

An Algorithmic Investigation of Hybrid Beamforming for 5G and Beyond Networks

Xianghao Yu



Jun Zhang



Outline

❖ Background and Motivation

❖ Preliminaries of Hybrid Beamforming

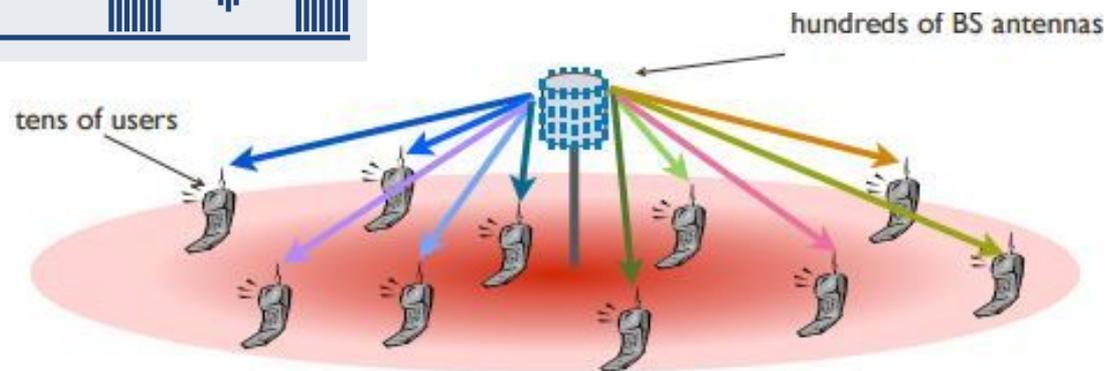
❖ Hybrid Beamforming Design

- **Improve Spectral Efficiency:** Approaching the Fully Digital
- **Boost Computational Efficiency:** Convex Relaxation
- **Fight for Hardware Efficiency:** How Many Phase Shifters Are Needed?

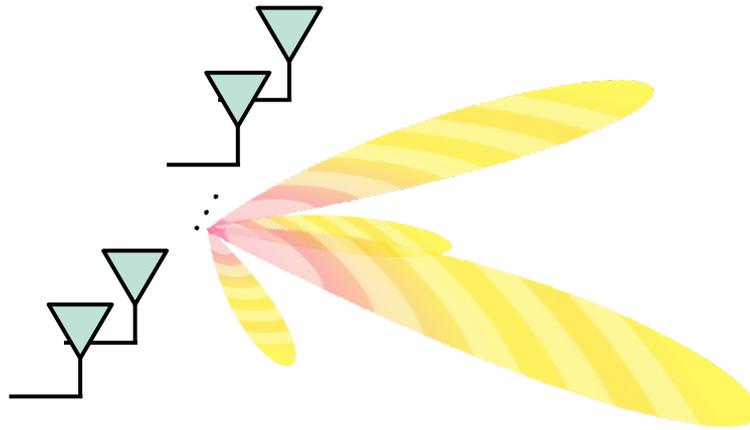
❖ Conclusions

Background and Motivation

❖ Key enabler for 5G and beyond: Massive MIMO



Background and Motivation



$$N_t > 100$$

Beamforming!

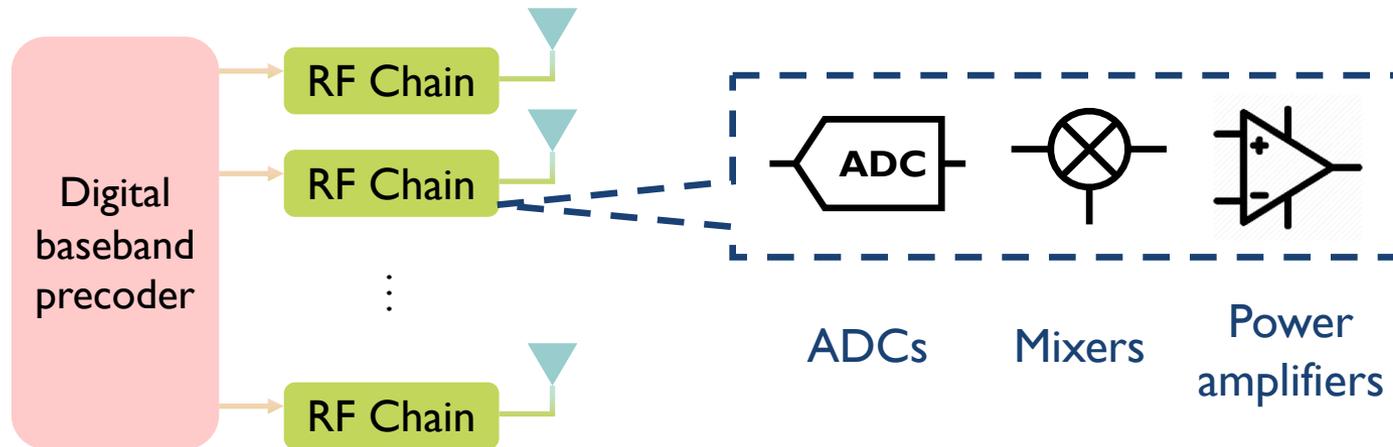
Higher array gains and narrower beams

- Higher spectral efficiency
- Higher energy efficiency
- Better interference management

Background and Motivation

❖ Conventional beamforming

- Performed **digitally** at the **baseband**
- Requires **an RF chain per antenna element**



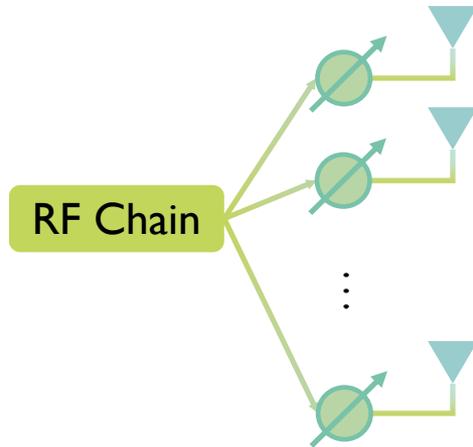
Costly and power hungry for large-scale antenna arrays



Background and Motivation

❖ Existing solution: **Analog** beamforming

➤ **One** RF chain only

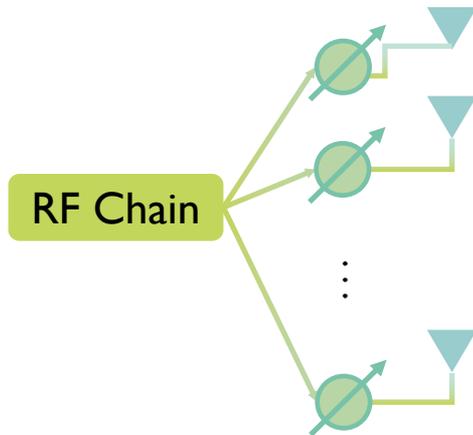


- Beams direction readily controlled by a series of **phase shifters** in the **RF domain**
- Low cost and low hardware complexity

Background and Motivation

❖ Existing solution: **Analog** beamforming

➤ Limitations



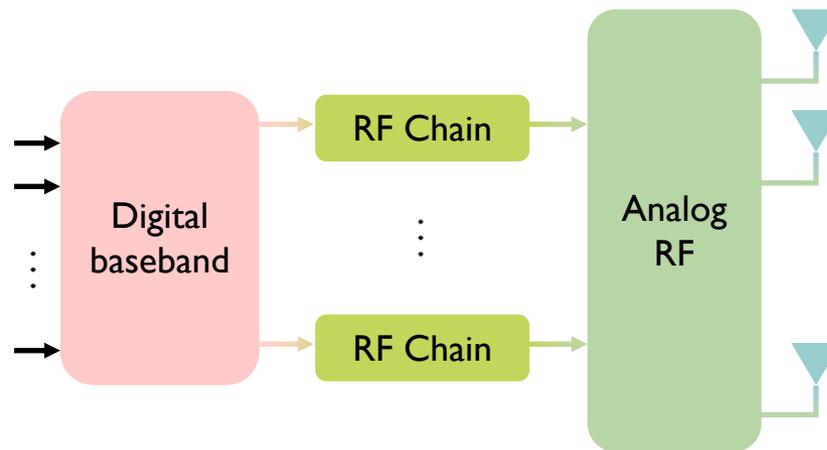
Benefits of MIMO

- Spatial multiplexing
- Support space-division multiple access (SDMA)

Analog beamforming can only support single-stream transmissions

Background and Motivation

❖ A new solution: Hybrid beamforming



- Multi-stream transmission, ability to support SDMA
- Number of RF chains **much smaller than # antennas**
- Combine the benefits of **digital and analog** beamforming

Background and Motivation

❖ Attentions on hybrid beamforming

- O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014.
 - **The 2017 Marconi Prize Paper Award in Wireless Communications**
- F. Sotiridis and W. Yu, “Hybrid digital and analog beamforming design for large-scale antenna arrays,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501-513, Apr. 2016.
 - **The 2017 IEEE Signal Processing Society Best Paper Award**
- A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., “Channel estimation and hybrid precoding for millimeter wave cellular systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831-846, Oct. 2014.
 - **The 2016 Signal Processing Society Young Author Best Paper Award**
- 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.
 - **The 2018 Signal Processing Society Young Author Best Paper Award**

