Distributed Interference Management and Identification for Wireless Networks

Sławomir Stańczak

joint work with
Holger Boche, Renato L. G. Cavalcante, Mario Goldenbaum

Technische Universität Berlin
Heinrich-Hertz-Lehrstuhl für Informationstheorie und theoretische Informationstechnik
and
Fraunhofer Heinrich Hertz Institute
Einsteinufer 37
10587 Berlin, Germany

May 29, 2013

Presenter

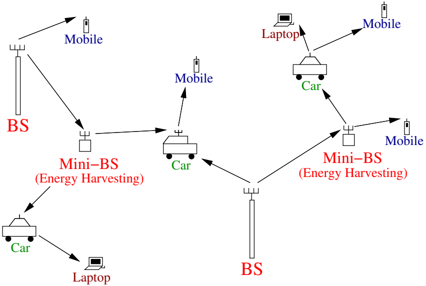
Slawomir Stanczak

stanczak@ieee.org
http://www.mk.tu-berlin.de/

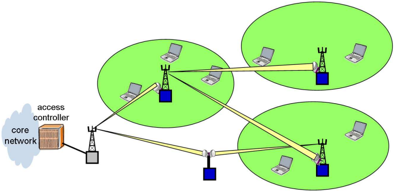
Fraunhofer Heinrich Hertz Institute
Berlin, Germany
and
Heinrich Hertz Chair for Information Theory and
Theoretical Information Technology
Faculty of EECS
Technical University of Berlin

Challenges

- Invented for P2P links and successful in cellular settings, information theory is inadequate for MANETs as it ignores:
  - bursty traffic, finite sessions/flows
  - queuing delay, energy limitations
  - overhead, hardware
- Communication networking theory
  - neglects the interference structure
  - does not utilize the broadcast property of the wireless channel

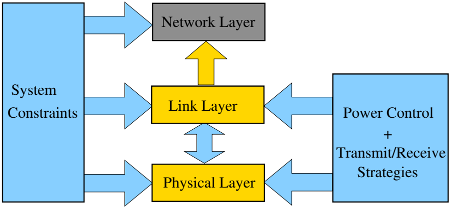


Wireless Networks



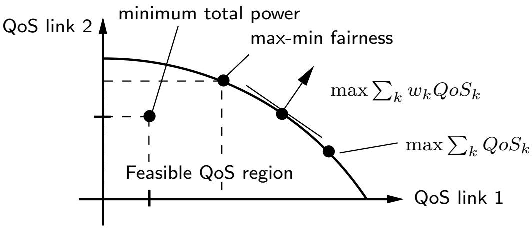
- Focus: High data rates, low or moderate channel dynamics
- unreliable shared radio channel
- limited resources

⇒ Resource allocation and interference management are necessary.



Quality of Service

- User-centric approaches (inelastic applications):
  - Satisfy strict QoS requirements of applications permanently.
- Network-centric approaches (elastic applications):
  - Maximize the aggregate utility as perceived by the network operator.
  - Address the issue of fairness.



The Focus

- Single-hop communication with $K > 1$ logical links (users)
- Concurrent transmissions
- Single-user decoding
- Individual power constraints $\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_K)$
- Multiple antenna elements
- Single data stream per link
- Combination with routing and network coding strategies possible.

- The focus of this talk: Joint power control and beamforming for resource allocation and interference management.

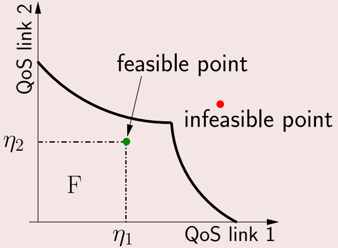
Part I

Arbitrary but Fixed Channels

Outline

Feasible Utility Set: $\mathbb{F}$

Given a channel, $\mathbb{F}$ is the set of all utility (or QoS) values that can be achieved by means of power control with all links being active concurrently.



- Assumption: $\eta_k \uparrow \Leftrightarrow$ QoS $\uparrow$
  - Downward-comprehensive
  - Upper-bounded
- may be non-convex.
- depends on power constraints $\mathbb{P}$.

- $\mathbb{F}$ depends on the physical-layer realization: Key properties of many multiuser systems are captured by interference functions.

Signal-to-Interference(+noise) Ratio (SIR)

Strictly monotonic Utility-SIR map: $\gamma : \mathbb{R} \to \mathbb{R}_+$

For any $\boldsymbol{\eta} \in \mathrm{F}$, there is a power vector $\mathbf{p} = (p_1, \ldots, p_K) \in \mathrm{P}$ such that

$$\gamma(\eta_k) = \mathrm{SIR}_k(\mathbf{p}) = \frac{p_k}{I_k(\mathbf{p})} \qquad \begin{array}{l} \leftarrow \text{transmit power} \\ \leftarrow \text{interference function} \end{array}$$

- e.g. Gaussian capacity (in nats/channel use): $\gamma(x) = e^x - 1, x \geq 0$.

Axiomatic Interference Framework

Standard Interference Functions (SIF), Yates'95

A1 Positivity: $I_k(\mathbf{p}) > 0$ for all $\mathbf{p} \geq 0$.

A2 Scalability: $I_k(\mu\mathbf{p}) < \mu I_k(\mathbf{p})$ for any $\mathbf{p} \geq 0$ and for all $\mu > 1$.

A3 Monotonicity: $I_k(\mathbf{p}^{(1)}) \geq I_k(\mathbf{p}^{(2)})$ if $\mathbf{p}^{(1)} \geq \mathbf{p}^{(2)}$.

- Interference functions depend on the choice of beamformers.
- The framework captures many practical interference scenarios.

Interference Functions: Examples and Focus

Linear interference function

$$I_k(\mathbf{p}) = (\mathbf{V}\mathbf{p} + \mathbf{z})_k$$

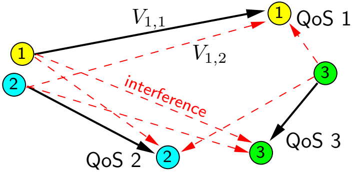
- Matched-filter receiver
- SIC receiver

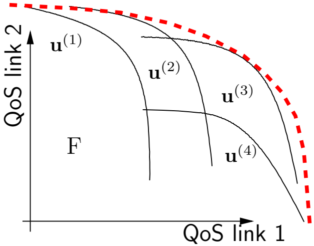
Minimum interference function

$$I_k(\mathbf{p}) = \min_{u_k \in Z_k} (\mathbf{V}(\mathbf{u})\mathbf{p} + \mathbf{z}(\mathbf{u}))_k$$

- MMSE receiver
- Optimal beamforming

$\mathbf{V} = (v_{k,l}) = (\frac{V_{k,l}}{V_{k,k}})$
$\mathbf{V} \geq 0$: non-symmetric, traceless





Beamforming

Signal-to-Interference(+noise) Ratio (SIR)

$$\mathrm{SIR}_k = \frac{p_k |\mathbf{r}_k^H \mathbf{H}^{(k,k)} \mathbf{x}_k|^2}{\sum_{l \neq k} |\mathbf{r}_k^H \mathbf{H}^{(k,l)} \mathbf{x}_l|^2 p_l + \|\mathbf{r}_k\|_2^2 \sigma^2} = \frac{p_k}{\sum_{l \neq k} \frac{V_{k,l}}{V_{k,k}} p_l + \sigma_k} = \frac{p_k}{I_k(\mathbf{p})}$$

- $\mathbf{x}_k$: TX beamformer of user k
- $\mathbf{r}_k$: RX beamformer of user k

Outline

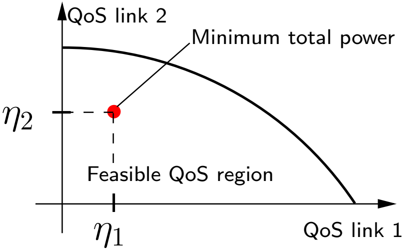
Outline

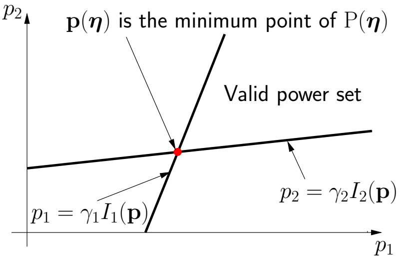
Problem Statement

Problem (QoS-based power control under SIFs)

$$\mathbf{p}(\boldsymbol{\eta}) = \arg\min_{\mathbf{p}\in \mathrm{P}(\boldsymbol{\eta})} \mathbf{w}^T \mathbf{p} \qquad \mathbf{w} > 0$$

$$\mathrm{P}(\boldsymbol{\eta}) := \left\{\mathbf{p} \in \mathbb{R}_+^K : \forall_k \, \mathrm{SIR}_k(\mathbf{p}) \geq \gamma(\eta_k)\right\}.$$





Zander'92, Foschini'94, Yates'95, Ulukus'98, Bambos'00, Boche&Schubert ...

Fixed-Point Iteration, Yates'95

Fixed-Point Existence, Uniqueness and Convergence

Let $\mathbf{I}(\mathbf{p}) = (\gamma_1 I_1(\mathbf{p}), \ldots, \gamma_K I_K(\mathbf{p}))$ be any SIF for some $\gamma_k \equiv \gamma(\eta_k) > 0$. If there is $\mathbf{p} > 0$ such that $\mathbf{p} \geq \mathbf{I}(\mathbf{p})$, then

1. $\text{Fix}(\mathbf{I}) = \{\mathbf{p} > 0 : \mathbf{I}(\mathbf{p}) = \mathbf{p}\} \neq \emptyset$ is singleton and
2. the fixed-point iteration

$$\mathbf{p}(n+1) = \mathbf{I}(\mathbf{p}(n)), \quad \text{for some} \quad \mathbf{p}(0) \geq 0$$

converges to the unique fixed point $\mathbf{p}(\boldsymbol{\eta}) = \bar{\mathbf{p}} = \mathbf{I}(\bar{\mathbf{p}})$.

- Component-wise increasing (decreasing) if $\mathbf{p}(0) = \mathbf{0}$ $(\mathbf{p}(0) \in \mathrm{P}(\boldsymbol{\eta}))$.
- Amenable to distributed implementation, scalable, works for any SIF.
- Asynchronous operation possible.
- But how should new users join the network without disrupting the connections of active users?

Fixed-Point Iteration, Yates'95

Fixed-Point Existence, Uniqueness and Convergence

Let $\mathbf{I}(\mathbf{p}) = (\gamma_1 I_1(\mathbf{p}), \ldots, \gamma_K I_K(\mathbf{p}))$ be any SIF for some $\gamma_k \equiv \gamma(\eta_k) > 0$. If there is $\mathbf{p} > 0$ such that $\mathbf{p} \geq \mathbf{I}(\mathbf{p})$, then

1. $\text{Fix}(\mathbf{I}) = \{\mathbf{p} > 0 : \mathbf{I}(\mathbf{p}) = \mathbf{p}\} \neq \emptyset$ is singleton and
2. the fixed-point iteration

$$\mathbf{p}(n+1) = \mathbf{I}(\mathbf{p}(n)), \quad \text{for some} \quad \mathbf{p}(0) \geq 0$$

   converges to the unique fixed point $\mathbf{p}(\boldsymbol{\eta}) = \bar{\mathbf{p}} = \mathbf{I}(\bar{\mathbf{p}})$.

- Component-wise increasing (decreasing) if $\mathbf{p}(0) = \mathbf{0}$ ($\mathbf{p}(0) \in \mathrm{P}(\boldsymbol{\eta})$).
- Amenable to distributed implementation, scalable, works for any SIF.
- Asynchronous operation possible.
- But how should new users join the network without disrupting the connections of active users?

Outline

Admission Control Scheme

- User k is called active at time n if $\mathrm{SIR}_k(\mathbf{p}(n)) \geq \gamma_k$.
- Define $\mathcal{A}_n$ to be the set of all active users at time n and $\mathcal{B}_n := \mathcal{K} \setminus \mathcal{A}_n$.

Power control with active link protection ($\delta > 1$)

$$p_k(n+1) = \begin{cases} \delta\,\gamma_k I_k(\mathbf{p}(n)) & k \in \mathcal{A}_n \text{ (active users)} \\ \delta\, p_k(n) = \delta^{n+1} p_k(0) & k \in \mathcal{B}_n \text{ (inactive users)} \end{cases}$$

- $\delta > 1$ can be interpreted as protection margin.
  - the larger δ, the faster power-up of the inactive users.
  - δ cannot be too large for all users to be fully admissible.

