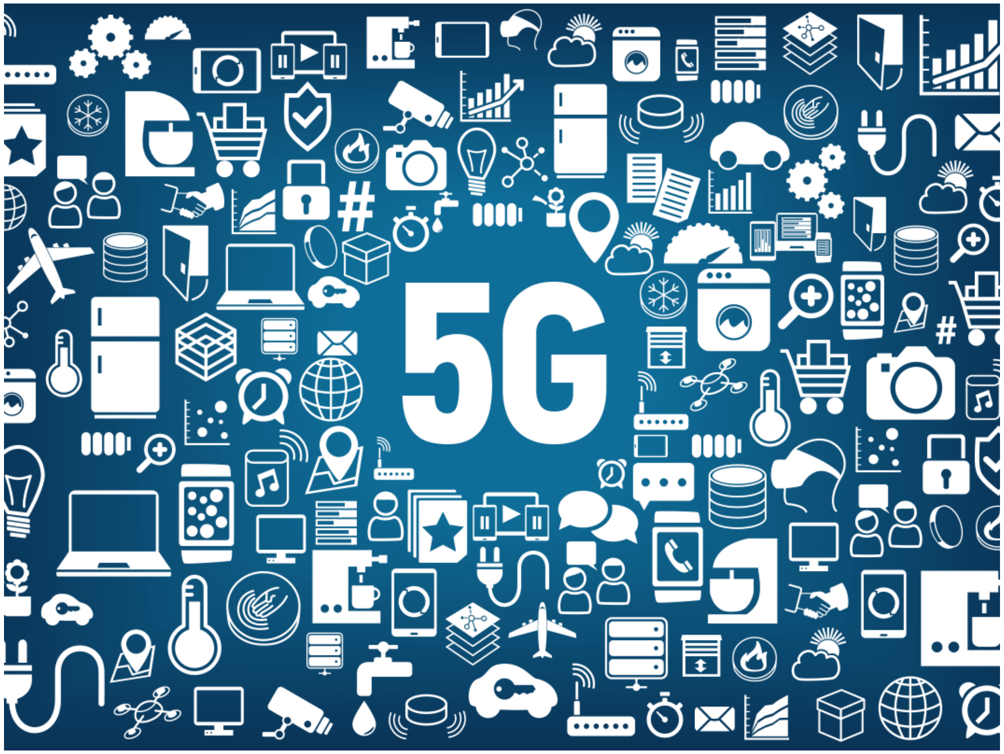
Dense Cooperative Wireless Networks for 5G

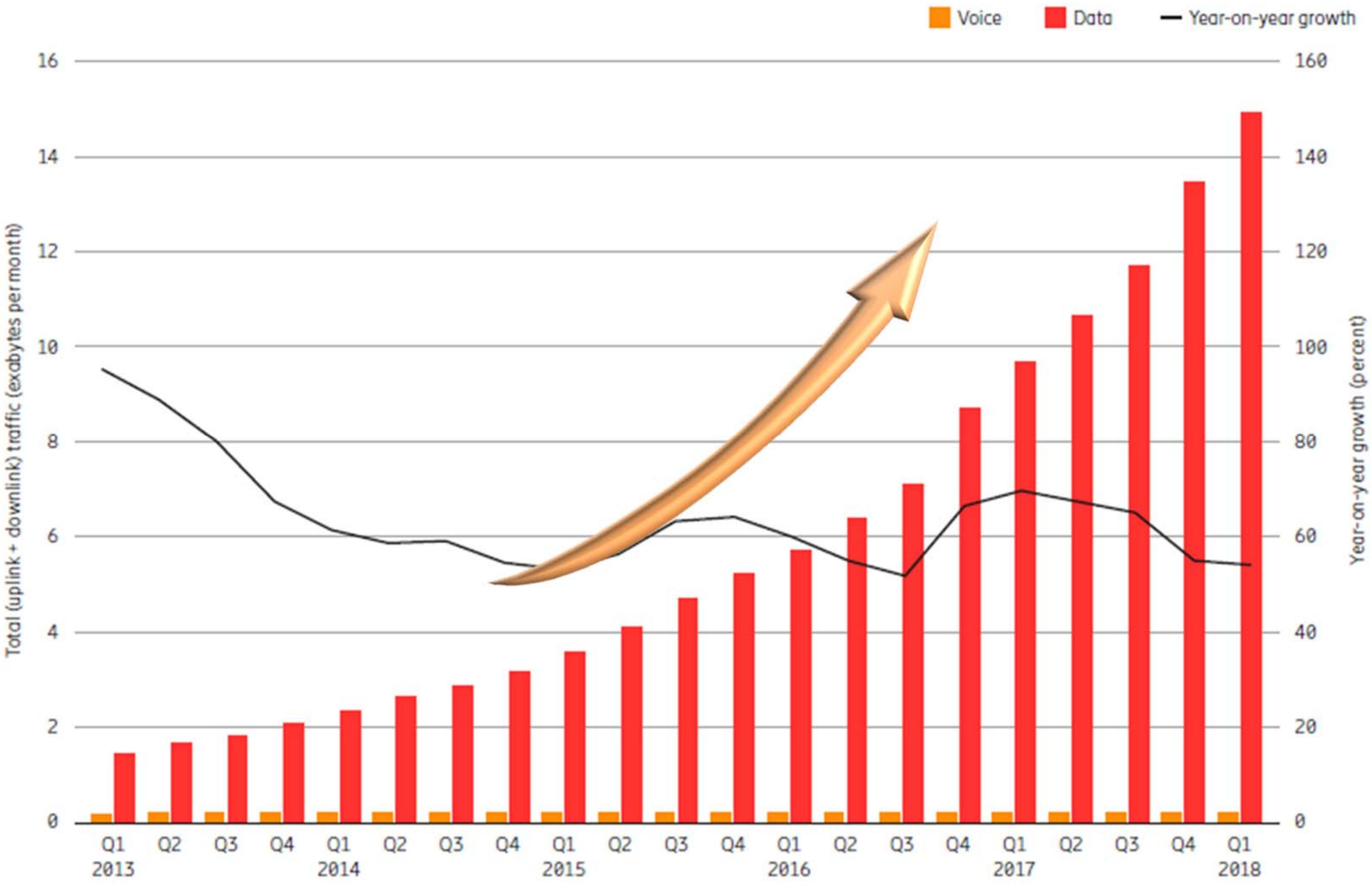
Jun ZHANG



Outline

- Background – Why we need 5G?
- Millimeter-wave communications
- Structured optimization for cooperation
- Large-scale network analysis
- Takeaways

Exponentially Growing Mobile Data



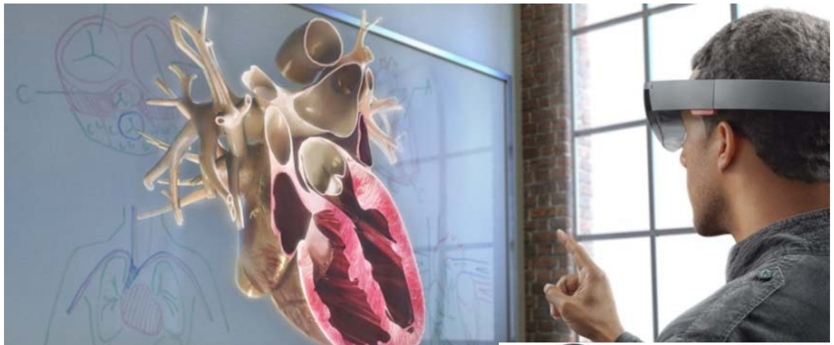
Ericsson traffic measurements (Q1 2018)

1 EB (Exabyte) = 10^{18} B
= 1 billion GB

Exploding Mobile Applications

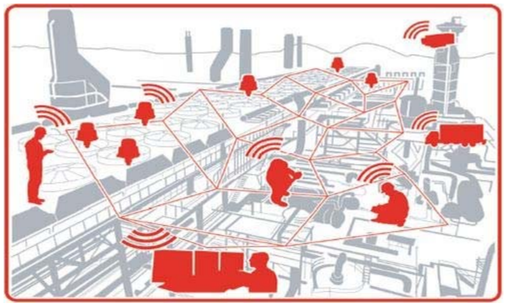
Source: Lux Review

Intelligent Mobile Applications



Higher data rate

Massive connections



Uniform coverage

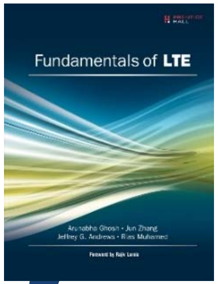
Security & privacy

Green communications

A (Personal) Journey From 4G to 5G

Cloud-RANs
HetNets
Small cells
D2D
Massive MIMO
Mm-wave Communications
Wireless Caching
Mobile Edge Computing
Cloud Computing

September 2010



Fundamentals of **LTE**

Arunabha Ghosh • Jun Zhang
Jeffrey G. Andrews • Rias Muhamed

THE UNIVERSITY OF TEXAS AT AUSTIN

Nov 17, 2009

2008

Dec 14 2009, first LTE deployment in Stockholm and Oslo

Q1 2017, 5G initial study

Dec 2017, 5G Rel-15 non-standalone (NSA)

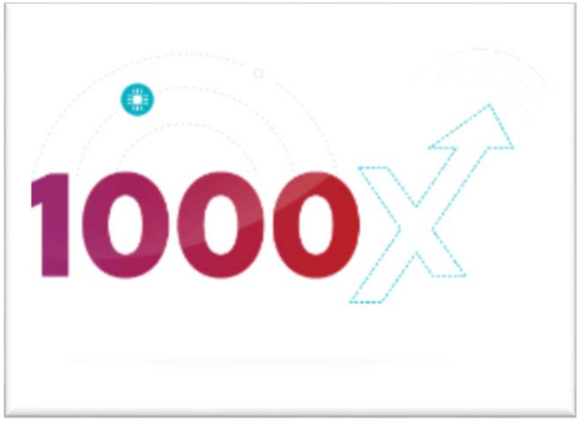
2018

Jun 2018, 5G Rel-15 standalone (SA)