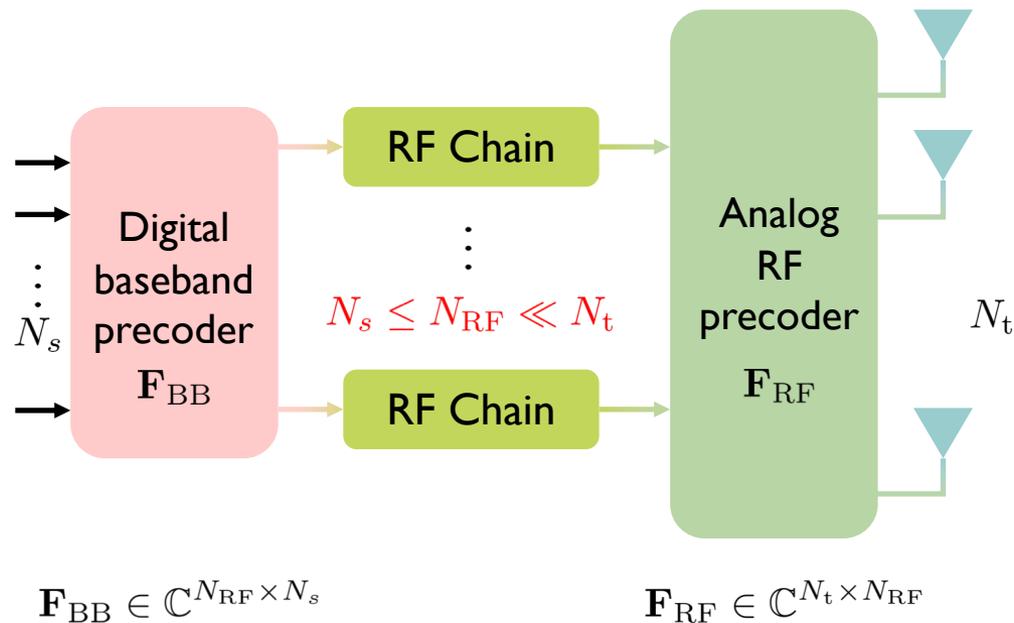
Preliminaries of Hybrid Beamforming

Preliminaries of Hybrid Beamforming

❖ Hybrid beamforming

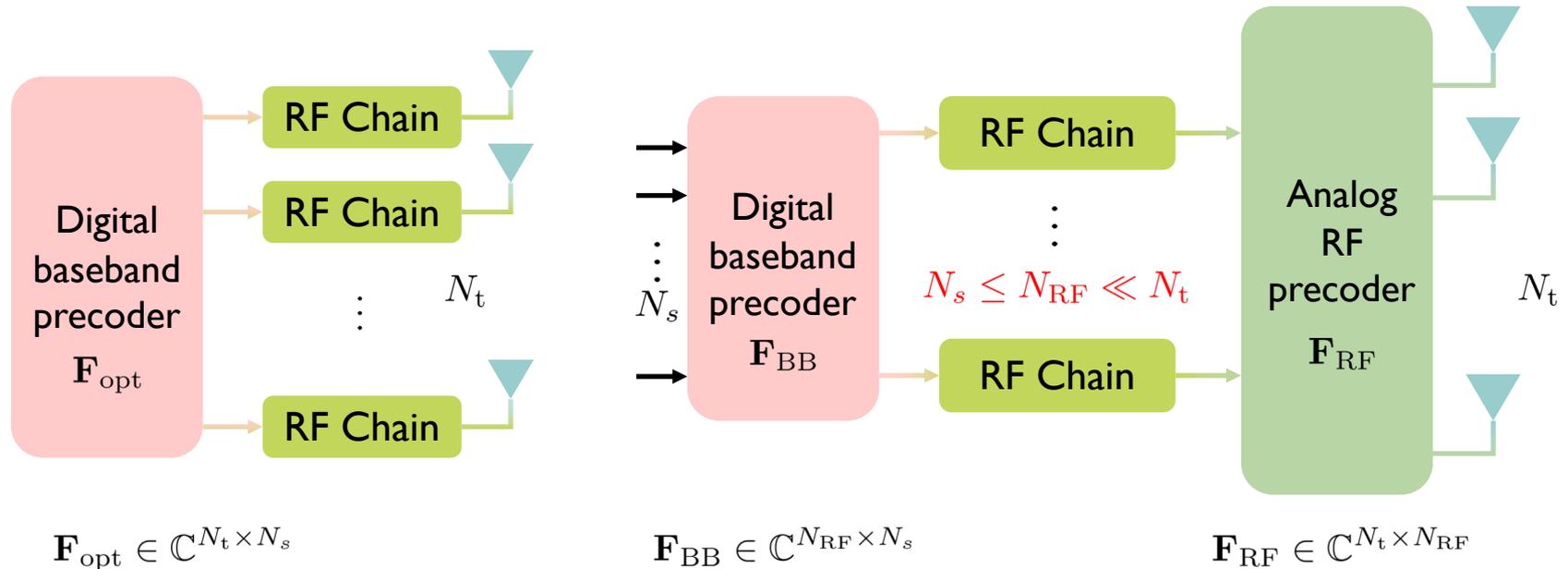
➤ Also called *Hybrid precoding*; *Analog/digital precoding*

➤ **Notations** in hybrid beamforming



Preliminaries of Hybrid Beamforming

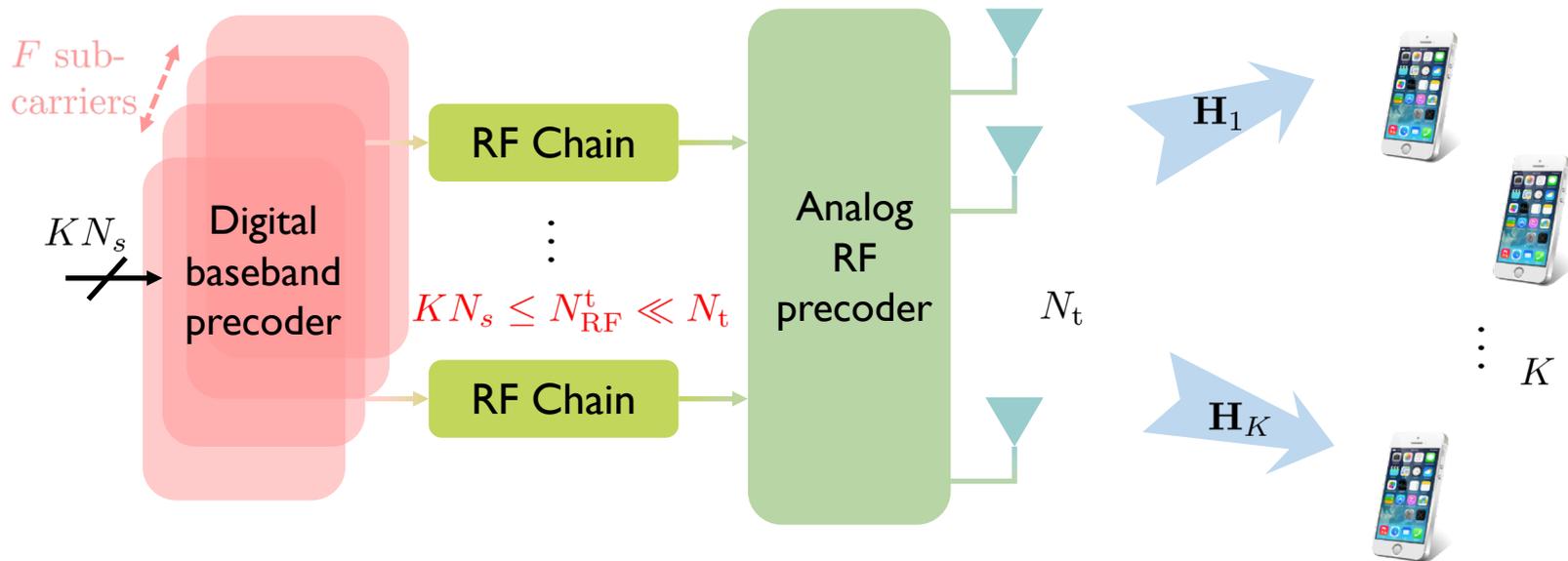
❖ Fully digital precoding vs. Hybrid precoding



- Main differentiating part: **Analog RF precoder**
- Mapping from low-dimensional RF chains to high-dimensional antennas, typically implemented by **phase shifters**

Preliminaries of Hybrid Beamforming

❖ General multiuser multicarrier (MU-MC) systems



- One separate digital precoder for each user on each subcarrier $\mathbf{F}_{BBk,f}$
- Analog precoder \mathbf{F}_{RF} is **shared by all the users and subcarriers**

Preliminaries of Hybrid Beamforming

❖ Generic hybrid beamforming problem

- Minimize the Euclidean distance between the hybrid precoders and the fully digital precoder [O. El Ayach *et al.*, 2014]

$$\begin{aligned} & \underset{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{minimize}} && \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \\ & \text{subject to} && \|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \leq P_{\text{max}} \\ & && \mathbf{F}_{\text{RF}} \in \mathcal{A}_x \end{aligned}$$

- This formulation applies with an arbitrary digital precoder.
- It is applicable to different hybrid beamforming structures.
- It facilitates beamforming algorithm design.
- The obtained algorithmic approaches also help other formulations.

Preliminaries of Hybrid Beamforming

❖ Generic hybrid beamforming problem

$$\begin{aligned} & \underset{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{minimize}} && \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \\ & \text{subject to} && \|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \leq P_{\text{max}} \\ & && \mathbf{F}_{\text{RF}} \in \mathcal{A}_x \end{aligned}$$

Main difficulties

- **Unit modulus constraints** for phases $|(\mathbf{F}_{\text{RF}})_{i,j}| = 1$
- **Structure constraints** for \mathcal{A}_x (different hybrid architectures)

Preliminaries of Hybrid Beamforming

❖ Unit modulus constraints

Common approaches

- Codebook based, e.g., OMP [O. El Ayach *et al.*, 2014]

The columns of the analog precoding matrix \mathbf{F}_{RF} selected from array response vectors

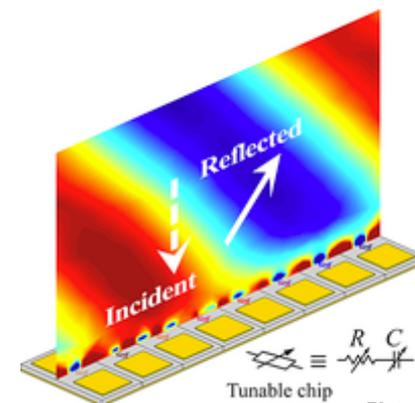
- Manifold optimization – directly tackle unit modulus constraints

[Yu *et al.*, 2016]

- Convex relaxation [Yu *et al.*, 2019]

Other applications

- Intelligent reflecting surfaces (IRSs)



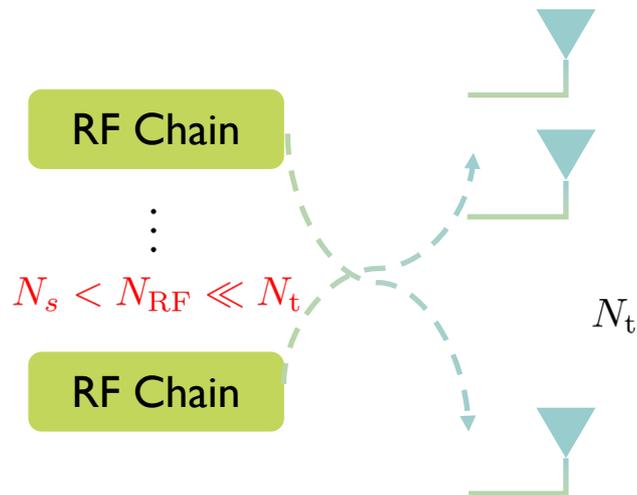
[Liu *et al.*, 2019]

Preliminaries of Hybrid Beamforming

❖ A taxonomy of hybrid beamforming structures

(I) Mapping strategy:

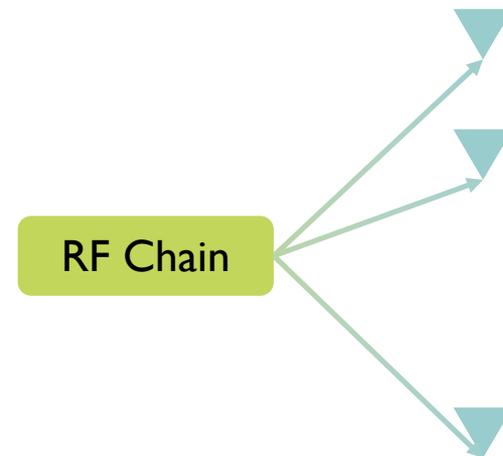
Which antennas should be connected to each RF chain?



