Greed is good: Leveraging Submodularity for Antenna Selection in Massive MIMO

Aritra Konar

&

Nikos Sidiropoulos

Dept. of ECE, University of Virginia

University of Minnesota
Driven to Discover℠

Introduction

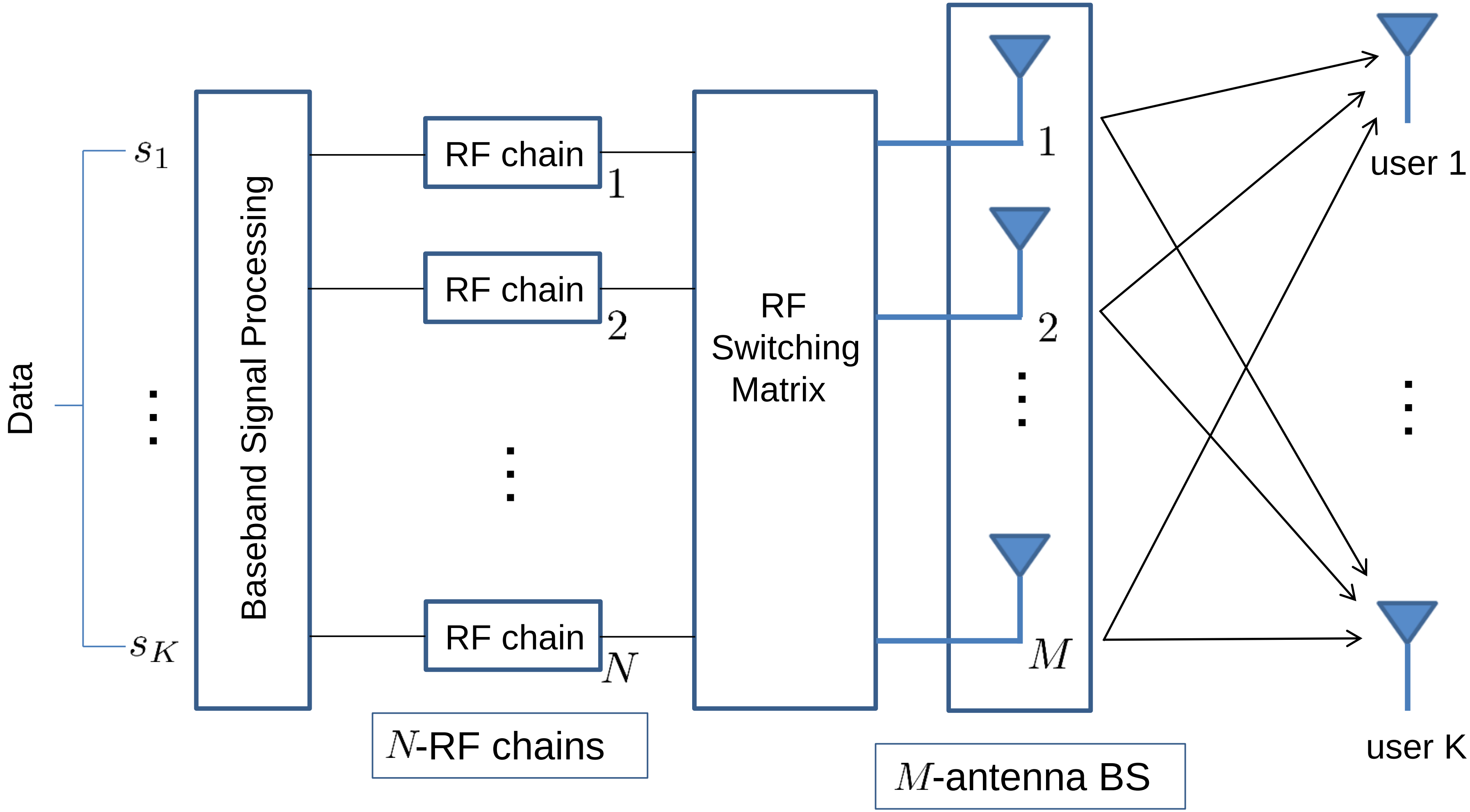
- Massive MIMO: [Marzetta 2010]
  - Large number of transmit antennas deployed at BS for serving users sharing same time-frequency resource
  - Orders of magnitude improvement in spectral and energy efficiency
  - Simple signal processing techniques exhibit near-optimal performance
  - A leading physical-layer technology candidate for 5G