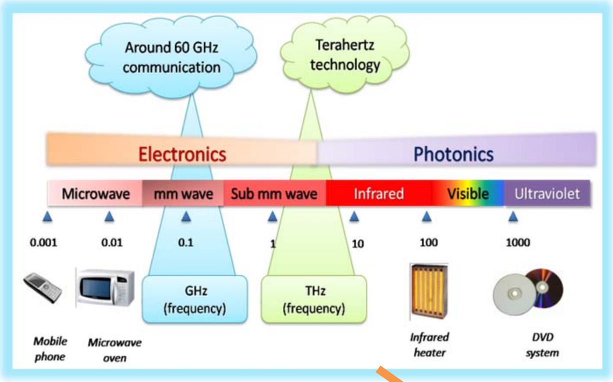
2020

The Holy Grail for 5G

-- The **1000X** Capacity Challenge

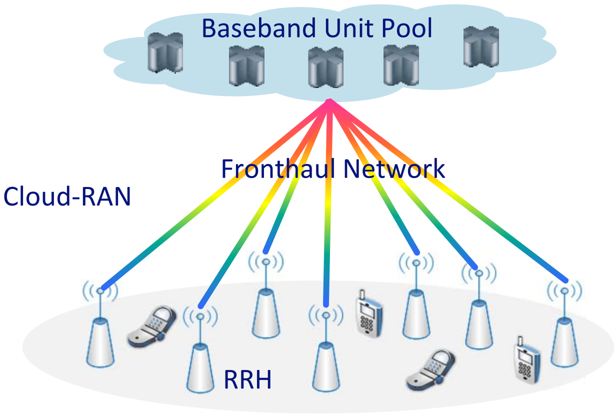


The 1000X Capacity Challenge

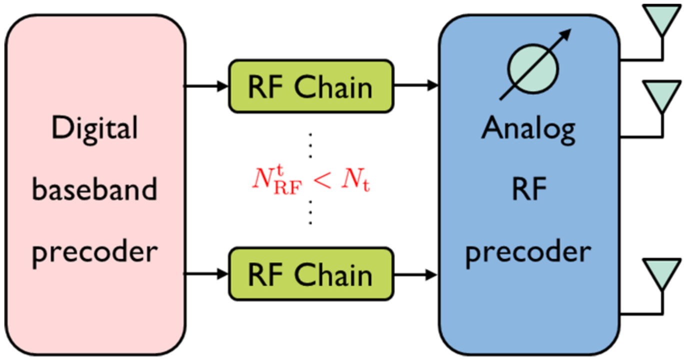


Capacity = Bandwidth (Hz) x Spectral Efficiency (bps/Hz) x # Links

1000 = 10 x 5 x 20



Hybrid Beamforming for Millimeter-Wave Communications



Digital baseband precoder

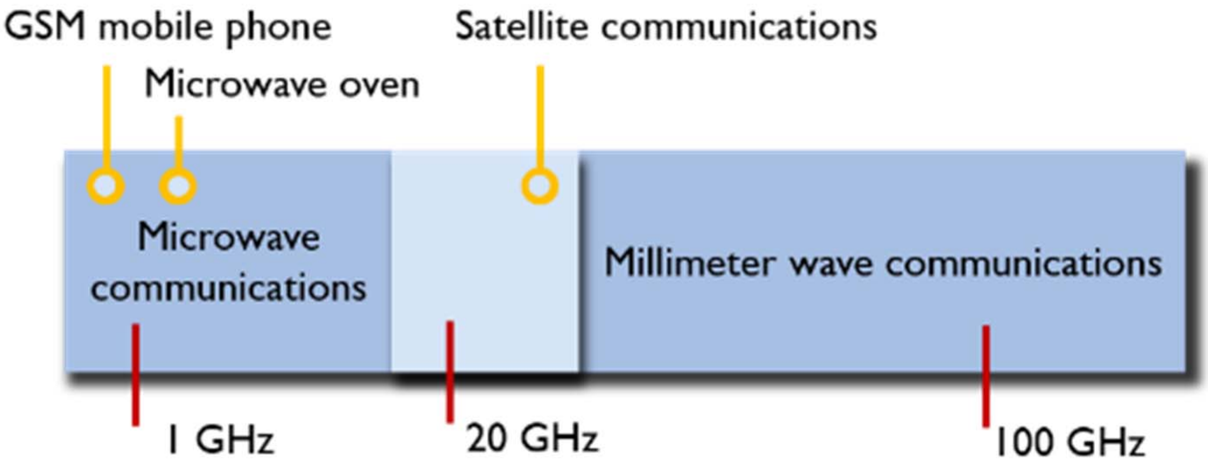
RF Chain

$N_{\mathrm{RF}}^{\mathrm{t}} < N_{\mathrm{t}}$

RF Chain

Analog RF precoder

Why mm-wave?



- Current cellular systems (microwave spectrum) ~ **600MHz**
- **Mm-wave spectrum ~ 29GHz**

5G = Millimeter-wave

[At least to someone]

Design Challenges

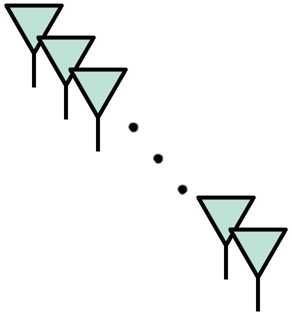
Data Rate = Spectral Efficiency x Bandwidth

Challenge I

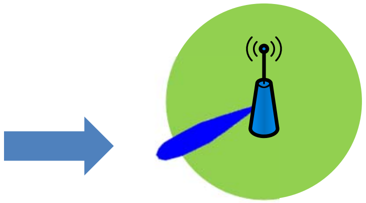
Receive Power: $P_r = \frac{P_t}{4\pi d^2} \frac{\lambda^2}{4\pi}$ ↓

Noise Power: $N_o = kT_e B$ ↑

SNR ↓



Large-scale antenna array



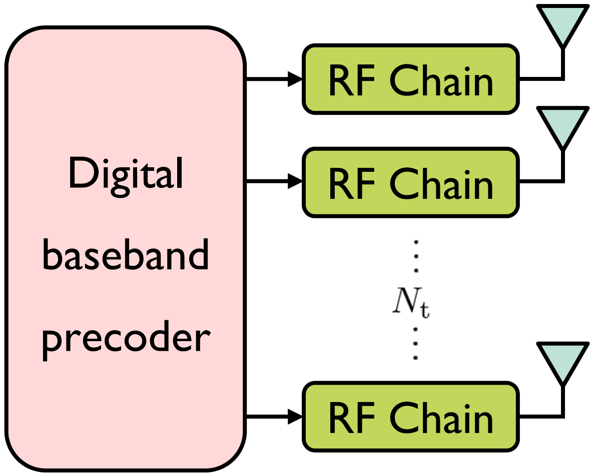
Beamforming

$$P_r = \frac{P_t}{4\pi d^2} \frac{\lambda^2}{4\pi} G_r G_t$$

Beamforming gains

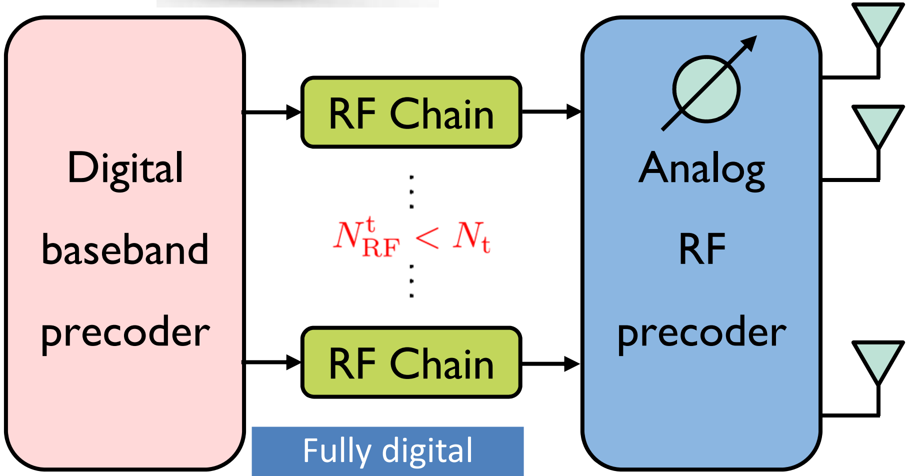
Design Challenges

Fully digital beamformer



Challenge II

Hybrid beamformer



Fully digital beamformer

$$\underset{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{BB}}}{\text{minimize}} \quad \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2$$

$$\text{subject to} \quad \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 \leq P_{\max}$$

$$|(\mathbf{F}_{\mathrm{RF}})_{i,j}| = 1, (i,j) \in \mathcal{A}$$

Challenge III

Non-convex, Large N_t

Main Design Objectives

- **High spectral efficiency**
  - Approach fully digital beamforming
- **High computational efficiency**
  - Overcome non-convexity, large problem sizes
- **High hardware efficiency**
  - With as few hardware components as possible

Spectral Efficiency: Manifold Optimization

- **Difficulty**: Analog beamformer with unit modulus constraints

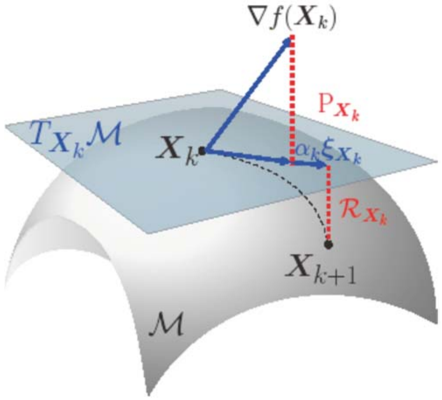
$$\begin{aligned} \underset{\mathbf{F}_{\mathrm{RF}}}{\text{minimize}} \quad & \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 \\ \text{subject to} \quad & |(\mathbf{F}_{\mathrm{RF}})_{i,j}| = 1, \forall i, j. \end{aligned}$$

Non-convex

But forms a complex circle manifold

Proposal

- Manifold optimization for analog beamformer
  - *Optimization over **non-Euclidean geometry***
- Alternating minimization for analog/digital

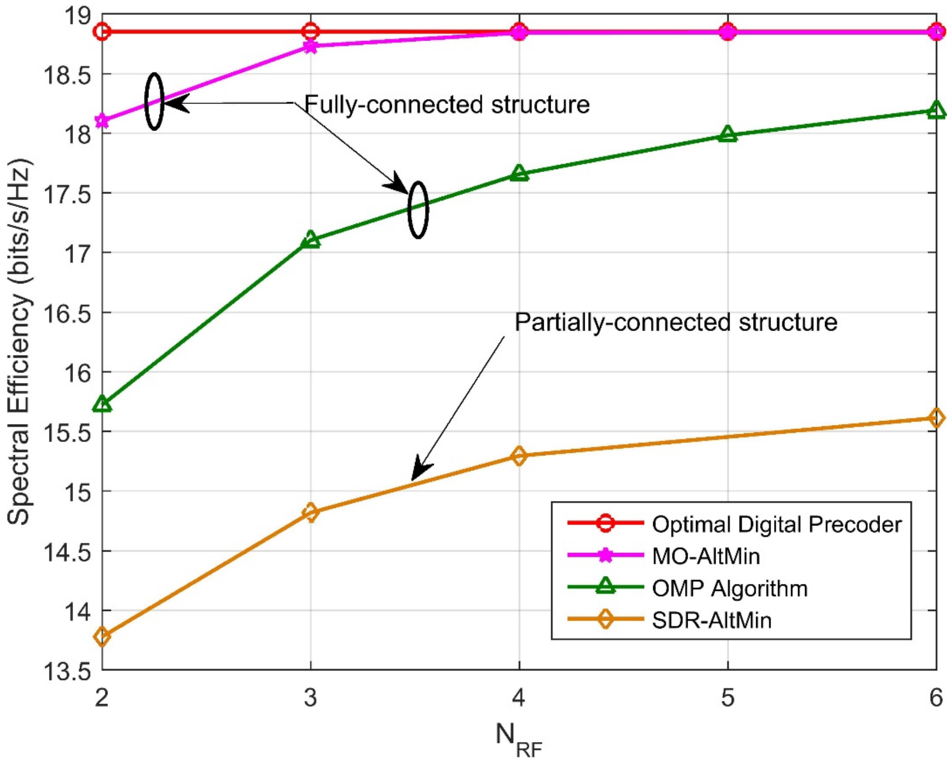


Spectral Efficiency: Manifold Optimization

- How many RF chains are needed?

data streams

$N_{\mathrm{t}} = 144,\ N_{\mathrm{r}} = 36,\ N_{\mathrm{RF}}^{\mathrm{t}} = N_{\mathrm{RF}}^{\mathrm{r}} = N_{\mathrm{RF}},\ N_s = 2,\ \mathrm{SNR} = 0\,\mathrm{dB}$



RF chains ~ # data streams enough to approach the fully digital beamformer

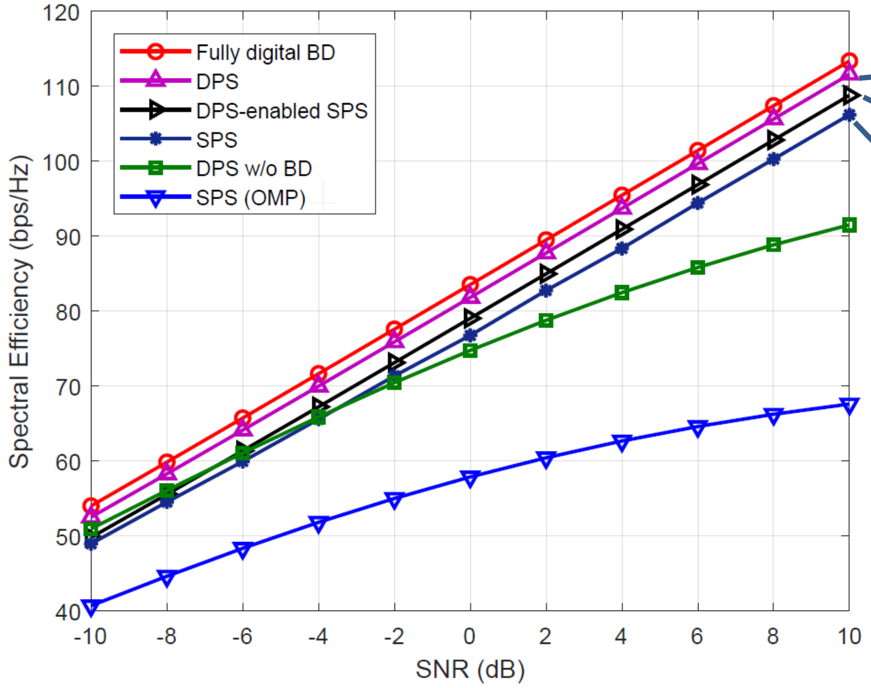
X. Yu, J.-C. Shen, **J. Zhang**, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485-500, Apr. 2016.