Bambos'00, Chee Wei Tan'09 (only $I_k(\mathbf{p}) = (\mathbf{V}\mathbf{p} + \mathbf{z})_k$)

Properties of the admission control scheme

Theorem

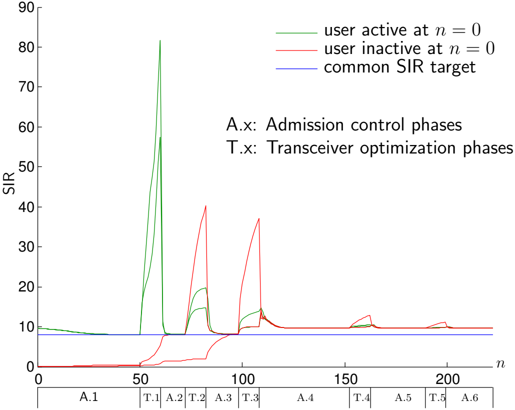
Let I_k be any standard interference function. Then,

- *All SIRs converge to some values.*
- *All users are admitted in finite time if $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_K)$ is feasible.*
- *Transmit powers are bounded if and only if $\delta \cdot \boldsymbol{\gamma}$ is feasible.*
- *Active users ($k \in \mathcal{A}_n$):*
  - *Preservation of active users: $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$.*
  - *Bounded power overshoot: $p_k(n+1) < \delta p_k(n)$.*

  Inactive users ($k \in \mathcal{B}_n$):
  - *SIRs of inactive users are increasing $\mathrm{SIR}_k(\mathbf{p}(n)) < \mathrm{SIR}_k(\mathbf{p}(n+1))$.*

Stanczak&Kaliszan&Bambos'09

No power constraints, TX and RX beamforming



- $K = 10, n_T = n_R = 4, \gamma = 8, \delta\gamma = 9.6, \mathcal{A}_n = \{1, \ldots, 5\}$
- The highest feasible SIR (example):
  - 0.88 (fixed beamformers), 1.37 (RX beamforming), 8 (TX/RX beamforming)

Incorporating power constraints

Theorem

Suppose that $\delta\gamma$ *is feasible and*

$$\mathbf{p}(m) \leq \beta\delta\mathbf{I}\big(\mathbf{p}(m)\big)$$

holds for some $m \in \mathbb{N}_0$ *and* $\beta \in [1, \beta_{\max}]$*. Then, there exists* $\beta_{\max} > 1$ *such that* $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ *for all* $n \geq m$*.*

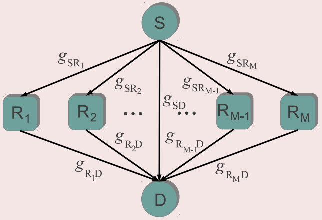
- Active users send distress signals until the condition is satisfied.

Stanczak&Kaliszan&Bambos'09

Outline

Energy-Efficient Relaying using Rateless Codes

Ravanshid&Lampe&Huber'11



- Half-duplex mode
- synchronous operation mode
- Frequency-flat AWGN channels
- For nodes $P, Q \in \{\mathsf{S}, \mathsf{D}, \mathsf{R}_1, \ldots, \mathsf{R}_M\}$, $g_{PQ} \in \mathbb{C}$ denotes the SNR at node Q when node P transmits with unit power.
- g_{PQ} are known at all nodes.

- n is the total number of time slots until the destination decodes the message
- $N \leq M$ is the number of relays that decoded the message during this time.
- n_m denotes the number of time slots until R_m decodes the message.
- The time fraction during which relay R_m listens to the source transmission is given by $\lambda_m = \frac{n_m}{n}$, whereas the transmission rate is $R = \frac{k}{n}$.

Optimal Rate

Let $\mathcal{C}(x) = \log(1+x)$, $C_0 = \mathcal{C}(p_{\mathsf{S}} g_{\mathsf{SD}})$ and $C_m = \mathcal{C}(p_{\mathsf{S}} g_{\mathsf{SR}_m})$ for $m \in \{1, \ldots, N\}$ for the capacities of the links originating from the source node S. Then,

$$J_i = \mathcal{C}\Big(p_{\mathsf{S}} g_{\mathsf{SD}} + \sum_{j=1}^{i} r_j p_j g_{\mathsf{R}_j \mathsf{D}}\Big) + \sum_{j=1}^{i} \mathcal{C}\left((1-r_j) p_j g_{\mathsf{R}_j \mathsf{D}}\right)$$

for the *collaborative capacity* of relays $\mathsf{R}_1, \ldots, \mathsf{R}_i$, the optimal rate is (given N)

$$R_N(\mathbf{P}) = \frac{J_N}{1 - \frac{C_0}{C_1} + \sum_{m=1}^{N-1} \left(\frac{1}{C_m} - \frac{1}{C_{m+1}}\right) J_m + \frac{J_N}{C_N}}.$$

Definition (Optimal rate)

For any $\mathbf{p}$, we define $R(\mathbf{p}) := R_{N(\mathbf{p})}(\mathbf{p})$ where

- $N(\mathbf{p}) = \max\{m \in \{1, \ldots, M\} : \lambda_m \leq 1\}$ is the optimal number of decoding relays (in the sense of maximizing the rate).

Problem Formulation

- For the power allocation $\mathbf{p}$ at the relays, the quantities

$$E_{\mathsf{R}_m}(\mathbf{p}) \quad = \quad \frac{(1-\lambda_m(\mathbf{p}))p_m}{R(\mathbf{p})} \text{ for } m \in \{1, \ldots, N(\mathbf{p})\}. \qquad (1)$$

  measure the energy spent for each transmitted bit at each relay node R_m.
- Given $\gamma_{\mathsf{R}_m} > 0, m \in \{1, \ldots, N\}$, maximize the transmission rate subject to the constraints on the energy-per-bit usage at the relay nodes R_m.

Problem (Problem 1)

$$\underset{\mathbf{p}\in\mathbb{R}^M_{++}}{\textit{maximize}}\ R(\mathbf{p})$$

$$\textit{subject to} \quad E_{R_m}(\mathbf{p}) \leq \gamma_{R_m}\ , m \in \{1, \ldots, N(\mathbf{p})\}.$$

Buehler&Stanczak'13

Relaxed Problem

Problem 1a

$$\underset{\mathbf{p}\in\mathbb{R}^M_{++}}{\text{maximize}}\ R(\mathbf{p})$$

$$\text{subject to} \quad \widehat{E}_{\mathsf{R}_m}(\mathbf{p}) \le \gamma_{\mathsf{R}_m}\ , m \in \{1,\dots,M\}.$$

where

$$\widehat{E}_{\mathsf{R}_m}(\mathbf{p}) = \frac{p_m}{R(\mathbf{p})} \text{ for } m \in \{1,\dots,M\}\,. \tag{2}$$

Proposition (Buehler&Stanczak)

The function R is a standard interference function in the relay powers $\mathbf{p}$.

Algorithmic Solution

Corollary

If Problem 1a is feasible for the Eb constraints $\gamma_{R_m} > 0, m \in \{1, \ldots, N\}$*, then the algorithm*

$$\mathbf{p}(n+1) = (\gamma_{R_1} R(\mathbf{p}(n)), \ldots, \gamma_{R_M} R(\mathbf{p}(n)))$$

converges to the optimal solution of Problem 1a, which is also a feasible (but generally suboptimal) power allocation for Problem 1.

Buehler&Stanczak'13

Outline

Positive Concave Functions

- It may be difficult to show the axioms, while concavity can be easily to verify.
- In such cases, the following proposition may be useful.

Proposition (Cavalcante'13)

If $\mathbf{I}(\mathbf{p}) > 0$ *is concave in* $\mathbf{p} > 0$*, then the map is a standard interference function.*

- Example: The load in LTE systems can be shown to be a fixed point of some positive and concave function. The axiomatic framework for standard interference functions can be used to
  - show the existence and uniqueness of a fixed point,
  - check feasibility of different SON configurations,
  - compute the load vector by means of the fixed-point algorithm.

Cavalcante&Pollakis&Stanczak'13

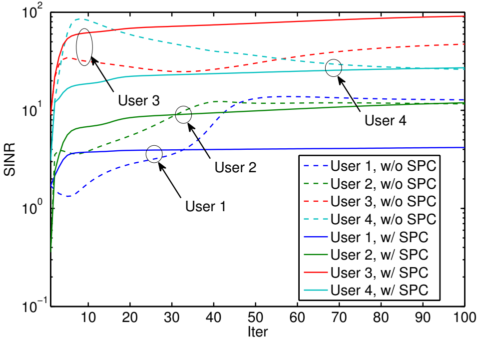
Outline

Problem statement

$$\max_{k} \min_{k} \mathrm{SIR}_k = \max_{k} \min_{k} \frac{p_k}{I_k(\mathbf{p})} \quad \forall_k p_k \leq \hat{p}_k$$

1. Power control for any fixed TX and RX beamformers [Aein '73,...]
   - $I_k(\mathbf{p}) = (\mathbf{V}\mathbf{p} + \boldsymbol{\sigma})_k, \mathbf{V} \geq 0, \mathbf{V} \neq \mathbf{V}^T$
2. Joint power control and RX beamforming [Zander '01,...]
   - $I_k(\mathbf{p}) = \min_{\mathbf{r}}(\mathbf{V}(\mathbf{r}) \cdot \mathbf{p} + \boldsymbol{\sigma}(\mathbf{r}))_k$
3. Joint power control and transceiver optimization [e.g. Chang et al. '02, Stanczak et al. '08]
   - Monotonicity and convergence due to power control
   - Similar approach achieves interference alignment [Gomadam et al. '08]
   - Extensions to general MIMO [Sezgin,Stanczak'11]

Max-Min SINR vs. Max SINR of [Gomadam et al. '08]



$$K = 4,\ N_{TX} = 3,\ N_{RX} = 2$$

- How to solve the power control part in a distributed manner?
  - Solution is a positive eigenvector of some irreducible matrix.

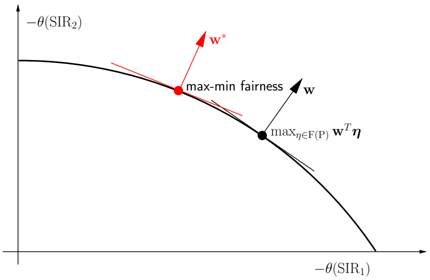
Outline

An alternative problem formulation

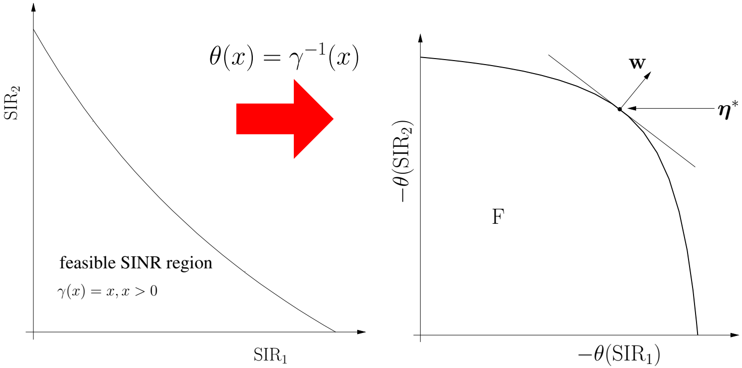
- Instead of maximizing $\min_k \mathrm{SIR}_k(\mathbf{p})$, consider

$$\mathbf{p}^* = \arg\min_{\mathbf{p}\in \mathrm{P}} \sum_k w_k \theta\Big(\frac{p_k}{I_k(\mathbf{p})}\Big)$$

- $\mathbf{u} = (\theta(\mathrm{SIR}_1), \dots, \theta(\mathrm{SIR}_K))$
- θ is concave and strictly decreasing
- Compute $\mathbf{w}^*$ in parallel to $\mathbf{p}^*$!



Convexity of Utility Set



- Specify a class of θ so that F is a convex set.

Characterization of Utility Set

Theorem

$\mathbf{u}$ *is feasible* ($\mathbf{u} \in \mathcal{U}$) *if and only if*

$$\max_k \rho(\mathbf{D}(\mathbf{u})\mathbf{G}_k) \leq 1$$

- $\mathbf{D}(\mathbf{u}) = \mathrm{diag}(\gamma(u_1), \ldots, \gamma(u_K))$
- $\gamma(x) = \theta^{-1}(x)$
- $\mathbf{G}_k = \mathbf{V} + \frac{1}{\hat{p}_k}\mathbf{z}\mathbf{e}_k^T$

➠ Perron-Frobenius theory.

One Slide Tutorial on Theory of Nonnegative Matrices

$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	• $\{0, 0\}$ • $\{(1,0), (0,0)\}$	• $\rho(\mathbf{B}) = 0$ • nonnegative eigenvector
$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	• $\{1, -1\}$ • $\{(1,1), (1,-1)\}$	• $\rho(\mathbf{B}) = 1$ (simple) • positive eigenvector
$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$	• $\{\frac{1}{2}(1+\sqrt{5}), \frac{1}{2}(1-\sqrt{5})\}$ • $\{(\lambda_1 - 1, 1), (\lambda_2 - 1, 1)\}$	• $\lvert\lambda_2\rvert < \rho(\mathbf{B})$ (simple) • positive eigenvector

$\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$ $\quad$ $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\quad$???

- Irreducibility plays a crucial role (closed under summation)
- Many applications: algebraic graph theory, FMC, stochastic matrices ...
  - Does $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{b}, \mathbf{b} \geq 0$, have a positive solution?
- Books of Seneta, Minc, Gantmacher,...

A Useful Spectral Radius Characterization

Theorem (Arnold, Gundlach, Demetrius '94)

Let $\mathbf{B} = (x_{k,l}) \geq 0$ *be irreducible. Then,*

$$\log \rho(\mathbf{B}) = \sup_{\mathbf{A} \in \mathcal{A}(\mathbf{B})} \left(\sum_{k,l} q_k a_{k,l} \log \frac{b_{k,l}}{a_{k,l}} \right)$$

where $\mathbf{q} = (q_1, \ldots, q_K) \in \Pi_K^+$ *is the left Perron eigenvector of* $\mathbf{A} = (a_{k,l})$.