- Challenge:
  - Cost and hardware complexity of large-scale antenna systems
  - Assigning one RF chain per antenna element infeasible
  - This talk: Use antenna selection to reduce the number of RF chains at BS

Prior Art

- Point-to-point case:
  - Maximize energy efficiency [Li-Song-Debbah 2014]
    - Heuristic selection; no theoretical guarantees
  - Maximize received SNR [Gkizeli-Karystinos 2014]
    - Optimally solvable in polynomial-time for ≤ 2 receive antennas
- Multi-user case:
  - Maximize downlink sum-rate capacity with fixed user power allocation [Gao *et. al* 2013]
    - Convex relaxation + rounding; no theoretical guarantees
    - Observed to work well empirically on certain measured massive MIMO channels
  - This work: Same scenario + criterion, different algorithmic approach

Problem Scenario



$$K \leq N \leq M$$

Problem Statement

- Signal Model:
  - For a given subset of antennas $\mathcal{S} \subseteq [M] := \{1, \cdots, M\}$

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}^{[\mathcal{S}]}\mathbf{x} + \mathbf{n}$$

$\mathbf{y} \in \mathbb{C}^K$: received signal across all users

$\mathbf{H}^{[\mathcal{S}]} \in \mathbb{C}^{K \times N}$: subset of columns of $\mathbf{H} \in \mathbb{C}^{K \times M}$

$\mathbf{x} \in \mathbb{C}^N$: transmit signal vector across selected antennas with $\mathbb{E}\{\|\mathbf{x}\|^2\} = 1$

$\rho > 0$: transmit power budget

$\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$

- Antenna Selection Criterion: [Gao *et. al* 2013]

$$\boxed{C(\mathbf{H}) = \max_{\substack{\mathbf{P}\in\mathcal{P}, \\ |\mathcal{S}|\leq N}} \log_2 \det\left(\mathbf{I}_{|\mathcal{S}|} + \rho(\mathbf{H}^{[\mathcal{S}]})^H \mathbf{P} \mathbf{H}^{[\mathcal{S}]}\right)}$$

$$\boxed{\mathcal{P} := \{\mathrm{diag}(\mathbf{p}) \mid \mathbf{p} >= \mathbf{0}, \mathbf{1}^T\mathbf{p} \leq 1\}}$$

Mixed-Integer problem, hard to solve

Problem Statement

- Problem "Simplification":
  - Fix user power allocations; e.g., optimal solution without selection
  - Obtain subset selection problem

$$\boxed{\mathcal{S}^* = \arg \max_{|\mathcal{S}| \leq N} \left\{ f(\mathcal{S}) := \log_2 \det \left(\mathbf{I}_{|\mathcal{S}|} + \rho (\mathbf{H}^{[\mathcal{S}]})^H \mathbf{P} \mathbf{H}^{[\mathcal{S}]} \right) \right\}}$$

  - NP-hard! [Ko-Lee-Queyranne 1995]

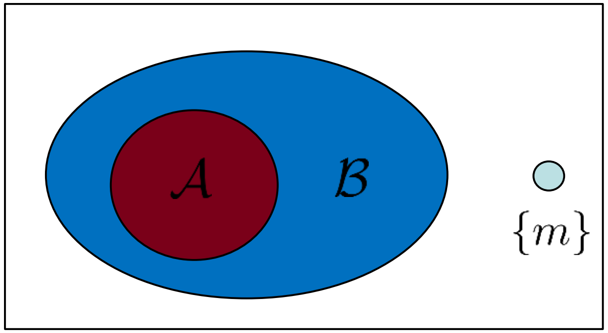
- Relax and Round: [Gao *et. al* 2013]
  - Relax discrete variables, solve convex optimization problem, perform rounding to select antennas
    - Computationally expensive: $O(M^{3.5})$ [M is large in massive MIMO]
    - Hard to quantify sub-optimality of obtained solution
  - Does there exist a more efficient and well-principled approach?