Computational Efficiency: Convex Relaxation

$$\begin{aligned} &\underset{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{BB}}}{\text{minimize}} && \left\|\mathbf{F}_{\mathrm{opt}}-\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F^2 \\ &\text{subject to} && \left|(\mathbf{F}_{\mathrm{RF}})_{i,j}\right|=1 \end{aligned} \Rightarrow \begin{aligned} &\underset{\mathbf{F}_{\mathrm{RF}},\mathbf{F}_{\mathrm{BB}}}{\text{minimize}} && \left\|\mathbf{F}_{\mathrm{opt}}-\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\right\|_F^2 \\ &\text{subject to} && \left|(\mathbf{F}_{\mathrm{RF}})_{i,j}\right|\le 1 \end{aligned}$$

- Solution of relaxed problem: **SVD**
- Analog beamformer: **Phase extraction**

$N_{\mathrm{t}}=256,\ N_{\mathrm{r}}=16,\ K=3,\ F=128,\ N_s=3,\ N_{\mathrm{RF}}^{\mathrm{t}}=9,$ and $N_{\mathrm{RF}}^{\mathrm{r}}=3$



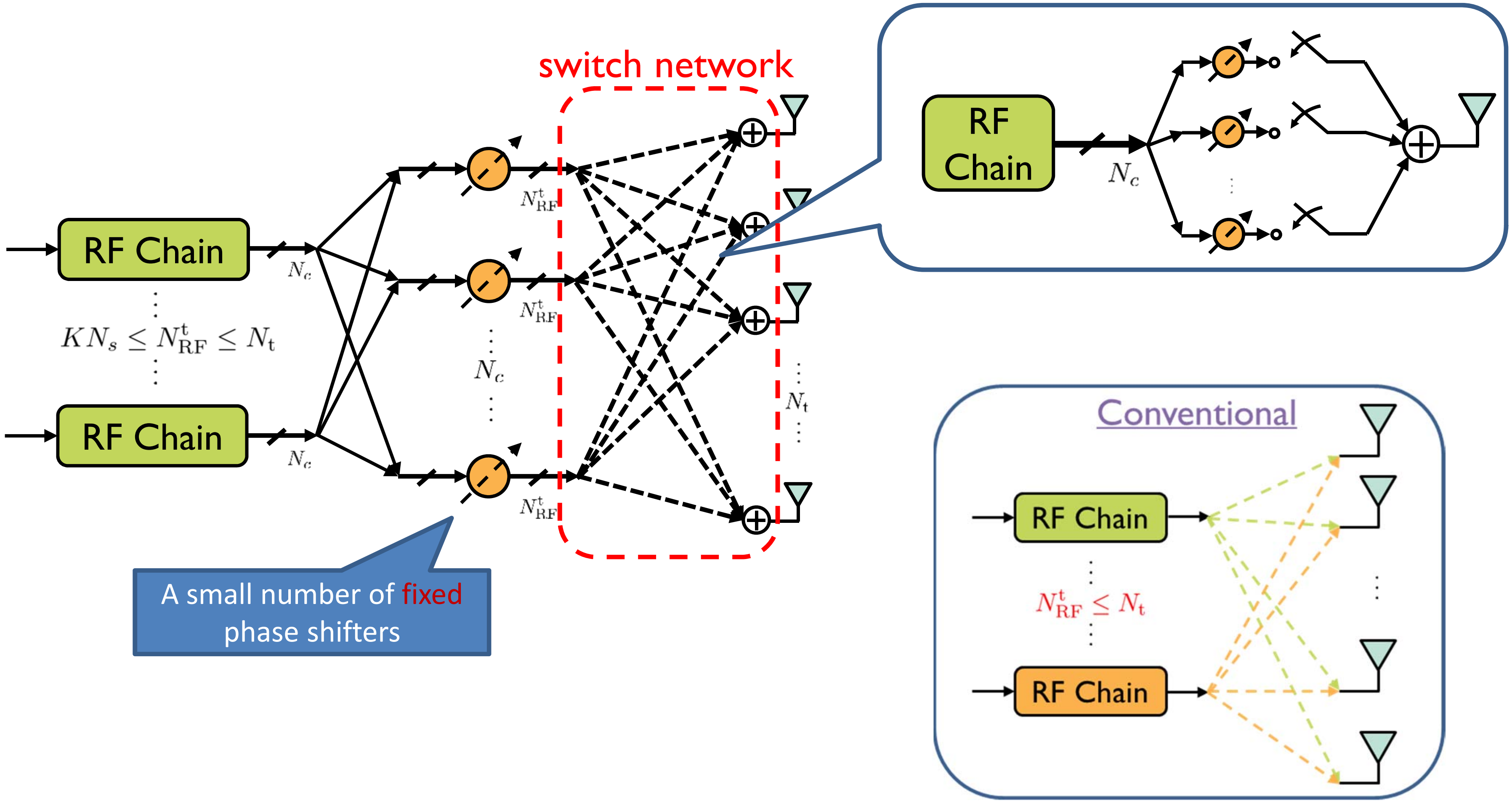
Relaxed solution

With phase extraction

Existing algorithm

X. Yu, **J. Zhang**, and K. B. Letaief, "Alternating minimization for hybrid precoding in multiuser OFDM mmWave Systems," *IEEE Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2016. (**Invited Paper**)

Hardware Efficiency: A New Analog Network

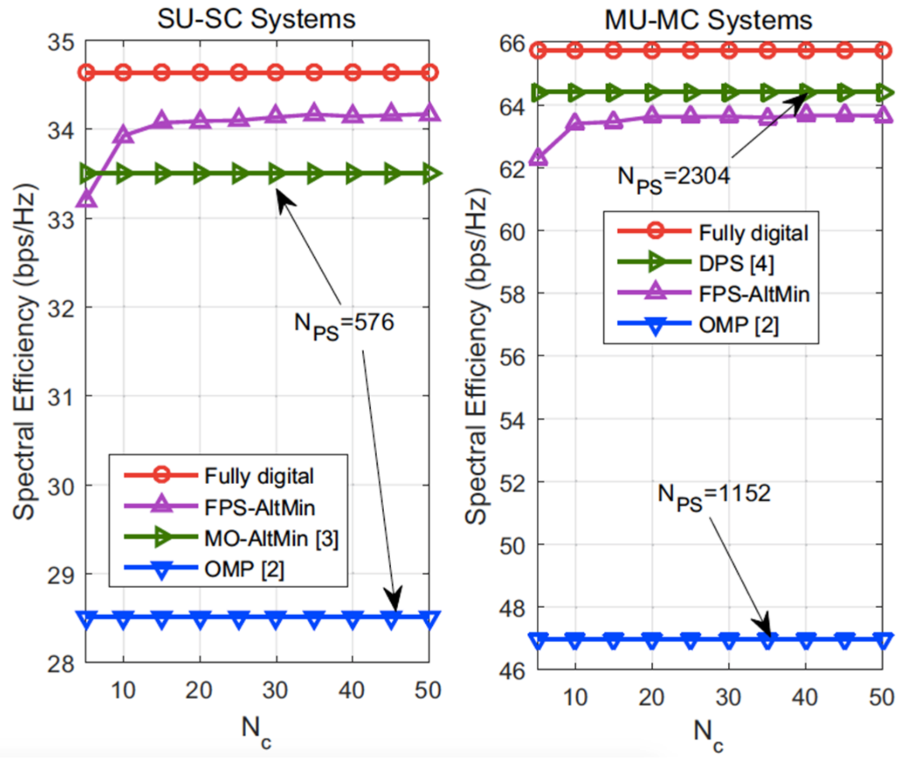


A small number of fixed phase shifters

Hardware Efficiency: A New Analog Network

- How many phase shifters are needed?

$N_{\mathrm{t}} = 256,\ N_{\mathrm{r}} = 16,\ K = 4,\ F = 128,\ N_s = 2,\ N_{\mathrm{RF}}^{\mathrm{t}} = 8$, and $N_{\mathrm{RF}}^{\mathrm{r}} = 2$

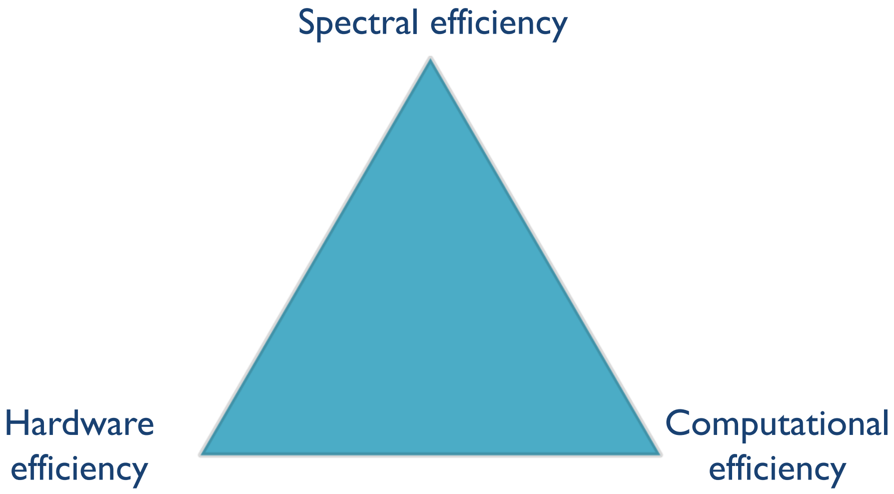


Only ~15 fixed phase shifters are sufficient!

