ACM MOBIHOC 2023

TEXAS
The University of Texas at Austin

Carnegie Mellon University

Microsoft

2ACE: SPECTRAL PROFILE-DRIVEN MULTI-RESOLUTIONAL COMPRESSIVE SENSING FOR MMWAVE CHANNEL ESTIMATION




Yiwen Song
Carnegie Mellon University

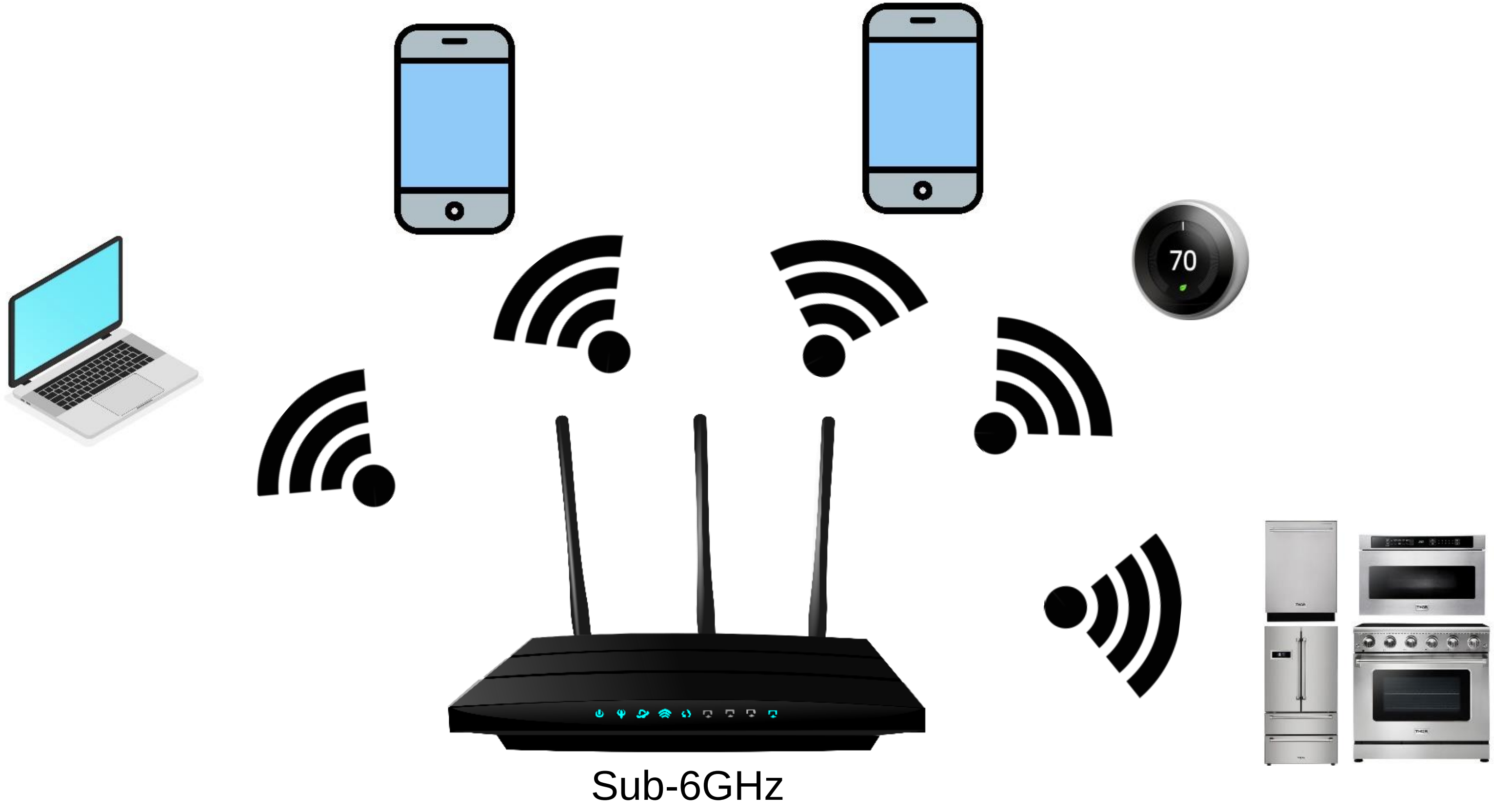
Changhan Ge
UT Austin

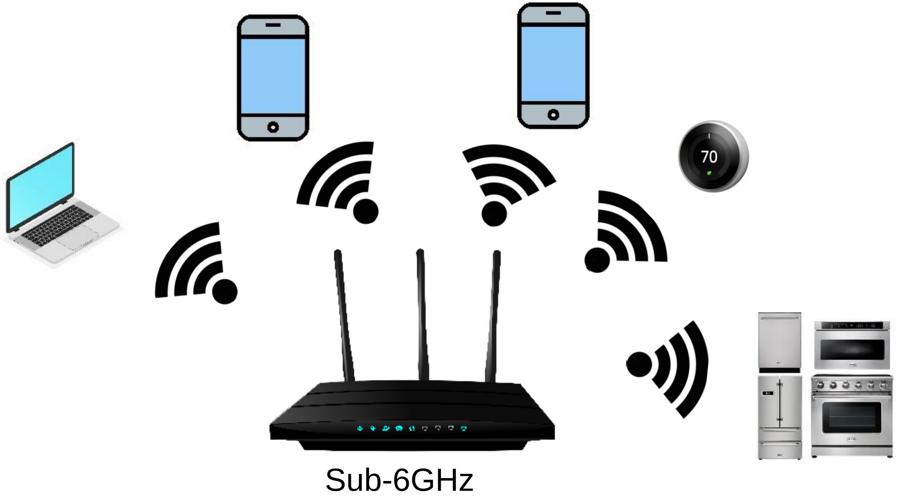
Lili Qiu
UT Austin &
Microsoft Research Asia

Yin Zhang

October 24 2023 @ Washington DC, United States of America

Sub-6GHz band is becoming more and more crowded…

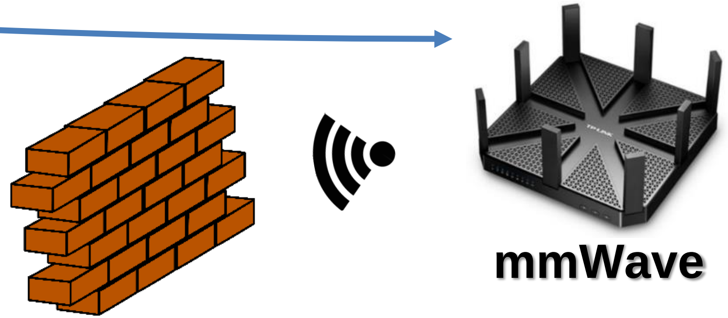




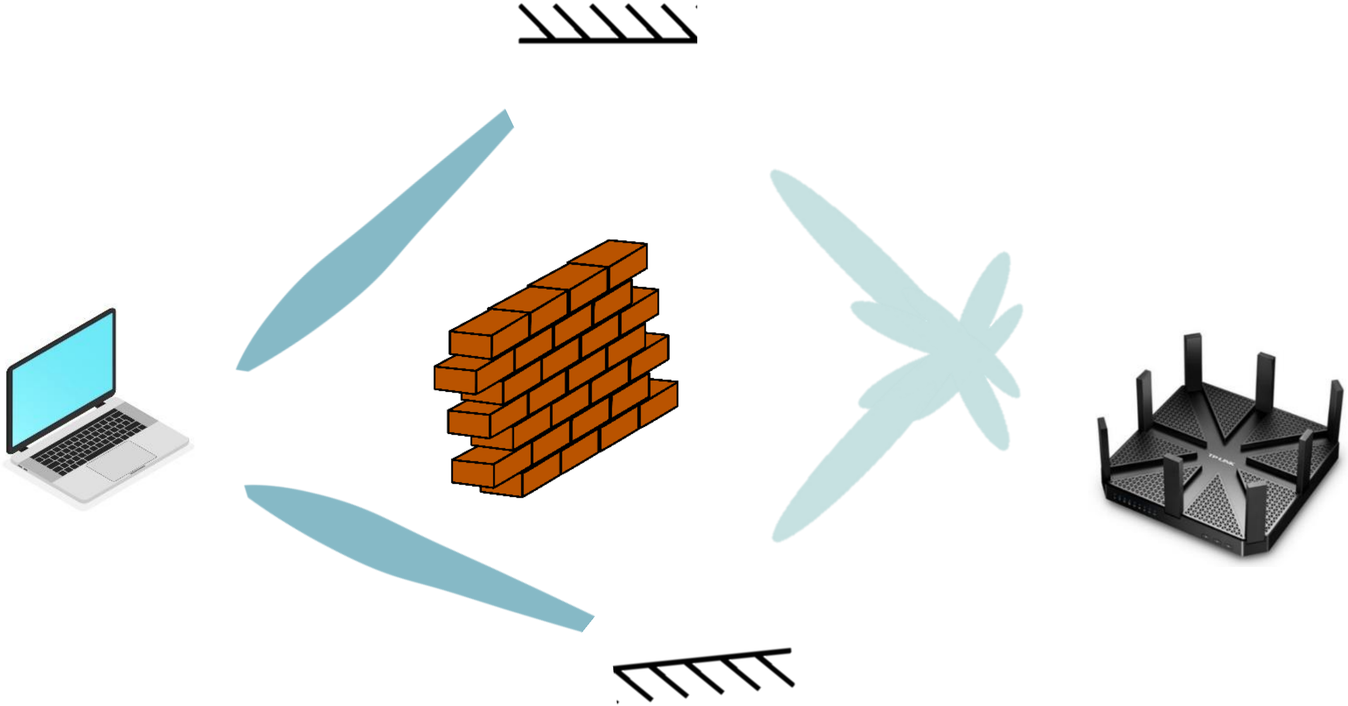
mmWave provides:

- ✓ Higher bandwidth, higher throughput.
- ✓ Higher directionality.
- ✓ More sensing opportunities.

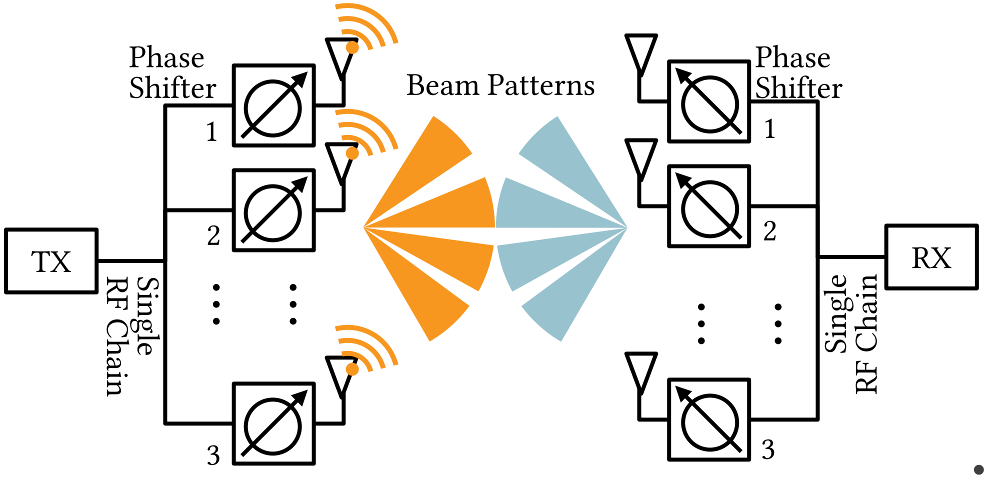
mmWave can be easily blocked. ✗



Beamforming to combat occlusions.



Channel estimation is critical to beamforming



- Adjust the **phase & amplitude** at each antenna to obtain optimal beam patterns.

Channel

$$b = \boldsymbol{w}^T \boldsymbol{H} \boldsymbol{f} \ \gamma + \sigma$$

Rx combiner

Tx precoder

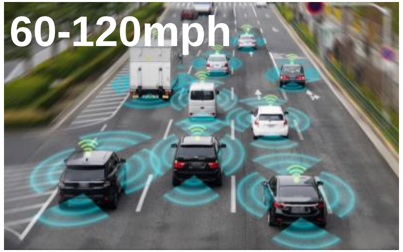
- To set the correct combiner and precoder, one need to estimate channel, i.e., How the wave propagates.

mmWave asks for fast and accurate channel estimation methods.

Complex indoor environment

Large antenna array

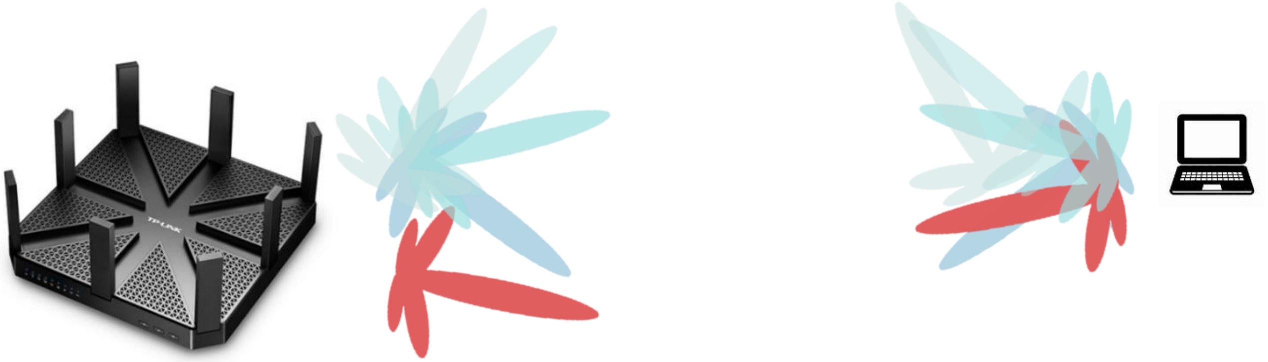
NOKIA AirScale [1] 64Tx 64Rx Massive MIMO mmWave antenna array



High Mobility

[1] Image: https://www.rfwireless-world.com/

Channel estimation is accomplished through a probing process.



known variables: precoders and combiners

RSS $|b|$ measured
& fed back by Rx

$$|b| = |\boldsymbol{w}^T \boldsymbol{H} \boldsymbol{f} \gamma + \sigma|$$

Variable needs to recover: CSI Matrix

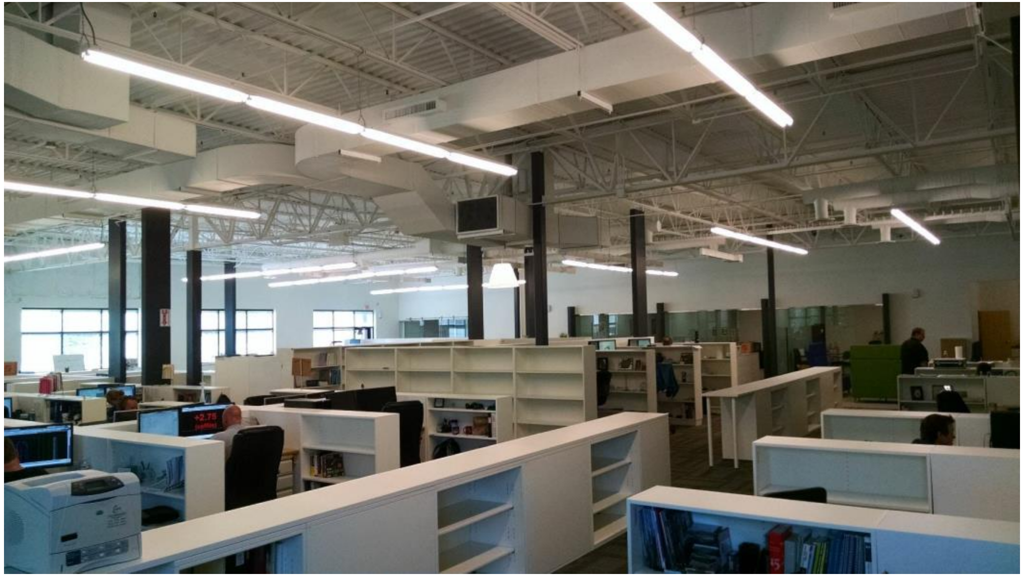
Phase retrieval & difficult problem

Existing approaches on channel estimation

802.11ad Sector Level Sweeping (SLS)	Sweep through pre-defined directions. Fast but inaccurate.
ACO [MobiCom'18]	Special codebook. Medium overhead, but coarse channel estimation.
PhaseLift [CPAM'13]	Compressive sensing recovery. Accurate, but slow & requires large number of probes.
PLGAMP/PLOMP [MobiHoc'19]	Low-rank CSI assumption-based compressive sensing. Fast when channel is sparse, otherwise inaccurate.

TEXAS The University of Texas at Austin Carnegie Mellon University Microsoft WHAT STARTS HERE CHANGES THE WORLD

**2ACE investigates how channel matrices looks like,
and use the matrix property to improve compressive sensing.**



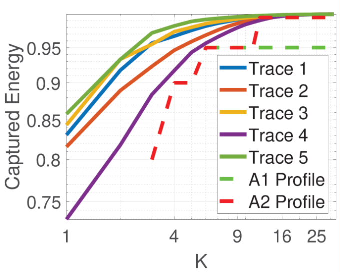
2ACE: Accelerated & Accurate Channel Estimation.