➠ Useful bounds like $\rho(\mathbf{X} \circ \mathbf{Y}) \leq \rho(\mathbf{X})\rho(\mathbf{Y})$.

A Sufficient Convexity Condition

Theorem

$\rho(\mathbf{D}(\mathbf{u})\mathbf{G}_i)$ is log-convex for each i if $\mathbf{D}(\mathbf{u})$ is log-convex in $\mathbf{u}$ or, equivalently, if γ is a log-convex function.

- $\gamma(x) = \theta^{-1}(x)$ is log-convex if and only if $\theta(e^x)$ is convex.

Corollary

The following holds:

- *$\rho(\mathbf{D}(\mathbf{u})\mathbf{G}_i)$ is log-convex for each i if $\theta(e^x)$ is convex.*
- *The utility set is convex if $\theta(e^x)$ is convex.*

A Necessary Convexity Condition

Theorem

If $\rho(\mathbf{D}(\mathbf{u})\mathbf{G})$ is convex for any irreducible nonnegative matrix $\mathbf{G}$, then $\theta(e^x)$ is convex.

- The theorem implies that if the spectral radius $\rho(\mathbf{D}(\mathbf{u})\mathbf{G}_i), 1 \leq i \leq K$, which determines the feasible utility set, is required to be a convex function of $\mathbf{u}$ for any interference coupling (channel realization), then θ must be chosen such that $\theta(e^x)$ is convex.
- Log-convexity of $\gamma(x)$ seems to be essential for the utility maximization problem to be tractable and solvable in an efficient way.

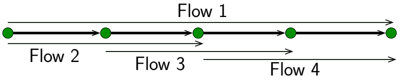
Examples of Function Classes

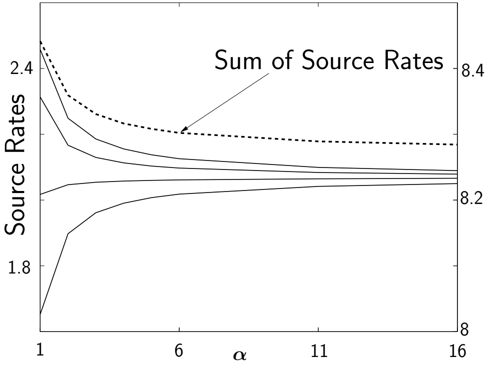
$$\Psi_\alpha(x) = \begin{cases} \frac{x^{1-\alpha}}{1-\alpha} & \alpha > 1 \\ \log(x) & \alpha = 1 \end{cases} \quad \tilde{\Psi}_\alpha(x) = \begin{cases} \log x & \alpha = 1 \\ \log \frac{x}{1+x} & \alpha = 2 \\ \log \frac{x}{1+x} + \sum_{j=1}^{\alpha-2} \frac{1}{j(1+x)^j} & \alpha > 2 \end{cases}$$

Max-Min Rate Allocation

Arbitrarily Close Approximation

Let $\eta_k^* = \Psi_\alpha(\mathrm{SIR}_k^*)$ and let $\nu_k^* = \log(1 + \mathrm{SIR}_k^*)$. Then, $\boldsymbol{\nu}^*$ converges to the max-min rate allocation as $\alpha \to \infty$.





Utility-Based Power Control

Equivalent minimization problem

$$\mathbf{p}^* = \arg\min_{\mathbf{p}\in\mathrm{P}} F(\mathbf{p}) = \arg\min_{\mathbf{p}\in\mathrm{P}} \sum\nolimits_k w_k \theta\big(\mathrm{SIR}_k(\mathbf{p})\big)\,.$$

- Positivity of minimizers: $\mathbf{p}^* > 0$
- Even if $\theta(e^x)$ is convex, the problem is not convex in general.

Convex Statement of the Problem

Theorem

If $I_k(e^{\mathbf{s}})$ is log-convex and $\theta(e^x)$ convex, the following problem is convex:

$$\mathbf{s}^* = \underset{\mathbf{s}\in \mathrm{S}}{\arg\min} F_e(\mathbf{s}) \qquad \begin{cases} \mathbf{s} := \log \mathbf{p}, \mathbf{p} > 0 \\ \mathrm{S} := \{\log \mathbf{x} : \mathbf{x} \in \mathrm{P}_+\} \\ F_e(\mathbf{s}) = F(e^{\mathbf{s}}) \end{cases}$$

- $I_k(e^{\mathbf{s}}) = \sum_l v_{k,l} e^{s_l} + z_k$ is log-convex (Hoelder inequality).
- Log-convexity is given in the worst-case design.

Gradient-Projection Algorithm

- Let $\tau > 0$ be constant step size (small enough), and let

$$\mathbf{s}(n+1) = \Pi_{\mathrm{S}}\Big[\mathbf{s}(n) - \tau \nabla F_e(\mathbf{s}(n))\Big] \qquad \mathbf{s}(0) \in \mathrm{S}$$

- $\nabla F_e(\mathbf{s}) = \mathrm{diag}(e^{s_1}, \ldots, e^{s_K}) \nabla F(e^{\mathbf{s}})$:

$$\nabla F(\mathbf{p}) = (\mathbf{I} - \mathbf{V}^T \boldsymbol{\Gamma}(\mathbf{p}))\mathbf{g}(\mathbf{p})$$

  - $g_k(\mathbf{p}) = w_k \theta'(\mathrm{SIR}_k(\mathbf{p}))\mathrm{SIR}_k(\mathbf{p})/p_k$ (locally available)
  - $\boldsymbol{\Gamma}(\mathbf{p}) = \mathrm{diag}(\mathrm{SIR}_1(\mathbf{p}), \ldots, \mathrm{SIR}_K(\mathbf{p}))$

Computation of the Gradient Vector

$$\nabla F(\mathbf{p}) = \underbrace{(\mathbf{I}+\mathbf{\Gamma}(\mathbf{p}))\mathbf{g}(\mathbf{p})}_{\text{local variable}} - \underbrace{(\mathbf{I}+\mathbf{V}^T)\mathbf{\Gamma}(\mathbf{p})\mathbf{g}(\mathbf{p})}_{\text{global variable}}$$

- Problem is to obtain $\Sigma_k(\mathbf{p}) = \sum_l v_{l,k} m_l(\mathbf{p}), m_l(\mathbf{p}) = g_l(\mathbf{p})\mathrm{SIR}_l(\mathbf{p})$.
  - Distribute m_l using a flooding protocol (What about $v_{l,k}$?).
  - Estimate the sum Σ_k using an adjoint network.

The Notion of an Adjoint Network

Definition

Two networks with K links and gain matrices $\mathbf{V}_1$ and $\mathbf{V}_2$ are referred to as being adjoint (to each other) if $\mathbf{V}_1 = \mathbf{V}_2^T$.

- Reverse the roles of transmitter and receivers is not sufficient

$$\underbrace{\mathbf{V}_1 = \mathbf{DG}}_{\text{primal network}} \qquad \underbrace{\mathbf{V}_2 = \mathbf{DG}^T}_{\text{reversed network}} \qquad \mathbf{V}_1 \neq \mathbf{V}_2^T$$

➠ In addition, each user in the reversed network needs to inverse its channel:

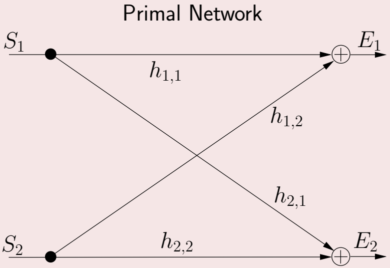
$$\underbrace{\mathbf{V}_1 = \mathbf{DG}}_{\text{primal network}} \qquad \underbrace{\mathbf{V}_2 = \mathbf{G}^T\mathbf{D}}_{\text{adjoint network}} \qquad \mathbf{V}_1 = \mathbf{V}_2^T$$

- $\mathbf{D} = \mathrm{diag}(\frac{1}{V_{1,1}}, \ldots, \frac{1}{V_{K,K}})$
- $\mathbf{G} = (V_{k,l})$ if $k \neq l$ and $(\mathbf{G})_{k,k} = 0$.

Adjoint Networks: A Simple Example

- Primal network

$$\mathbf{V}_1 = \begin{pmatrix} 0 & \frac{|h_{1,2}|^2}{|h_{1,1}|^2} \\ \frac{|h_{2,1}|^2}{|h_{2,2}|^2} & 0 \end{pmatrix}$$



Adjoint Networks: A Simple Example

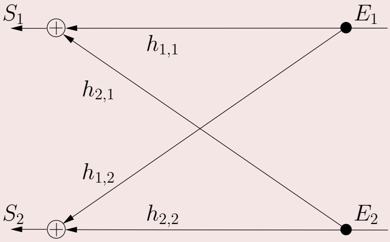
- Primal network

$$\mathbf{V}_1 = \begin{pmatrix} 0 & \frac{|h_{1,2}|^2}{|h_{1,1}|^2} \\ \frac{|h_{2,1}|^2}{|h_{2,2}|^2} & 0 \end{pmatrix}$$

- Reversed network + $X_k/h_{k,k}$

$$\mathbf{V}_2 = \begin{pmatrix} 0 & \frac{|h_{2,1}|^2}{|h_{2,2}|^2} \\ \frac{|h_{1,2}|^2}{|h_{1,1}|^2} & 0 \end{pmatrix}$$

Adjoint Network



The message m_k is contained in the transmit power

Alternate Use of Primal and Adjoint Networks

1. Concurrent transmission of training sequences at powers $p_k(n), k \in \mathcal{K}$.
2. Receiver side estimation of SIRs and interferences. The receivers calculate $g_k(\mathbf{p}(n)), k \in \mathcal{K}$, and feed the SIRs back to the transmitters using a control channel. Transmitter-side computation of $g_k(\mathbf{p}(n))$.
3. Concurrent transmission over the adjoint network of zero-mean random symbols at powers $|\mathrm{SIR}_k(\mathbf{p}(n)) \cdot g_k(\mathbf{p}(n))|, k \in \mathcal{K}$.
4. Transmitter side estimation of the received power and subtraction of noise variances from the estimates. The transmitters compute

$$\nabla_k F(\mathbf{p}(n)) = g_k(\mathbf{p}(n)) - (\mathbf{V}^T \mathbf{\Gamma}(\mathbf{p}(n))\mathbf{g}(\mathbf{p}(n)))_k$$

5. Update of transmit powers with $\mathbf{s}(n) = \log \mathbf{p}(n)$;$n \to n+1$.

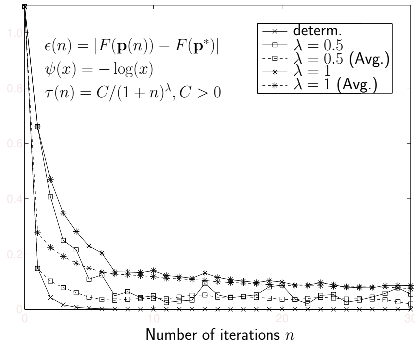
Noisy Measurements

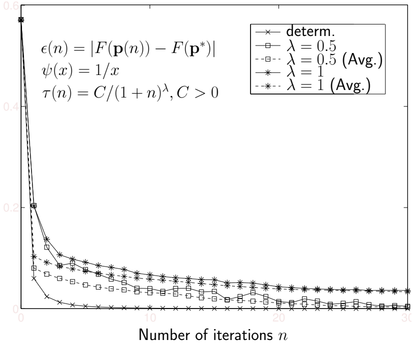
- Only noisy observations are available:

$$\Delta_k(n) = \nabla_k F(\mathbf{p}(n)) \quad + \quad \underbrace{\delta M_k(n)}_{\text{estimation noise}}$$

- Analysis in the framework of stochastic approximation.
- Use a diminishing step size $\{\tau(n)\}, \tau(n) > 0$ [Kushner'03]:
  - non-increasing sequence with $\lim_{n\to\infty} \tau(n) = 0$
  - $\sum_{n=0}^{\infty} \tau(n) = \infty$.
  - Common assumption: $\sum_{n=0}^{\infty} \tau(n)^2 < +\infty$.
- To improve the initial convergence rate, one may utilize averaging of iterates in parallel to the stochastic recursion [Polyak'92].

Noisy Measurements: Simulations





An Alternative Algorithmic Approach

- Gradient projection algorithm
  - Advantages: simplicity, monotonicity
  - Disadvantages: step size, only linear convergence rate, projection
- An alternative approach: Conditional Newton iteration for finding stationary points of a modified Lagrangian function.

Min-max reformulation

$$\min_{\mathbf{s},\mathbf{t}} \max_{\mathbf{u}} \sum_k w_k \theta\Big(\frac{e^{s_k}}{u_k}\Big) \quad \text{subject to} \begin{cases} e^{\mathbf{s}} - \hat{\mathbf{p}} \leq 0 \Leftrightarrow \mathbf{s} \in \mathrm{S} \\ \mathbf{u} - \mathbf{t} \leq 0 \\ \forall_k \, I_k(e^{\mathbf{s}}) - t_k = 0 \,. \end{cases}$$

- Linear interference function: $I_k(e^{\mathbf{s}}) = (\mathbf{V}e^{\mathbf{s}} + \mathbf{z})_k$.
    - The Hessian is diagonal and its diagonals are given by the gradient.
- Use theory of max-min/convex-concave functions and theory of non-linear Lagrangians to improve the convergence speed/rate.