Signal flow

(II) Hardware implementation:

What kind of hardware should be used to realize each connection?

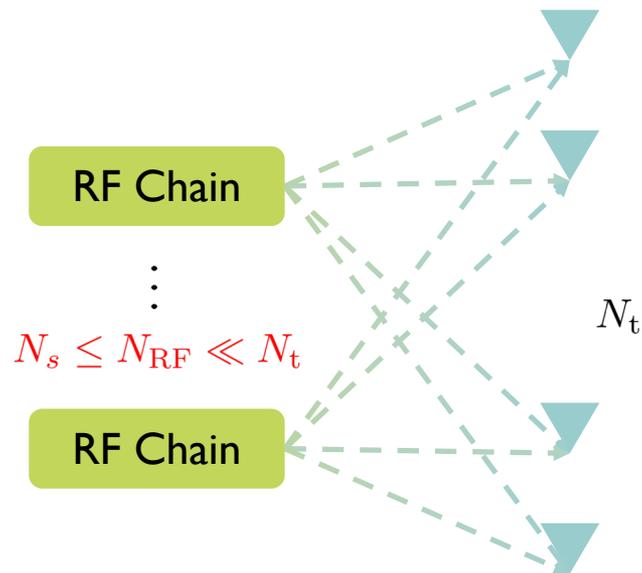


Adopted hardware

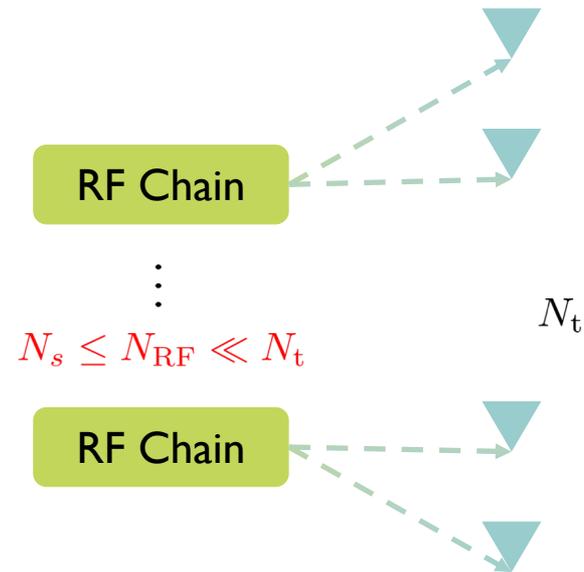
Preliminaries of Hybrid Beamforming

❖ The state-of-the-art hybrid beamforming structures

➤ Mainly focus on different mapping strategies



Fully-connected

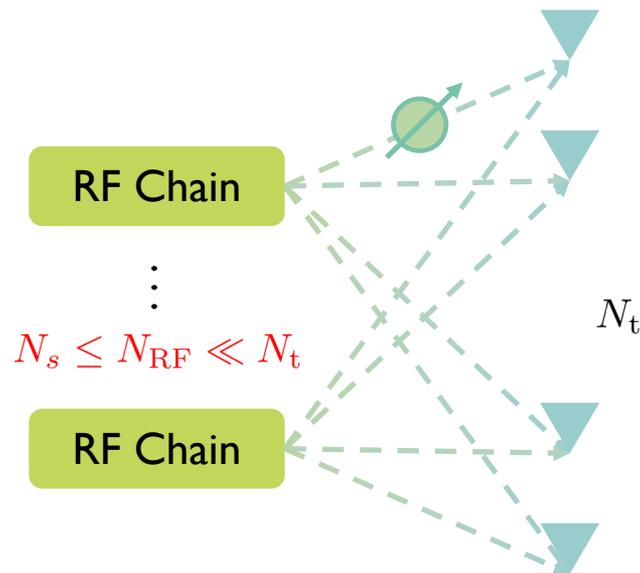


Partially-connected

Preliminaries of Hybrid Beamforming

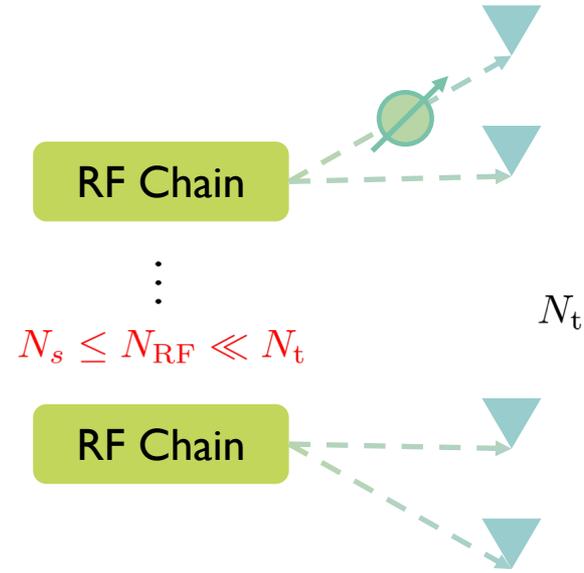
❖ The state-of-the-art hybrid beamforming structures

- One prevalent hardware implementation: **Single phase shifter (SPS)**



SPS Fully-connected

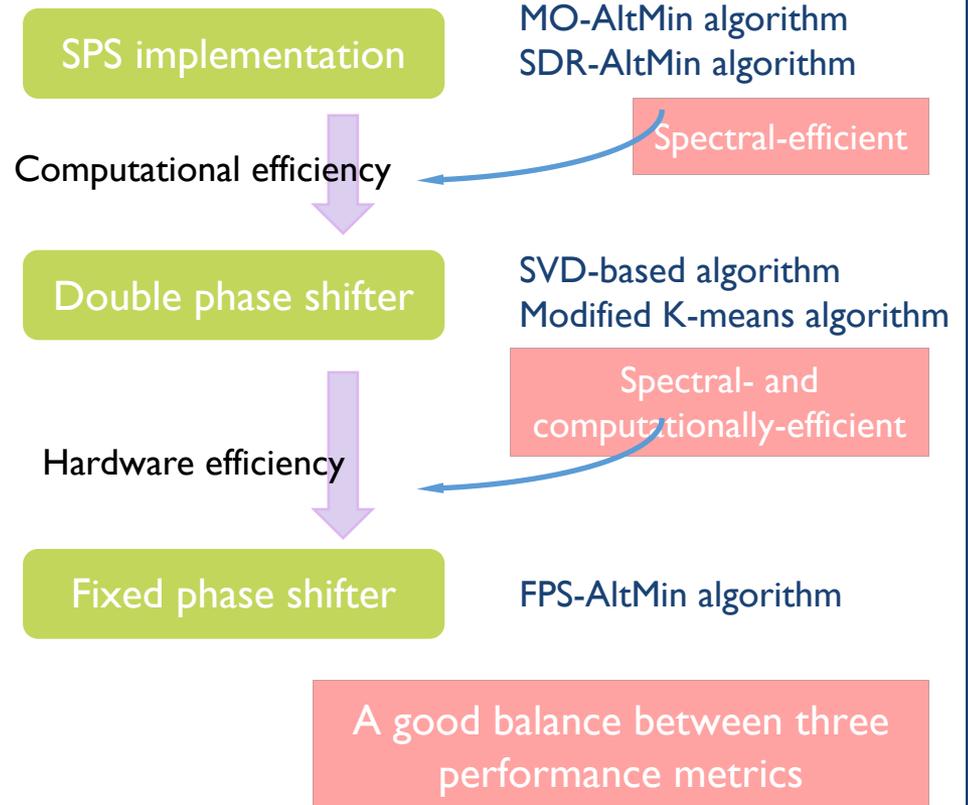
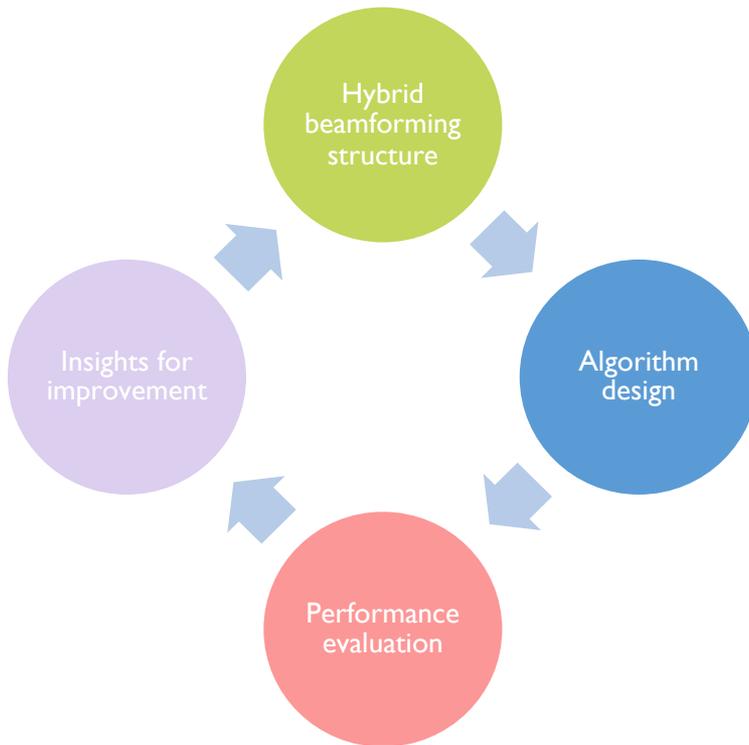
$$N_{PS} = N_t N_{RF}$$



SPS Partially-connected

$$N_{PS} = N_t$$

Preliminaries of Hybrid Beamforming



Effective algorithms are required to reveal system insights

Preliminaries of Hybrid Beamforming

❖ Three key aspects to investigate

Spectral efficiency

- **Q1:** Can hybrid beamforming provide performance close to the fully digital one?

