ➢ An F-Lipschitz optimization problem

Simulation Example: Peer-to-peer estimation

- Network with 30 nodes randomly deployed.



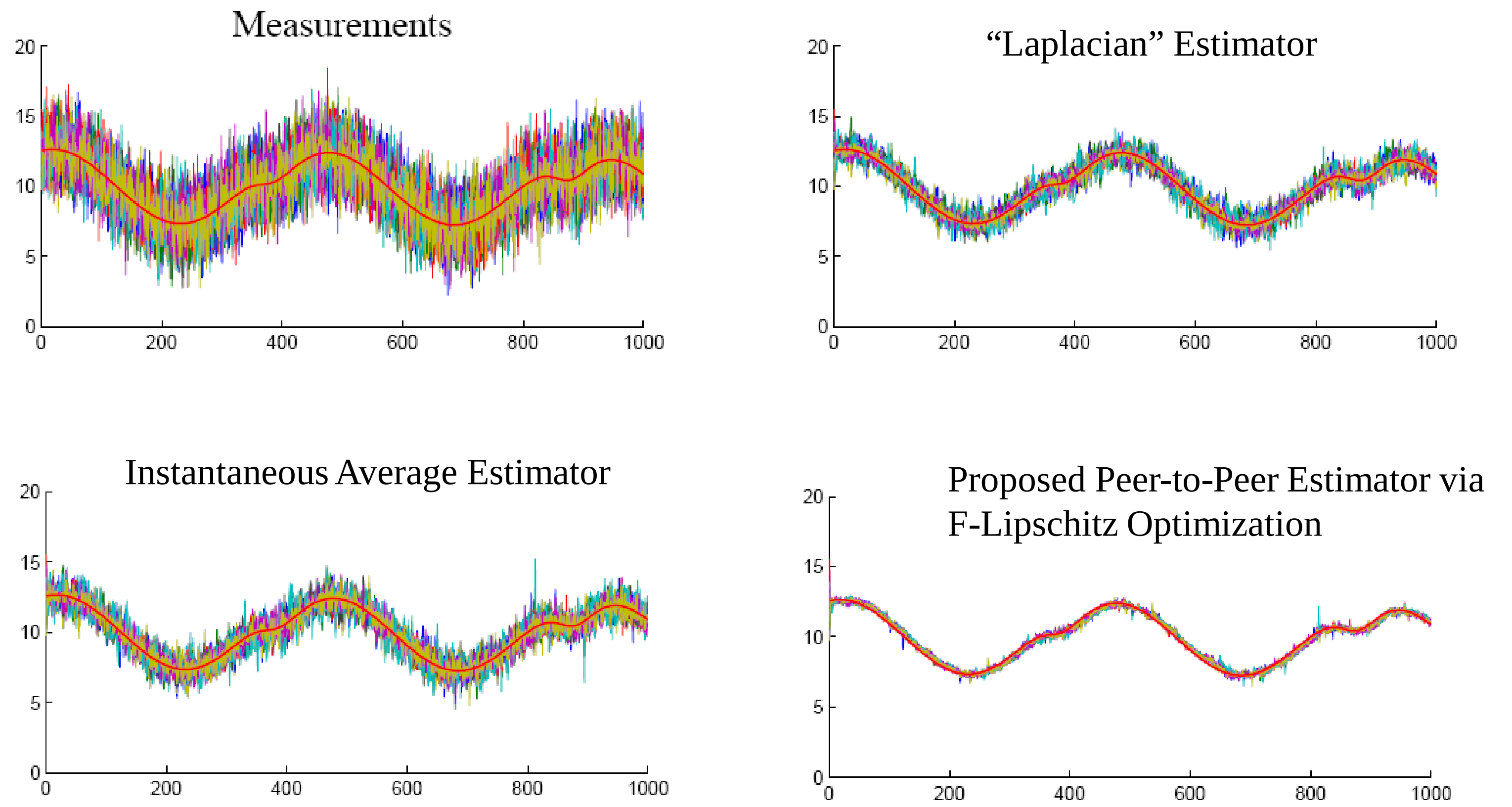
- Signal to track:

$$d(t) = 3\sin(2\pi t/1500) - 8\cos(2\pi t/1800)\sin(2\pi t/800)$$

- Variance of the additive noise:

$$\sigma^2 = 1.2 \qquad\qquad u_i(t) = d(t) + v_i(t)$$

Simulation Example: Peer-to-peer estimation



Packet loss probability $q_{ij} = 10\% \pm 5\%$

Conclusions

- F-Lipschitz optimization enables fast computations of the solution of a class of convex and non-convex optimization problems.
  - Central idea: optimal solution achieved when all the constraints are active.
  - F-Lipschitz optimization solve several problems much more efficiently than traditional Lagrangian methods.
- The interference function theory optimization is a particular case of F-Lipschitz optimization.
- Perhaps, in many situations, it is better to "F-Lipschitzfy" than "convexify".
- More info on http://www.ee.kth.se/~carlofi/

Acknowledgements

- S. Boyd (Stanford Univ.)
- K. H. Johansson, M. Johansson, U. Jönsson, A. Möller (KTH)
- A. Speranzon (UTRC)