Conditional Newton Algorithm

$$\begin{cases} \begin{pmatrix} \mathbf{s}(n+1) \\ \boldsymbol{\mu}(n+1) \end{pmatrix} = \begin{pmatrix} \mathbf{s}(n) \\ \boldsymbol{\mu}(n) \end{pmatrix} - \left(\nabla^2_{(\mathbf{s},\boldsymbol{\mu})} L(\mathbf{z}(n))\right)^{-1} \nabla_{(\mathbf{s},\boldsymbol{\mu})} L(\mathbf{z}(n)) \\ \nabla_{(\mathbf{u},\boldsymbol{\lambda}^u,\boldsymbol{\lambda},\mathbf{t})} L(\mathbf{z}(n+1)) = 0 \quad \text{can be solved explicitely} \end{cases}$$

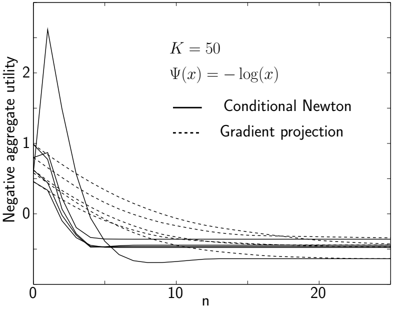
- $L(\mathbf{z}) = L(\mathbf{s}, \mathbf{u}, \boldsymbol{\mu}, \boldsymbol{\lambda}^u, \boldsymbol{\lambda}, \mathbf{t})$: A nonlinear Lagrangian with no nonnegative constraints on dual variables
  - $\mathbb{R} \to \mathbb{R}_+ : \lambda \to \psi(\lambda)$ with $\psi(x) = x^2, x \in \mathbb{R}$.

Theorem

The algorithm converges to a global optimum if $\theta(x) = -\log(x)$ and $\theta(x) = 1/x$. The convergence rate is quadratic.

A Comparison with a Gradient-Projection Algorithm

- Advantages:
  - No step size control
  - Quadratic convergence rate
  - Unconstrained iteration
  - Distributed implementation possible via adjoint network
- Disadvantages:
  - Monotonicity is not guaranteed



- Significantly better local convergence rate but no monotonicity.

Incorporation of QoS requirements

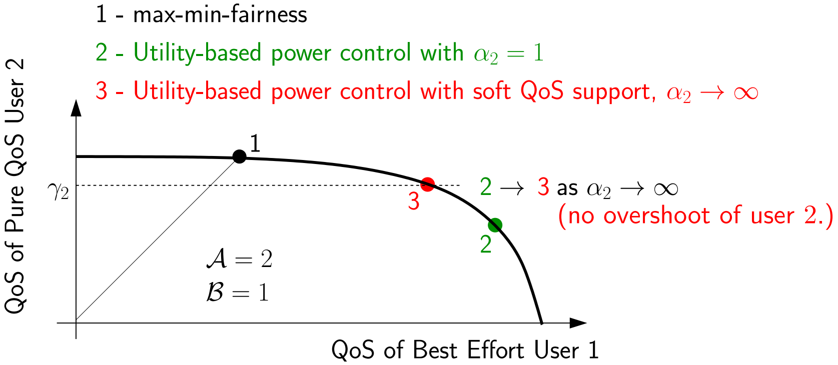
Soft QoS Support

$$\tilde{F}_\alpha(\mathbf{p}) = \underbrace{\sum_{k\in\mathcal{A}} a_k\theta_\alpha\Big(\frac{\mathrm{SIR}_k(\mathbf{p})}{\gamma_k}\Big)}_{\text{penalty term}} + \underbrace{\sum_{k\in\mathcal{B}} b_k\theta\big(\mathrm{SIR}_k(\mathbf{p})\big)}_{\text{aggregate utility}} \,.$$

- $\mathcal{A}$: QoS users need to satisfy $\mathrm{SIR}_k \geq \gamma_k, k \in \mathcal{A}$
- $\mathcal{B}$: best-effort users
  - $\mathcal{A} \setminus \mathcal{B}$: pure QoS users (voice)
  - $\mathcal{A} \cap \mathcal{B}$: best-effort users with QoS requirements (video)
  - $\mathcal{B} \setminus \mathcal{A}$: pure best-effort users (data)

- Each user, say user k, determines its utility by choosing $\alpha_k \geq 1$.

Soft QoS Support: Example



- Alternative: Suitable weighting achieves max-min fairness for any utilities.

How to choose/compute the weights?

Theorem (Friedland, Karlin '75, Boche, Stanczak '04)

Let $\mathbf{B} \geq 0$ *be irreducible, and let* $\mathbf{w} = \mathbf{x} \circ \mathbf{y} \in \Pi_K^+$. *If* $\theta(e^x)$ *is convex, then*

$$\forall_{\mathbf{s} \in \mathbb{R}_{++}^K} \ \sum_k w_k \theta\left(\frac{s_k}{(\mathbf{Bs})_k}\right) \geq \theta(1/\rho(\mathbf{B}))$$

Equality holds if $\mathbf{s} = \mathbf{x} > 0$. *Moreover, if the equality holds in the optimum for some weight vector* $\mathbf{z}$, *then* $\mathbf{z} = \mathbf{w}$.

- The approach may not be meaningful for an algorithmic solution.
- Compute $\mathbf{w}^*$ iteratively in parallel to the power iteration.

Collatz-Wielandt Characterization of the Perron Root

Theorem

Let $\mathbf{B} \geq 0$ *be irreducible, and let* $\theta(e^x)$ *be convex. Then,* $(\mathbf{x}, \mathbf{w}) \in \mathbb{R}^K_{++} \times \Pi^+_K$ *is a saddle point:*

$$\theta(1/\rho(\mathbf{B})) = \min_{\mathbf{s}\in\mathbb{R}^K_{++}} \max_{\mathbf{z}\in\Pi^+_K} \sum_k z_k \theta\left(\frac{s_k}{(\mathbf{Bs})_k}\right) = \max_{\mathbf{z}\in\Pi^+_K} \min_{\mathbf{s}\in\mathbb{R}^K_{++}} \sum_k z_k \theta\left(\frac{s_k}{(\mathbf{Bs})_k}\right)$$

$$= \sum_k w_k \theta\left(\frac{x_k}{(\mathbf{Bx})_k}\right)$$

- $\mathbf{x} > 0$ *is unique up to positive multiples,*
- $\mathbf{w} = \mathbf{y} \circ \mathbf{x}$ *is a unique vector in* $\Pi^+_K = \{\mathbf{x} > 0, \sum_k x_k = 1\}$.

Max-min weight characterization

Theorem

Let $\bar{\mathbf{p}}$ and $\bar{\mathbf{q}}$ are principal right and left eigenvectors of $\mathbf{B}^{(k_0)}$. Let

- $\mathbf{w}^* = \bar{\mathbf{q}} \circ \bar{\mathbf{p}} > 0$ *and*
- $\theta(e^x)$ *be convex in* $x \in \mathbb{R}$.

Then $\mathbf{p}^* = \bar{\mathbf{p}}$ *(max-min fair power control).*

- The approach may not be meaningful for an algorithmic solution.
- Compute $\mathbf{w}^*$ iteratively in parallel to the power iteration.

Saddle-point algorithm

- Design a saddle-point power control algorithm that converges to $(\mathbf{w}^*, \bar{\mathbf{p}})$.
  - $\mathbf{p}, \mathbf{w}$ and the dual variables are iterated simultaneously.

Min-max formulation

$$\min_{\mathbf{s}} \max_{\mathbf{w}} \sum\nolimits_k w_k \theta\Big(\frac{e^{s_k}}{I_k(e^{\mathbf{s}})}\Big) \quad \text{subject to} \begin{cases} e^{\mathbf{s}} - \hat{\mathbf{p}} \le 0 \\ \|\mathbf{w}\|_1 - 1 = 0, \mathbf{w} \ge 0\,. \end{cases}$$

Theorem

A saddle-point algorithm operating on a classical linear constrained Lagrangian converges to a global optimum that is a saddle point given by

$$\max_{\mathbf{u}\in\Pi_K} \min_{\mathbf{s}\in\mathrm{S}} \sum_{k\in\mathcal{K}} w_k \theta\Big(\frac{e^{s_k}}{I_k(e^{\mathbf{s}})}\Big) = \min_{\mathbf{s}\in\mathrm{S}} \max_{\mathbf{u}\in\Pi_K} \sum_{k\in\mathcal{K}} w_k \theta\Big(\frac{e^{s_k}}{I_k(e^{\mathbf{s}})}\Big)$$

Outline

Joint Power Control and RX Beamforming

- Optimize RX beamformers alternately with $(\mathbf{w}, \mathbf{p})$

Alternating optimization

Require: $n = 0, \mathbf{r}_k(0) \in \mathrm{U}_k, k \in \mathcal{K}, \mathbf{p}(0) \in \mathrm{P}, \mathbf{u}(0) \in \Pi_K^+$

repeat

$\quad n = n + 1$

$\quad \mathbf{p}(n) = \arg\min_{\mathbf{p} \in \mathrm{P}} \max_{\mathbf{u} \in \Pi_K^+} G(\mathbf{u}, \mathbf{p}, \mathbf{R}(n-1))$

$\quad \mathbf{r}_k(n) = \arg\max_{\mathbf{r} \in \mathrm{U}_k} \mathrm{SIR}_k(\mathbf{p}(n), \mathbf{r}) \quad \forall_k$

until termination condition is satisfied

Convergence of the algorithm

Theorem

The proposed algorithm converges to a global optimum of the max-min SIR balancing problem over the joint space of transmit powers and receive beamformers.

- What about transmit beamforming?
  - A big challenge mainly due to the lack of the uplink-downlink duality

Outline

Power with TX and RX Beamforming

Swapping roles

(i) Given TX beamformers, perform joint power control and RX beamforming

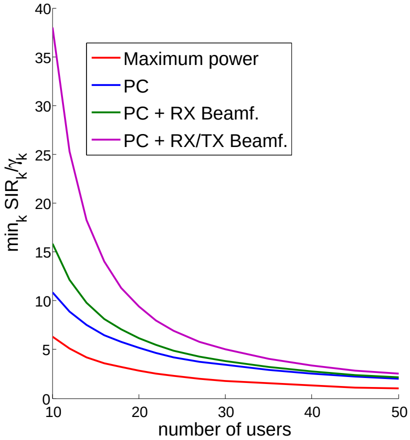
(ii) Swap the roles of transmitters and receivers, and go to (i)

- If QoS power control is applied $p_k(n+1) = I_k(\mathbf{p}(n))$, then
  - the scheme amenable to distributed implementation
  - each SIR increases monotonically
  - provides interference alignment for SNR $\to \infty$.

Outline

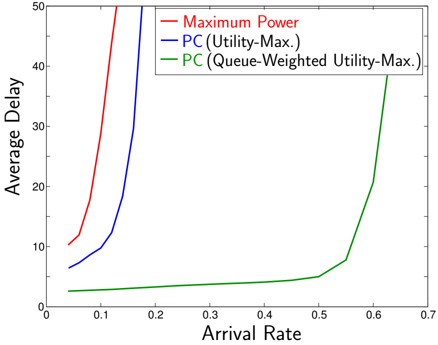
1. Some Definitions
2. Applications of Standard Interference Functions
   - Fixed-Point Power Control Algorithm
   - Admission Control
   - Other Applications of Interference Functions
   - A Useful Sufficient Condition
3. Affine Interference Functions
   - Power control with fixed beamformers
     - Fixed weights
     - Improving Convergence Rate
     - Utility-based optimization with QoS support
     - Max-min fair weights
   - Joint power control and receive beamforming
   - Joint power control and transceiver optimization
   - Simulations

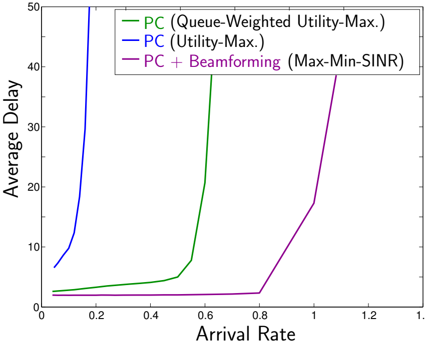
The Worst-Case Performance



- 4 RX and TX antennas,
- SNR$=$30 dB,
- relatively strong interference
- 4 iterations on average

Average Delay





Part II

Time-Varying Channels

Outline

Motivation

- Typically, algorithms for distributed optimization assume that the cost-function is fixed (i.e., it does not change during the iterations of the algorithm).
- However, in large-scale networks it may be unrealistic to solve (approximately) each instance of the optimization problem whenever the cost function changes.
  - Too much energy may be spent in coordination.
  - By the time one instance of the optimization problem is solved, the optimization problem may have changed substantially.
- Changing the cost function at each iteration in an ad-hoc fashion may lead to unexpected results (even if there is a time-invariant solution).

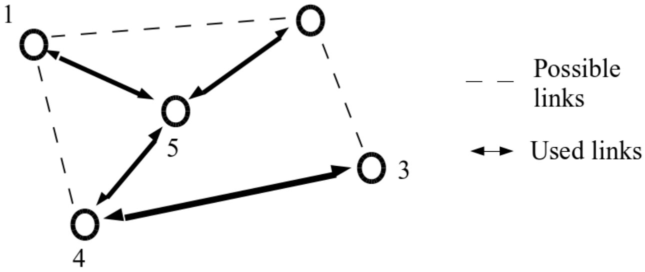
Basic Idea

Two-step Algorithmic Solution

1) **Local optimization step:** All nodes estimates the (same) parameter of interest by optimizing their local functions
2) **Consensus step:** Nodes exchange and fuse some information to improve their estimates. In each node the estimate resulting from cooperation should be better than that obtained with the node working alone.