X. Yu, **J. Zhang**, and K. B. Letaief, "Hybrid precoding in millimeter wave systems: How many phase shifters are needed?" in *Proc. IEEE Globecom*, Singapore, Dec. 2017. **(Best Paper Award)**

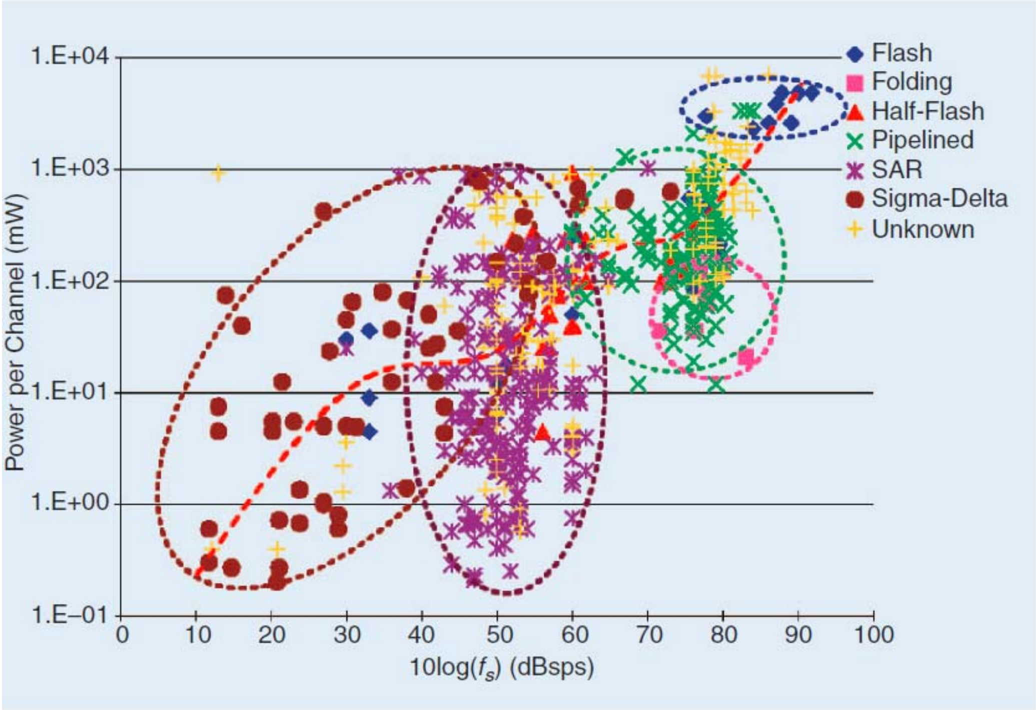
Conclusions

- A new design aspect
  - **Hardware-aware communications**
- An effective design approach
  - **Hardware-algorithm co-design**
- Three key design aspects



Future Direction I

- Mm-wave with low-precision ADCs



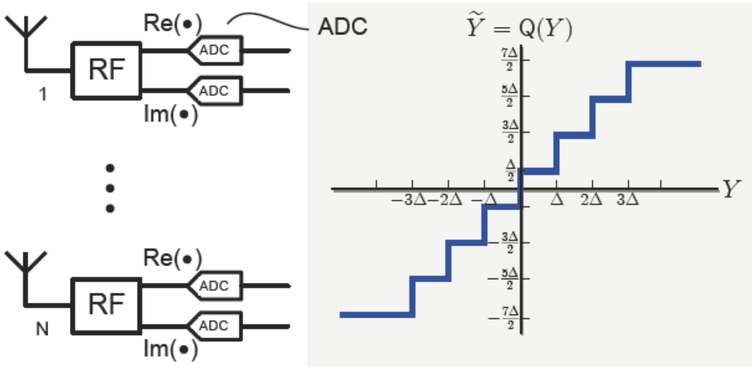
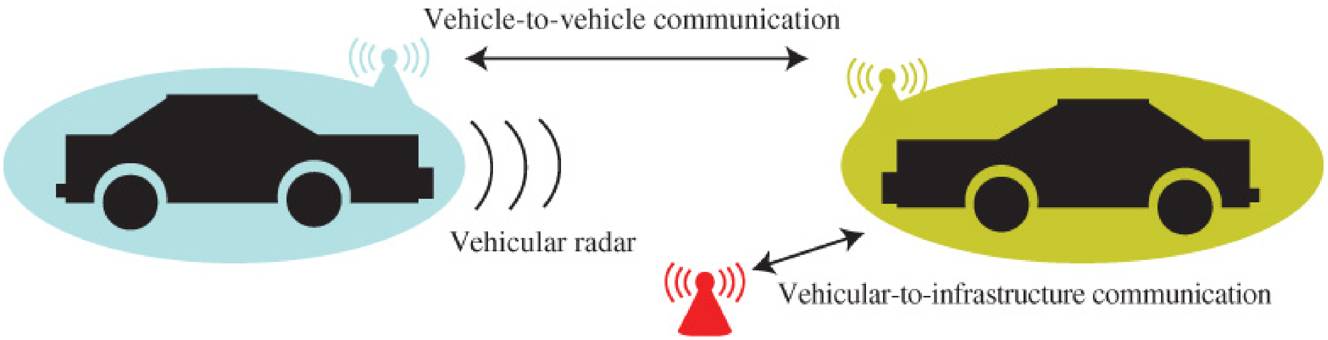


Fig. 1. The MIMO system with low precision ADCs.

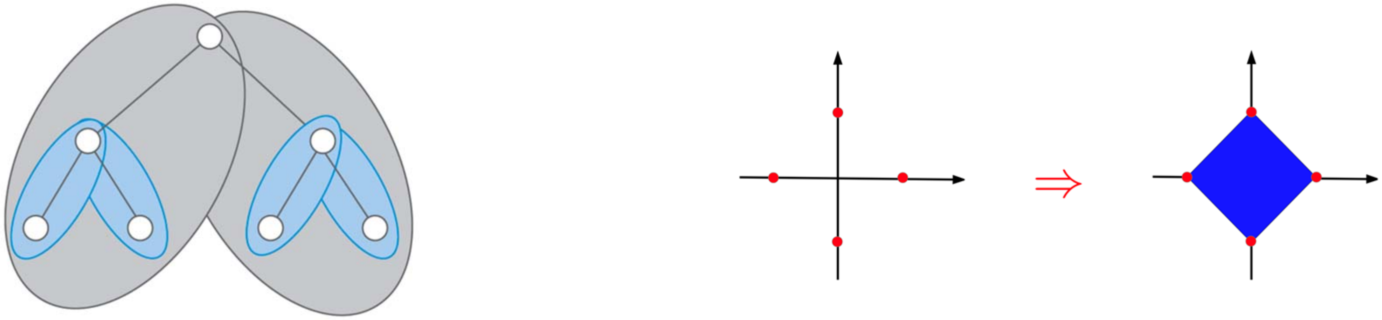
Future Direction II

- Joint mm-wave communication-radar for connected vehicles



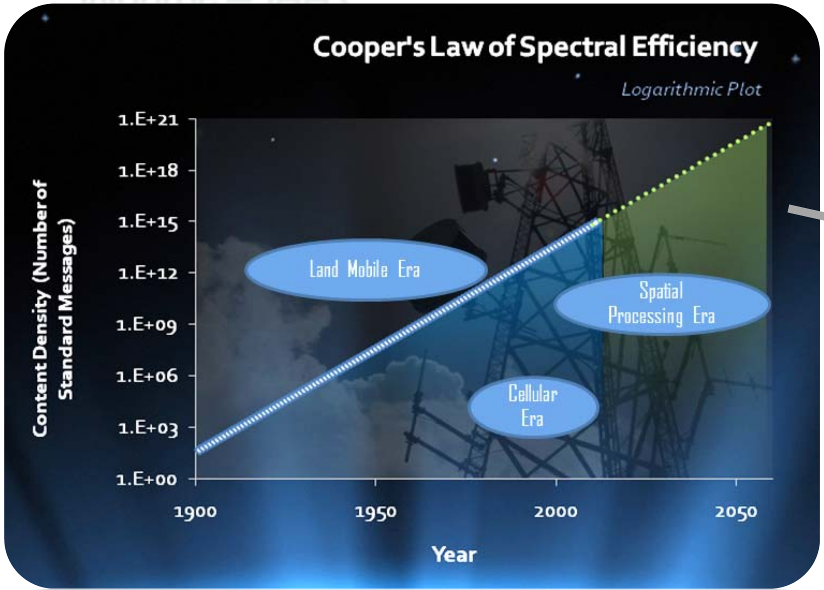
  - **Radar:** Continuous automatic detection/ranging, even for non-communication-enabled users
  - **V2V**: Collaborative communication for real-time cooperative detection and ranging

Structured Optimization for Cooperative DenseNets

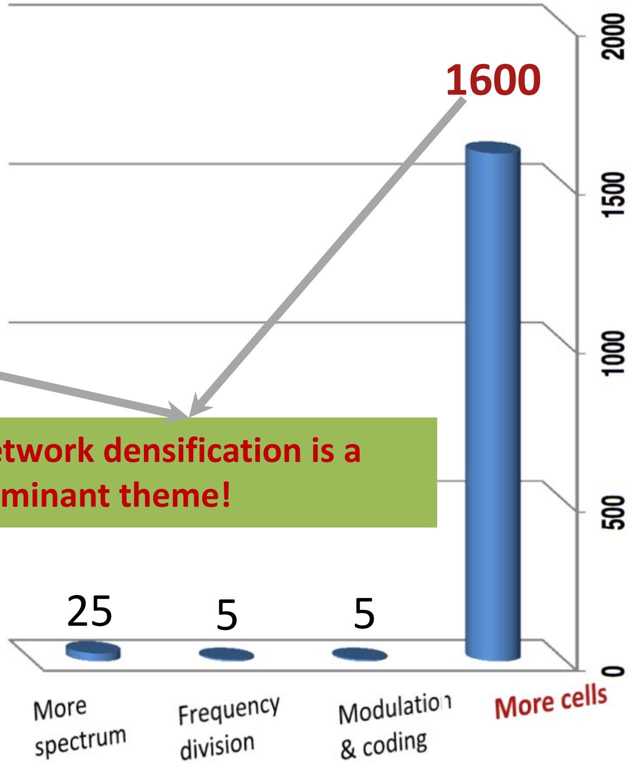


Network Densification: The Dominant Theme for 5G

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



Factor of Capacity Increase since 1950

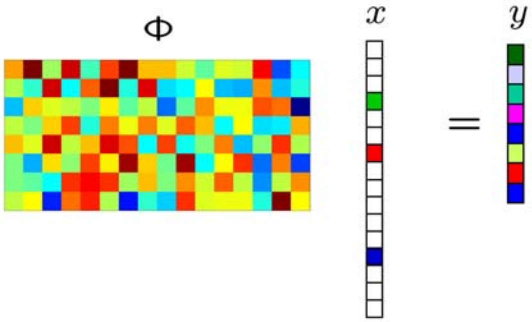


Network densification is a dominant theme!