Submodularity

- Definition:
  - A set function $f : 2^M \rightarrow \mathbb{R}$ is submodular if for any $\mathcal{A}, \mathcal{B} \subseteq [M]$

$$\boxed{f(\mathcal{A} \cup \mathcal{B}) + f(\mathcal{A} \cap \mathcal{B}) \leq f(\mathcal{A}) + f(\mathcal{B})}$$

  - Equivalently, for all $\mathcal{A} \subseteq \mathcal{B} \subseteq [M] \setminus m$:



$$\boxed{f(\mathcal{A} \cup \{m\}) - f(\mathcal{A}) \geq f(\mathcal{B} \cup \{m\}) - f(\mathcal{B})}$$

A diminishing returns property

  - A set function is monotone if $\mathcal{A} \subseteq \mathcal{B} \implies f(\mathcal{A}) \leq f(\mathcal{B})$
    - Equivalently, for submodular functions,

$$f([M]) \geq f([M] \setminus m), \forall\, m \in [M]$$

Submodularity

- Proposition:
  - Objective function of antenna selection criterion is monotone submodular
  - Express $f(\mathcal{S}) = \log_2 \det\left(\mathbf{I}_{|\mathcal{S}|} + \rho(\mathbf{H}^{[\mathcal{S}]})^H \mathbf{P}\mathbf{H}^{[\mathcal{S}]}\right)$

$$= \log_2 \det\left(\mathbf{I}_M[\mathcal{S},\mathcal{S}] + \rho\tilde{\mathbf{G}}[\mathcal{S},\mathcal{S}]\right) \qquad \boxed{\tilde{\mathbf{G}} = \mathbf{H}^H\mathbf{P}\mathbf{H}}$$

$$= \log_2 \det(\mathbf{\Sigma}[\mathcal{S},\mathcal{S}]) \qquad \boxed{\mathbf{\Sigma} := \mathbf{I}_M + \rho\tilde{\mathbf{G}}}$$

  - Consider the Gaussian random vector $\mathbf{z} \sim \mathcal{N}(\mathbf{0},\mathbf{\Sigma})$ with differential entropy

$$h(\mathbf{z}) = \frac{1}{2}\log_2 \det(\mathbf{\Sigma})$$ (Up to additive constants)

  - For a given subset of random variables $\mathcal{S} \subseteq [M]$

$$h(\mathbf{z}^{[\mathcal{S}]}) = \frac{1}{2}\log_2 \det(\mathbf{\Sigma}[\mathcal{S},\mathcal{S}]) + c|\mathcal{S}|$$

Submodularity

Proof of submodularity:

- Differential entropy is submodular [Fujishige 1978, Kelmans-Kimelfeld 1983, Krause-Guestrin 2005, Shamaiah *et al.* 2010, Bach 2013]
- Given two arbitrary subsets $\mathcal{A}, \mathcal{B} \subseteq [M]$

$$I(\mathbf{z}^{[\mathcal{A}\setminus\mathcal{B}]}; \mathbf{z}^{[\mathcal{B}\setminus\mathcal{A}]} | \mathbf{z}^{[\mathcal{A}\cap\mathcal{B}]}) = h(\mathbf{z}^{[\mathcal{A}]}) + h(\mathbf{z}^{[\mathcal{B}]}) - h(\mathbf{z}^{[\mathcal{A}\cup\mathcal{B}]}) - h(\mathbf{z}^{[\mathcal{A}\cap\mathcal{B}]}) \geq 0$$