Overcoming low-rank structure

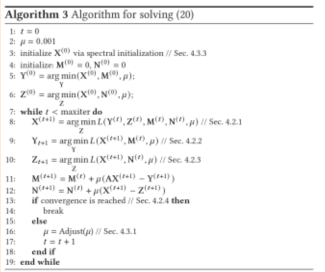
Speed up estimation

Support different probing budgets

Support dynamic environment



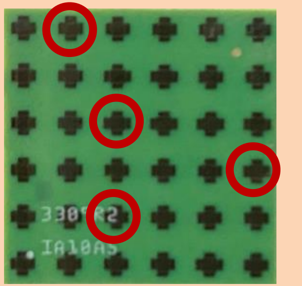
Spectral profile: describe CSI matrix structure.



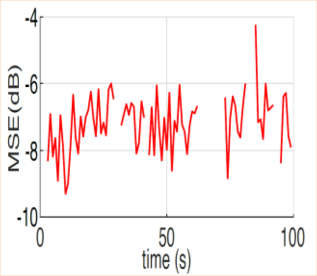
Algorithm 3 Algorithm for solving (20)

$t = 0$
$\mu = 0.001$
initialize $\mathbf{X}^{(0)}$ via spectral initialization // Sec. 4.3.3
initialize: $\mathbf{M}^{(0)} = 0, \mathbf{N}^{(0)} = 0$
$\mathbf{Y}^{(0)} = \arg\min_{\mathbf{Y}}(\mathbf{X}^{(0)}, \mathbf{M}^{(0)}, \mu)$;
$\mathbf{Z}^{(0)} = \arg\min_{\mathbf{Z}}(\mathbf{X}^{(0)}, \mathbf{N}^{(0)}, \mu)$;
while $t <$ maxiter **do**
  $\mathbf{X}^{(t+1)} = \arg\min_{\mathbf{Z}} L(\mathbf{Y}^{(t)}, \mathbf{Z}^{(t)}, \mathbf{M}^{(t)}, \mathbf{N}^{(t)}, \mu)$ // Sec. 4.2.1
  $\mathbf{Y}_{t+1} = \arg\min_{\mathbf{Y}} L(\mathbf{X}^{(t+1)}, \mathbf{M}^{(t)}, \mu)$ // Sec. 4.2.2
  $\mathbf{Z}_{t+1} = \arg\min_{\mathbf{Z}} L(\mathbf{X}^{(t+1)}, \mathbf{N}^{(t)}, \mu)$ // Sec. 4.2.3
  $\mathbf{M}^{(t+1)} = \mathbf{M}^{(t)} + \mu(\mathbf{A}\mathbf{X}^{(t+1)} - \mathbf{Y}^{(t+1)})$
  $\mathbf{N}^{(t+1)} = \mathbf{N}^{(t)} + \mu(\mathbf{X}^{(t+1)} - \mathbf{Z}^{(t+1)})$
  if convergence is reached // Sec. 4.2.4 **then**
    break
  else
    $\mu = \text{Adjust}(\mu)$ // Sec. 4.3.1
    $t = t + 1$
  end if
end while

Enhancements.



Multi-resolution algorithm.



Confidence indicator.

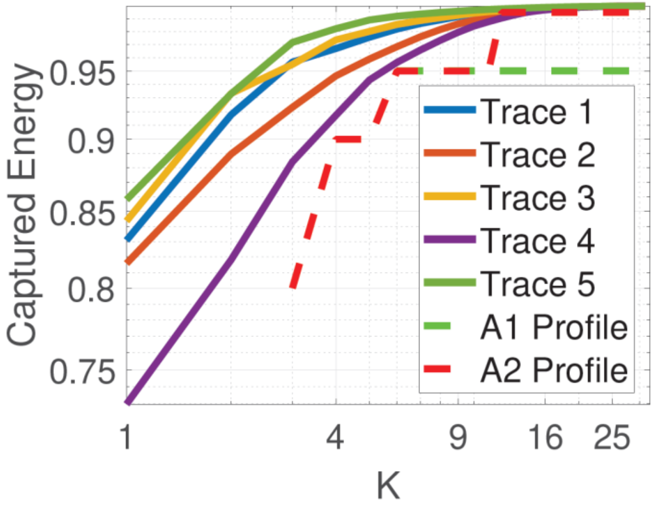
How does actual channel matrix look like?

$$\begin{bmatrix} \blacksquare & \cdots & \blacksquare \\ \vdots & \ddots & \vdots \\ \blacksquare & \cdots & \blacksquare \end{bmatrix} \xrightarrow{\text{SVD}} U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_s \end{bmatrix} V^T, \sigma_1 \geq \cdots \geq \sigma_s$$

$$\frac{\sum_{k=1}^{K} \sigma_k}{\sum_{k=1}^{s} \sigma_k}$$

Energy captured by the first K singular values

Total energy



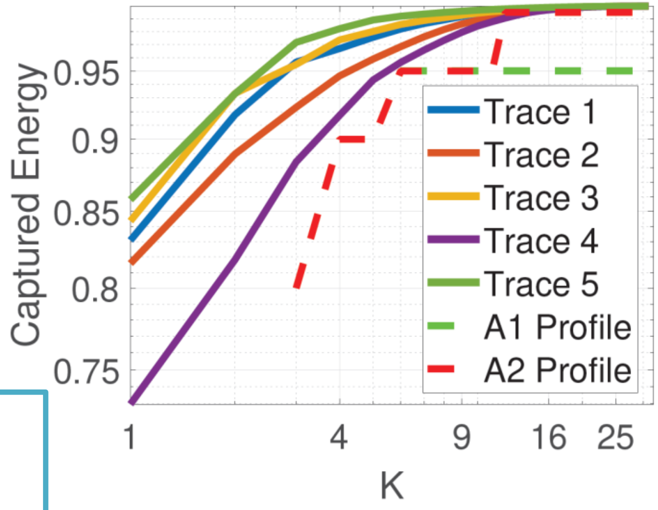
We use a lower-bound to characterize **the energy captured by the first K singular values–** Called **Spectral Profile**.

$$\frac{\sum_{k=1}^{K} \sigma_k}{\sum_{k=1}^{s} \sigma_k}$$

Energy captured by the first K singular values

Total energy

Spectral Profile: First K eigenvalues account $x\%$ energy.
K, x can be defined according to different matrices.



Use the spectral profile as a regularization – 2ACE

Minimize the squared error.

$$\min_{\mathbf{X}} \quad \frac{1}{2}\||\mathbf{Y}| - \mathbf{b}\|_2^2 + I(\mathbf{Z}, P)$$
$$\text{subject to} \quad \mathbf{AX} = \mathbf{Y} \quad \text{and} \quad \mathbf{X} = \mathbf{Z}$$

Regularization through spectral profile P.

The problem can be solved via Alternating Direction Method of Multiplier (ADMM).

We then have the Augmented Lagrangian as follow:

$$L(\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{M}, \mathbf{N}, \mu) = \frac{1}{2}\||\mathbf{Y}| - \mathbf{b}\|_2^2 + I(\mathbf{Z}, P) + < \mathbf{M}, \mathbf{AX} - \mathbf{Y} > + < \mathbf{N}, \mathbf{X} - \mathbf{Z} > + \frac{\mu}{2}\|\mathbf{AX} - \mathbf{Y}\|_2^2 + \frac{\mu}{2}\|\mathbf{X} - \mathbf{Z}\|_2^2$$

(See our paper for step-by-step math)

2ACE: Enhancements

Dynamic Update of μ

Define **primal residue**

$$r_{\text{prim}} = \sqrt{\|\mathbf{AX}^{(t+1)} - \mathbf{Y}^{(t+1)}\|_2^2 + \|\mathbf{X}^{(t+1)} - \mathbf{Z}^{(t+1)}\|_2^2}$$

