Scalable Sparse Optimization in Dense Wireless Cooperative Networks

Yuanming Shi
School of Information Science and Technology
ShanghaiTech University
Outline

- Introduction

- Two Vignettes:
  - Group Sparse Beamforming for Green Cloud-RAN
  - Low-Rank Matrix Completion for Topological Interference Management

- Summary
Part I: Introduction
Ultra Mobile Broadband

- **Era of mobile data traffic deluge**

![Sphere of photos]

**10x**
Data growth by 2019

![Tablet and character]

**497 M**
Mobile devices added in 2014

![Smartphone and globe]

**72%**
Video traffic by 2019

Source: Cisco VNI Mobile, 2015
Solution?

Marty Pens Cooper’s Law: Data Over Usable Spectrum Doubles Every 30 Months – 1997

Network densification is a dominated theme!
Challenges: Green, Flexibility, Scalability

- **Networking issues:**
  - Huge network power consumption
  - Massive channel state information acquisition

- **Computing issues:**
  - Large-scale performance optimizations
  - Critical for latency

*Credit: Alcatel-Lucent, 2013*
Part II: Two Vignettes

Low-Rank Matrix Completion

Group Sparse Beamforming
Vignette A: *Group Sparse Beamforming for Green Cloud-RAN*
**Dense Cloud Radio Access Networks**

- **Dense Cloud-RAN**: A cost-effective way for network densification and cooperation

**Cost-effective cooperative wireless networks to improve the network capacity and network energy efficiency**

1. Centralized signal processing and resource allocation
2. Dense deployment of low-cost low-power RRHs
3. Real-time cloud infrastructure with BS virtualization
**Network Power Consumption**

- **Goal:** Design a green dense Cloud-RAN

- **Prior works:** Physical-layer transmit power consumption
  - Wireless power control: [Chiang, et al., FT 08], [Qian, et al., TWC 09], [Sorooshyari, et al., TON 12], …
  - Transmit beamforming: [Sidiropoulos and Luo, TSP 2006], [Yu and Lan, TSP 07], [Gershman, et al., SPMag 10], …

- **Unique challenge:**
  - Network power consumption:
    - RRHs, fronthaul links, etc.
Network Adaptation

- **Question:** Can we provide a holistic approach for network power minimization?

- **Key observation:** Spatial and temporal mobile data traffic variation

- **Approach:** Network adaptation
  - Adaptively switch off network entities to save power
**Problem Formulation**

- **Goal:** Minimize network power consumption in Cloud-RAN

\[
\begin{align*}
\text{minimize} & \quad f_1(v) + f_2(v) \\
\text{subject to} & \quad \frac{|h_k^H v_k|^2}{\sum_{i \neq k} |h_k^H v_i|^2 + \sigma_k^2} \geq \gamma_k, \forall k.
\end{align*}
\]

- **Fronthaul power:** \( f_1(v) = \sum_{l=1}^L P_c^l I(T(v) \cap V_l \neq \emptyset) \)
- **Transmit power:** \( f_2(v) = \sum_{l=1}^L \sum_{k=1}^K \frac{1}{c_l} ||v_{lk}||_2^2 \)

- **Prior algorithms:** heuristic or computationally expensive: [Philipp, et. al, TSP 13], [Luo, et. al, JSAC 13], [Quek, et. al, TWC 13], …
Finding Structured Solutions

- **Proposal**: Group sparse beamforming framework [1]

- Switch off the \( l \)-th RRH \( \tilde{v}_l = 0 \), i.e., group sparsity structure in \( v \)

Proposed Algorithm

- **Proposition [1]:** The *tightest* convex positively homogeneous lower bound of the combinatorial composite objective function

\[ \Omega(v) = 2 \sum_{i=1}^{L} \sqrt{\frac{P_{i}^{c}}{\eta_{i}}} \| \tilde{v}_{i} \|_{2} \]

minimize \( \Omega(v) \)

induce group sparsity

- Adaptive RRH selection: switch off the RRHs with smallest coefficients in the aggregative beamformers

Minimize the Group Sparsity Inducing Norm for RRH Ordering

\( \theta^{*} \)