Hardware efficiency

- **Q2:** How many RF chains are needed?
- **Q3:** How many phase shifters are needed?

Computational efficiency

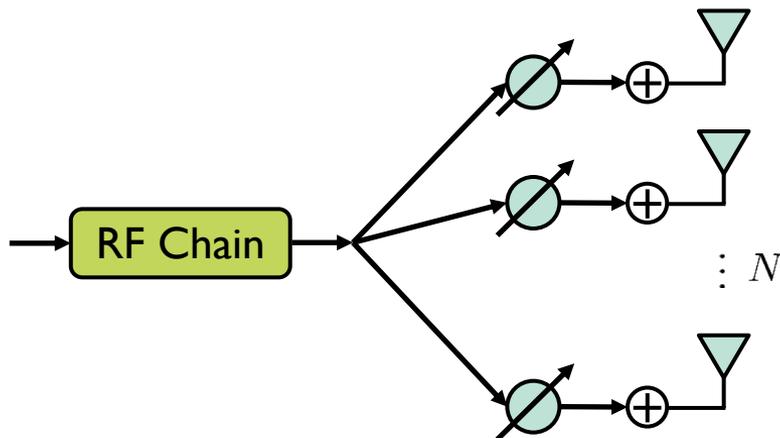
- **Q4:** How to efficiently design hybrid beamforming algorithms?

Improve Spectral Efficiency: Approaching the Fully Digital Beamforming

[Ref] 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., Special Issue on Signal Process. for Millimeter Wave Wireless Commun.*, vol. 10, no. 3, pp. 485-500, Apr. 2016. (**The 2018 IEEE Signal Processing Society Young Author Best Paper Award**)

Improve Spectral Efficiency

❖ Single phase shifter (SPS) implementation

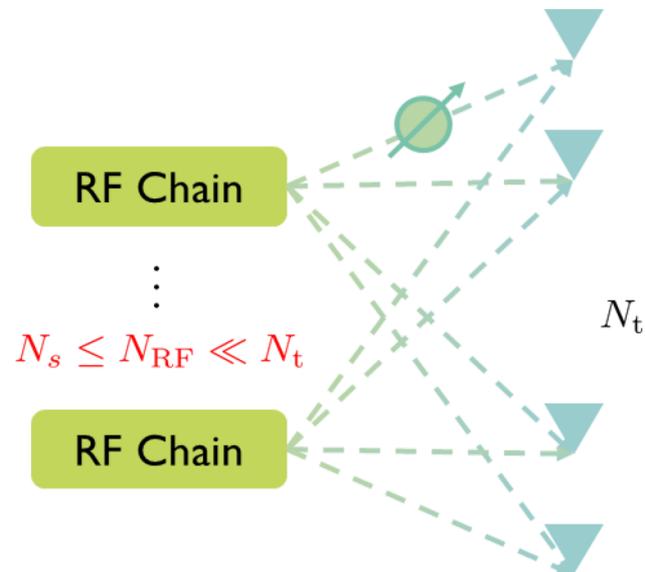


$$N = \begin{cases} N_t & \text{fully-connected} \\ N_t/N_{\text{RF}}^t & \text{partially-connected} \end{cases}$$

➤ Fully digital achieving condition: $N_{\text{RF}}^t \geq 2KN_s$, $N_{\text{RF}}^r \geq 2N_s$

Q: Can we further reduce the number of RF chains?

(I) Fully-Connected Mapping



Improve Spectral Efficiency

(I) Fully-Connected Mapping

❖ Start from single-user systems

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

➤ Alternating minimization

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

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

➤ Digital precoder: $\mathbf{F}_{\text{BB}} = \mathbf{F}_{\text{RF}}^\dagger \mathbf{F}_{\text{opt}}$

➤ Difficulty: Analog precoder design with the **unit modulus constraints**

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

➤ The vector $\mathbf{x} = \text{vec}(\mathbf{F}_{\text{RF}})$ forms a complex circle manifold

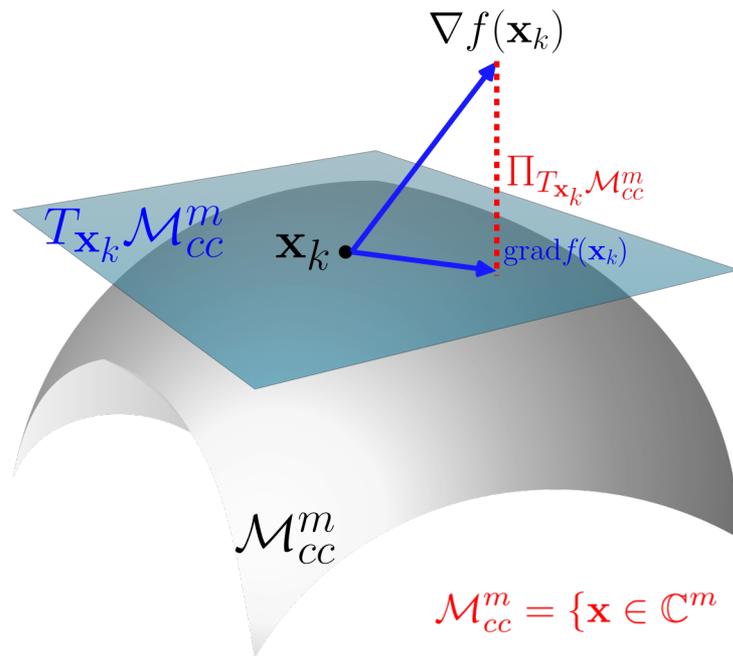
$$\mathcal{M}_{cc}^m = \{\mathbf{x} \in \mathbb{C}^m : |\mathbf{x}_1| = |\mathbf{x}_2| = \dots = |\mathbf{x}_m| = 1\}, \quad m = N_t N_{\text{RF}}^t.$$

Improve Spectral Efficiency

(I) Fully-Connected Mapping

❖ Manifold optimization (cont.)

- Euclidean space: **gradient descent**
- Similar approaches on manifolds?

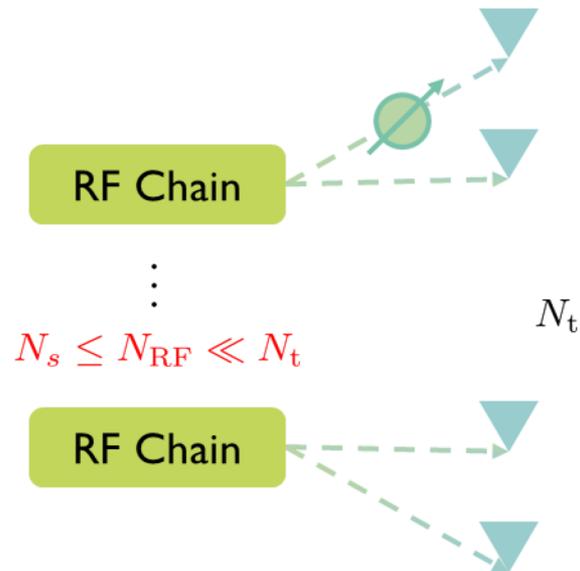


- Tangent space
- Riemannian gradient
- Retraction

Gradient-based algorithm on manifolds

$$\mathcal{M}_{cc}^m = \{\mathbf{x} \in \mathbb{C}^m : |\mathbf{x}_1| = |\mathbf{x}_2| = \dots = |\mathbf{x}_m| = 1\}, \quad m = N_t N_{RF}^t.$$

(II) Partially-Connected Mapping



Improve Spectral Efficiency

(II) Partially-Connected Mapping

❖ SPS partially-connected

- \mathcal{A}_x : Block diagonal \mathbf{F}_{RF} with unit modulus non-zero elements

$$\mathbf{F}_{\text{RF}} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{p}_2 & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{p}_{N_{\text{RF}}^t} \end{bmatrix}$$

$$\mathbf{p}_i = \left[\exp \left(j\theta_{(i-1)\frac{N_t}{N_{\text{RF}}^t} + 1} \right), \cdots, \exp \left(j\theta_{i\frac{N_t}{N_{\text{RF}}^t}} \right) \right]^T$$

phase shifters connected to the i -th RF chain

- Problem decoupled for each RF chain

- Closed-form solution for \mathbf{F}_{RF}

$$\arg \{(\mathbf{F}_{\text{RF}})_{i,l}\} = \arg \left\{ (\mathbf{F}_{\text{opt}})_{i,:} (\mathbf{F}_{\text{BB}})_{l,:}^H \right\}, \quad 1 \leq i \leq N_t, \quad l = \left\lceil i \frac{N_{\text{RF}}^t}{N_t} \right\rceil$$

Improve Spectral Efficiency

(II) Partially-Connected Mapping

❖ SPS partially-connected (cont.)

➤ Optimization of \mathbf{F}_{BB}

$$\begin{aligned} & \underset{\mathbf{F}_{\text{BB}}}{\text{minimize}} && \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \\ & \text{subject to} && \|\mathbf{F}_{\text{BB}}\|_F^2 = \frac{N_{\text{RF}}^t N_s}{N_t}. \end{aligned}$$