Cooperation is critical for interference management and resource allocation in wireless dense networks (DenseNets)

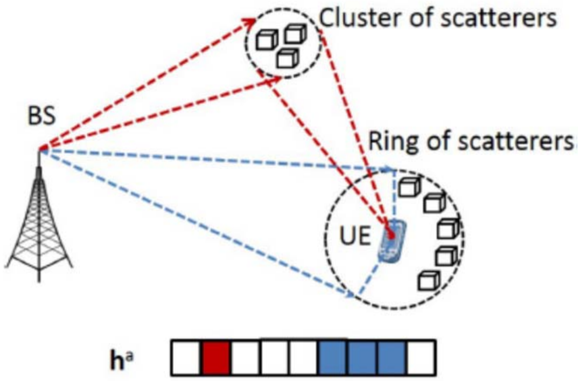
Structured Optimization/Estimation

- Previous successes



1		0	0	
	1	0	0	
0		1		0
0			1	0
	0			1

- Emerging applications in wireless networks



Case Study: Network Power Minimization in Cloud-RAN

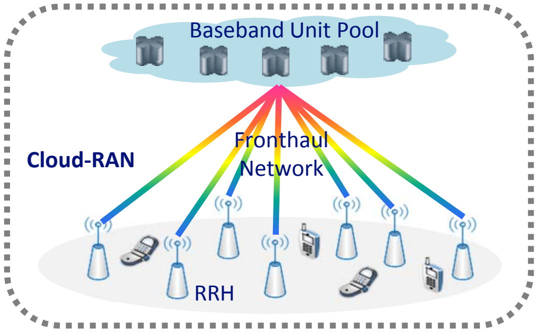
- **Problem Formulation**

$$\underset{\mathbf{v}\in\mathcal{C}}{\text{minimize}} \underbrace{\sum_{l=1}^{L} P_l^c I(\text{Supp}(\mathbf{v}) \cap \mathcal{G}_l \neq \emptyset)}_{\text{fronthaul power}} + \underbrace{\sum_{l=1}^{L} \sum_{k=1}^{K} \frac{1}{\eta_l} \|\mathbf{v}_{lk}\|_2^2}_{\text{transmit power}}$$

NP-hard

Jon Lee
Sven Leyffer *Editors*
Mixed Integer
Nonlinear
Programming

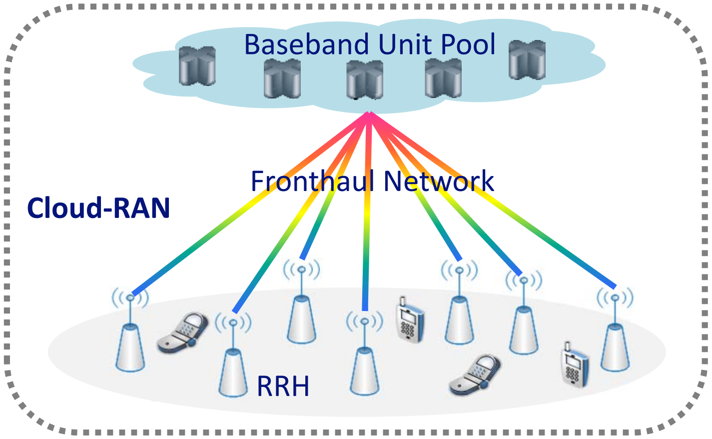
Large Scale



- **Prior works: *Heuristic or computationally expensive*** [Philipp, et. al, TSP 13], [Luo, et. al, JSAC 13], [Quek, et. al, TWC 13]

Innovation I: Group Sparse Beamforming

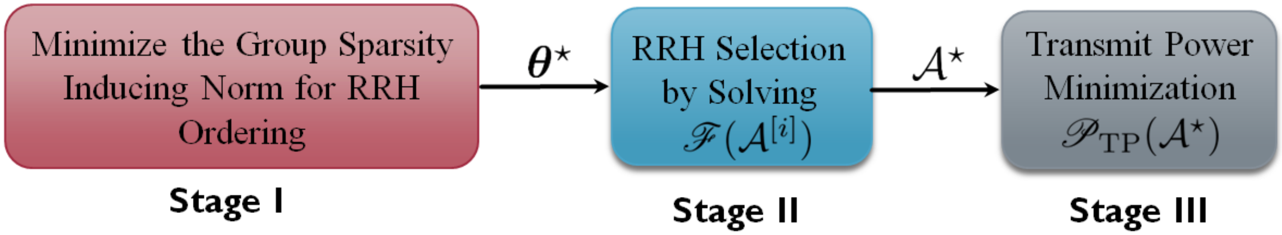
- Proposal: Identify solution structures



$$\mathbf{v} = [\underbrace{\mathbf{v}_{11}^T, \dots, \mathbf{v}_{1K}^T}_{\tilde{\mathbf{v}}_1^T}, \dots, \underbrace{\mathbf{v}_{L1}^T, \dots, \mathbf{v}_{LK}^T}_{\tilde{\mathbf{v}}_L^T}]^T$$

Beamforming coefficients of the first RRH, forming a **group**

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

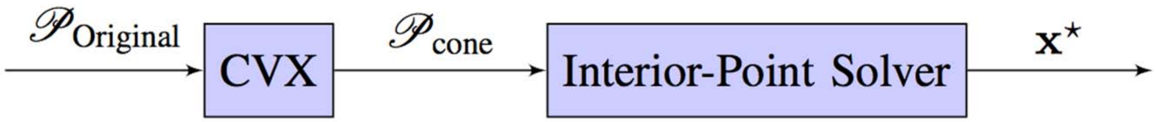


Computational Challenge

- Prior works: Mainly focus on small-size networks
- Unique challenges in dense Cloud-RANs

  1) High dimension; 2) Large number of constraints; 3) complicated structures

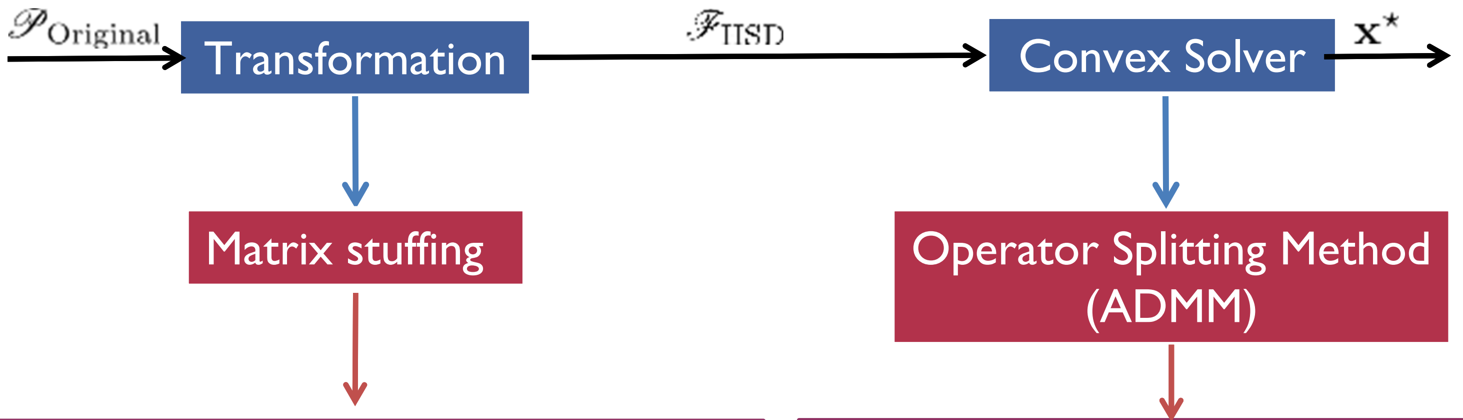
- **Existing approach**: Disciplined convex programming framework [Grant & Boyd '08]



$\mathscr{P}_{\text{Original}} \rightarrow$ CVX $\xrightarrow{\mathscr{P}_{\text{cone}}}$ Interior-Point Solver $\xrightarrow{\mathbf{x}^\star}$

- General purpose, inefficient

Innovation II: A Two-stage Optimization Framework



$\mathscr{P}_{\text{Original}}$ → Transformation → $\mathscr{F}_{\text{HSD}}$ → Convex Solver → $\mathbf{x}^\star$

Transformation → Matrix stuffing

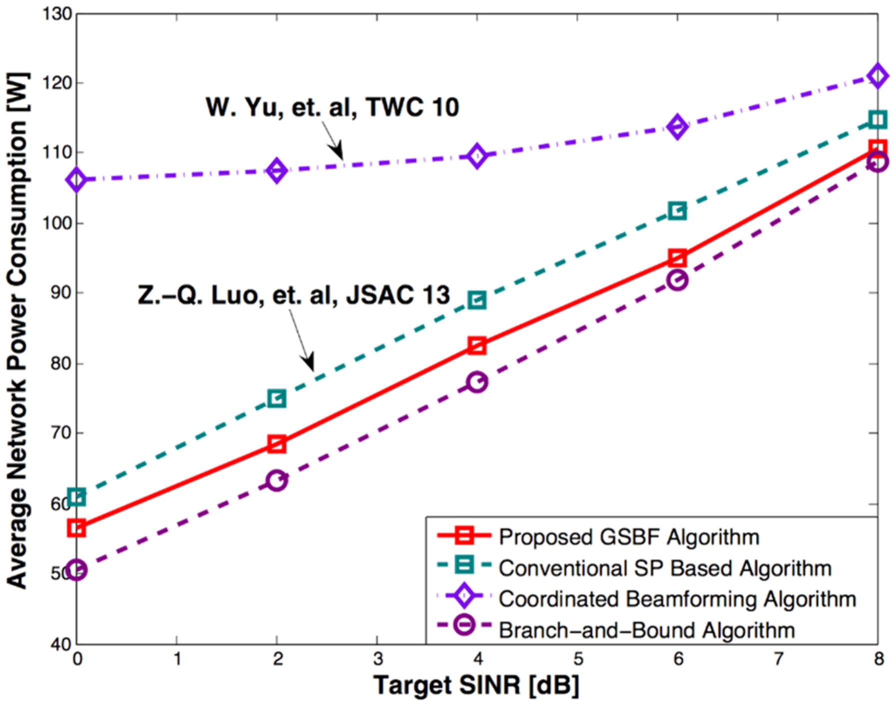
Convex Solver → Operator Splitting Method (ADMM)

- ➢ Given a network, generate and keep a homogeneous self-dual (HSD) embedding form structure.
- ➢ Only copy the problem data to the HSD embedding data.
- ➢ Avoid repeated transforming problems.

- ➢ Each iteration can be solved with closed-form and in parallel.
- ➢ Utilize cloud computing environment in the BBU pool with computational resources sharing among BBUs.

The Power of Group Sparse Beamforming

- Group spare beamforming (10 RRHs, 15 MUs)



Further improvement

- Sparsity enhancement [JSAC 2016];
- Random matrix approach [TWC 2018]

Y. Shi, **J. Zhang**, and K. B. Letaief, "Group sparse beamforming for green Cloud-RAN," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2809-2823, May 2014. 2014. **(The 2016 Marconi Prize Paper Award)**

The Efficacy of Large-scale Optimization

- **Example:** Power minimization coordinated beamforming problem

Network Size (*L*=*K*)		20	50	100	150
CVX+SDPT3	Modeling Time [sec]	**0.7563**	4.4301	N/A	N/A
	Solving Time [sec]	4.2835	**326.2513**	N/A	N/A
	Objective [W]	12.2488	6.5216	N/A	N/A
Matrix Stuffing+ADMM	Modeling Time [sec]	**0.0128**	0.2401	2.4154	9.4167
	Solving Time [sec]	0.1009	**2.4821**	23.8088	81.0023
	Objective [W]	12.2523	6.5193	3.1296	2.0689

Matrix stuffing can speedup 60x over CVX

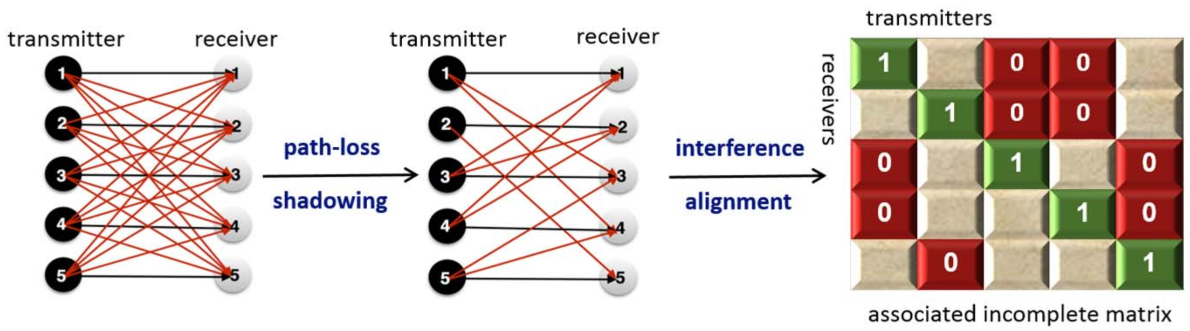
ADMM can speedup 130x over the interior-point method

Y. Shi, **J. Zhang**, B. O'Donoghue, and K. B. Letaief, "Large-scale convex optimization for dense wireless cooperative networks," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4729-4743, Sept. 2015. **(The 2016 IEEE Signal Processing Society Young Author Best Paper Award)**

Conclusions

- Structured optimization is a powerful tool for designing cooperative DenseNets
- Large-scale convex optimization algorithms are necessary

- Other applications
  - Sparse estimation for channel information acquisition
    - J.-C. Shen, **J. Zhang**, et. al, "Compressed CSI acquisition in FDD massive MIMO: How much training is needed?" IEEE Trans. Wireless Commun., 2016.
  - Low-rank matrix completion for topological interference management
    - Y. Shi, **J. Zhang**, and K. B. Letaief, "Low-rank matrix completion for topological interference management by Riemannian pursuit," IEEE Trans. Wireless Commun., 2016.