How the constraints are satisfied

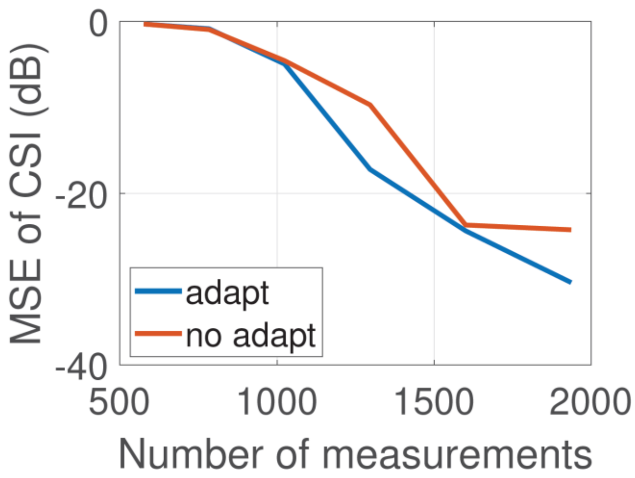
Define **combined residue**

$$r_{\text{comb}}^{(t+1)} = \mu r_{\text{prim}}^2 + \mu(\|\mathbf{Y}^{(t+1)} - \mathbf{Y}^{(t)}\|_2^2 + \|\mathbf{Z}^{(t+1)} - \mathbf{Z}^{(t)}\|_2^2)$$

How large is the step length

Algorithm 2 Adaptation of μ

1: **if** $r_{\text{comb}}^{(t+1)} > 0.8 r_{\text{comb}}^{(t)}$ **then**
2: $\quad \mu^{(t+1)} = 1.03\mu^{(t)}$
3: **else**
4: $\quad \mu^{(t+1)} = \mu^{(t)}$
5: **end if**



2ACE: Enhancements

$$\min_{\mathrm{X}} \quad \frac{1}{2}\||\mathrm{Y}| - \mathbf{b}\|_2^2 + I(\mathrm{Z}, P)$$

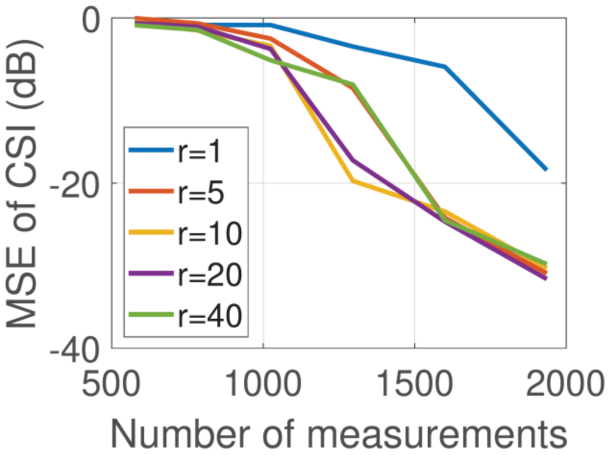
$$\text{subject to} \quad \mathrm{AX} = \mathrm{Y} \quad \text{and} \quad \mathrm{X} = \mathrm{Z}$$

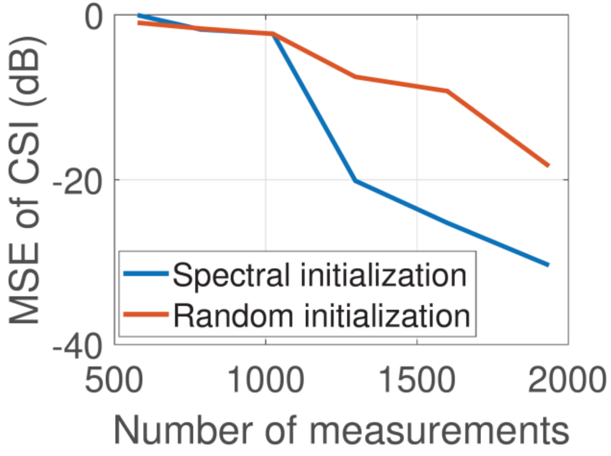
Parallel refinement

- Solving r candidate $\boldsymbol{X}$ in parallel

Spectral Initialization

- Initialize the candidate $\boldsymbol{X}$ according to the best rank-r estimation.





2ACE: Enhancements

Choice of Spectral Profiles

Algorithm 4 2ACE Algorithm to incorporate dynamic profile

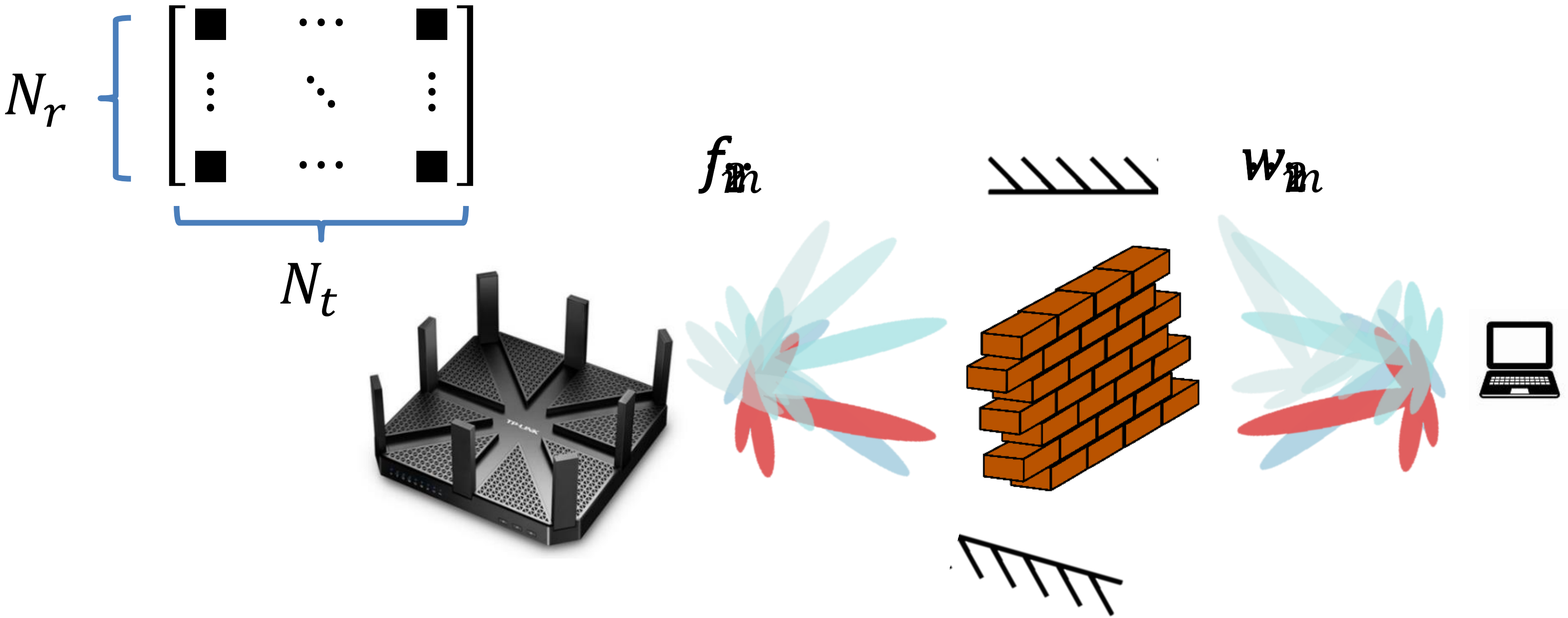
if $m >= 3n$ **then**
  // no need to use spectral profile w/ enough constraints
  $P = \{\}$
else if $m < n$ **then**
  // focus on estimating 1st singular vector w/ too few constraints
  $P = \{(r_1, 0.95)\}$
else
  $P = \{(r_1, f_1), (r_2, f_2), (r_3, f_3), (r_4, f_4)\}$
end if

Large probe number: no spectral profile

Small probe number: Coarse spectral profile

Medium probe number: Detailed spectral profile

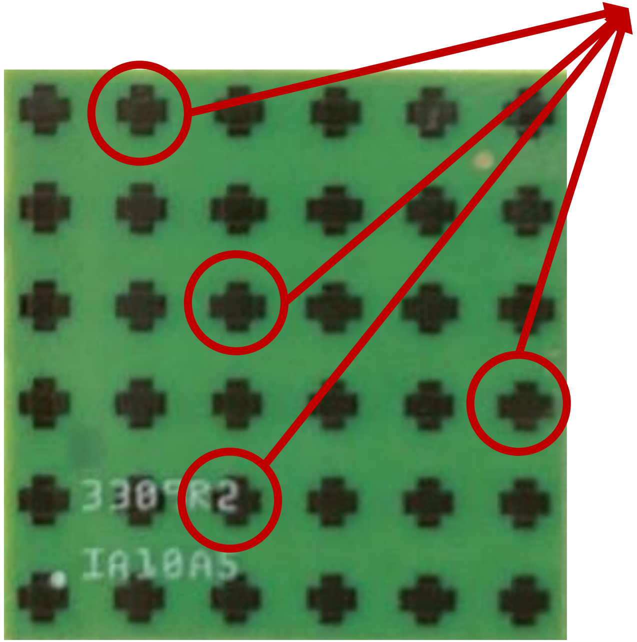
#Probes m is dependent on the size of the channel matrix.