➤ Reformulate as a non-convex problem

$$\begin{aligned} & \underset{\mathbf{Y} \in \mathbb{H}^n}{\text{minimize}} && \text{Tr}(\mathbf{C}\mathbf{Y}) \\ & \text{subject to} && \begin{cases} \text{Tr}(\mathbf{A}_1\mathbf{Y}) = \frac{N_{\text{RF}}^t N_s}{N_t} \\ \text{Tr}(\mathbf{A}_2\mathbf{Y}) = 1 \\ \mathbf{Y} \succeq 0, \text{rank}(\mathbf{Y}) = 1 \end{cases} \end{aligned}$$

convex

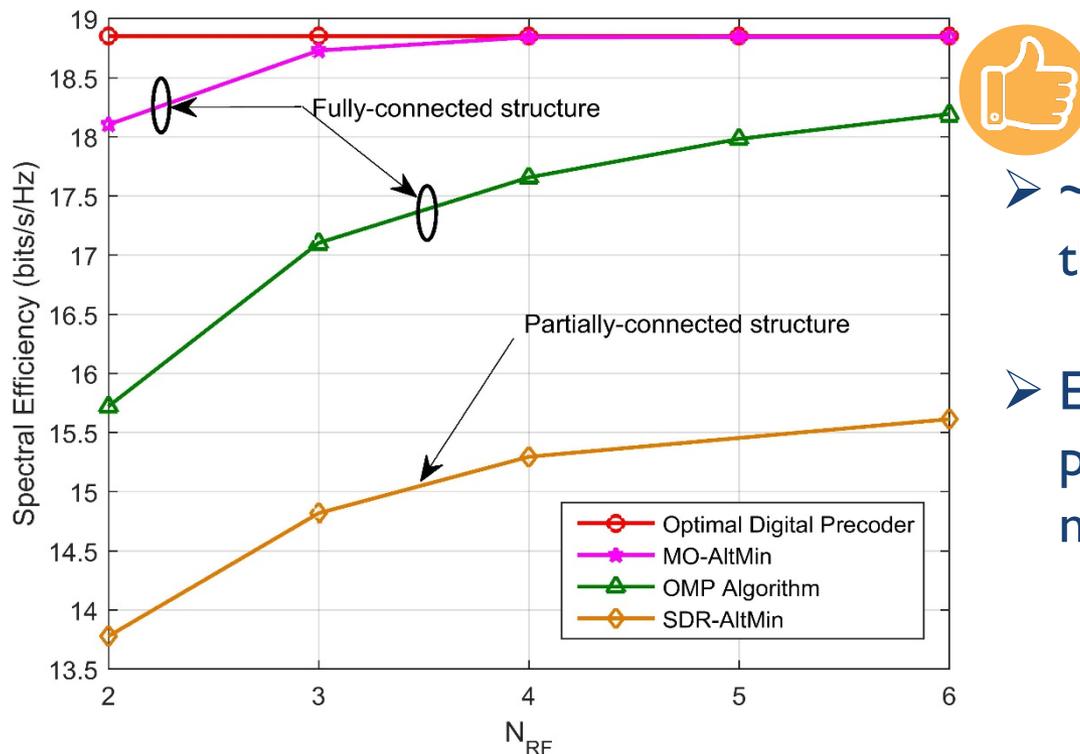
$$\begin{aligned} n &= N_{\text{RF}}^t N_s + 1, \mathbf{y} = [\text{vec}(\mathbf{F}_{\text{BB}}) \quad t]^T, \\ \mathbf{Y} &= \mathbf{y}\mathbf{y}^H, \mathbf{f} = \text{vec}(\mathbf{F}_{\text{opt}}), \\ \mathbf{A}_1 &= \begin{bmatrix} \mathbf{I}_{n-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} \mathbf{0}_{n-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}, \\ \mathbf{C} &= \begin{bmatrix} (\mathbf{I}_{N_s} \otimes \mathbf{F}_{\text{RF}})^H (\mathbf{I}_{N_s} \otimes \mathbf{F}_{\text{RF}}) & -(\mathbf{I}_{N_s} \otimes \mathbf{F}_{\text{RF}})^H \mathbf{f} \\ -\mathbf{f}^H (\mathbf{I}_{N_s} \otimes \mathbf{F}_{\text{RF}}) & \mathbf{f}^H \mathbf{f} \end{bmatrix}. \end{aligned}$$

➤ Semidefinite relaxation (SDR) is tight for this case so globally optimal solution is obtained [Z.-Q. Luo et al., 2010]

Improve Spectral Efficiency

❖ Simulation results

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



- $\sim N_s$ RF chains are sufficient for the fully-connected mapping
- Employing fewer PSs, the partially-connected mapping needs more RF chains

Limitation: Computational efficiency of the MO-AltMin is not good, thus difficult to extend to MU-MC settings

Boost Computational Efficiency: Convex Relaxation

[Ref] X. Yu, J. Zhang, and K. B. Letaief, “Alternating minimization for hybrid precoding in multiuser OFDM mmWave Systems,” in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2016. **(Invited Paper)**

[Ref] X. Yu, J. Zhang, and K. B. Letaief, “Doubling phase shifters for efficient hybrid precoding in millimeter-wave multiuser OFDM systems,” *J. Commun. Inf. Netw.*, vol. 4, no. 2, pp. 51-67, Jul. 2019.

Boost Computational Efficiency

- ❖ Main difficulty in designing the SPS implementation
 - Analog precoder with the **unit modulus constraints**

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

- An intuitive way to boost computational efficiency is to relax this highly non-convex constraint as a convex one

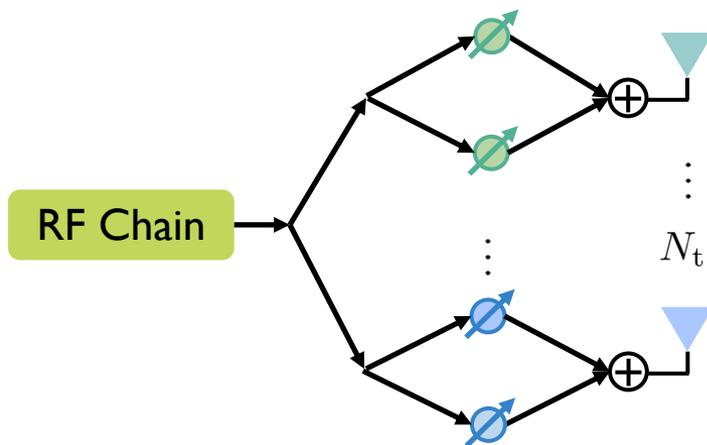
$$\begin{aligned} & \underset{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{minimize}} && \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \\ & \text{subject to} && |(\mathbf{F}_{\text{RF}})_{i,j}| \leq \gamma, \forall i, j. \end{aligned}$$

- The value of γ does not affect the hybrid beamformer design
- We shall choose $\gamma=2$ instead of keeping it as 1. **Why?**

Boost Computational Efficiency

❖ Double phase shifter (DPS) implementation

- The relaxed solution with $\gamma=2$ can be realized by a hardware implementation



- Unit modulus constraint is eliminated

- Sum of two phase shifters

$$|e^{j\theta_1} + e^{j\theta_2}| \leq 2$$

Boost Computational Efficiency

(I) Fully-Connected Mapping

❖ Fully-connected mapping

➤ RF-only precoding

$$\begin{aligned} & \underset{\mathbf{F}_{\text{RF}}}{\text{minimize}} && \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \\ & \text{subject to} && |(\mathbf{F}_{\text{RF}})_{i,j}| \leq 2 \end{aligned} \quad \longleftrightarrow \quad \begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} && \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + 2\|\mathbf{x}\|_1 \\ & && \text{LASSO} \end{aligned}$$

➤ Closed-form solution for semi-unitary codebooks $\mathbf{F}_{\text{BB}}\mathbf{F}_{\text{BB}}^H = \mathbf{I}_{N_{\text{RF}}^t}$

$$\mathbf{F}_{\text{RF}}^* = \mathbf{F}_{\text{opt}}\mathbf{F}_{\text{BB}}^H - \exp\{j\angle(\mathbf{F}_{\text{opt}}\mathbf{F}_{\text{BB}}^H)\} \circ (|\mathbf{F}_{\text{opt}}\mathbf{F}_{\text{BB}}^H| - 2)^+.$$

➤ Hybrid precoding

$$\begin{aligned} & \underset{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}}{\text{minimize}} && \|\mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 \\ & \text{subject to} && |(\mathbf{F}_{\text{RF}})_{i,j}| \leq 2 \end{aligned} \quad \longrightarrow \quad \text{Matrix factorization}$$

Redundant

Boost Computational Efficiency

(I) Fully-Connected Mapping

❖ Fully-connected mapping (cont.)

➤ Optimality in **single-carrier** systems

$$\mathbf{F}_{\text{opt}} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}} \text{ with } \underline{N_{\text{RF}}^t = KN_s} \text{ and } \underline{N_{\text{RF}}^r = N_s} \text{ when } F = 1$$


Minimum number of RF chains

➤ It reduces the required number of RF chains **by half** for achieving the fully digital precoding

➤ **Multi-carrier** systems

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

➤ Low-rank matrix approximation: SVD, **globally optimal solution**

Boost Computational Efficiency

(I) Fully-Connected Mapping

❖ Fully-connected mapping (cont.)

➤ **Q: How to use this relaxed result for SPS implementation?**

➤ Optimal solution:

$$\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^H$$

➤ Some clues: The unitary matrix \mathbf{U}_1 fully extracts the information of the column space of $\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$, whose basis are the orthonormal columns in \mathbf{F}_{RF}

➤ Phase extraction

$$\mathbf{F}_{\text{RF}} = \exp\{j\angle(\mathbf{U}_1)\}, \quad \mathbf{F}_{\text{BB}} = \mathbf{S}_1\mathbf{V}_1^H$$

unit modulus constraint

**Convex relaxation-enabled
(CR-enabled) SPS**

Boost Computational Efficiency

(II) Partially-Connected Mapping

❖ Partially-connected mapping

➤ Block diagonal structure

$$\mathbf{F}_{\text{RF}} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{p}_2 & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{p}_{N_{\text{RF}}^t} \end{bmatrix} \quad \mathbf{p}_j = \left[a_{(j-1)\frac{N_t}{N_{\text{RF}}^t} + 1}, \dots, a_{j\frac{N_t}{N_{\text{RF}}^t}} \right]^T$$

➤ Decoupled for each RF chain

$$\mathcal{P}_j : \underset{\{a_i\}, \mathbf{x}_j}{\text{minimize}} \sum_{i \in \mathcal{F}_j} \|\mathbf{y}_i - a_i \mathbf{x}_j\|_2^2,$$

$$\mathcal{F}_j = \left\{ i \in \mathbb{Z} \mid (j-1)\frac{N_t}{N_{\text{RF}}^t} + 1 \leq i \leq j\frac{N_t}{N_{\text{RF}}^t} \right\}, \mathbf{y}_i = \mathbf{F}_{\text{opt}}^T(i, :), \text{ and } \mathbf{x}_j = \mathbf{F}_{\text{BB}}^T(j, :)$$

➤ Eigenvalue problem

$$\mathbf{x}_j^* = \lambda_1 \left(\sum_{i \in \mathcal{F}_j} \mathbf{y}_i \mathbf{y}_i^H \right), \quad a_i^* = \frac{\mathbf{x}_j^H \mathbf{y}_i}{\|\mathbf{x}_j\|_2^2}$$

Boost Computational Efficiency

(II) Partially-Connected Mapping

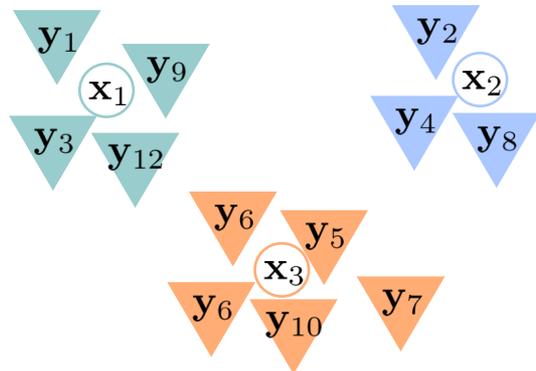
❖ DPS partially-connected mapping (cont.)