RRH Selection by Solving \( \mathcal{F}(\mathcal{A}^{[i]}) \)

\( \mathcal{A}^{*} \)

Transmit Power Minimization \( \mathcal{P}_{TP}(\mathcal{A}^{*}) \)

**Stage I**

**Stage II**

**Stage III**
The Power of Group Sparse Beamforming

- **Example:** Group sparse beamforming for green Cloud-RAN [1] (10 RRHs, 15 MUs)

Advantages:
1) Enabling flexible network adaptation;
2) Offering efficient algorithm design via convex programming
3) Empowering wide applications

Scalability in Dense Cloud-RAN?

Minimize the Group Sparsity Inducing Norm for RRH Ordering

Stage I

RRH Selection by Solving $\mathcal{F}(\mathcal{A}[\hat{i}])$

Stage II

Transmit Power Minimization $\mathcal{P}_{TP}(\mathcal{A}^*)$

Stage III

High computational complexity: a sequence of convex optimization and feasibility problems needs to be solved.
Solution: **Large-Scale Convex Optimization for Dense Cloud-RAN**
Large-Scale Convex Optimization

- **Large-scale convex optimization**: A powerful tool for system design in dense wireless networks

- **Prior works**: Mainly focus on small-size networks or well-structured problems
  - Limitations: scalability [Luo, et al., SPMag 10], parallelization [Yu and Lan, TWC 10], infeasibility detection [Liao, et al., TSP 14], …

- **Unique challenges in dense Cloud-RAN**:  
  - Design problems: 1) A high dimension; 2) a large number of constraints; 3) complicated structures
Matrix Stuffing and Operator Splitting

- **Goal:** Design a unified framework for general large-scale convex optimization problem $\mathcal{P}_{\text{Original}}$?

- **Disciplined convex programming framework** [Grant & Boyd '08]

  \[ \mathcal{P}_{\text{Original}} \xrightarrow{\text{CVX}} \mathcal{P}_{\text{cone}} \xrightarrow{\text{Interior-Point Solver}} x^* \]

  Time consuming: modeling phase & solving phase

- **Proposal:** Two-stage approach for large-scale convex optimization

  \[ \mathcal{P}_{\text{Original}} \xrightarrow{\text{Matrix Stuffing}} \mathcal{P}_{\text{HSD}} \xrightarrow{\text{ADMM Solver}} x^* \]

  - **Matrix stuffing:** Fast homogeneous self-dual embedding (HSD) transformation
  - **Operator splitting (ADMM):** Large-scale homogeneous self-dual embedding
Stage One: Fast Transformation

- **Example:** Coordinated beamforming problem family (with transmit power constraints and QoS constraints)

\[ \mathcal{P}_{\text{Original}} : \text{minimize } \|v\|^2_2 \]

subject to

\[ \|D_l v\|_2 \leq \sqrt{P_l}, \forall l, \]

\[ \|C_k v + g_k\|_2 \leq \beta_k r_k^T v, \forall k. \]

- **Smith form reformulation** [Smith ’96]

  - **Key idea:** Introduce a new variable for each subexpression in \( \mathcal{P}_{\text{Original}} \)

\( \mathcal{G}_1(l) : \)

\[ \left\{ \begin{array}{l}
(\mathbf{y}_0^l, \mathbf{y}_1^l) \in \mathbb{Q}^{KN_l+1} \\
\mathbf{y}_0^l = \sqrt{P_l} \in \mathbb{R} \\
\mathbf{y}_1^l = \mathbf{D}_l v \in \mathbb{R}^{KN_l}
\end{array} \right. \]

- Second-order cone
- Linear constraint

The Smith form is ready for standard cone programming transformation.
Stage One: Fast Transformation