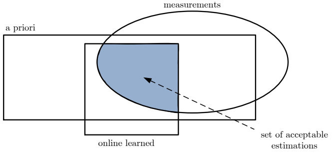
- The information exchange is limited by the network topology.

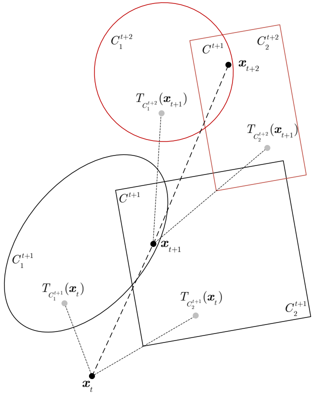
Diffusion Networks



- The diffusion mode of cooperation
  - The nodes exchange information with their neighbors
  - No node has access to all information available to the network
  - Links are possibly unreliable

- Each node constructs/updates one or multiple sets in online fashion from measurements and information received from other nodes.
- Additional sets are constructed from a-priori knowledge.
- The sets are constructed so that their intersection contains an optimal estimate.
- Each node projects its estimate on its local sets.
- Information is exchanged for the process to converge to or track a common value in the intersection.





Outline

Fundamental theorem of POCS (Gubin *et al.* '67)

Let C_i $(i = 1, \ldots, M)$ denote $M < \infty$ closed convex sets in a Hilbert space. Assume that

$$C_0 := \bigcap_{i=1}^{M} C_i \neq \emptyset.$$

Then for every $\boldsymbol{x} \in \mathcal{H}$, the sequence

$$\{\boldsymbol{x}_n := T^n(\boldsymbol{x})\},$$

where $T := T_{C_1} T_{C_2} \cdots T_{C_M}$, converges weakly to a point $\boldsymbol{p}$ of C_0.

$$\lim_{n \to \infty} < \boldsymbol{x}_n, \boldsymbol{z} > = < \boldsymbol{p}, \boldsymbol{z} >, \quad \forall \boldsymbol{z} \in \mathcal{H}$$

Fundamental theorem of POCS (Gubin *et al.* '67)

Let C_i $(i = 1, \ldots, M)$ denote $M < \infty$ closed convex s

In a finite-dimensional Hilbert space

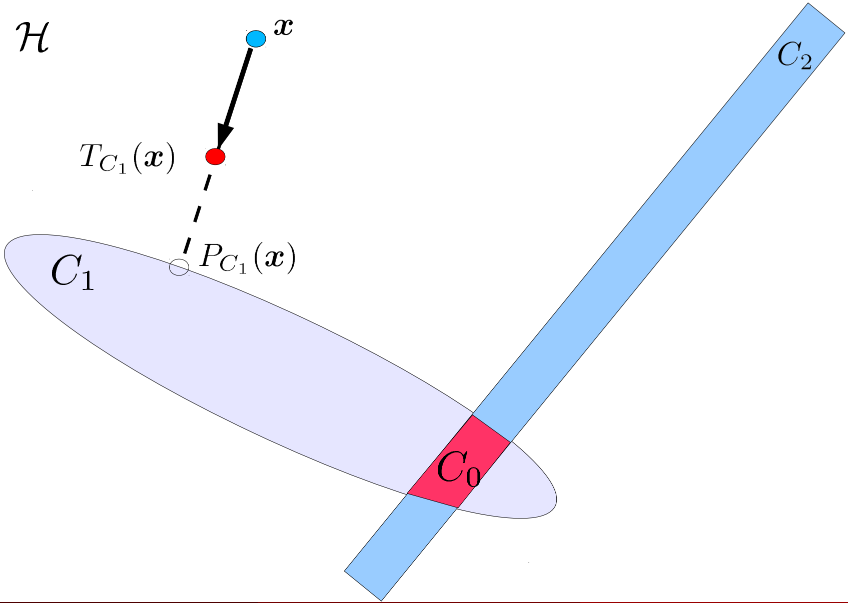
$$\lim_{n \to \infty} < \boldsymbol{x}_n, \boldsymbol{z} > = < \boldsymbol{p}, \boldsymbol{z} >, \quad \forall \boldsymbol{z} \in \mathcal{H}$$

$$\Updownarrow$$

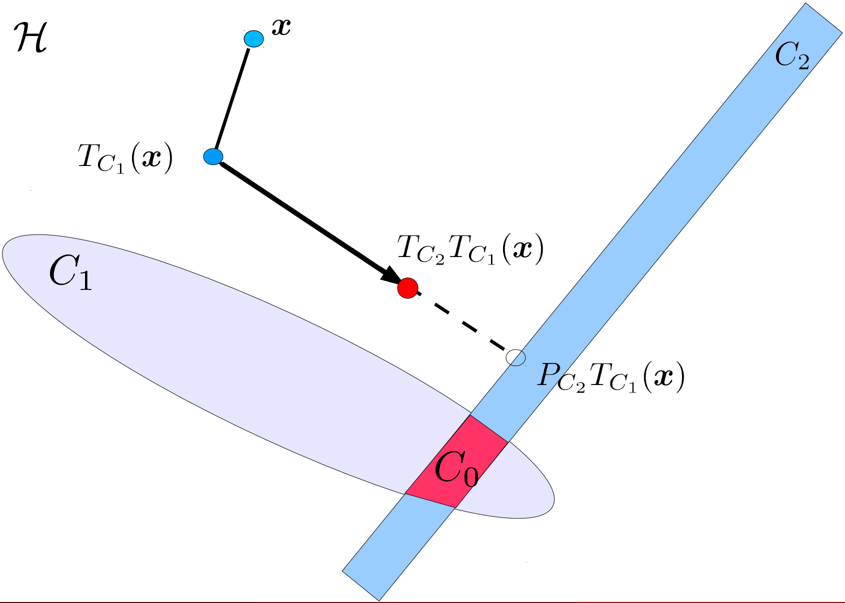
$$\lim_{n \to \infty} \|\boldsymbol{x}_n - \boldsymbol{p}\| = 0$$

$$\lim_{n \to \infty} < \boldsymbol{x}_n, \boldsymbol{z} > = < \boldsymbol{p}, \boldsymbol{z} >, \quad \forall \boldsymbol{z} \in \mathcal{H}$$

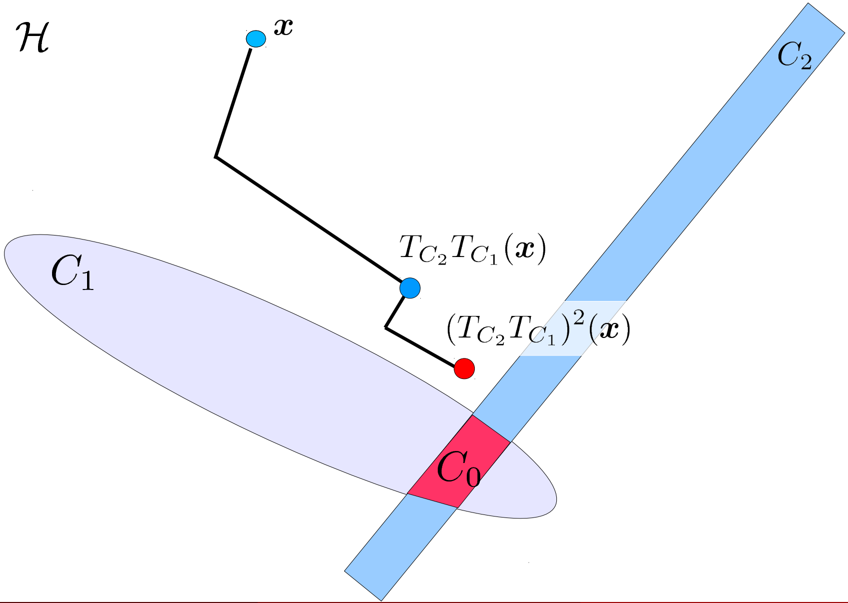
Geometric interpretation



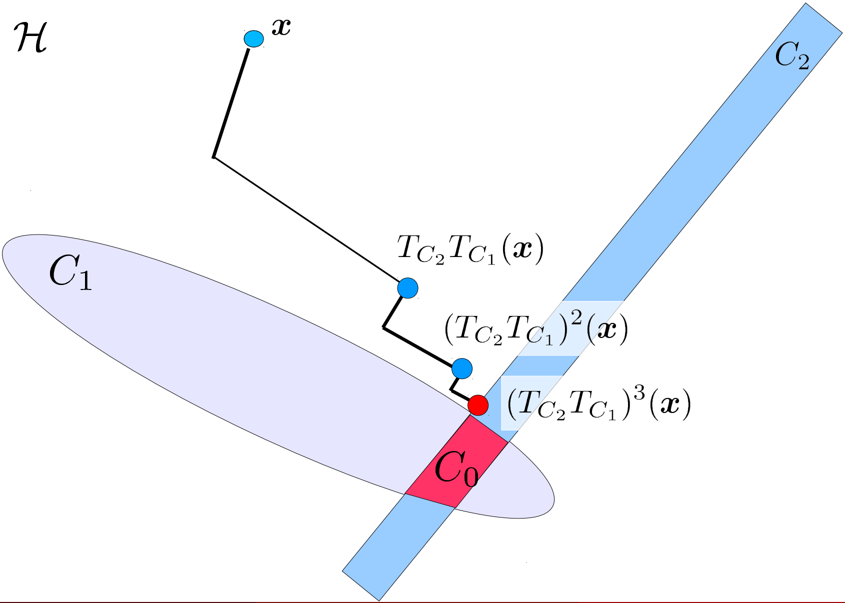
Geometric interpretation



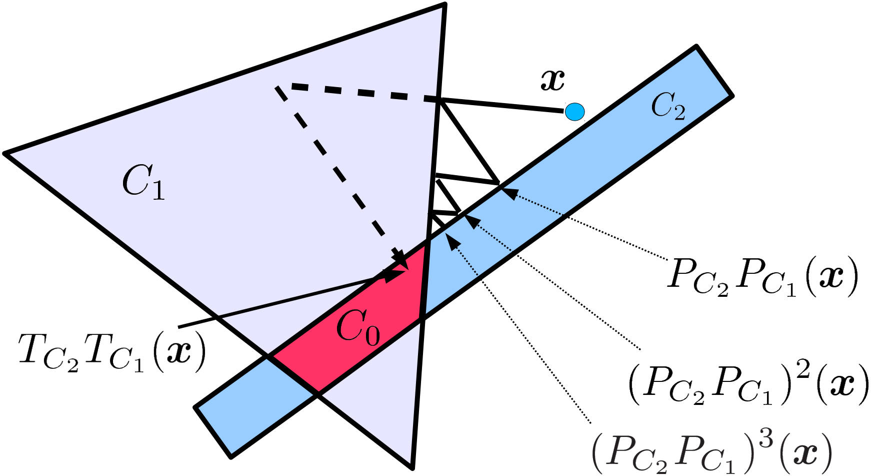
Geometric interpretation



Geometric interpretation



Projections vs relaxed projections



Limitations of the POCS algorithms

- The POCS theory assumes a finite number of closed convex sets.
- The fundamental theorem of POCS cannot be applied to cases with time-varying closed convex sets.
  - Time-varying sets are used to capture scenarios with time-varying channels.
- In such cases we can use the adaptive projected subgradient method.