- Alternatively, given $\mathcal{A} \subseteq \mathcal{B} \subseteq [M] \setminus m$

$$h(\mathbf{z}^{[\mathcal{A}\cup\{m\}]}) - h(\mathbf{z}^{[\mathcal{A}]}) = h(z_m|\mathbf{z}^{[\mathcal{A}]}) \geq h(z_m|\mathbf{z}^{[\mathcal{B}]}) = h(\mathbf{z}^{[\mathcal{B}\cup\{m\}]}) - h(\mathbf{z}^{[\mathcal{B}]})$$

Proof of monotonicity:

- Required to show

$$\begin{aligned} f([M]) - f([M] \setminus m) &= \log_2 \det(\mathbf{\Sigma}) - \log_2 \det(\mathbf{\Sigma}[[M] \setminus m, [M] \setminus m]) \\ &\geq 0, \forall\, m \in [M] \end{aligned}$$

- Follows as a consequence of Cauchy's Theorem of interlacing eigen-values

Submodularity

- Antenna selection problem:
  - Equivalent to maximizing a monotone submodular function subject to cardinality constraint on number of selected antennas
- The upshot:
  - Problem is well posed
    - Few antennas can possibly capture significant fraction of downlink capacity
- The catch:
  - Still need to perform subset selection! (NP-hard)
  - Exploit submodularity to obtain bumper-to-bumper insurance?

Greed is good for Antenna Selection

- Greedy Algorithm:
  - Start with $\mathcal{S}_0 = \emptyset$
  - At iteration $i \in [N] := \{1, \cdots, N\}$

$$\boxed{\mathcal{S}_i = \mathcal{S}_{i-1} \cup \left\{ \arg\max_{m \notin \mathcal{S}_{i-1}} f(\mathcal{S}_{i-1} \cup \{m\}) - f(\mathcal{S}_{i-1}) \right\}}$$

  - Guaranteed $(1 - 1/e)$-factor approximation for all instances! [Nemhauser-Fisher-Wolsey 1978]
    - Independent of all system parameters
  - Provably optimal approximation factor
    - Cannot be improved in polynomial-time [Nemhauser-Wolsey1978]

Greed is good for Antenna Selection

- Running time:
  - Evaluate $f(.)$ on $O(MN)$ sets
  - Cost of evaluation
    - Define $\mathcal{S}_i^{(v)} := \mathcal{S}_{i-1} \cup \{v\}, \forall v \notin \mathcal{S}_{i-1}, \forall i \in [N]$
    - Then

$$\boxed{f(\mathcal{S}_i^{(v)}) = \log_2 \det\left(\mathbf{I}_K + \rho \tilde{\mathbf{H}}^{[\mathcal{S}_i^{(v)}]}(\tilde{\mathbf{H}}^{[\mathcal{S}_i^{(v)}]})^H\right)} \qquad \boxed{\tilde{\mathbf{H}} = \mathbf{P}^{1/2}\mathbf{H}}$$

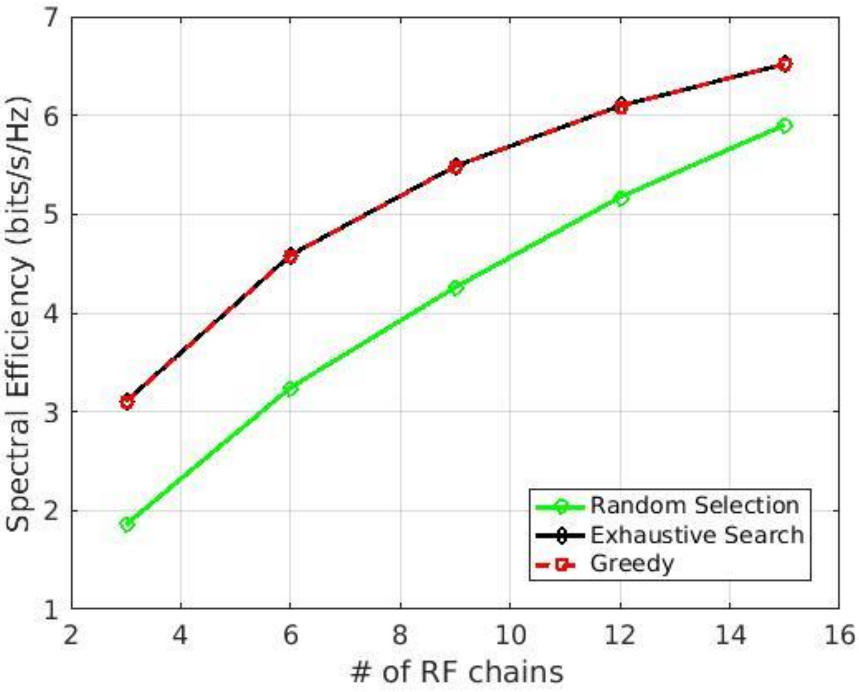
  - Overall complexity: $O(MNK^3)$
  - Can be improved to: $O(MNK^2)$
    - Evaluating $f(\mathcal{S}_i^{(v)})$ requires rank-1 updates of the form