- **HSD embedding** of the **primal-dual pair** of transformed standard cone program (based on KKT conditions)

\[
\begin{align*}
\text{minimize} & \quad c^T \nu \\
\text{subject to} & \quad A \nu + \mu = b \\
& \quad (\nu, \mu) \in \mathbb{R}^n \times \mathcal{K},
\end{align*}
\]

\[
\begin{align*}
\text{maximize} & \quad -b^T \eta \\
\text{subject to} & \quad -A^T \eta + \lambda = c \\
& \quad (\lambda, \eta) \in \{0\}^n \times \mathcal{K}^*
\end{align*}
\]

\[\mathcal{P}_{\text{HSD}}: \text{find } (x, y) \quad \text{subject to } y = Qx \quad x \in \mathcal{C}, y \in \mathcal{C}^*\]

**Certificate of infeasibility:** \( \tau = 0, \kappa > 0 \)

- **Matrix stuffing for fast transformation:**
  - Generate and keep the structure \( Q \)
  - Copy problem instance parameters to the pre-stored structure \( Q \)
Stage Two: Parallel and Scalable Computing

- **HSD embedding in consensus form:**

\[
\mathcal{F}_{\text{HSD}} : \text{find } (x, y) \\
\text{subject to } y = Qx \\
x \in \mathcal{C}, y \in \mathcal{C}^* 
\]

\[
\mathcal{P}_{\text{ADMM}} : \text{minimize } \mathcal{I}_{C \times C^*}(x, y) + \mathcal{I}_{Qx=y}(\tilde{x}, \tilde{y}) \\
\text{subject to } (x, y) = (\tilde{x}, \tilde{y}) 
\]

- **Final algorithm:** Apply the operating splitting method (ADMM) [Donoghue, Chu, Parikh, and Boyd '13]

\[
\tilde{x}^{[i+1]} = (I + Q)^{-1}(x^{[i]} + y^{[i]}) \quad \text{subspace projection} \\
x^{[i+1]} = \Pi_C(\tilde{x}^{[i+1]} - y^{[i]}) \quad \text{parallel cone projection} \\
y^{[i+1]} = y^{[i]} - \tilde{x}^{[i+1]} + x^{[i+1]} \quad \text{computationally trivial} 
\]
Proximal Algorithms for Cone Projection

- Proximal algorithms for parallel cone projection [Parikh & Boyd, FTO 14]

- Projection onto the second-order cone: \( C = \{(y, x) \in \mathbb{R} \times \mathbb{R}^{p-1} \mid \|x\| \leq y\} \)

\[
\Pi_C(\omega, \tau) = \begin{cases} 
0, \|\omega\|_2 \leq -\tau \\
(\omega, \tau), \|\omega\|_2 \leq \tau \\
\left(1/2\right)(1 + \tau/\|\omega\|_2)(\omega, \|\omega\|_2), \|\omega\|_2 \geq |\tau|.
\end{cases}
\]

- Projection onto positive semidefinite cone: \( C = S^+_n \)

\[
\Pi_C(V) = \sum_{i=1}^{n}(\lambda_i) + u_i u^T_i
\]

SVD is computationally expensive
### Numerical Results (I)

**Example:** Power minimization coordinated beamforming problem [2]

<table>
<thead>
<tr>
<th>Network Size ($L=K$)</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVX+SDPT3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modeling Time [sec]</td>
<td>0.7563</td>
<td>4.4301</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Solving Time [sec]</td>
<td>4.2835</td>
<td>326.2513</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Objective [W]</td>
<td>12.2488</td>
<td>6.5216</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Matrix Stuffing+ADMM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modeling Time [sec]</td>
<td>0.0128</td>
<td>0.2401</td>
<td>2.4154</td>
<td>9.4167</td>
</tr>
<tr>
<td>Solving Time [sec]</td>
<td>0.1009</td>
<td>2.4821</td>
<td>23.8088</td>
<td>81.0023</td>
</tr>
</tbody>
</table>

Matrix stuffing can speedup **60x** over CVX

ADMM can speedup **130x** over the interior-point method

---

Numerical Results (II)

- Group sparse beamforming for network power minimization [2]
Can We do Better?