Yamada '03,

Yamada and Ogura '04

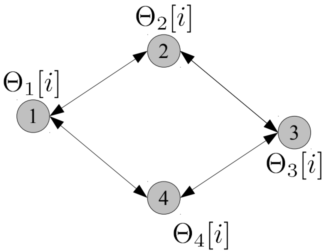
Slavakis, Yamada, and Ogura '06

Outline

Problem Statement

- N agents generate a sequence of optimization problems indexed by i.

$$\min. \Theta[i](\boldsymbol{h})$$



where

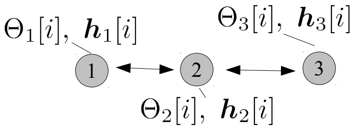
$$\Theta[i](\boldsymbol{h}) = \sum_{k=1}^{N} \Theta_k[i](\boldsymbol{h})$$

$\Theta_k[i] : \mathbb{R}^M \to [0, \infty)$ ***Local*** convex function known by agent k (private information)

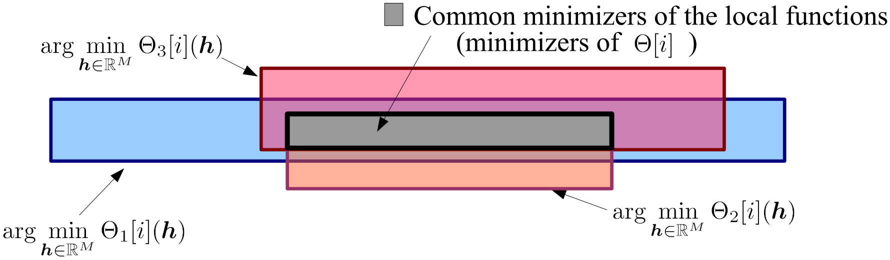
Assumptions (1)

- **Each agent has its own estimate $\boldsymbol{h}_k[i]$ of a minimizer of the *global function*:**

$$\Theta[i](\boldsymbol{h}) = \sum_{k=1}^{N} \Theta_k[i](\boldsymbol{h})$$

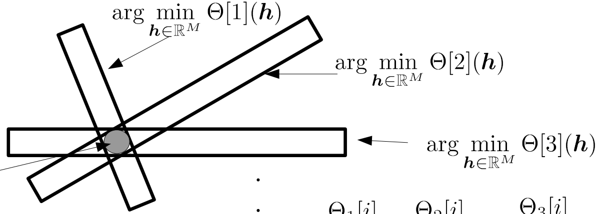


- **At each time i, the local functions have a common (nonempty) set of minimizers**



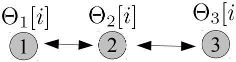
Assumptions (2)

- **(*time structure*) The global functions $\Theta[1],\ \Theta[2],\ \ldots$ have minimizers in common**



Time-invariant minimizers

$$\cap_{i\geq 0} \cap_{k=1}^{N} \Upsilon_k[i] \neq \emptyset$$



$$\Upsilon_k[i] := \left\{ h \in \mathbb{R}^M \mid \Theta_k[i](h) = \Theta_k^\star[i] := \inf_{h\in\mathbb{R}^M} \Theta_k[i](h) \right\}$$

$$\Theta[i](h) = \sum_{k=1}^{N} \Theta_k[i](h)$$

As a result, there exists $h \in \mathbb{R}^M$ that is a minimizer of every local function and every global function $\Theta[i]$ at any time instant i...
===> <u>We should find, at every node, a point that minimizes infinitely many global functions</u>

Adaptive Projected Subgradient Method (1)

- **Step 1) Apply a particular version of the adaptive projected subgradient method in each node [Yamada and Ogura' 04]**

$(k = 1, \ldots, N)$ $\qquad \inf_{\boldsymbol{h}} \Theta_k[i](\boldsymbol{h})$ ***(I assume to be 0 in the following)***

Agent

A subgradient

$$\boldsymbol{h}'_k[i+1] = \boldsymbol{h}_k[i] - \mu_k[i] \frac{(\Theta_k[i](\boldsymbol{h}_k[i]) - \Theta^\star_k[i])}{(\|\Theta'_k[i](\boldsymbol{h}_k[i])\|^2 + \delta_k[i])} \Theta'_k[i](\boldsymbol{h}_k[i]),$$

Step size $\mu_k[i] \in (0, 2)$

Small constant

- **Step 2) Information exchange**

$$\boldsymbol{h}_k[i+1] = \sum_{j \in \mathcal{N}_k[i]} \boldsymbol{W}_{kj}[i] \boldsymbol{h}'_j[i+1], \quad k = 1, \ldots, N,$$

Neighbors of agent k

Random matrices satisfying properties of consensus matrices (very easy to construct!)

$$\boldsymbol{h}_k[i+1] = \sum_{j \in \mathcal{N}_k[i]} \boldsymbol{W}_{kj}[i]\boldsymbol{h}'_j[i+1], \quad k = 1, \ldots, N,$$

Random matrices

Neighbors of agent k

$$\boldsymbol{P}[i] := \begin{bmatrix} \boldsymbol{W}_{11}[i] & \ldots & \boldsymbol{W}_{1N}[i] \\ \vdots & \ddots & \vdots \\ \boldsymbol{W}_{N1}[i] & \ldots & \boldsymbol{W}_{NN}[i] \end{bmatrix}$$

(a projection matrix)

- $\left\| E\left[\boldsymbol{P}[i]^T(\boldsymbol{I} - \boldsymbol{J})\boldsymbol{P}[i]\right] \right\|_2 \leq (1-\epsilon), \quad 0 < \epsilon \leq 1$
- $\left\| E\left[\boldsymbol{P}[i]^T\boldsymbol{P}[i]\right] \right\|_2 = 1$
- $\boldsymbol{P}[i][\boldsymbol{h}_1[i]^T \ \ldots \ \boldsymbol{h}_N[i]^T]^T = [\boldsymbol{h}_1[i]^T \ \ldots \ \boldsymbol{h}_N[i]^T]^T$

If $\boldsymbol{h}_1[i] = \boldsymbol{h}_2[i] = \ldots = \boldsymbol{h}_N[i]$

39

Adaptive Projected Subgradient Method: Convergence

Theorem (Cavalcante et al. 2009, Cavalcante et al. 2011)

If $\mu_k[i] \in (0,2),\ (k \in \mathcal{N})$,

then $\boxed{E[\|\boldsymbol{\psi}[i+1] - \boldsymbol{\psi}^\star\|^2] < E[\|\boldsymbol{\psi}[i] - \boldsymbol{\psi}^\star\|^2]}$ for every

$$\boldsymbol{\psi}^\star \in C^\star := \{\boldsymbol{\psi} = [\boldsymbol{h}^T\ \boldsymbol{h}^T\ \ldots \boldsymbol{h}^T]^T \in \mathbb{R}^{MN} \mid \boldsymbol{h} \in \Upsilon^\star\}$$

$$\Upsilon^\star := \cap_{i \geq 0} \cap_{k=1}^{N} \Upsilon_k[i] \neq \emptyset$$

(set of time-invariant solutions)

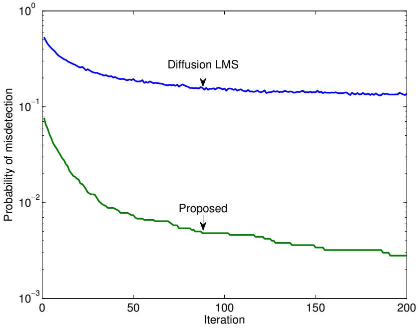
Noisy Information Exchange

- So far we have assumed noiseless information exchange.
  - This ignores the presence of noise in the case of analog computation schemes.
- Recently we have extended the APSM to cope with noisy communication in the consensus step (corrupted by additive noise).
  - The framework of stochastic approximation was used to dampen the effects of noise (by a sequence of decreasing step sizes).
- Under some mild conditions, the algorithm was shown to converge almost surely.

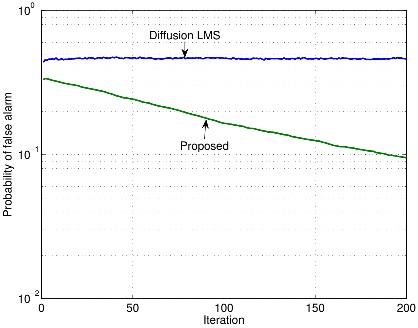
Cavalcane and Stanczak '13

Empirical Evaluation (Signal Detection)

- 20 nodes distributed uniformly at random in a unit grid
- Nodes are neighbors if their Euclidean distance is less or equal than $\sqrt{\log(N)/N}$
- Noise i.i.d with pdf $f = (1-\beta)\mathcal{N}(0,\nu^2) + \beta\mathcal{N}(0,\kappa\nu^2)$



(a)



(b)

Outline

Motivation

- Distributed signal detection methods use consensus algorithms as a building block.
- Traditional consensus algorithms (gossip algorithms) can be too slow in practice.
  - Existing acceleration techniques are often insufficient.
  - Most algorithms deal with the average consensus problem.
- Other non-linear functions are of interest (e.g. maximum, geometric mean).
- Approach
  - harness interference for computations.
  - Find appropriate function representations.

f-Consensus

f-Consensus Problem

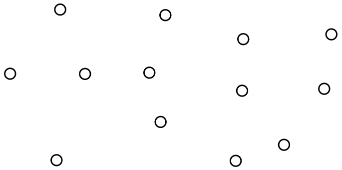
Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

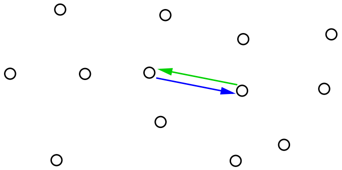


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

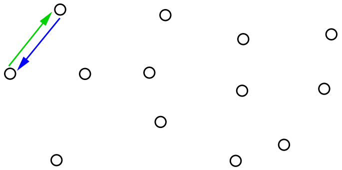


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

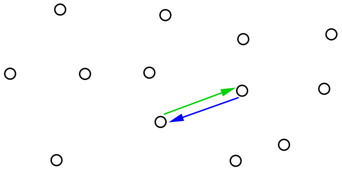


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

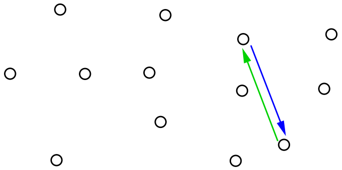


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

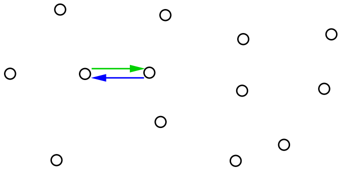


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

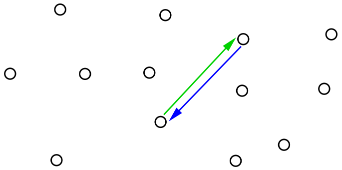


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

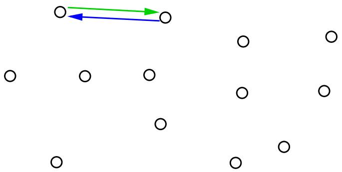


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

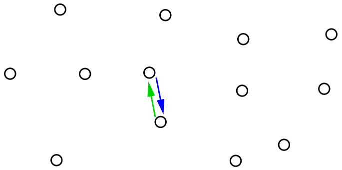


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

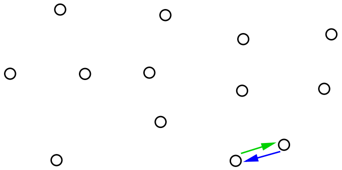


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

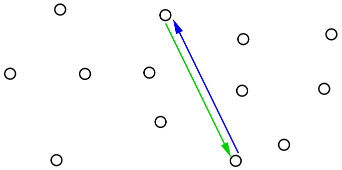


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

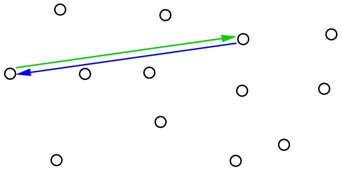


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

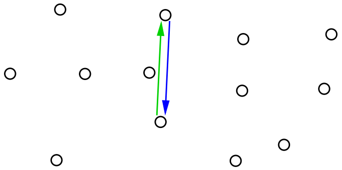


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

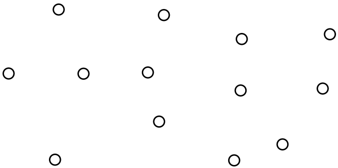


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Randomized Gossiping

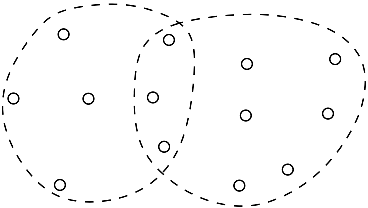


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (2 Cluster)

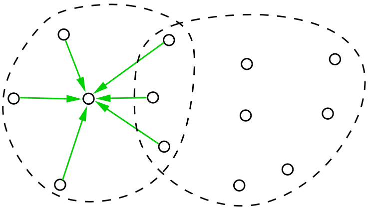


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (2 Cluster)

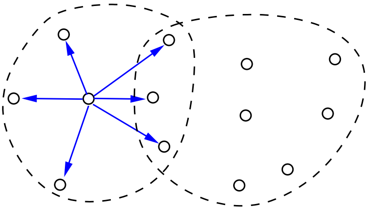


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (2 Cluster)

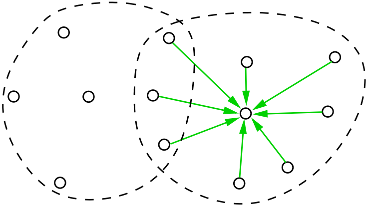


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (2 Cluster)

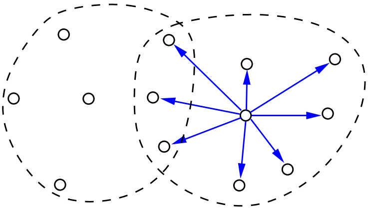


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (2 Cluster)

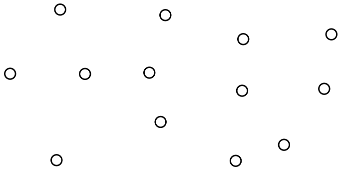


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \dots, x_n)$ as fast as possible.

Cluster-Based Consensus (2 Cluster)

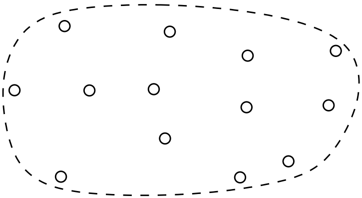


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (1 Cluster)

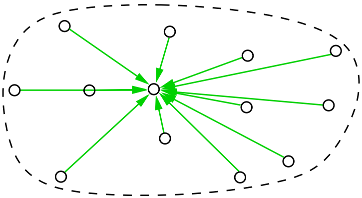


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (1 Cluster)

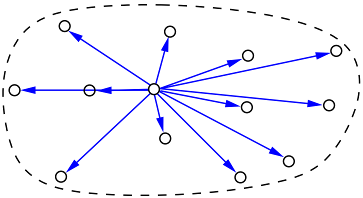


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \ldots, x_n)$ as fast as possible.