- Not much performance gain obtained by simply adopting the DPS implementation 

- **Dynamic mapping:**

Adaptively separate all N_t antennas into N_{RF} groups

$$\underset{\{\mathcal{D}_j\}_{j=1}^{N_{\text{RF}}^t}}{\text{maximize}} \quad \sum_{j=1}^{N_{\text{RF}}^t} \lambda_1 \left(\sum_{i \in \mathcal{D}_j} \mathbf{y}_i \mathbf{y}_i^H \right)$$



- **Modified K-means algorithm**

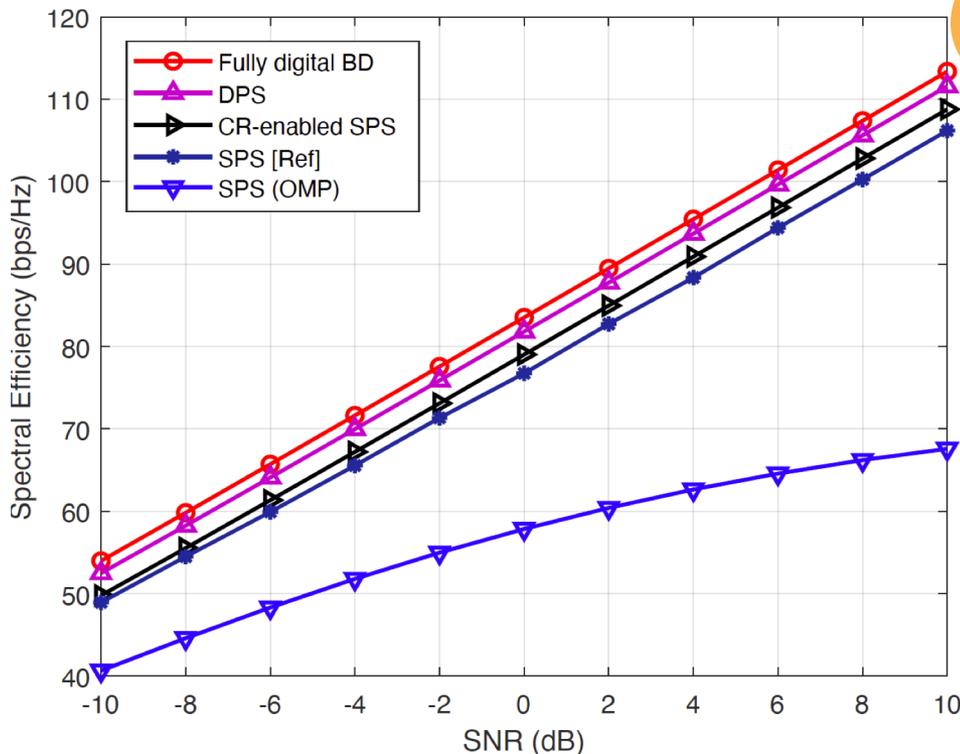
- **Centroid:** $\mathbf{x}_j^* = \lambda_1 \left(\sum_{i \in \mathcal{D}_j} \mathbf{y}_i \mathbf{y}_i^H \right)$

- **Clustering:** $j^* = \arg \max_j \left| \mathbf{y}_i^H \mathbf{x}_j \right|^2$

Boost Computational Efficiency

❖ Simulation results (Fully-connected)

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



➤ Achieve near-optimal spectral efficiency and optimal multiplexing gain with low-complexity algorithms

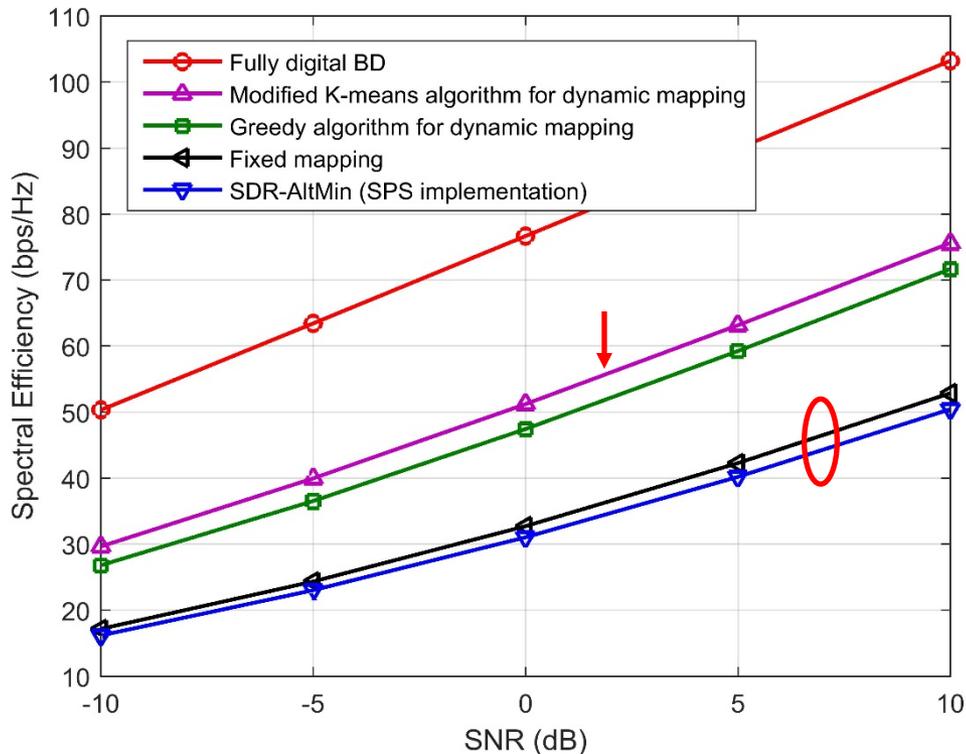
➤ Effectiveness of the proposed CR-enabled SPS method

[Ref] F. Sotriani and W. Yu, "Hybrid Analog and Digital Beamforming for mmWave OFDM Large-Scale Antenna Arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432-1443, July 2017.

Boost Computational Efficiency

❖ Simulation results (Partially-connected)

$$N_t = 256, N_r = 16, K = 4, F = 128, N_s = 2 \quad N_{\text{RF}}^t = KN_s, \text{ and } N_{\text{RF}}^r = N_s$$



- Simply doubling PSs in the partially-connected mapping is far from satisfactory
- Superiority of the modified K-means algorithm with lower computational complexity than the greedy algorithm

Boost Computational Efficiency

❖ Discussions

➤ Comparison of computational complexity

Implementation	Structure	Design approach	Hardware complexity (No. of phase shifters)	Computational complexity	Performance
SPS	Fully-connected	MO-AltMin	$N_{\text{RF}}^t N_t$	Extremely high	✓✓✓
	Partially-connected	SDR-AltMin	N_t	High	✓
DPS	Fully-connected	Matrix decomposition	$2N_{\text{RF}}^t(N_t - N_{\text{RF}}^t)$	$\mathcal{O}(N_{\text{RF}}^t{}^2 N_t F)$	✓✓✓✓
	Partially-connected	Modified K-means	$2N_t$	$\mathcal{O}(N N_{\text{RF}}^t{}^2 N_t F)$	✓✓

➤ The proposed DPS implementation enables low-complexity design for hybrid beamforming

Boost Computational Efficiency

❖ Discussions

- The number of RF chains has been reduced to the minimum

$$N_{\text{RF}}^t = KN_s$$

- A large number of high-precision phase shifters are still needed

	Fully-connected	Partially-connected
SPS	$N_t N_{\text{RF}}$	N_t
DPS	$2N_t N_{\text{RF}}$	$2N_t$

- Need to adapt the phases to channel states

- ❖ Practical phase shifters are typically with coarsely **quantized** phases

How to reduce # phase shifters?

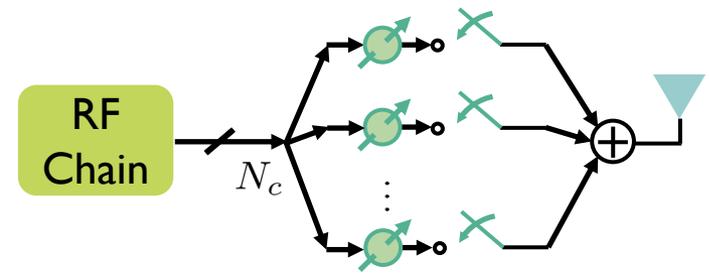
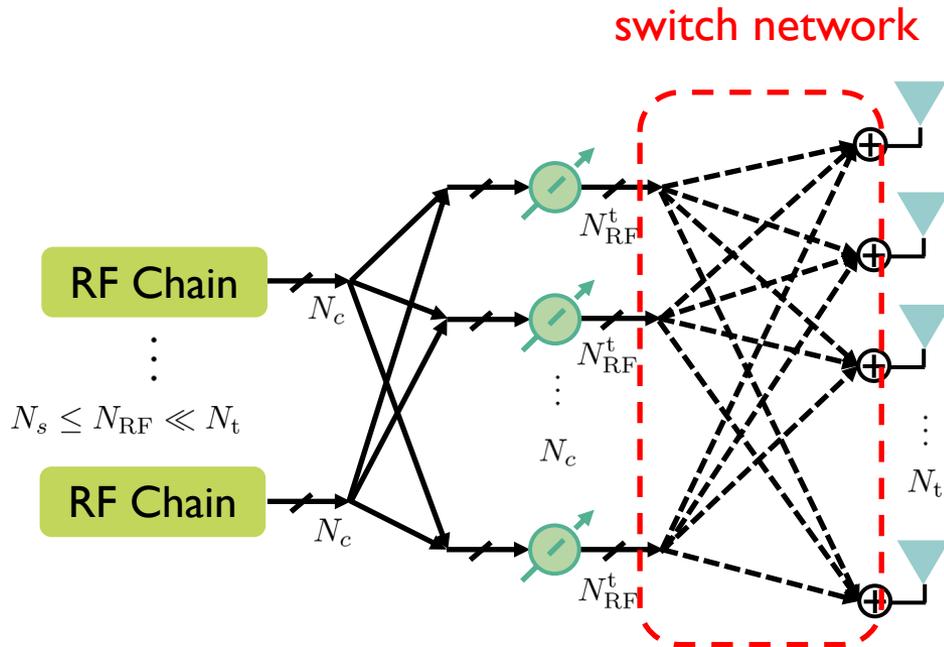
Fight for Hardware Efficiency: How Many Phase Shifters Are Needed?