1. High computational complexity (computing depends on the instantaneous CSI)
2. Limited capability to enhance group sparsity
Solution: Large System Analysis for Enhanced Group Sparse Beamforming
Proposed Algorithm: **Iterative Reweighted-$l_2$ Algorithm**
Proposed Method

- Smoothed $\ell_p$-minimization approach to induce group sparsity

\[
\begin{align*}
\text{minimize} & \quad g_p(v; \epsilon) := \sum_{l=1}^{L} \nu_l \left( \| \tilde{v}_l \|^2_2 + \epsilon^2 \right)^{p/2} \\
\text{subject to} & \quad \frac{|h_k^H v_k|^2}{\sum_{i \neq k} |h_k^H v_i|^2 + \sigma_k^2} \geq \gamma_k, \forall k
\end{align*}
\]

- Enhance sparsity:

\[
\|z\|_0 = \lim_{p \to 0} \|z\|_p^p = \lim_{p \to 0} \sum_i |z_i|^p
\]
Majorization-Minimization Algorithm

- Solve the following (nonconvex) smoothed $\ell_p$-minimization problem

$$\min_{z \in \mathcal{C}} f(z) := \sum_{i=1}^{m} (z_i^2 + \epsilon^2)^{p/2}$$

- **MM algorithm:** the successive upper-bound minimization method

1: Find a feasible point $z^{[0]} \in \mathcal{C}$ and set $k = 0$
2: **repeat**
3: $z^{[k+1]} = \arg\min_{z \in \mathcal{C}} g(z | z^{[k]})$ (**global minimum**)  
4: $k \leftarrow k + 1$
5: **until** some convergence criterion is met

- An upper bound for the objective function $f(z)$ can be constructed as

$$Q(z; \omega^{[k]}) := \sum_{i=1}^{m} \omega_i^{[k]} z_i^2$$

$$\omega_i^{[k]} = \frac{p}{2} \left[ \left( z_i^{[k]} \right)^2 + \epsilon^2 \right]^{\frac{p}{2} - 1}, \forall i$$
**Enhanced Group Sparse Beamforming**

- **Final algorithm**: iterative reweighted-$\ell_2$ algorithm

\[
\begin{aligned}
\text{minimize} & \quad \sum_{l=1}^{L} \omega_l^{[n]} \left\| \tilde{v}_l \right\|_2^2 \\
\text{subject to} & \quad \frac{|\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{i \neq k} |\mathbf{h}_k^H \mathbf{v}_i|^2 + \sigma_k^2} \geq \gamma_k, \forall k \\
\end{aligned}
\]

weights:
\[
\omega_l^{[n]} = \frac{p_n}{2} \left[ \left\| \tilde{v}_l^{[n]} \right\|_2^2 + \epsilon^2 \right]^{\frac{p}{2} - 1}, \forall l
\]

**Advantageous:**
1. Enhance sparsity
2. Lead to closed form solution via duality theory
Simple Solution Structures

- Optimal beamforming vectors $\mathbf{v}_1^*, \ldots, \mathbf{v}_K^*$ are given by
  \[
  \mathbf{v}_k^* = \sqrt{\frac{p_k}{LN}} \frac{\left( Q[n] + \sum_{i=1}^{K} \frac{\lambda_i}{LN} \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \mathbf{h}_k}{\left( Q[n] + \sum_{i=1}^{K} \frac{\lambda_i}{LN} \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \mathbf{h}_k}, \quad \forall k
  \]

- The $K$ powers are given by
  \[
  \begin{bmatrix}
  p_1 \\
  \vdots \\
  p_K
  \end{bmatrix} = \mathbf{M}^{-1}
  \begin{bmatrix}
  \sigma_1^2 \\
  \vdots \\
  \sigma_K^2
  \end{bmatrix}
  \]

- The Lagrange multipliers can be computed from the fixed-point equations
  \[
  \lambda_k = LN \left[ \left( 1 + \frac{1}{\gamma_k} \right) \mathbf{h}_k^H \left( Q[n] + \sum_{i=1}^{K} \frac{\lambda_i}{LN} \mathbf{h}_i \mathbf{h}_i^H \right)^{-1} \mathbf{h}_k \right]^{-1}
  \]

The first step to reduce computational complexity.
Optimality

**Theorem 1:** Let \( \{v[n]\}_{n=1}^\infty \) be the sequence generated by the iterative reweighted- \( \ell_2 \) algorithm. Then, every limit point \( \tilde{v} \) of \( \{v[n]\}_{n=1}^\infty \) has the following properties:

1) \( \tilde{v} \) is a KKT point of the smoothed \( L_p \)-minimization problem

2) \( g_p(v[n]; \epsilon) \) converges monotonically to \( g_p(v^*; \epsilon) \) for some KKT point \( v^* \)

**RRH ordering criteria to determine which RRHs should be switched off**

\[
\theta_l = \kappa_l \|\tilde{v}_l\|_2^2 = \kappa_l \sum_{k=1}^K v_k^H Q_{lk} v_k, \forall l = 1, \ldots, L
\]

**Challenges** to compute the ordering criteria

- Massive instantaneous CSI
- High computation cost
Random Matrix Theory: *Large System Analysis*
Modern Applications

Random Matrix Theory

Robust Statistics
- Since 2012
  - Detection in impulsive noises

Wireless Communication
- Since 1990
  - Performance analysis
  - Optimal transceiver design

Machine Learning
- Since 2015
  - Subspace clustering
  - Community detection

Signal Processing
- Since 2007
  - Estimation
  - Detection
Deterministic Equivalent of Optimal Parameters (I)

- Channel models in Cloud-RAN with distributed RRHs:
  \[ h_k = \Theta_k^{1/2} g_k, g_k \sim \mathcal{CN}(0, I_{NL}), \Theta_k = \text{diag}\{d_{k1}, \ldots, d_{kL}\} \otimes I_N \]

- Optimal Lagrange multipliers
  \[ \lambda_k = L N \left[ \left(1 + \frac{1}{\gamma_k}\right) h_k^H \left(Q[n] + \sum_{i=1}^{K} \frac{\lambda_i}{L N} h_i h_i^H\right)^{-1} h_k \right]^{-1} \]

- Lemma 1 (Deterministic Equivalent of the \( \lambda \)-Parameter):

  Assume \( 0 < \lim \inf_{N \to \infty} K/N \leq \lim \sup_{N \to \infty} K/N < \infty \). Let \( \{d_{kl}\} \) and \( \{\gamma_k\} \) satisfy \( \lim \sup_{N} \max_{k,l} \{d_{kl}\} < \infty \) and \( \lim \sup_{N} \max_{k} \gamma_k < \infty \), respectively. We have

  \[
  \max_{1 \leq k \leq K} |\lambda_k - \lambda_k^0| \xrightarrow{N \to \infty} 0 \quad \text{almost surely}
  \]

  where

  \[
  \lambda_k^0 = \gamma_k \left(\frac{1}{L} \sum_{i=1}^{L} d_{kl} \eta_l\right)^{-1}, \quad \eta_l = \left(\frac{1}{NL} \sum_{i=1}^{K} \frac{d_{il}}{\frac{1}{L} \sum_{j=1}^{L} d_{ij} \eta_j} \frac{\gamma_i}{1 + \gamma_i} + \omega[n] \right)^{-1}
  \]
Lemma 2 (Asymptotic Result for the Optimal Powers): Let $\Delta \in \mathbb{R}^{K \times K}$ be such that $[\Delta]_{k,i} := \frac{1}{NL} \frac{\gamma_i}{(1+\gamma_i)^2} \psi_{ik}$. If and only if $\limsup_K \|\Delta\|_2 < 1$, then

$$\max_k |p^*_k - p^o_k| \xrightarrow{N \to \infty} 0 \quad \text{almost surely}$$

where $p^o_k = \gamma_k \frac{\psi'_k}{\psi'^2_k} \left( \frac{\tau_k}{(1+\gamma_k)^2} + \sigma_k^2 \right)$