$$\tilde{\mathbf{H}}^{[\mathcal{S}_i^{(v)}]}(\tilde{\mathbf{H}}^{[\mathcal{S}_i^{(v)}]})^H = \tilde{\mathbf{H}}^{[\mathcal{S}_{i-1}]}(\tilde{\mathbf{H}}^{[\mathcal{S}_{i-1}]})^H + \tilde{\mathbf{h}}^{[v]}(\tilde{\mathbf{h}}^{[v]})^H$$

  - Can be improved further via lazy evaluations [Minoux 1978]
    - Scales linearly with M in practice

Preliminary Results

BS with 20 antennas, 3 users, single sub-carrier, Rayleigh fading, 500 MC trials, $\rho = -2\text{dB}$



N	3	6	9	12	15
Greedy	99.86	99.97	99.99	100	100
Random	60.15	70.94	77.69	84.85	90.54

Average approximation quality of obtained solutions (in %)

N	3	6	9	12	15
Greedy	94.29	98.60	99.67	100	100
Random	23.63	39.87	52.25	62.30	76.03

Worst-case approximation quality of obtained solutions (in %)

Greedy algorithm provides near-optimal solution in all cases

Experimental Setup:

❑ Channel Model

- ➢ BS equipped with ULA with following channel model

$$\mathbf{h}_k^H = \sqrt{\frac{N}{L_m}} \sum_{l=1}^{L_m} \alpha_k^{(l)} \mathbf{a}_t\left(\phi_{t_k}^{(l)}\right)^H, \ \forall\ k \in [K]$$