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

[Ref] X. Yu, J. Zhang, and K. B. Letaief, “A hardware-efficient analog network structure for hybrid precoding in millimeter wave systems,” *IEEE J. Sel. Topics Signal Process., Special Issue on Hybrid Analog-Digital Signal Processing for Hardware-Efficient Large Scale Antenna Arrays*, vol. 12, no. 2, pp. 282-297, May 2018.

Fight for Hardware Efficiency

❖ Fixed phase shifter (FPS) implementation



Q: How to design these adaptive switches?

➤ N_c multi-channel **fixed PSs** [Z. Feng et al., 2014]

Fight for Hardware Efficiency

❖ Problem formulation

➤ $\mathcal{A}_x: \mathbf{F}_{\text{RF}} = \mathbf{S}\mathbf{C}$

➤ FPS matrix $\mathbf{C} = \text{diag}(\overbrace{\mathbf{c}, \mathbf{c}, \dots, \mathbf{c}}^{N_{\text{RF}}^t}), \quad \mathbf{c} = \frac{1}{\sqrt{N_c}} [e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_{N_c}}]^T$

Phases are fixed

➤ Binary switch matrix $\mathbf{S} \in \{0, 1\}^{N_t \times N_c N_{\text{RF}}^t}$

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

$$\text{subject to} \quad \mathbf{S} \in \{0, 1\}^{N_t \times N_c N_{\text{RF}}^t}$$

NP-hard

❖ An objective upper bound enables a low-complexity algorithm

➤ Enforce a semi-orthogonal constraint on \mathbf{F}_{BB} [X.Yu et al., 2016]

$$\mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{BB}} = \alpha^2 \mathbf{F}_{\text{DD}}^H \mathbf{F}_{\text{DD}} = \alpha^2 \mathbf{I}_{KN_s}$$

$$\|\mathbf{F}_{\text{opt}} - \mathbf{S}\mathbf{C}\mathbf{F}_{\text{BB}}\|_F^2 \leq \|\mathbf{F}_{\text{opt}}\|_F^2 - 2\alpha \Re \text{Tr}(\mathbf{F}_{\text{DD}} \mathbf{F}_{\text{opt}}^H \mathbf{S}\mathbf{C}) + \alpha^2 \|\mathbf{S}\|_F^2$$

Fight for Hardware Efficiency

❖ Alternating minimization

➤ Digital precoder

$$\begin{aligned} & \underset{\mathbf{F}_{\text{DD}}}{\text{maximize}} && \Re \text{Tr} (\mathbf{F}_{\text{DD}} \mathbf{F}_{\text{opt}}^H \mathbf{S} \mathbf{C}) \\ & \text{subject to} && \mathbf{F}_{\text{DD}}^H \mathbf{F}_{\text{DD}} = \mathbf{I}_{KN_s} \end{aligned}$$

➤ Semi-orthogonal Procrustes solution $\mathbf{F}_{\text{DD}} = \mathbf{V}_1 \mathbf{U}^H$

$$\alpha \mathbf{F}_{\text{opt}}^H \mathbf{S} \mathbf{C} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}_1^H$$

➤ Switch matrix optimization

$$\begin{aligned} & \underset{\alpha, \mathbf{S}}{\text{minimize}} && \left\| \Re (\mathbf{F}_{\text{opt}} \mathbf{F}_{\text{DD}}^H \mathbf{C}^H) - \alpha \mathbf{S} \right\|_F^2 \\ & \text{subject to} && \mathbf{S} \in \{0, 1\}^{N_t \times N_c N_{\text{RF}}^t} \end{aligned}$$

➤ Once α is optimized, the optimal \mathbf{S} is determined correspondingly

$$\mathbf{S}^* = \begin{cases} \mathbf{1} & \left\{ \Re (\mathbf{F}_{\text{opt}} \mathbf{F}_{\text{DD}}^H \mathbf{C}^H) > \frac{\alpha}{2} \mathbf{1}_{N_t \times N_c N_{\text{RF}}^t} \right\} & \alpha > 0 \\ \mathbf{1} & \left\{ \Re (\mathbf{F}_{\text{opt}} \mathbf{F}_{\text{DD}}^H \mathbf{C}^H) < \frac{\alpha}{2} \mathbf{1}_{N_t \times N_c N_{\text{RF}}^t} \right\} & \alpha < 0 \end{cases}$$

Fight for Hardware Efficiency

❖ Alternating minimization (cont.)

➤ Optimization of α

$$\alpha^* = \arg \min_{\{\tilde{x}_i, \bar{x}_i\}_{i=1}^n} \{f(\tilde{x}_i), f(\bar{x}_i)\}$$

$$\tilde{\mathbf{x}} = \text{vec}(\Re(\mathbf{F}_{\text{opt}} \mathbf{F}_{\text{DD}}^H \mathbf{C}^H))$$
$$\tilde{\mathbf{x}} \in \mathbb{R}^n, \quad n = N_t N_{\text{RF}}^t N_c$$
$$\bar{x}_i \triangleq \begin{cases} \frac{\sum_{j=1}^i \tilde{x}_j}{i} & \alpha < 0 \text{ and } \frac{\sum_{j=1}^i \tilde{x}_j}{i} \in [2\tilde{x}_i, 2\tilde{x}_{i+1}] \\ \frac{\sum_{j=i+1}^n \tilde{x}_j}{n-i} & \alpha > 0 \text{ and } \frac{\sum_{j=i+1}^n \tilde{x}_j}{n-i} \in [2\tilde{x}_i, 2\tilde{x}_{i+1}] \\ +\infty & \text{otherwise} \end{cases}$$

➤ Search dimension: $|\mathcal{X}| = 2N_t N_{\text{RF}}^t N_c$

➤ **Acceleration:** Optimal point can only be obtained at \bar{x}_i

$$\alpha^* = \arg \min_{\bar{x}_i} f(\bar{x}_i)$$

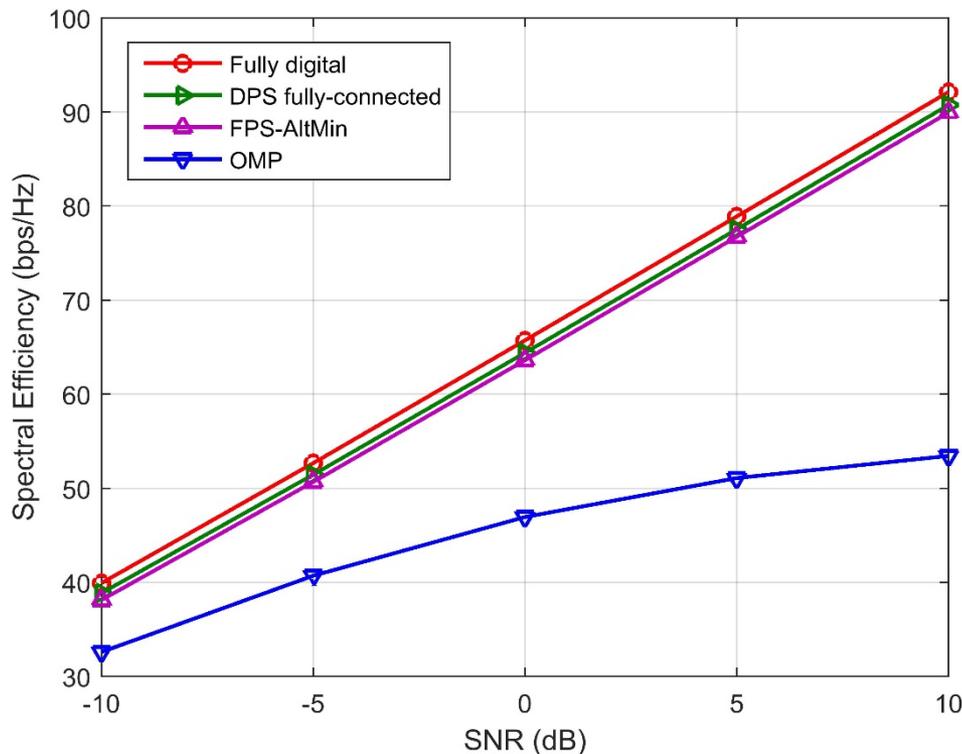
➤ Search dimension $\ll 2N_t N_{\text{RF}}^t N_c$

➤ Convergence guarantee

Fight for Hardware Efficiency

❖ Simulation results: MU-MC systems

$N_t = 144$, $N_r = 16$, $K = 4$, $F = 128$, $N_s = 2$, $N_{\text{RF}}^t = 8$, and $N_{\text{RF}}^r = 2$

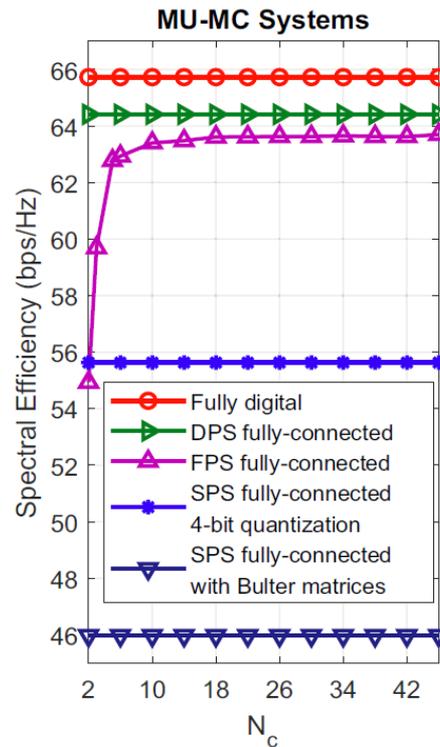
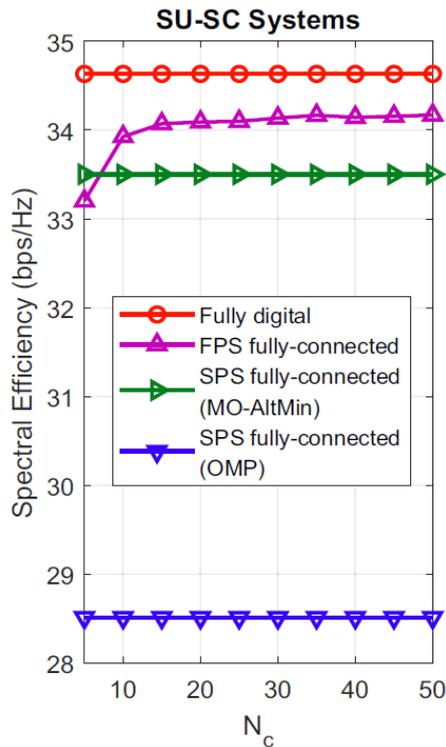


- Slightly inferior to the DPS fully-connected mapping with much fewer PSs
- Significant improvement over the OMP algorithm

Fight for Hardware Efficiency

❖ Simulation results: How many PSs are needed?