Here $\psi'_k = \frac{1}{L} \sum_{l=1}^L d_{kl} \eta_l$, $\psi'_{ik}$ and $\psi'_i$ are given by

$$\psi'_k = \frac{1}{L} \sum_{l=1}^L d_{kl} \eta_l^2 + \frac{1}{NL} \sum_{j=1}^K \frac{\lambda_{j}^{\psi'_i}}{(1+\gamma_j)^2} \frac{1}{L} \sum_{l=1}^L d_{il}d_{jl} \eta_l^2$$

$$\psi'_{ik} = \frac{1}{NL} \sum_{j=1}^K \frac{\lambda_{j}^{\psi'_i}}{(1+\gamma_j)^2} \frac{1}{L} \sum_{l=1}^L d_{il}d_{jl} \eta_l^2 + \frac{1}{L} \sum_{l=1}^L d_{il}d_{kl} \eta_l^2$$

$$\begin{align*}
\tau &= \sigma^2 (I_K - \Delta)^{-1} \delta \\
\delta_k &= \frac{1}{NL} \sum_{i=1}^K \gamma_i \frac{\psi'_{ik}}{\psi'^2_i}
\end{align*}$$
Theorem 2 (Asymptotic Result for RRH Ordering Criteria):

\[
\max_l |\theta_l - \theta_l^o| \xrightarrow{N \to \infty} 0 \quad \text{almost surely}
\]

where

\[
\theta_l^o = \frac{\kappa_l}{NL} \sum_{k=1}^{K} p_k^o \frac{\psi_{kl}}{\psi_k'}
\]

\[
\psi_{kl} = \frac{1}{NL} d_{kl} \eta_l^2 + \frac{1}{NL} \sum_{j=1}^{K} \frac{\lambda_j^{o2} \psi_{jl}}{(1+\gamma_j)^2} \frac{1}{L} \sum_{l=1}^{L} d_{il} d_{jl} \eta_l^2
\]

The ordering criteria will change only when the long-term channel attenuation is updated!

The second step to reduce computational complexity
Simulation Results (I)

- Convergence results (5 30-antenna RRHs and 5 single antenna MUs) [3]

Simulation Results (II)

- Network power minimization (5 10-antenna RRHs and 6 single antenna MUs) [3]

![Graph showing average network power consumption vs target SINR](image)

Conclusions

- **Network power minimization**: A difficult non-convex mixed combinatorial optimization problem

- **Key techniques (scalable algorithms design)**:
  - **GSBF**: convexify the combinatorial composite network power consumption function using the mixed $\ell_1/\ell_2$-norm
  - **Large-Scale Convex Optimization**:
    - Matrix stuffing: fast transformation
    - Operator splitting method (ADMM): large-scale HSD embedding
  - **Enhanced GSBF**:
    - Smoothed $\ell_p$-minimization with iterative reweighted-$\ell_2$ algorithm
    - Large random matrix theory: low computational complexity of RRH selection

- **Results**: group sparse optimization offers a principled way to design a dense green Cloud-RAN
Vignette B: Low-Rank Matrix Completion for Topological Interference Management
Issue B: Interference Management

- **Goal:** Interference mitigation in dense wireless networks

- **Prior works:** Abundant CSIT → Relaxed CSIT
  - **Perfect CSIT** [Cadambe and Jafar, TIT 08]
  - **Delayed CSIT** [Maddah-Ali and Tse, TIT 12]
  - **Alternating CSIT** [Tandon, et al., TIT 13], partial and imperfect CSIT [Shi, et al., TSP 14], …

- **Curses:** CSIT is rarely abundant (due to training & feedback overhead)

Start here? → Applicable? ← Prior works

No CSIT → CSIT → Perfect CSIT
Topological Interference Management

- **Blessings:** Partial connectivity in dense wireless networks

- **Approach:** Topological interference management (TIM) [Jafar, TIT 14]
  - **Maximize the achievable DoF:** Only based on the network topology information (no CSIT)

- Degrees of Freedom?

\[
\text{DoF} = \lim_{\text{SNR} \to \infty} \frac{C(\text{SNR})}{\log(\text{SNR})}
\]
TIM via Index Coding

- **Theorem [Jafar, TIT 14]:** Under linear (vector space) solutions, TIM problem and index coding problem are *equivalent*.

- **Bottleneck:** the only finite-capacity link

- **Only a few index coding problems have been solved!**
**TIM via LRMC**

- **Goal:** Deliver one data stream per user over $N$ time slots
  - $v_i \in \mathbb{C}^N$: tx. beamformer at the $i$-th tx.
  - $u_j \in \mathbb{C}^N$: rx. beamformer at the $j$-th rx.