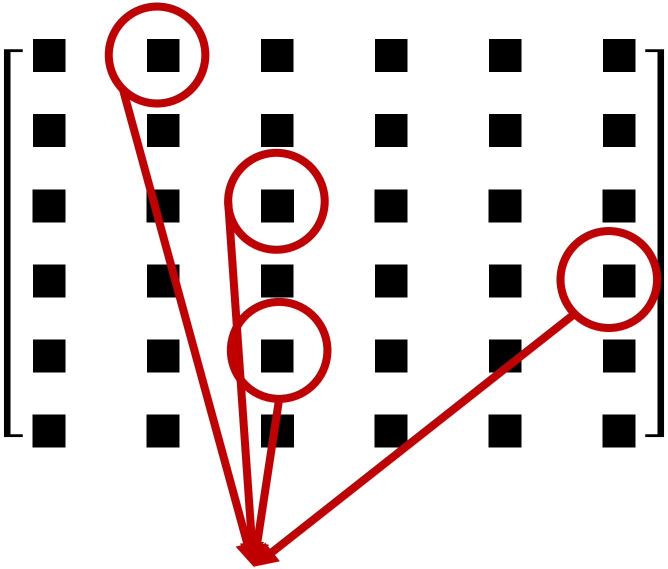
$$m \sim 1.5 N_r N_t$$

What if there is no enough probing budget? Probing budget < $N_r N_t$

Multiple antennas can be grouped as one “virtual” antenna.

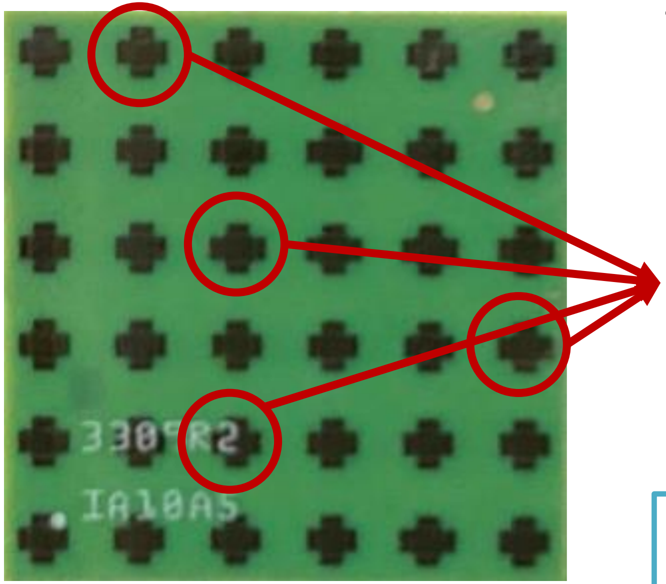


$\equiv$



The beamforming weights on these elements stay the same.
The elements of the channel matrix are assumed to be the same.

2ACE: Multi-resolution Channel Estimation



- By grouping N_t/N_r antennas into K groups, we recover a CSI matrix of size $\frac{N_t}{K} \times \frac{N_r}{K}$ instead.

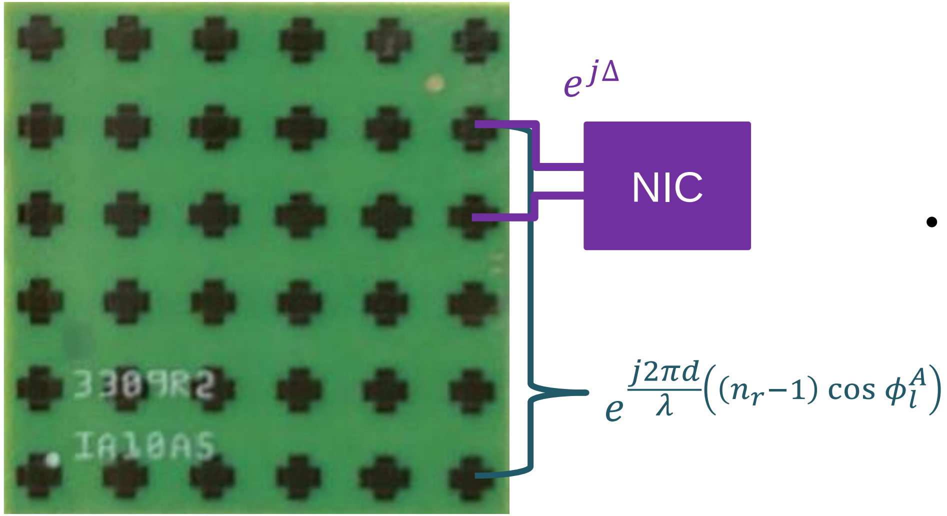
- Challenge: Minimize grouping error.

-- Selecting antennas with similar channels.

How to **identify antenna with similar channels without channel probing**?

2ACE: Multi-resolution Channel Estimation

- Phase offset comes from two parts:
  - Hardware offset due to differences in length of transmission line.
    - Calibrate through the method in M-cube [2]
  - Phase difference in array response.
    - Estimate through a rough angle estimation.

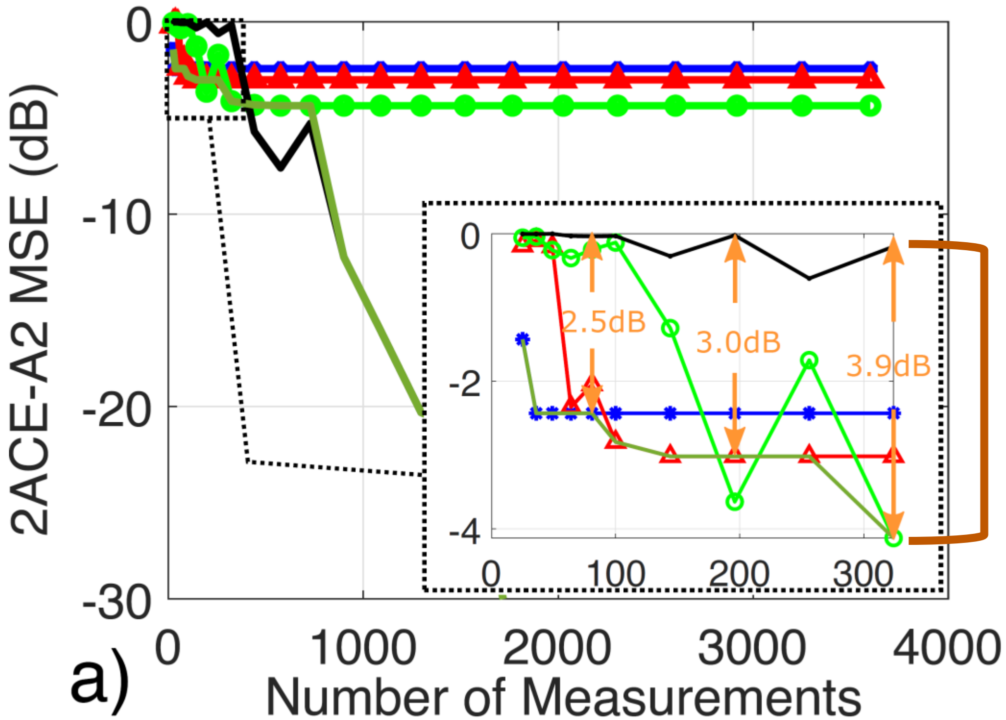


- We group the antennas with minimum sum of phase offset difference.

[2] Zhao, Renjie, et al. "M-cube: A millimeter-wave massive MIMO software radio." *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 2020.