1		0	0	
	1	0	0	
0		1		0
0			1	0
	0			1

Future Direction I

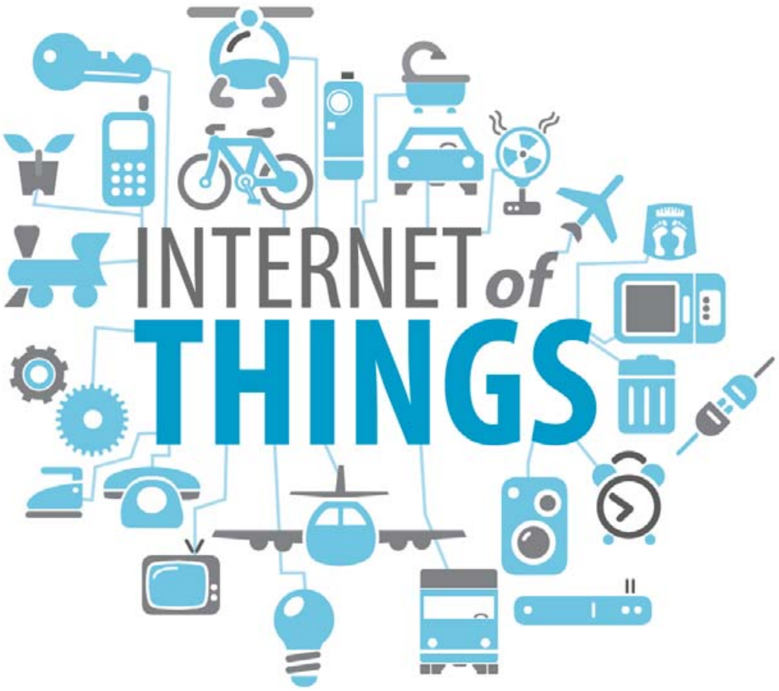
- Massive connectivity for IoT applications

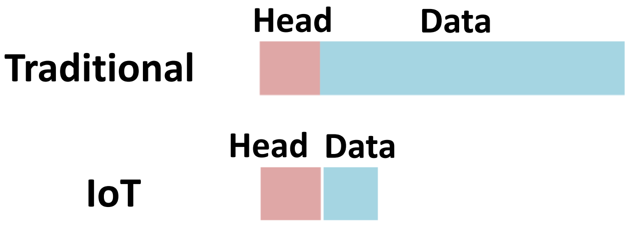
Sporadic traffic: only a small fraction of potentially large number of devices are active

Sparse estimation for user activity detection and channel estimation

Future Direction II

- Overhead-free communications for IoT





Overhead-free communication

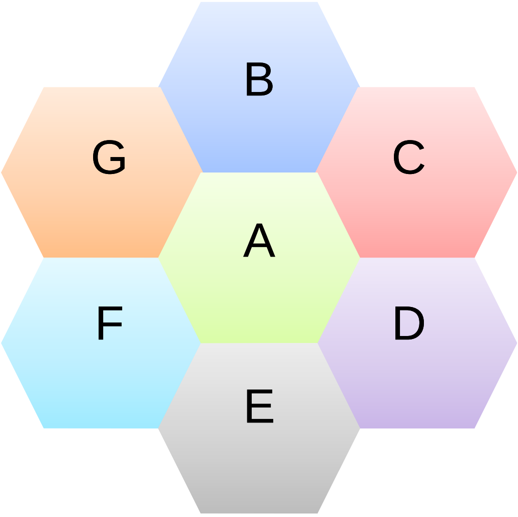
$$y = \sum_{i=1}^{K} h_i * x_i + w$$

Target: Simultaneous channel estimation and signal detection via blind deconvolution/demixing

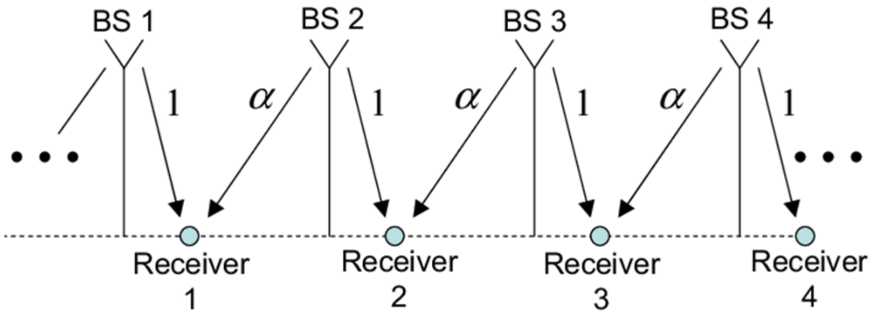
Large-scale Network Analysis via Stochastic Geometry

Traditional Approaches for Network Analysis

- Grid model



- Wyner model

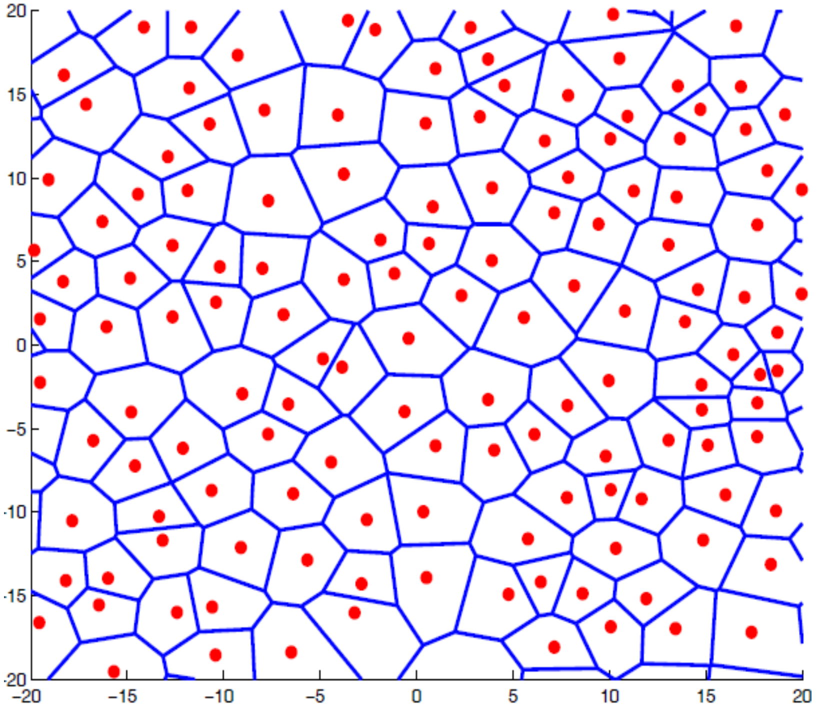


- Either less accurate or weakens the analytical tractability

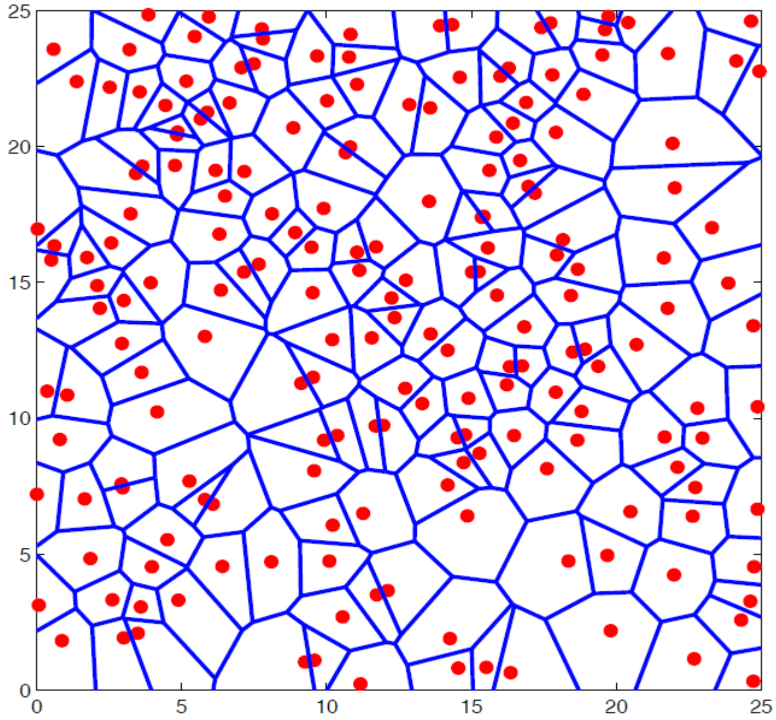
J. Xu, **J. Zhang**, and J. G. Andrews, "On the accuracy of the Wyner model in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3098-3109, Sept. 2011.