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



➤ Only ~10 fixed phase shifters are sufficient!

➤ 200 times reduction compared with the DPS implementation

Fight for Hardware Efficiency

❖ Simulation results: How much power can be saved?

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

TABLE II

POWER CONSUMPTION OF THE ANALOG NETWORK FOR DIFFERENT HYBRID PRECODER STRUCTURES IN MU-MC SYSTEMS

	Phase shifter		Other hardware		Total power [‡]
	Number N_{PS}	Type	Hardware	Number N_{OC}	P_{total}
DPS fully-connected	2304	Adaptive	N/A	N/A	115.2 W
FPS fully-connected	10	Fixed [§]	Switch	11520	59.2 W
SPS fully-connected 4-bit quantization	1152	Adaptive	N/A	N/A	57.6 W
FPS fully-connected	2	Fixed	Switch	2304	11.84 W
SPS fully-connected with Butler matrices	3456	Fixed	Coupler	4032	109.44 W

Conclusions

Conclusions

❖ Questions answered

- **Q1:** Can hybrid beamforming provide performance close to the fully digital one? **YES**
- **Q2:** How many RF chains are needed? KN_s
- **Q3:** How many phase shifters are needed? **~10 FPSs**
- **Q4:** How to efficiently design hybrid beamforming algorithms?

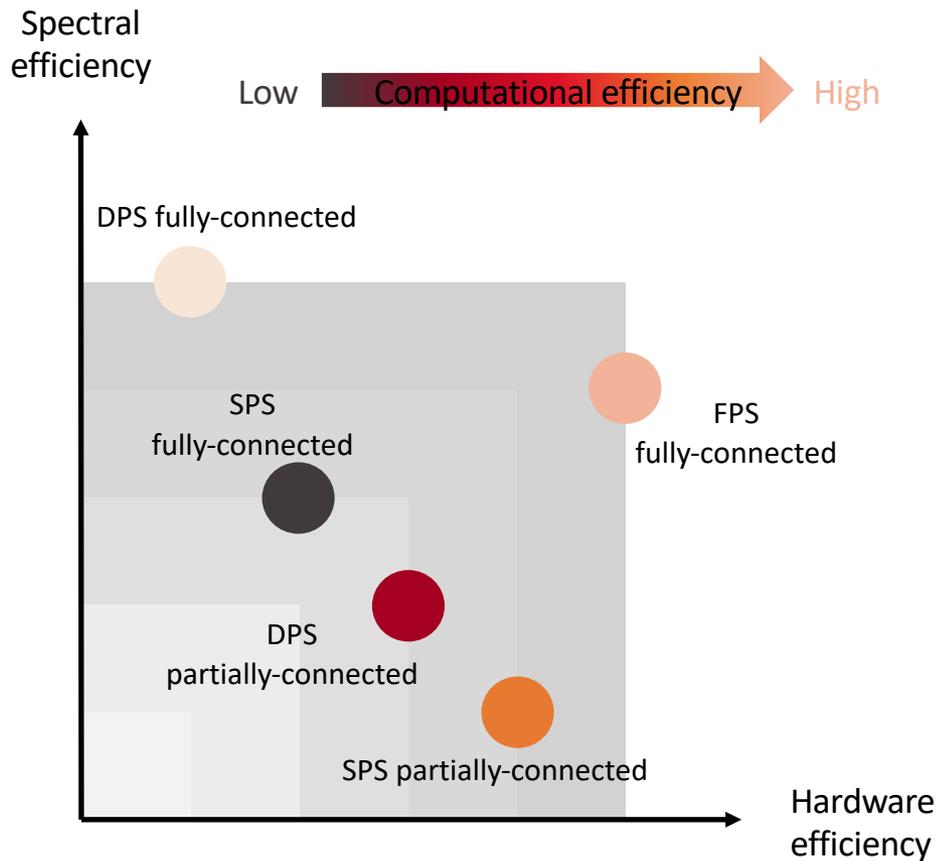
Alternating minimization provides the basic principle

Manifold optimization provides good benchmark

Convex relaxation enables low-complexity algorithms

Conclusions

❖ Comparisons between different hybrid precoder structures



- **SPS**: May not be a good choice
- **DPS**: An excellent candidate for low-complexity algorithms
- **FPS**: Trade-off between the hardware and computational complexity, with satisfactory performance

Conclusions

❖ Our own results

- **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., Special Issue on Signal Process. for Millimeter Wave Wireless Commun.*, vol. 10, no. 3, pp. 485-500, Apr. 2016. (**The 2018 SPS Young Author Best Paper Award**)
- **J. Zhang**, **X. Yu**, and K. B. Letaief, “Hybrid beamforming for 5G and beyond millimeter-wave systems: A holistic view,” *IEEE Open J. Commun. Society*, vol. 1, no. 1, pp. 77–91, Dec. 2019.
- **X. Yu**, **J. Zhang**, and K. B. Letaief, “Doubling phase shifters for efficient hybrid precoding in millimeter-wave multiuser OFDM systems,” *J. Commun. Inf. Netw.*, vol. 4, no. 2, pp. 51-67, Jul. 2019.
- **X. Yu**, **J. Zhang**, and K. B. Letaief, “Hybrid precoding in millimeter wave systems: How many phase shifters are needed?” in *Proc. IEEE Global Commun. Conf. (Globecom)*, Singapore, Dec. 2017. (**Best Paper Award**)
- **X. Yu**, **J. Zhang**, and K. B. Letaief, “A hardware-efficient analog network structure for hybrid precoding in millimeter wave systems,” *IEEE J. Sel. Topics Signal Process., Special Issue on Hybrid Analog-Digital Signal Processing for Hardware-Efficient Large Scale Antenna Arrays*, vol. 12, no. 2, pp. 282-297, May 2018.

Conclusions

❖ References in this webinar

- X. Zhang, A. F. Molisch, and S.-Y. Kung, “Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection,” *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091-4103, Nov. 2005.
- E. Zhang and C. Huang, “On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems,” in *Proc. IEEE 80th Veh. Technol. Conf. (VTC2014-Fall)*, Vancouver, BC, Sep. 2014, pp. 1-5.
- O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., “Spatially sparse precoding in millimeter wave MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014.
- N. Boumal, B. Mishra, P. A. Absil, R. Sepulchre, “Manopt, a Matlab toolbox for optimization on manifolds,” *The Journal of Machine Learning Research* vol. 15, no, 1, pp. 1455-1459, 2014.
- X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, Jr., “Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays,” *IEEE J. Sel. Areas in Commun.*, vol. 34, no. 4, pp. 998-1009, Apr. 2016.

Conclusions

❖ References in this webinar

- F. Sofrabi and W. Yu, “Hybrid digital and analog beamforming design for large-scale antenna arrays,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501-513, Apr. 2016.
- Y. Lee, C.-H. Wang, and Y.-H. Huang, “A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems,” *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 305–317, Jan. 2015.
- L. Liang, W. Xu, and X. Dong, “Low-complexity hybrid precoding in massive multiuser MIMO systems,” *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec. 2014.
- F. Sofrabi and W. Yu, “Hybrid Analog and Digital Beamforming for mmWave OFDM Large-Scale Antenna Arrays,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432-1443, July 2017.
- Z. Feng, S. Fu, T. Ming, and D. Liu, “Multichannel continuously tunable microwave phase shifter with capability of frequency doubling,” *IEEE Photon. J.*, vol. 6, no. 1, pp. 1–8, Feb. 2014.

Conclusions

❖ References in this webinar

- A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., “Channel estimation and hybrid precoding for millimeter wave cellular systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831-846, Oct. 2014.
- Z. Gao, C. Hu, L. Dai, and Z. Wang, “Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels,” *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259-1262, June 2016.
- A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., “Hybrid precoding for millimeter wave cellular systems with partial channel knowledge,” in *Proc. Inf. Theory Applications Workshop (ITA)*, San Diego, CA, 2013, pp. 1-5.
- M. N. Kulkarni, A. Ghosh, and J. G. Andrews, “A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming,” *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1952-1967, May 2016.
- X. Gao, L. Dai, Y. Sun, S. Han, and C.-L. I, “Machine learning inspired energy-efficient hybrid precoding for mmWave massive MIMO systems,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May, 2017, pp. 1-6.

Conclusions

❖ References in this webinar

- A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, “Deep learning coordinated beamforming for highly-mobile millimeter wave systems,” *IEEE Access*, vol. 6, pp. 37328-37348, 2018.
- W. Roh et al., “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106-113, Feb. 2014.
- J. Mo, A. Alkhateeb, S. Abu-Surra and R. W. Heath, “Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2274-2287, Apr. 2017.
- Y. Zeng and R. Zhang, “Millimeter Wave MIMO With Lens Antenna Array: A New Path Division Multiplexing Paradigm,” *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557-1571, Apr. 2016.
- X. Gao, L. Dai, S. Han, C.-L. I, and X. Wang, “Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010-6021, Sep. 2017.

Conclusions

❖ Other references

- A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, “Hybrid beamforming for massive MIMO: A survey,” *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.
- J. Brady, N. Behdad, and A. M. Sayeed, “Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.
- J. Singh and S. Ramakrishna, “On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2670–2683, May 2015.
- S. Park, A. Alkhateeb, and R.W. Heath, Jr., “Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.
- T. E. Bogale, L. B. Le, A.Haghighat, and L.Vandendorpe, “On the number of RF chains and phase shifters, and scheduling design with hybrid analog digital beamforming,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.

Conclusions

❖ Other references

- A. Alkhateeb and R.W. Heath, Jr., “Frequency selective hybrid precoding for limited feedback millimeter wave systems,” *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
- R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, Jr., “Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?” *IEEE Access*, vol. 4, pp. 247–267, Jan. 2016.
- W. Ni and X. Dong, “Hybrid block diagonalization for massive multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, Jan. 2016.
- T. E. Bogale and L. B. Le, “Beamforming for multiuser massive MIMO systems: Digital versus hybrid analog-digital,” in *Proc. IEEE Global Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 4066–4071.

Thanks

For more information and **Matlab codes**:

<https://yuxianghao.github.io/>

<http://www.eie.polyu.edu.hk/~jeiezhang>