Path loss $\sim \mathcal{CN}(0,1)$

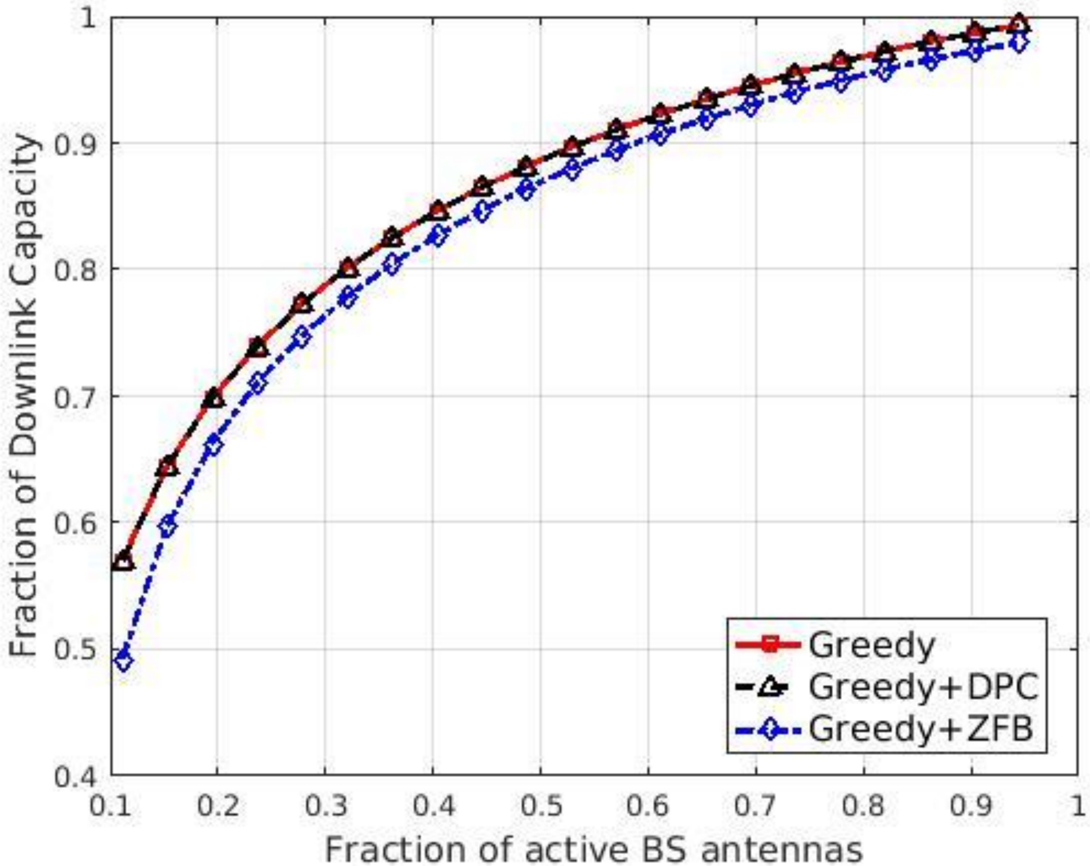
AoD $\sim \mathcal{U}[-\frac{\pi}{2}, \frac{\pi}{2}]$

❑ Setup

- ➢ After selection, design zero-forcing beamformer (ZFB) for reduced MIMO broadcast channel
- ➢ All results averaged across 500 MC trials

Results

Scenario with 144 Tx antennas, 12 users, 5-15 (randomly chosen) scattering paths per user, $\rho = 10\text{dB}$



Greedy selection + ZFB can indeed capture significant fraction of total downlink capacity using few RF chains (**50% with 11% of active antennas**)

Conclusions

- Submodularity for Antenna Selection in Massive MIMO
  - Greedy selection + ZFB works well at low complexity
  - Extensions
    - Multiple receive antennas per user
    - Multiple sub-carriers
    - Partially connected switching architectures
  - Paves the way for significant reduction of hardware complexity in large-scale antenna systems

Greed is good for Antenna Selection

- Extensions:
  - Multiple receive antennas per user
    - Straightforward; $(1-1/e)$-approximation factor
  - Multiple sub-carriers

$$\boxed{\max_{|\mathcal{S}|\leq N}\left\{F(\mathcal{S}) := \sum_{l=1}^{L}\log_2\det\left(\mathbf{I}_{|\mathcal{S}|}+\rho(\mathbf{H}^{[\mathcal{S}]}(\ell))^H\mathbf{P}(\ell)\mathbf{H}^{[\mathcal{S}]}(\ell)\right)\right\}}$$