Stochastic Geometry for Wireless Networks

Base station locations in a 4G network

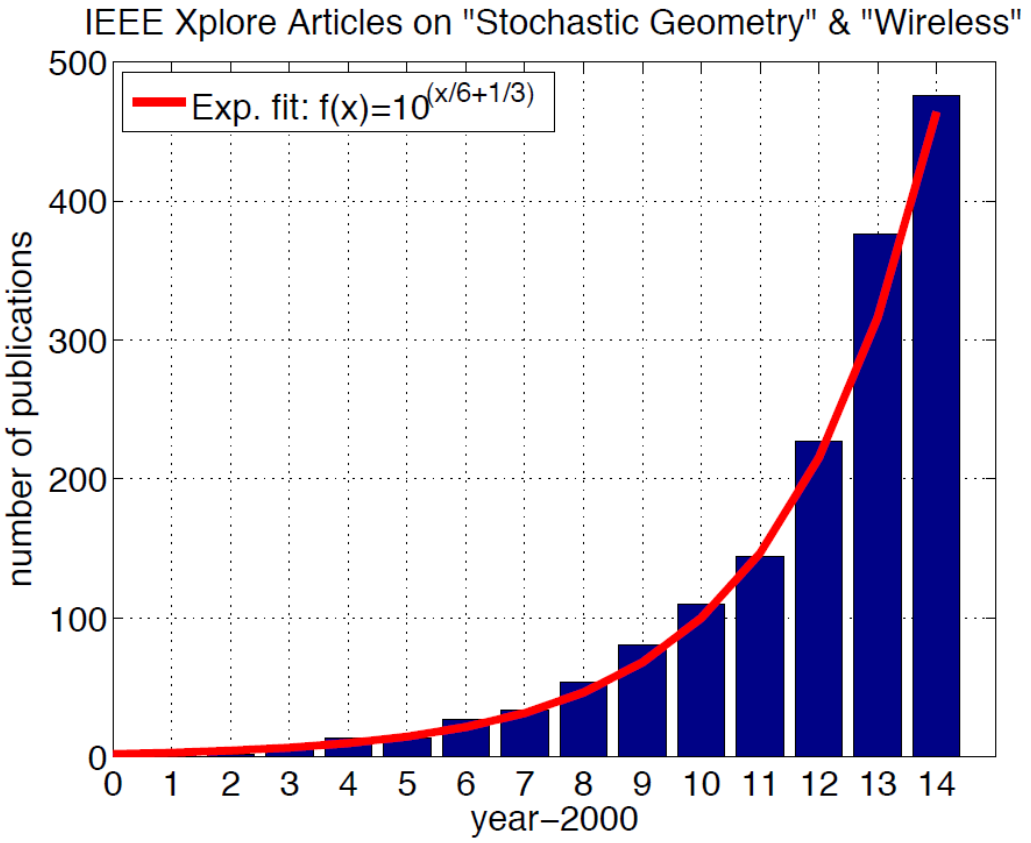


Poisson distributed base stations

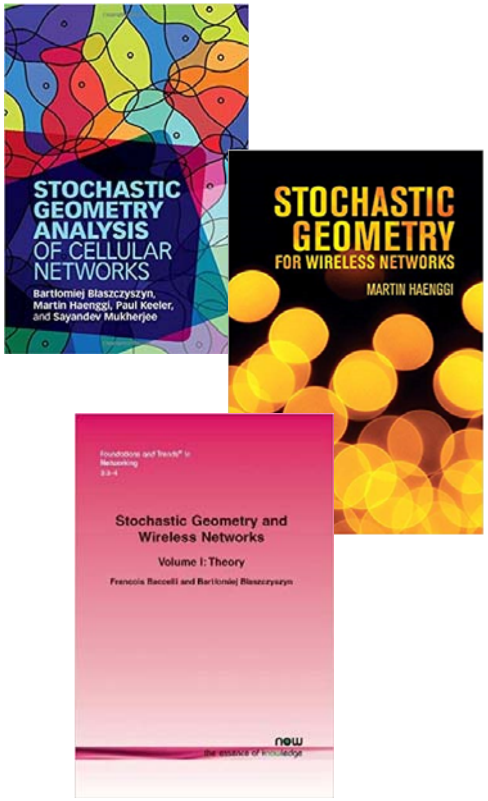


[Ref] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122-3134, Nov. 2011.

Stochastic Geometry for Wireless Networks



From Martin Haenggi



Why Random Spatial Network Models?

- **Advantages**
  - Avoid building and running system-level simulations
  - Quickly evaluate different PHY/MAC techniques
  - Expose salient network properties

- **Limitations** of existing studies
  - Tractable results mainly for single-antenna networks
  - Highly complicated expressions for multi-antenna networks
  - **A sample result:**

Essential for 5G

$$p_c(\tau) = \lambda \sum_{n=0}^{M-1} \frac{1}{n!} \sum_{p=0}^{n} B_{n,p}(x_1, \cdots, x_{n-p+1}) \frac{p!(2\tau^{\delta})^p}{(\lambda\tau^{\delta}\mathcal{C} + \lambda)^{p+1}}$$

where

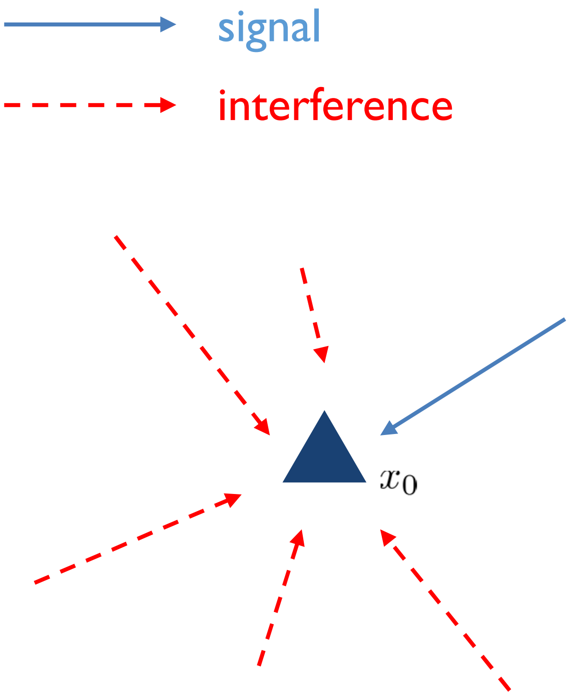
$$x_i = \frac{\delta\lambda\,(\kappa+i-1)!}{\text{[illegible]}(\kappa-1)!} B'\left(\kappa+\delta, i-\delta, \frac{1}{1+\tau}\right)$$

$$\mathcal{C} = \delta \sum_{i=1}^{\kappa} \binom{\kappa}{i} B'\left(\kappa-i+\delta, i-\delta, \frac{1}{1+\tau}\right)$$

Need a more tractable approach for characterizing multi-antenna 5G networks!

$B_{n,p}(x_1, \cdots, x_{n-p+1})$ is the incomplete exponential Bell polynomials, and the $B'(a,b,c)$ is the complementary in complete Beta function.

Network Model



- Transmitters: a Poisson Point Process
- Desired signal: $S = P_t g_{x_0} r_0^{-\alpha}$
- Interference power:

$$I = \sum_{x \in \Phi'} P_t g_x ||x||^{-\alpha}$$

$$\mathrm{SINR} = \frac{g_{x_0} r_0^{-\alpha}}{\sigma_{\mathrm{n}}^2 + \sum_{x \in \Phi'} g_x ||x||^{-\alpha}}$$

Coverage Probability: $p_{\mathrm{c}}(\tau) = \mathbb{P}\left[\mathrm{SINR} > \tau\right]$

Our Result – A General Matrix Representation

Toeplitz matrix

$$p_c(\tau) = \mathbb{E}_{r_0}\left[\left\|e^{\mathbf{T}_M}\right\|_1\right] \qquad \mathbf{T}_M = \begin{bmatrix} t_0 & & & & \\ t_1 & t_0 & & & \\ t_2 & t_1 & t_0 & & \\ \vdots & & & \ddots & \\ t_{M-1} & \cdots & t_2 & t_1 & t_0 \end{bmatrix}$$

$t_k = \frac{(-s)^k}{k!}\eta^{(k)}(s)$ $\eta(s) = \log \mathcal{L}(s)$ (Log-Laplace transform) $\|\mathbf{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|$

- A general framework for multi-antenna networks
    - A *compact expression* for numerical evaluation
    - *Closed-form expressions* for many cases
- Further analysis via properties of l_1-Toeplitz matrices
    - Investigate effects of *densification/antenna size*

X. Yu, C. Li, **J. Zhang**, and K. B. Letaief, "A tractable framework for performance analysis of dense multi-antenna networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017.

Applications

- **Area spectral efficiency and energy efficiency analysis**
  - C. Li, **J. Zhang**, and K. B. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2502-2517, May 2014.
- **Millimeter-wave network analysis**
  - X. Yu**, J. Zhang**, M. Haenggi, and K. B. Letaief, "Coverage analysis for millimeter wave networks: The impact of directional antenna arrays," *IEEE J. Select. Areas Commun,* vol. 35, no. 7, pp. 1498-1512, Jul. 2017.
- **Optimize success probability vs link reliability tradeoff**
  - C. Li, **J. Zhang**, J. G. Andrews, and K. B. Letaief, "Success probability and area spectral efficiency in multiuser MIMO HetNets," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1544-1556, Apr. 2016.
- **Develop interference management techniques**
  - C. Li, **J. Zhang**, M. Haenggi, and K. B. Letaief, "User-centric intercell interference nulling for downlink small cell networks," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1419-1431, Apr. 2015.

Case Study: Energy Efficiency

- Network energy efficiency $\eta_{\mathrm{EE}} = \frac{R_{\mathrm{a}}}{P_{\mathrm{a}}}$

Area spectral efficiency (ASE)

- Total power consumption of a unit area

Non-transmission power consumption

$$P_{\mathrm{a}} = \lambda_b p_{\mathrm{a}} \left(\frac{1}{\eta} P_{\mathrm{t}} + M P_{\mathrm{c}} \right) + \lambda_b P_0$$

Transmit power

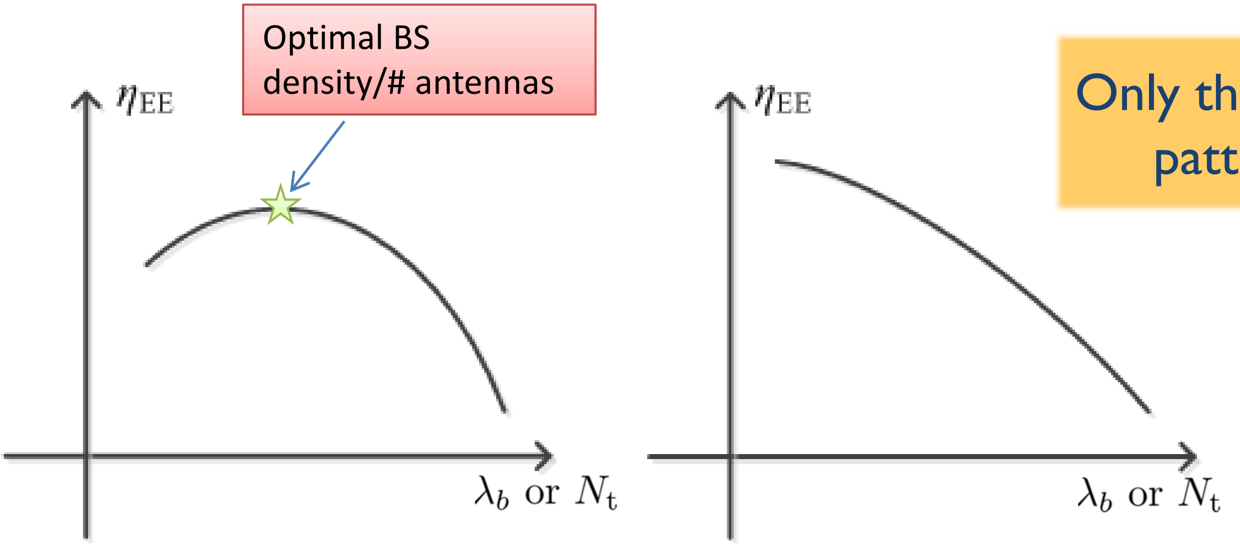
Circuit power

- ASE increases when deploying more BSs
- Power consumption will also increase

➢ How will the energy efficiency change with network densification?

➢ How will the energy efficiency change with # of BS antennas?