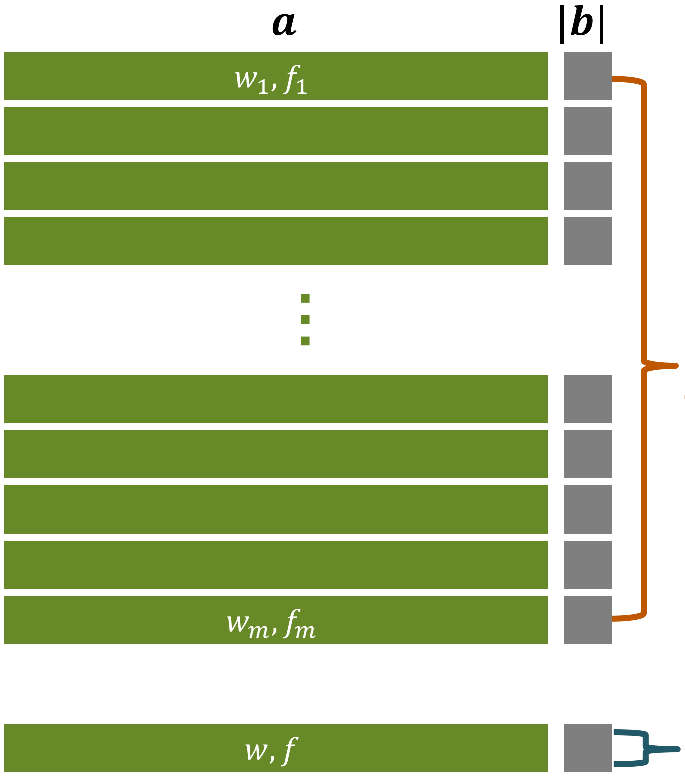
TEXAS The University of Texas at Austin Carnegie Mellon University Microsoft WHAT STARTS HERE CHANGES THE WORLD

Effectiveness of Multi-resolution (Simulation)



Multi-resolution improves 2ACE by **2.5 dB – 4 dB Mean Squared Error (MSE).**

2ACE: Confidence indicator



95% probes:
channel inference

Channel estimation

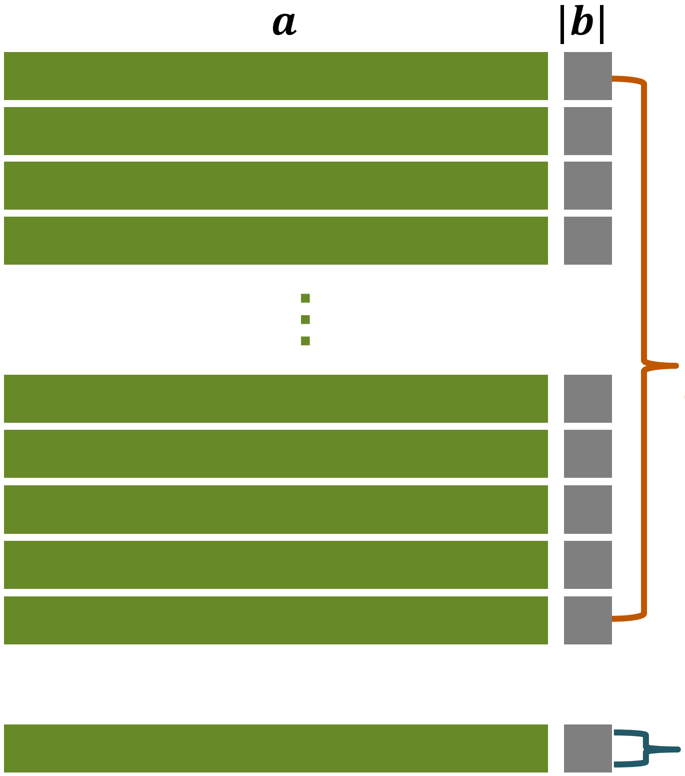
$\boldsymbol{H}$

$$|b| = |\boldsymbol{w}^T \boldsymbol{H} \boldsymbol{f} \gamma + \sigma|$$

5% probes:
validation

Verify with 5% probes,
Calculate confidence as 1-error.

2ACE: Confidence indicator - High confidence



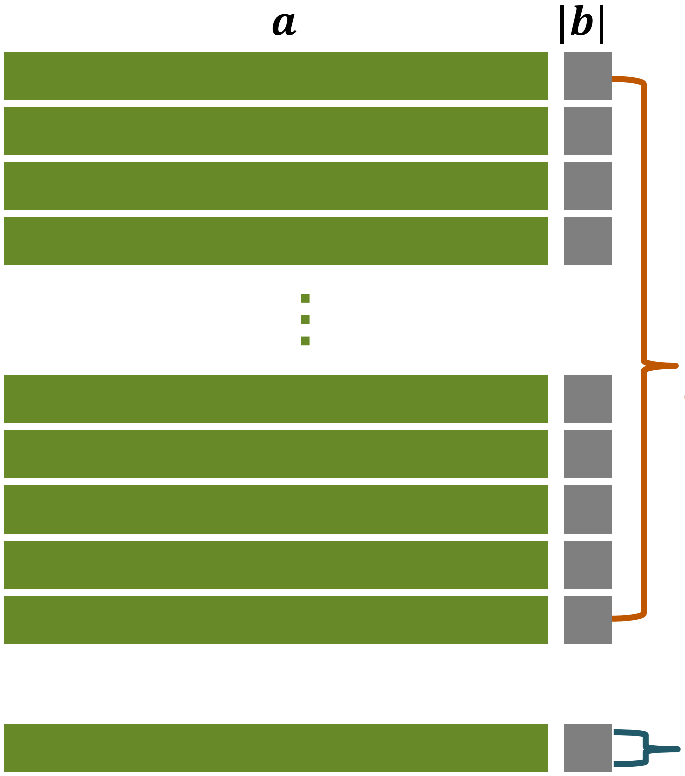
95% measurement:
channel inference

5% measurement:
validation

Adapt estimation,
Use all data to refine result.

Obtain more accurate channel H

2ACE: Confidence indicator - Low confidence



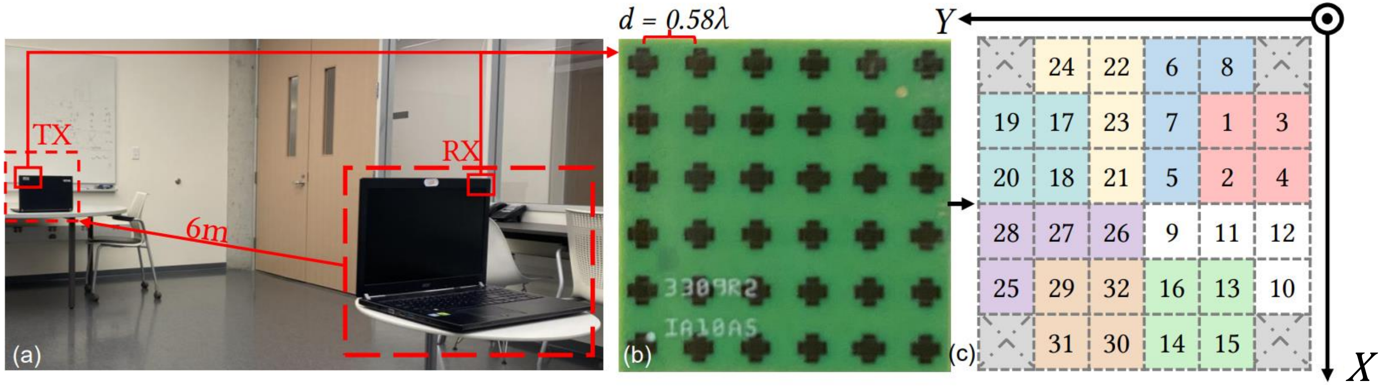
95% measurement:
channel inference

5% measurement:
validation

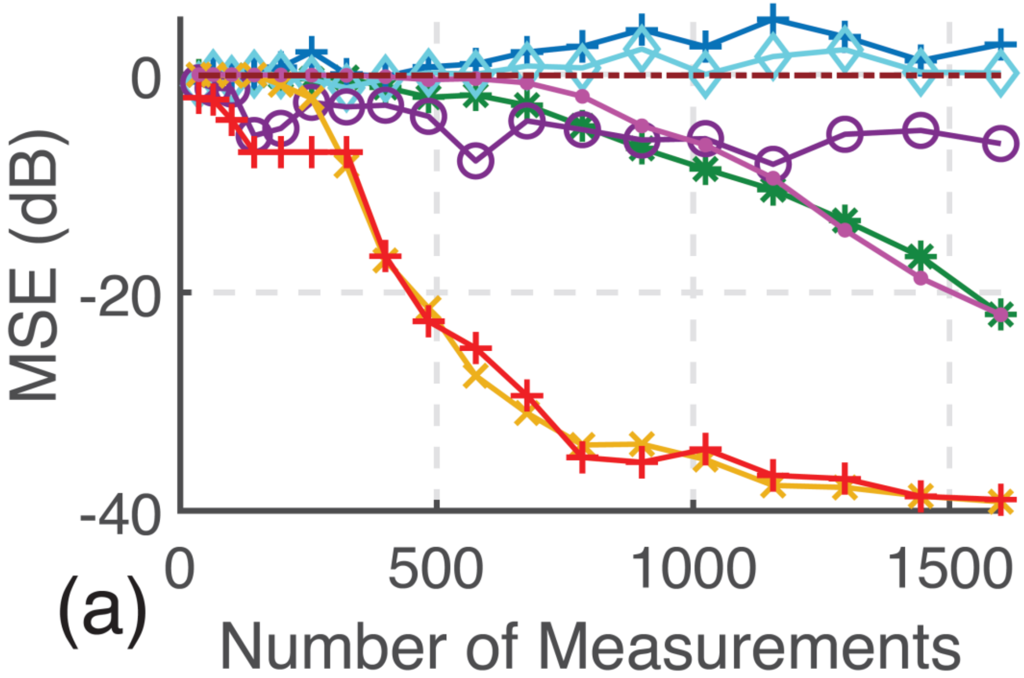
- Case (a): Not enough measurements.
  - Query for more measurement.
- Case (b): Enough measurements but channel changes.
  - Drop past measurements,
  - Then query for more measurements

Evaluation

- Simulation
  - Synthesize CSI matrix using multipath model.
  - Generate CSI matrix using Wireless Insite ray-tracing.
- Testbed
  - 2 laptops with Qualcomm QCA6320-based Baseband NIC
  - QCA6210-based 32-element antenna array.