Cluster-Based Consensus (1 Cluster)

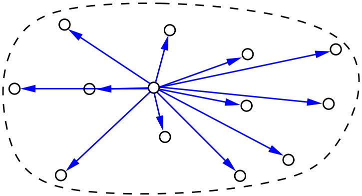


f-Consensus

f-Consensus Problem

Each sensor node attempts to compute $f(x_1, \dots, x_n)$ as fast as possible.

Cluster-Based Consensus (1 Cluster)



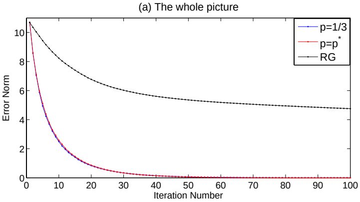
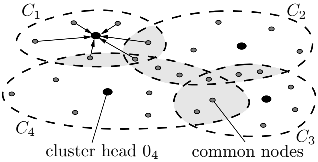


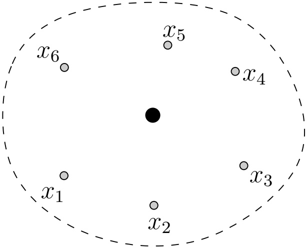
Figure: 3 Clusters, 16 nodes, "Arithmetic Mean"

M. Zheng and M. Goldenbaum and S. Stanczak and H. Yu, "Fast Average Consensus in Clustered Wireless Sensor Networks by Superposition Gossiping", IEEE WCNC, 2012.

Computation over MAC (CoMAC)

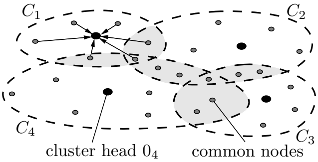
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

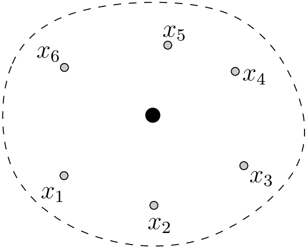




Computation over MAC (CoMAC)

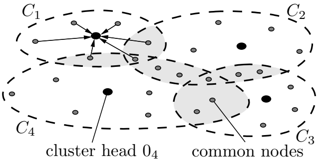
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

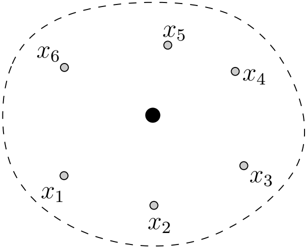




Computation over MAC (CoMAC)

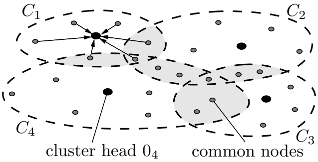
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

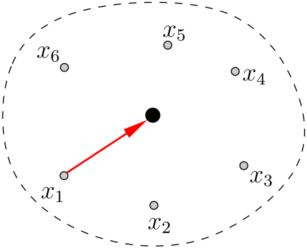




Computation over MAC (CoMAC)

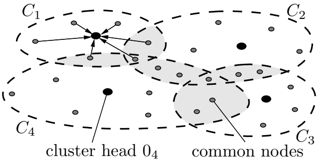
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

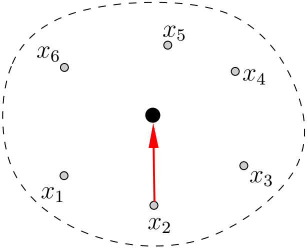




Computation over MAC (CoMAC)

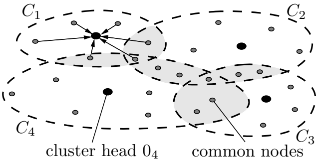
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

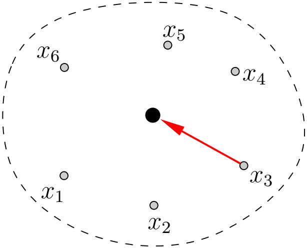




Computation over MAC (CoMAC)

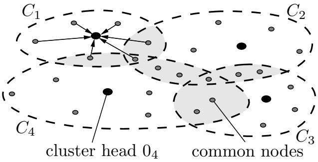
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

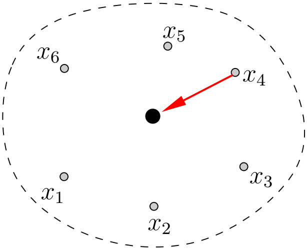




Computation over MAC (CoMAC)

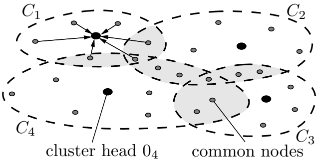
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

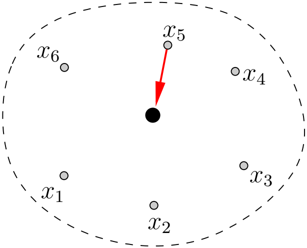




Computation over MAC (CoMAC)

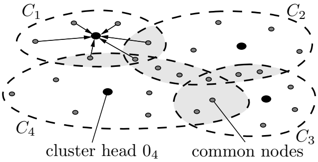
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

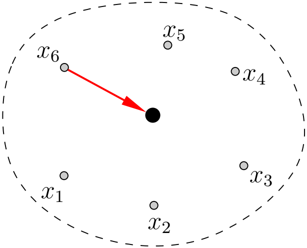




Computation over MAC (CoMAC)

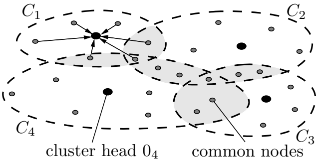
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

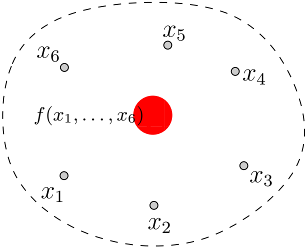




Computation over MAC (CoMAC)

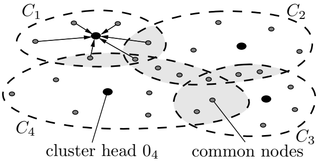
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

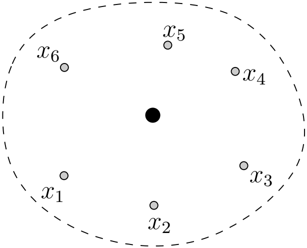




Computation over MAC (CoMAC)

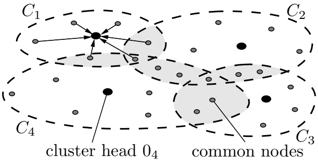
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

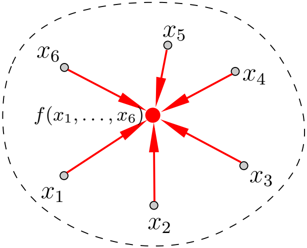




Computation over MAC (CoMAC)

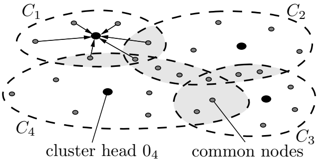
- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]

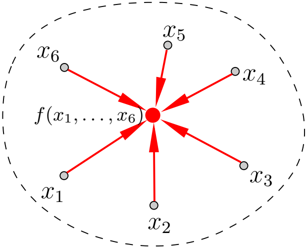




Computation over MAC (CoMAC)

- **Problem:** Compute a function f
- Typically transmissions are orthogonalized to combat interference
- Cluster head reconstructs each sensor signal **separately** and **subsequently** computes f
  - but it also destroys information!
  - Too much information sent!
- Separating **communication** and **computation** can be highly inefficient. [Nazer,Gastpar '07]
  - Can we merge data transmission and function computation into one step?
- Yes, as the broadcast property can be exploited to merge the processes of computation and communication. [Stanczak et al. '06]





- How to compute nonlinear functions?
- How to cope with practical impairments (asynchronism, fading)?

Outline

Analog Computation via MAC

- We developed a coding scheme to efficiently compute some functions. It
  - requires only coarse synchronization and channel magnitudes at the transmitters
  - calls for little CSI if multiple antenna elements are used at the sink.

M. Goldenbaum, S. Stanczak, "Robust Analog Function Computation via Wireless Multiple-Access Channels", submitted journal version, 2012 (available at arXiv).

Analog Computation via MAC

- We developed a coding scheme to efficiently compute some functions. It
  - requires only coarse synchronization and channel magnitudes at the transmitters
  - calls for little CSI if multiple antenna elements are used at the sink.
- Each sensor node transmits concurrently a sequence of independent random symbols
  - transmit power is controlled by sensor reading
  - fading effects are corrected through channel inversion, scaling and multiple antenna techniques at the sink.
- Sink estimates the function value from the received energy.
- The scheme can be used in connection with any cluster-based consensus algorithm to reach average consensus. (no proof of convergence)
- If average consensus can be reached, then it is possible to attain consensus on any function, provided that
  1. the sensor reading are suitably pre-processed and
  2. some post-processing is applied to the average consensus.

M. Goldenbaum, S. Stanczak, "Robust Analog Function Computation via Wireless Multiple-Access Channels", submitted journal version, 2012 (available at arXiv).

Analog Computation via MAC

- We developed a coding scheme to efficiently compute some functions. It
  - requires only coarse synchronization and channel magnitudes at the transmitters
  - calls for little CSI if multiple antenna elements are used at the sink.
- Each sensor node transmits concurrently a sequence of independent random symbols
  - transmit power is controlled by sensor reading
  - fading effects are corrected through channel inversion, scaling and multiple antenna techniques at the sink.
- Sink estimates the function value from the received energy.
- The scheme can be used in connection with any cluster-based consensus algorithm to reach average consensus. (no proof of convergence)
- If average consensus can be reached, then it is possible to attain consensus on any function, provided that
  1. the sensor reading are suitably pre-processed and
  2. some post-processing is applied to the average consensus.

M. Goldenbaum, S. Stanczak, "Robust Analog Function Computation via Wireless Multiple-Access Channels", submitted journal version, 2012 (available at arXiv).

Analog Computation via MAC

- We developed a coding scheme to efficiently compute some functions. It
  - requires only coarse synchronization and channel magnitudes at the transmitters
  - calls for little CSI if multiple antenna elements are used at the sink.
- Each sensor node transmits concurrently a sequence of independent random symbols
  - transmit power is controlled by sensor reading
  - fading effects are corrected through channel inversion, scaling and multiple antenna techniques at the sink.
- Sink estimates the function value from the received energy.
- The scheme can be used in connection with any cluster-based consensus algorithm to reach average consensus. (no proof of convergence)
- If average consensus can be reached, then it is possible to attain consensus on any function, provided that
  1. the sensor reading are suitably pre-processed and
  2. some post-processing is applied to the average consensus.

M. Goldenbaum, S. Stanczak, "Robust Analog Function Computation via Wireless Multiple-Access Channels", submitted journal version, 2012 (available at arXiv).

Outline

Wireless MAC

$$y = \sum_{i=1}^{N} h_i x_i + v \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1,\dots,x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N}y$
  - Geometric Mean: $f(x_1,\dots,x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1,\dots,x_N) = \sqrt{x_1^2+\cdots+x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$
- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

Analog Computability

Ideal Wireless MAC

$$y = \sum_{i=1}^{N} x_i \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1, \ldots, x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N} y$
  - Geometric Mean: $f(x_1, \ldots, x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1, \ldots, x_N) = \sqrt{x_1^2 + \cdots + x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$
- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

Ideal Wireless MAC

$$y = \sum_{i=1}^{N} \varphi_i(x_i) \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1, \ldots, x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N}y$
  - Geometric Mean: $f(x_1, \ldots, x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1, \ldots, x_N) = \sqrt{x_1^2 + \cdots + x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$
- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

Analog Computability

Ideal Wireless MAC

$$\psi(y) = \psi\left(\sum_{i=1}^{N} \varphi_i(x_i)\right) \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1, \ldots, x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N}y$
  - Geometric Mean: $f(x_1, \ldots, x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1, \ldots, x_N) = \sqrt{x_1^2 + \cdots + x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$
- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

Analog Computability

Ideal Wireless MAC

$$\psi(y) = \psi\left(\sum_{i=1}^{N} \varphi_i(x_i)\right) \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1, \ldots, x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N}y$
  - Geometric Mean: $f(x_1, \ldots, x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1, \ldots, x_N) = \sqrt{x_1^2 + \cdots + x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$
- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

Ideal Wireless MAC

$$\psi(y) = \psi\left(\sum_{i=1}^{N} \varphi_i(x_i)\right) \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1, \ldots, x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N} y$
  - Geometric Mean: $f(x_1, \ldots, x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1, \ldots, x_N) = \sqrt{x_1^2 + \cdots + x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$

- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

Analog Computability

Ideal Wireless MAC

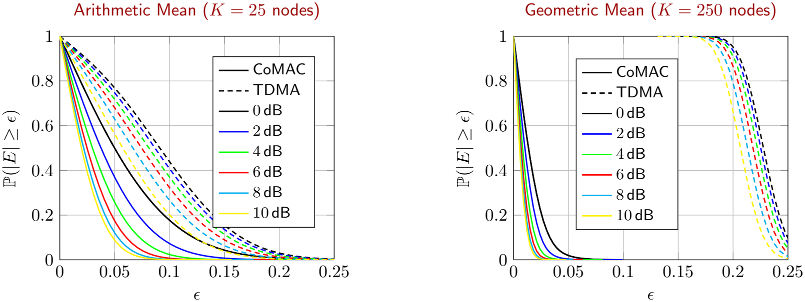
$$\psi(y) = \psi\left(\sum_{i=1}^{N} \varphi_i(x_i)\right) \qquad (*)$$

- $\Rightarrow$ We can estimate any function having a representation $(*)$.
- Examples:
  - Arithmetic Mean: $f(x_1, \ldots, x_N) = \frac{1}{N}\sum_i x_i$, $\varphi_i(x) = x$, $\psi(y) = \frac{1}{N}y$
  - Geometric Mean: $f(x_1, \ldots, x_N) = \left(\prod_i x_i\right)^{1/N}$, $\varphi_i(x) = \log(x)$, $\psi(y) = \exp(y/N)$
  - Euclidean Norm: $f(x_1, \ldots, x_N) = \sqrt{x_1^2 + \cdots + x_N^2}$, $\varphi_i(x) = x^2$, $\psi(y) = \sqrt{y}$
- Every function is universally computable via an ideal MAC, since every $f : [0,1]^N \to \mathbb{R}$ has a representation $(*)$.