- **We need:**
  \[
  X_{ij} = \begin{cases} 
  u_i^H v_i = 1, & \forall i, \\
  u_i^H v_j = 0, & \forall i \neq j, (i, j) \in \Omega, \\
  \ast, & \text{otherwise.}
  \end{cases}
  \]

- **Approach:** Low-rank matrix completion (LRMC) [4]

  \[
  \begin{align*}
  \text{minimize} & \quad \text{rank}(X) \\
  \text{subject to} & \quad \mathcal{P}_\Omega(X) = I_K
  \end{align*}
  \]

**Key conclusion:** DoF = $1/\text{rank}(X)$  

**Any network topology:** $\Omega$

---

LRMC & TIM & Index Coding

LRMC offers a new way to investigate these problems!
Riemannian Pursuit Algorithm

- **NP-hard**: Non-convex rank objective function

  \[
  \begin{aligned}
  &\text{minimize} & \quad \text{rank}(X) \\
  &\text{subject to} & \quad P_\Omega(X) = I_K \\
  & & \quad |\text{Tr}(X)| \leq \|X\|_*
  \end{aligned}
  \]

- **Poorly structured affine constraint**:
  - Nuclear-norm relaxation [Candes & Recht, FCM 09]: \(X^* = I_K\) (full rank)

- **Riemannian pursuit [4]**: Alternatively perform the fixed-rank optimization and rank increase

  \[
  \begin{aligned}
  &\text{minimize} & \quad \|P_\Omega(X) - I_K\|_F^2 \\
  &\text{subject to} & \quad \text{rank}(X) = r
  \end{aligned}
  \]

- **Riemannian optimization**: address convergence issues in fixed-rank methods
Riemannian Optimization for Fixed-Rank Problems

- Solve **fixed-rank problems** by Riemannian optimization [Absil, et al., 08]
  - Generalize Euclidean gradient (Hessian) to **Riemannian gradient (Hessian)**

\[
\nabla_{\mathcal{M}} f(X^{(k)}) = P_{X^{(k)}}(\nabla f(X^{(k)}))
\]

Riemannian Gradient  Euclidean Gradient

\[
X^{(k+1)} = \mathcal{R}_{X^{(k)}}(-\alpha^{(k)} \nabla_{\mathcal{M}} f(X^{(k)}))
\]

Retraction Operator

**Quotient manifold geometry of fixed rank matrices**

\[
[X] = \{(UQ_U, Q_U^T \Sigma Q_V, VQ_V) : Q_U, Q_V \in \mathcal{O}(r)\}
\]
Numerical Results (1): Convergence Rate

- Riemannian optimization over the quotient matrix manifold [4].

Riemannian algorithms:
1. Exploit the rank structure in a principled way
2. Develop second-order algorithms systematically
3. Scalable, SVD-free

Numerical Results (II): Symmetric DoF

Advantages:
1. Recover all the optimal DoF results for the special TIM problems in [Jafar ’14]
2. Provide numerical insights (optimal/lower-bound) for the general TIM problems
Conclusions

- **Topological interference management** significantly improves DoFs only based on the network topology information

- **Key techniques:**
  - Low-rank matrix completion
  - Riemannian optimization

- **Results:** **Low-rank matrix completion** provides a first algorithmic and systematic approach to investigate the TIM problem for any network topology.

- **Extensions:**
  - *User admission control, network topology design, finite SNR, …*
  - More applications: index coding, distributed storage and caching, …
  - Optimality: Riemannian pursuit algorithm, LRMC approach
Concluding Remarks

- Future network design: dense, cooperative, scalable, unified

1. **Structured models**: Sparsity, low-rankness
2. **Scalable algorithms**: Convex optimization, Riemannian optimization, ADMM
3. **Theory**: Global optimality?
Further Information: Sparse Optimization


Further Information: Low-Rank Optimization


Personal Webpage: http://shiyuanming.github.io/home.html
Thanks