Case Study: Energy Efficiency



Optimal BS density/# antennas

Only these two patterns

Effect of densification ← Non-transmission power consumption

Effect of more antennas ← Circuit power consumption

Conclusions

- A unified analytical framework for tractable analysis of large-scale multi-antenna networks

- Abundant applications in wireless networks
  - Coverage/outage, ASE, EE …
  - Small cells, HetNets, mm-wave …
  - ASE vs link reliability tradeoff
  - Interference coordination
  - PHY-layer security

Future Direction: New Scenarios

- UAVs: 3D coverage, different propagation properties


Flying cameras


Delivery

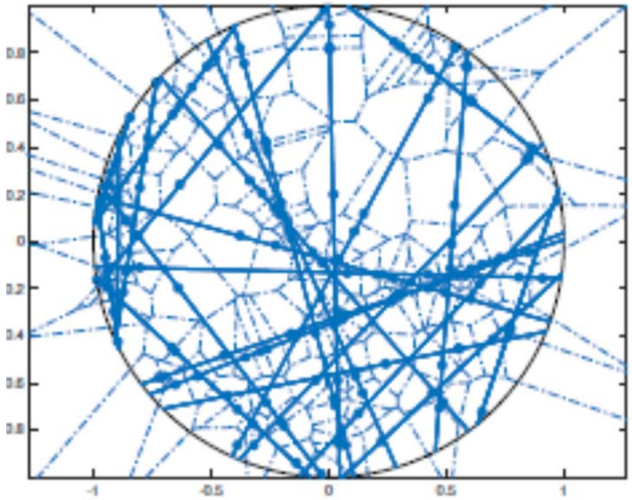

Public safety


Agricultural

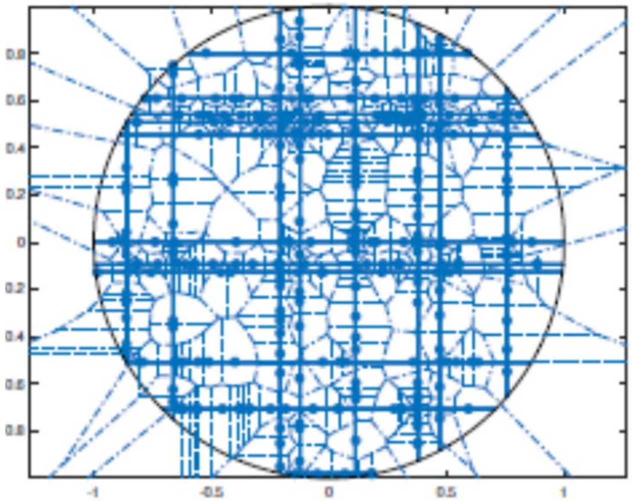

Inspection

- V2X: mobility, spatial modeling of roads

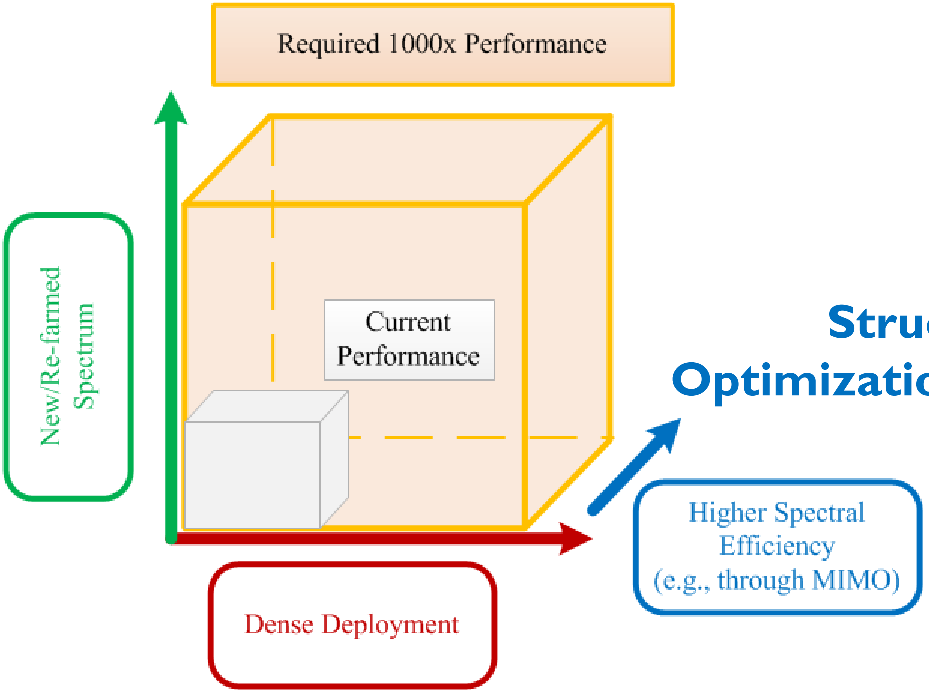
Isotropic Poisson line process



Manhattan Poisson line process



Takeaways



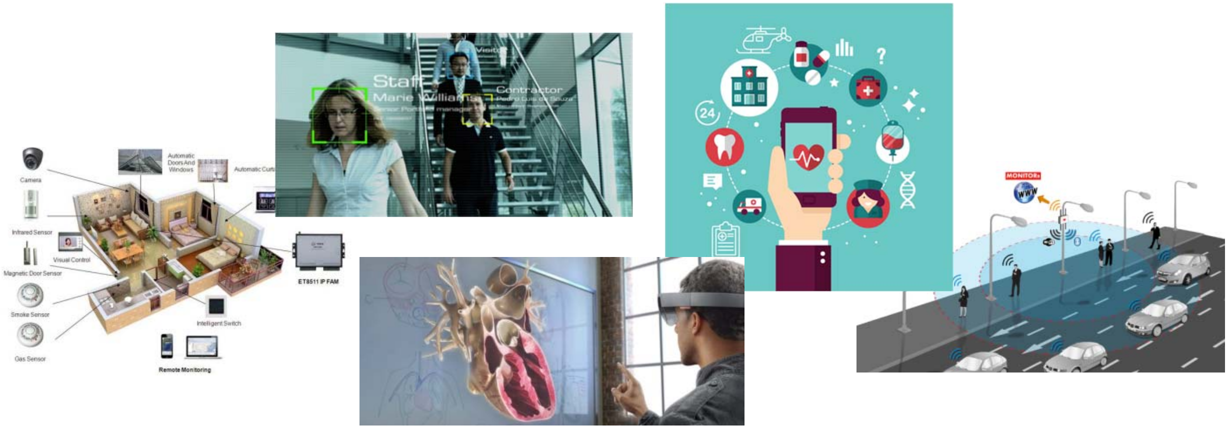
Hardware-algorithm Co-design

Structured Optimization\Estimation

Network Analysis via Stochastic Geometry

Additional Remarks

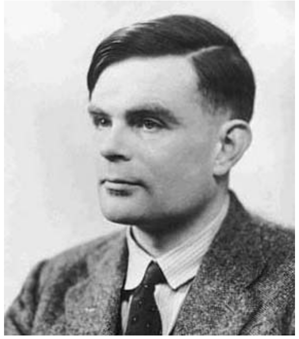
The Era of Mobile Intelligence



Communications

Computing

C. E. Shannon
(1916—2001)



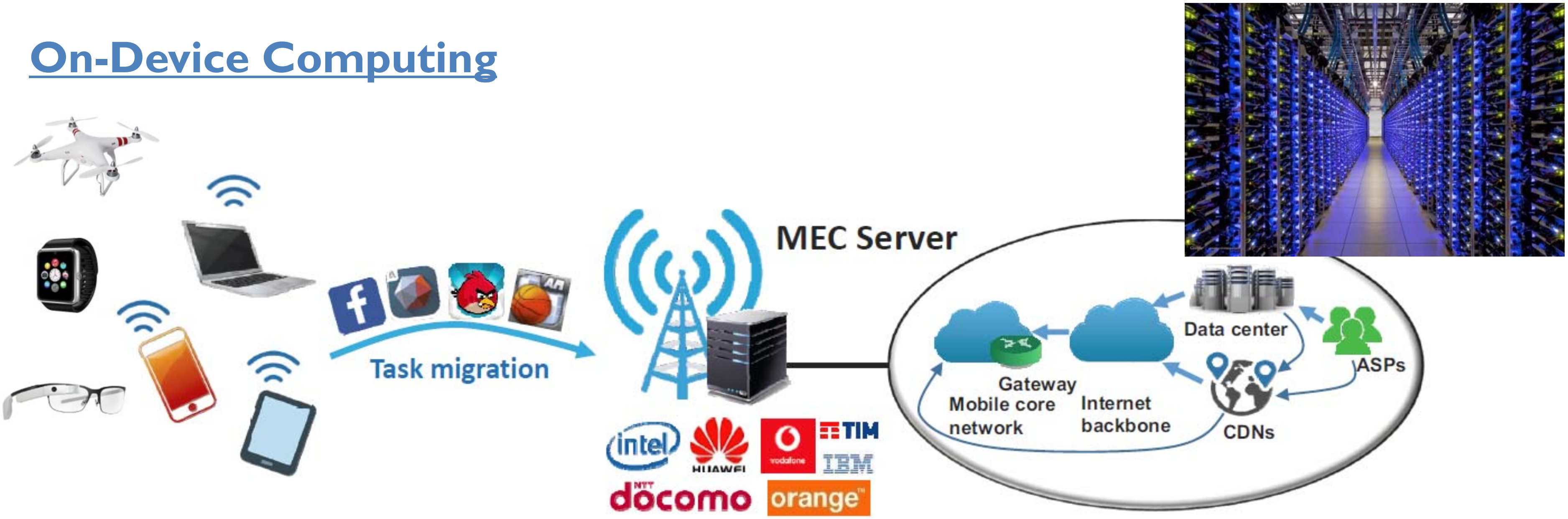
A. Turing
(1912—1954)

To Meet the Computation Demand
– A 3-Layer Approach

Cloud Computing

Mobile Edge Computing

On-Device Computing



Distributed Computing Systems

- **Mobile Edge Computing**
  - Y. Mao, C. You, **J. Zhang**, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, 2017.
  - Y. Mao, **J. Zhang**, S.H. Song, and K. B. Letaief, "Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5994-6009, Sept. 2017.
  - Y. Mao, **J. Zhang**, and K. B. Letaief, "Dynamic computation offloading for mobile-edge computing with energy harvesting devices," *IEEE J. Select. Areas Commun. - Series on Green Commun. and Networking*, vol. 34, no. 12, pp. 3590-3605, Dec. 2016.
- **Cloud Computing**
  - Y. Yu, R. Huang, W. Wang, **J. Zhang**, and K. B. Letaief, "SP-Cache: Load-balanced, Redundancy-free Cluster Caching with Selective Partition," *SC18*, Dallas, TX, November, 2018. (**Acceptance Rate: 19%**)
  - Y. Yu, W. Wang, **J. Zhang**, Q. Weng, and K. B. Letaief, "OpuS: Fair and efficient cache sharing for in-memory data analytics," *ICDCS* 2018. (**Acceptance Rate: 20%**)
  - Y. Yu, W. Wang, **J. Zhang**, and K. B. Letaief, "LRC: Dependency-aware cache management in data analytics clusters," *IEEE INFOCOM* 2017. (**Acceptance Rate: 21%**)

Wireless Edge Caching

- R. Wang, **J. Zhang**, S.H. Song, and K. B. Letaief, "Exploiting mobility in cache-assisted D2D networks: Performance analysis and optimization," *IEEE Trans. Wireless Commun.*, to appear.
- X. Peng, Y. Shi, **J. Zhang**, and K. B. Letaief, "Layered group sparse beamforming for cache-enabled green wireless networks," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5589-5603, Dec. 2017.
- R. Wang, **J. Zhang**, S.H. Song, and K. B. Letaief, "Mobility-aware caching in D2D networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5001-5015, Aug. 2017.
- R. Wang, X. Peng, **J. Zhang**, and K. B. Letaief, "Mobility-aware caching for content-centric wireless networks: Modeling and methodology," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 77-83, Aug. 2016.
- J. Liu, B. Bai, **J. Zhang**, and K. B. Letaief, "Content caching at the wireless network edge: A distributed algorithm via brief propagation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016. (**Best Paper Award**)
- X. Peng, J.-C. Shen, **J. Zhang**, and K. B. Letaief, "Joint data assignment and beamforming for backhaul limited caching networks," in *Proc. IEEE Int. Symp. on Personal Indoor and Mobile Radio Comm. (PIMRC)*, Washington, DC, Sept. 2014. (**Best Paper Award**)