M. Goldenbaum, H. Boche, S. Stanczak, "Harnessing Interference for Analog Function Computation in Wireless Sensor Networks", submitted, 2012.

Analog Computation via Signal Powers (Numerical Examples)

Performance Measure: $|E| = \left| \frac{\hat{f} - f}{f_{\max} - f_{\min}} \right|$

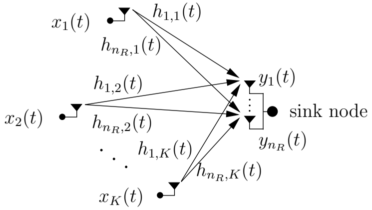


$K = 25$, Arithmetic Mean $\quad$ $K = 250$, Geometric Mean

- Estimation/Computation of non-linear functions by transmission-side data pre-processing and receiver-side signal post-processing.

Eliminating Fading Effects using Multiple Antenna Systems

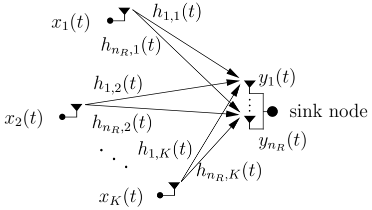
- SIMO: A correction of fading effects at the sink is possible.
- Simulation: $K = 25$, uncorrelated Rician fading: $H_{nk}^{(m)} \sim \mathcal{N}_{\mathbb{C}}(0.5, 0.75)$

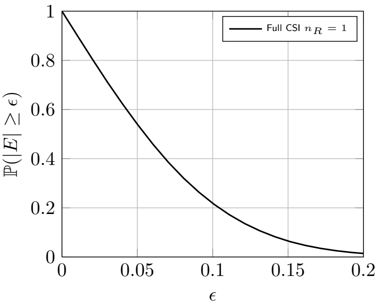


M. Goldenbaum, S. Stanczak, "On the Channel Estimation Effort for Analog Computation Over Multiple-Access Channels", ICASSP, 2010.

Eliminating Fading Effects using Multiple Antenna Systems

- SIMO: A correction of fading effects at the sink is possible.
- Simulation: $K = 25$, uncorrelated Rician fading: $H_{nk}^{(m)} \sim \mathcal{N}_{\mathbb{C}}(0.5, 0.75)$

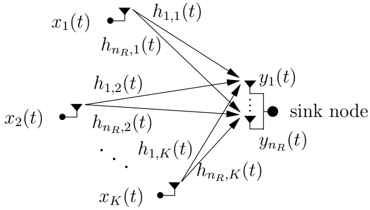


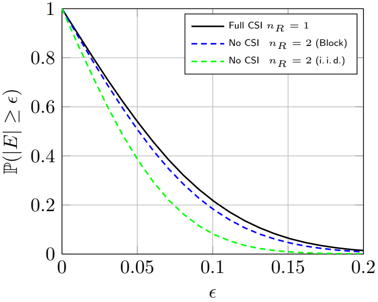


M. Goldenbaum, S. Stanczak, "On the Channel Estimation Effort for Analog Computation Over Multiple-Access Channels", ICASSP, 2010.

Eliminating Fading Effects using Multiple Antenna Systems

- SIMO: A correction of fading effects at the sink is possible.
- Simulation: $K = 25$, uncorrelated Rician fading: $H_{nk}^{(m)} \sim \mathcal{N}_{\mathbb{C}}(0.5, 0.75)$

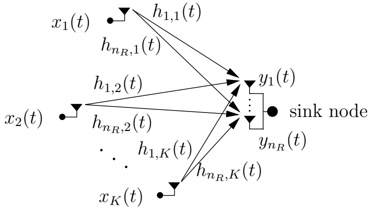


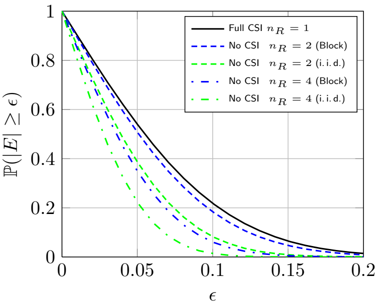


M. Goldenbaum, S. Stanczak, "On the Channel Estimation Effort for Analog Computation Over Multiple-Access Channels", ICASSP, 2010.

Eliminating Fading Effects using Multiple Antenna Systems

- SIMO: A correction of fading effects at the sink is possible.
- Simulation: $K = 25$, uncorrelated Rician fading: $H_{nk}^{(m)} \sim \mathcal{N}_{\mathbb{C}}(0.5, 0.75)$

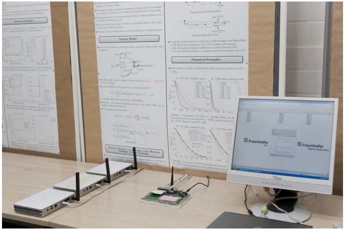




M. Goldenbaum, S. Stanczak, "On the Channel Estimation Effort for Analog Computation Over Multiple-Access Channels", ICASSP, 2010.

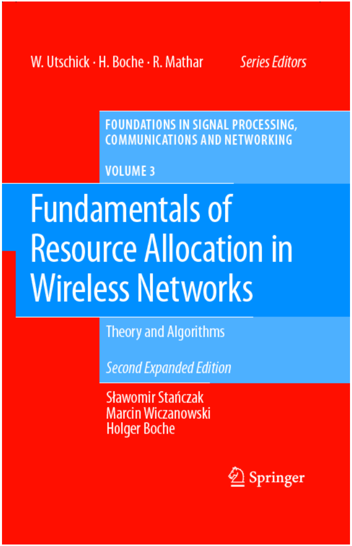
Ongoing work/Open problems

- Construction/offline computation of pre- and post-processing functions.
- Robustness against practical impairments like fading and noise (in networks)



Thank you!

References


W. Utschick · H. Boche · R. Mathar
Series Editors
FOUNDATIONS IN SIGNAL PROCESSING, COMMUNICATIONS AND NETWORKING
VOLUME 3
Fundamentals of Resource Allocation in Wireless Networks
Theory and Algorithms
Second Expanded Edition
Sławomir Stańczak
Marcin Wiczanowski
Holger Boche
Springer

References

D.P. Bertsekas.
Nonlinear Programming.
Athena Scientific, Belmont, Massachusetts, 1995.

H. Boche and S. Stanczak.
Convexity of some feasible QoS regions and asymptotic behavior of the minimum total power in CDMA systems.
IEEE Trans. on Comm., 52(12):2190 – 2197, December 2004.

H. Boche and S. Stanczak.
Log-convexity of the minimum total power in CDMA systems with certain quality-of-service guaranteed.
IEEE Trans. on Inform. Theory, 51(1):374–381, January 2005.

M. Chiang.
Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control.
IEEE Journal on Selected Areas in Communications, 23(1):104–116, Jan. 2005.

Stanczak S, Wiczanowski M, Boche H.
Distributed Utility-Based Power Control: Objectives and Algorithms.
IEEE Trans Signal Processing. 2007 Oct;55(10):5058–5068.

References

F.P. Kelly, A.K. Maulloo, and D.K.H. Tan.
Rate control for communication networks: Shadow prices, proportional fairness and stability.
Journal of Operations Research Society, 49(3):237–252, March 1998.

H.J. Kushner.
Stochastic Approximation and Recursive Algorithms and Applications.
Springer, 2003.

Seneta E.
Non-Negative Matrices and Markov Chains.
Springer, Berlin; 1981.

Arnold L, Gundlach V, Demetrius L.
Evolutionary Formalism for Products of positive Random Matrices.
Ann Appl Probab. 1994;4(3):859–901.

Friedland S, Karlin S.
Some Inequalities for the Spectral Radius of Non-Negative Matrices and Applications.
Duke Math J. 1975;42(3):459–490.

Kingman JFC.
A Convexity Property of Positive Matrices.
Quart, J Math Oxford Ser. 1961;12(2):283–284.

References

Wiczanowski M, Stanczak S, Boche H.
Providing quadratic convergence of decentralized power control in wireless networks – The method of min-max functions.
IEEE Trans Signal Processing. 2008 Aug;56(8):4053–4068.

J. Mo and J. Walrand.
Fair end-to-end window-based congestion control.
IEEE/ACM Trans. on Networking, 8(5), October 2000.

M. Xiao, N.B. Schroff, and E.K.P. Chong.
A utility-based power control scheme in wireless cellular systems.
IEEE/ACM Trans. Networking, 11(2):210–221, April 2003.

R.D. Yates.
A framework for uplink power control in cellular radio systems.
IEEE J. Select. Areas Commun., 13(7):1341–1347, September 1995.

Stanczak S, Kaliszan M, Bambos N, Wiczanowski M.
A Characterization of Max-Min SIR-Balanced Power Allocation with Applications.
In: Proc. IEEE International Symposium on Information Theory (ISIT). Seoul, Korea; 2009.

R.L.G. Cavalcante and S. Stanczak
A Distributed Subgradient Method for Dynamic Convex Optimization Problems under Noisy Information Exchange
IEEE Journal of Selected Topics in Signal Processing, 7(2): 243–256, April 2013

References

- **T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Broadcast gossip algorithms for consensus," IEEE Trans. Signal Processing, vol. 57, no. 7, pp. 2748–2761, July 2009.**
- **D. Blatt and A. O. Hero III, "Energy-based sensor network source localization via projection onto convex sets," IEEE Trans. Signal Processing, vol. 54, no. 9, pp. 3614–3619, Sept. 2006.**
- **S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," IEEE Trans. Inform. Theory, vol. 52, no. 6, pp. 2508–2530, June 2006**
- **R. L. G. Cavalcante and S. Stanczak, "Robust Set-Theoretic Distributed Detection in Diffusion Networks," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mar. 2012**
- **R.L.G. Cavalcante, A. Rogers, and N. R. Jennings, "Consensus acceleration in multi-agent systems with the Chebyshev semi-iterative method," in Proc. 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011), pp. 165-172, May 2011.**
- **R. L. G. Cavalcante, A. Rogers, N. R. Jennings, and I. Yamada, "Distributed Asymptotic Minimization of Sequences of Convex Functions by a Broadcast Adaptive Subgradient Method," IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 4, pp. 739-753, Aug. 2011**
- **R.L.G.Cavalcante and B. Mulgrew, "Adaptive filter algorithms for accelerated discrete-time consensus," IEEE Trans. Signal Processing, vol. 58, no. 3, pp. 1049-1058, March 2010.**
- **R.L.G.Cavalcante, I. Yamada, and B. Mulgrew, "An Adaptive Projected Subgradient Approach to Learning in Diffusion Networks," IEEE Trans. Signal Processing,vol. 57, no. 7, pp. 2762-2774, July 2009.**

References

- **H. H. Bauschke and P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, 2011.**
- **R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” Proc. IEEE, vol. 95, no. 1, pp. 215–233, Jan. 2007**
- **B. T. Polyak, “Minimization of unsmooth functionals,” USSR Comput. Math. Phys., vol. 9, pp. 14–29, 1969.**
- **A. H. Sayed and C. G. Lopes, "Adaptive processing over distributed networks," IEICE Transactions on Fundamentals of Electronics,Communications and Computer Sciences, vol. E90-A, no. 8, pp. 1504-1510, 2007**
- **K. Slavakis, I. Yamada, and N. Ogura, “The adaptive projected subgradient method over the fixed point set of strongly attracting nonexpansive mappings,” Numerical Functional Analysis and Optimization, vol. 27, no. 7-8, pp. 905–930, Dec. 2006.**
- **H. Stark and Y. Yang, Vector Space Projections – A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics. New York: Wiley, 1998**
- **I. Yamada, “Adaptive projected subgradient method: A unified view for projection based adaptive algorithms,” J. IEICE, vol. 86, no. 8, pp. 654–658, Aug. 2003, in Japanese.**
- **I. Yamada and N. Ogura, “Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions,” Numerical Functional Analysis and Optimization, vol. 25, no. 7/8, pp. 593–617, 2004**
- **M. Zheng, M. Goldenbaum, S. Stanczak, and H. Yu, “Fast Average Consensus in Clustered Wireless Sensor Networks by Superposition Gossiping,” Proc. IEEE Wireless Communications and Networking Conference (WCNC '12)**