CSI Estimation – NMSE (Simulation)

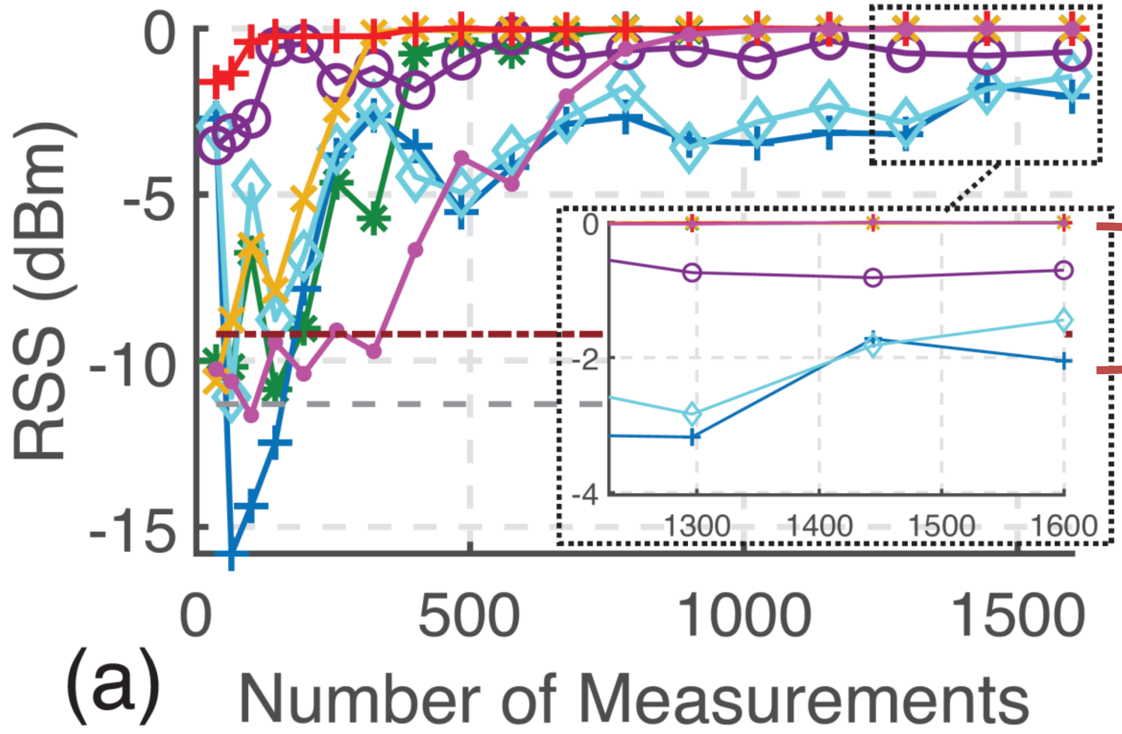


PLOMP and **PLGAMP** suffers from **over-fitting**, as reported.

PhaseLift and **Nuclear** converges **much slower**.

2ACE w/ Multi-resolution performs **optimally across baselines**.

Beamforming – RSS (Simulation)



2ACE w/ Multi-resolution

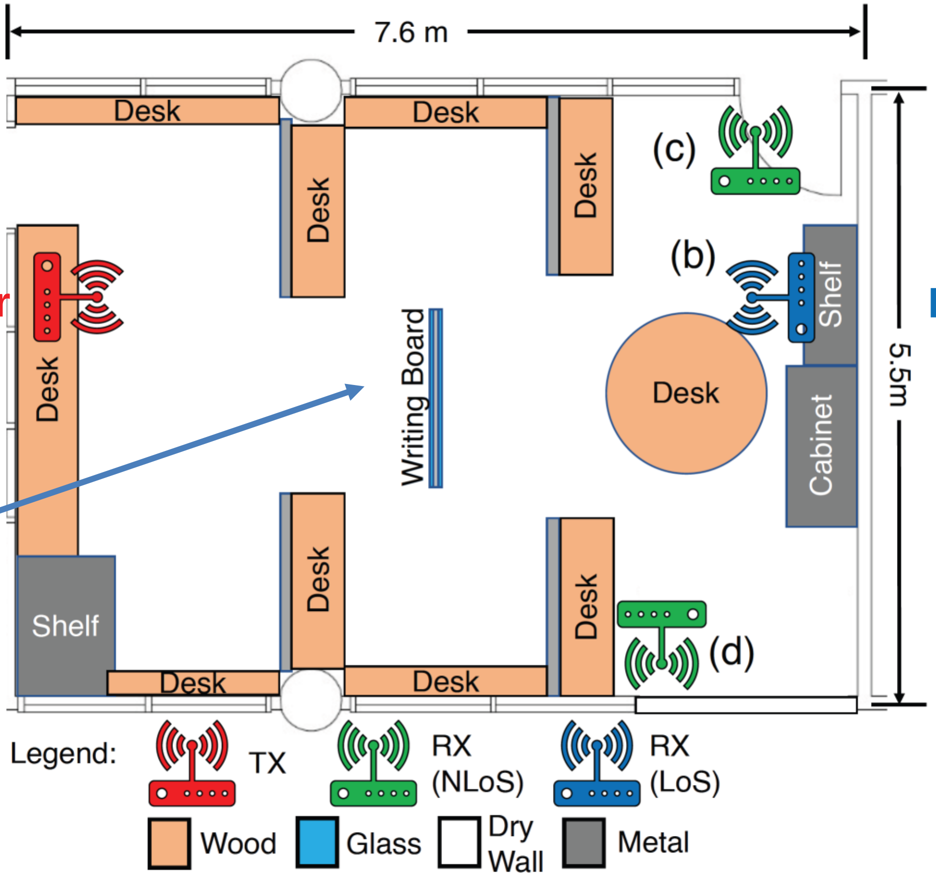
gives **2-3 dB increment** compared to baselines.

We evaluate beamforming RSS in an indoor environment.

Transmitter

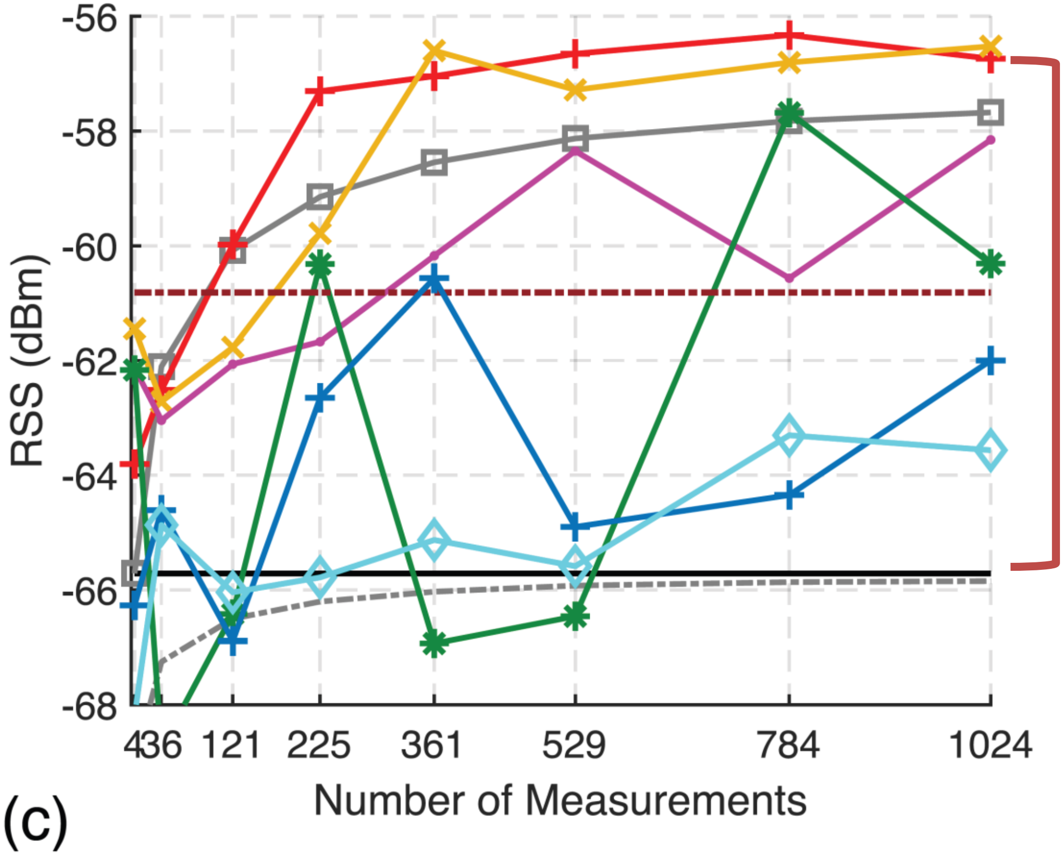
Writing board blocks line-of-sight

Desks, walls, shelves act as reflectors.