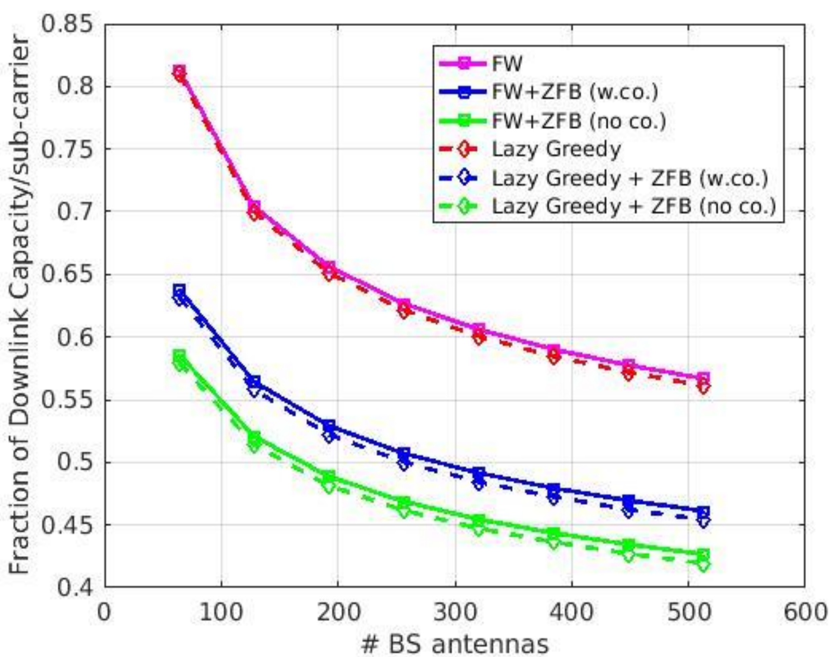
    - Monotonicity and submodularity preserved under non-negative sums; $(1-1/e)$-approximation factor
  - Partially connected switching architectures
    - Define array partition $[M]=\cup_{b\in[B]}\mathcal{M}_b$ into B disjoint sub-arrays; allocate N_b RF chains per sub-array
    - Feasible selection sets:

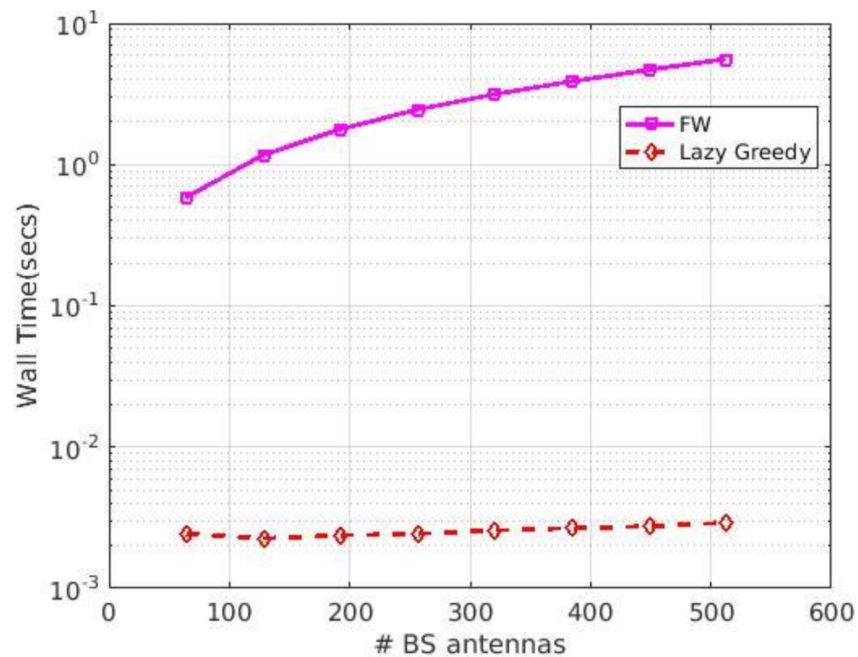
$$\mathcal{I}=\{\mathcal{S}\subseteq[M] : |\mathcal{S}\cap\mathcal{M}_b|\leq N_b, \forall\, b\in[B]\}$$

    - 0.5-approximation factor [Fisher-Nemhauser-Wolsey 1978]

Sneak peek………

N = 32 RF chains in a PC RF switching network with B = 32 sub-arrays of equal size, L = 32 sub-carriers, K = 12 users with 2 receive antennas, $\rho = 20\text{dB}$





Greedy with lazy evaluations demonstrates significantly better performance-complexity trade-off compared to convex relaxation; ZFB can still attain a significant portion of the sum-rate