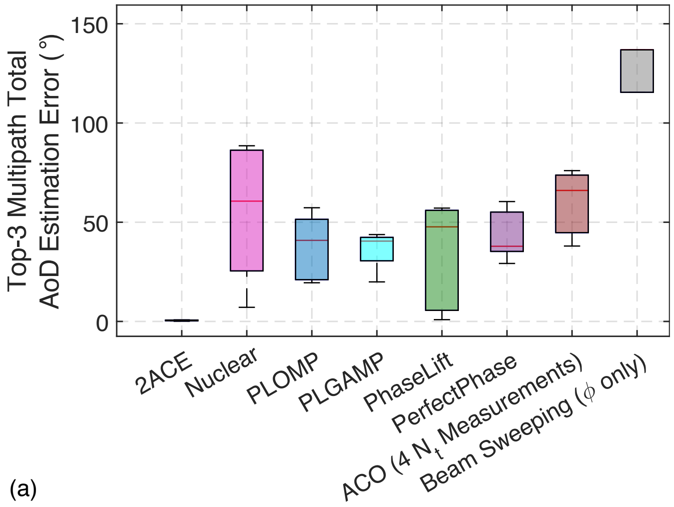
Receiver

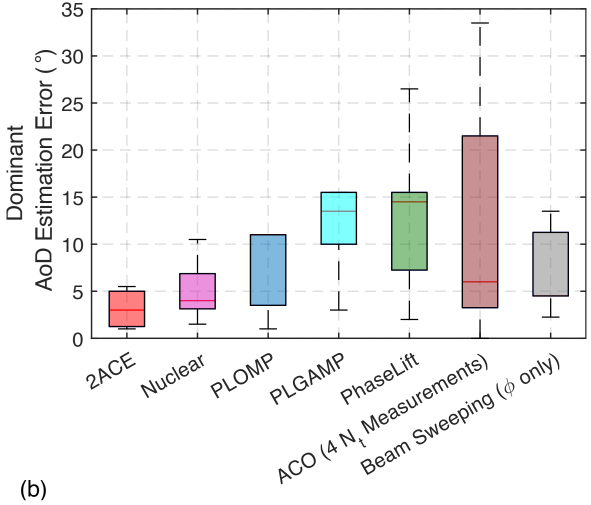
Beamforming – RSS (Testbed)



2ACE w/ Multi-resolution gives **1-9 dB increment** compared to different baselines.

Sensing – AoD Estimation





Simulated Top-3 AoD Estimation Error is **nearly 0**.

Testbed dominant AoD Estimation Error is average **2.5 degree**.

Conclusion & Discussion

- Propose **spectral profile** to drive channel estimation
  - Spectral profile **can also be applied to other domains** besides channel estimation.
- Various optimization techniques for **accelerating convergence.**
- **Multi-resolution** for low measurement budget.
  - Multi-resolution **can also be used for other compressive sensing algorithms.**
- Simulation and testbed experiments show optimality on **channel estimation**, **beamforming gain** and **angle estimation**.

Acknowledgement

This work is supported in part by NSF Grant CNS-2008824 and CNS-2107037. We appreciate the insightful feedback from ACM MobiHoc 2023 anonymous reviewers.

Ethical Concern

The personnel involved in the experiment are fully insured and paid. No personally identifiable information (PII) was collected during the exploration. This work does not raise any ethical concern.

Authors




Yiwen Song
Carnegie Mellon University

Changhan Ge
UT Austin

Lili Qiu
UT Austin &
Microsoft Research Asia

Yin Zhang

The University of Texas at Austin
Department of Computer Science
College of Natural Sciences

Carnegie Mellon University
Electrical & Computer Engineering

Microsoft

Thank You!

Questions?

Beam-training on commercial 802.11ad/ay devices

- Override Rx-side beam-training, i.e., Rx uses a quasi-omnidirectional beam.
- AP performs sector-level sweeping (SLS) and selects the precoder yielding strongest received signal strength (RSS)
- Pros: Simple and fast
- Cons:
  - Coarse and not optimal – impossible to exhaustively try codebooks
  - No CSI estimation – Only reports RSS

60 GHz WiFi
Access Point (AP)

60 GHz WiFi
Station (STA)

Channel Estimation Problem

- Recall that the received signal can be formulated as

known variables: precoders and combiners

RSS $|b|$ measured
& fed back by Rx

$$b = \boldsymbol{w}^T \boldsymbol{H} \boldsymbol{f} \gamma + \sigma$$

Variable needs to recover: CSI Matrix

- Hence, by vectorizing $\boldsymbol{H}$ as $\boldsymbol{x}$ and define $\boldsymbol{a} = \mathrm{w} \otimes f$ (Kronecker Product),
we formulate channel estimation problem as

Known $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m]$ and corresponding $\boldsymbol{b} = [|b_1|, |b_2|, \cdots, |b_m|]$
recover $\boldsymbol{x}$ that $\min_{x} \left||\boldsymbol{Ax}| - \boldsymbol{b}\right|^2$

Channel Estimation Methods – ACO & PhaseLift

- PhaseLift [2]: Compressive sensing-based recovery。
  - Pros: Relatively accurate given enough measurements.
  - Cons:
    - Large measurement overhead - takes $\geq 4N_t N_r$ measurements.
    - Computationally heavy – long algorithm running time.
    - Sharp phase transition – arbitrarily bad estimation given to few measurements.

- Adaptive Codebook Optimization (ACO) [3]: Leverage signal property
  - Pros: medium overhead $4(N_t + N_r)$, relative accurate given simple environment
  - Cons:
    - Requires a special codebook – a few bad probe is fatal.
    - Low resolution – the channel recovered is either $\boldsymbol{w}^T\boldsymbol{H}$ or $\boldsymbol{H}\boldsymbol{f}$, which is rank 1.

[3] Candes, Emmanuel J., Thomas Strohmer, and Vladislav Voroninski. "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming." Communications on Pure and Applied Mathematics 66.8 (2013): 1241-1274.
[4] Palacios, Joan, et al. "Adaptive codebook optimization for beam training on off-the-shelf IEEE 802.11 ad devices." Proceedings of the 24th Annual International Conference on Mobile Computing and Networking. 2018.

Channel Estimation Methods – PLOMP & PLGAMP

- PLOMP & PLGAMP [4]: Find the complex channel from the dominant AoAs and AoDs.

$$H_{n_t,n_r} = \sum_{l=1}^{\overbrace{L}^{L \text{ paths}}} \underbrace{h_l}_{\text{Complex response}} \cdot \overbrace{e^{\frac{j2\pi d}{\lambda}\left((n_t-1)\cos\phi_l^D + (n_r-1)\cos\phi_l^A\right)}}^{\text{Array response (Azimuth AoD } \phi_l^D \text{ and AoA } \phi_l^A)} \underbrace{e^{j\Delta}}_{\text{Hardware offset}}$$

- Pros: Fast convergence in certain low-rank scenarios.
- Cons:
  - Needs hardware-offset calibration – need to measure $e^{j\Delta}$ first.
  - Fail when CSI matrix is not low-rank - assume L is very small but this is not always true.
  - Fail if Tx and Rx not on the same elevation - Only models signals on azimuth plane.
  - Still computationally heavy – use PhaseLift as the first step.

[5] Zhang, Yi, et al. "Side-information-aided noncoherent beam alignment design for millimeter wave systems." Proceedings of the Twentieth ACM International Symposium on Mobile Ad Hoc Networking and Computing. 2019.