UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

ANTS LAB

# Over-the-Air Computation Systems: Optimization, Analysis and Scaling Laws

Wanchun Liu, et al

Presented by Chenpei Huang

Published in IEEE Transactions on Wireless Communications

# Outline

- Motivation
- AirComp and Federated Learning
- Computational-optimal AirComp
- Scaling Performance and conclusion

# Motivation

Over-the-air consensus

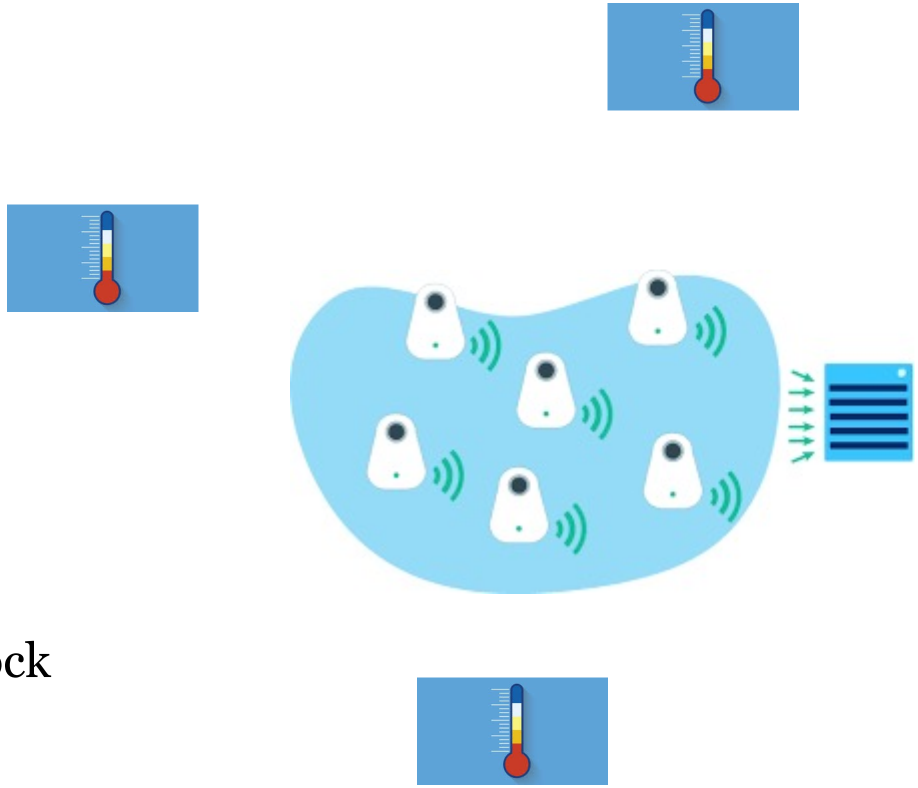

Wireless distributed machine learning

# Motivation – Over-the-air Consensus

*"Only the summary of data is of interest…"*

- Average
- Sum
- Min or Max
- …

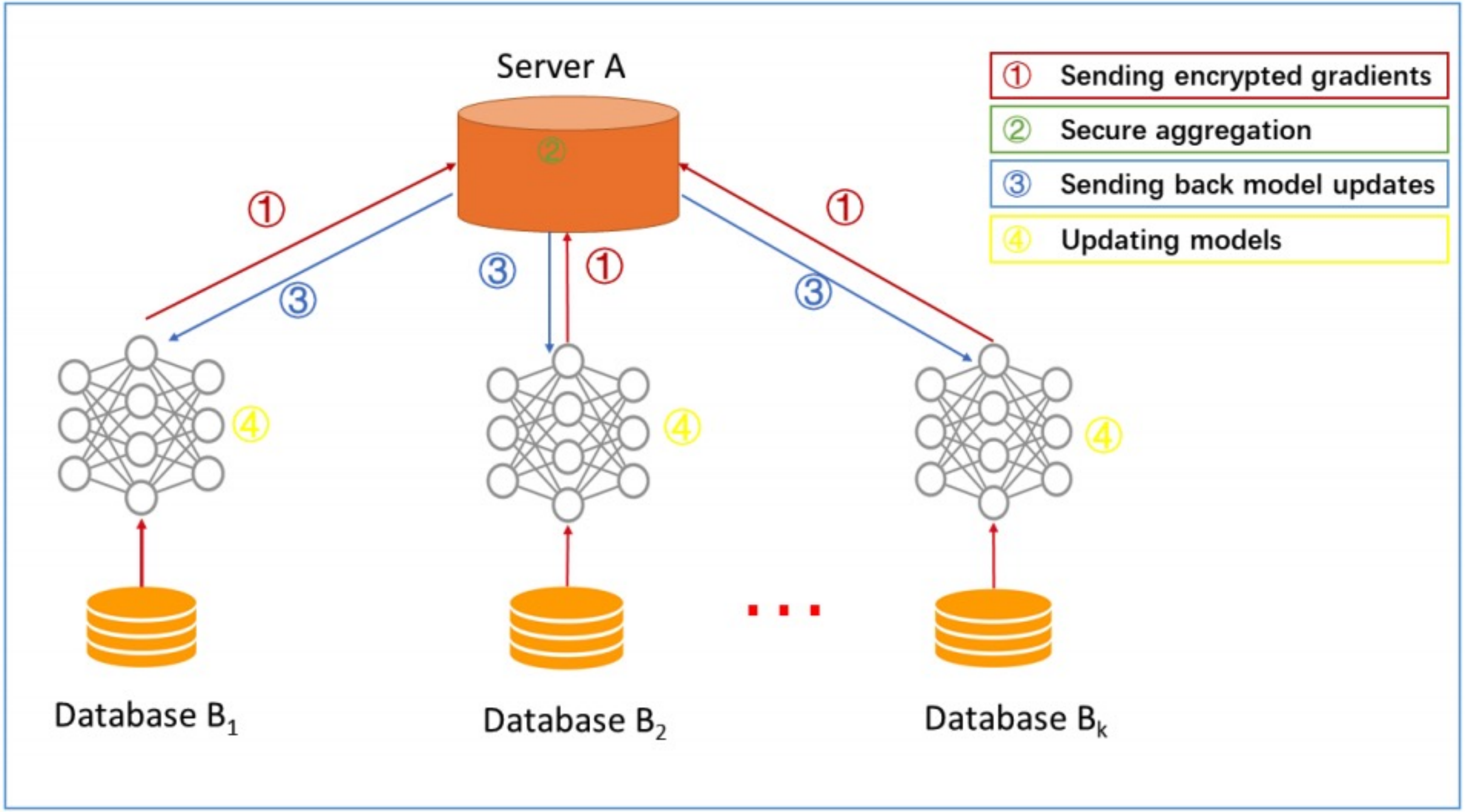
Abari, O., Rahul, H., & Katabi, D. (2016). *Over-the-air Function Computation in Sensor Networks*. http://arxiv.org/abs/1612.02307



Option 1: transmit all the data to the cluster head

Option 2: Jointly transmit using the same time-frequency block

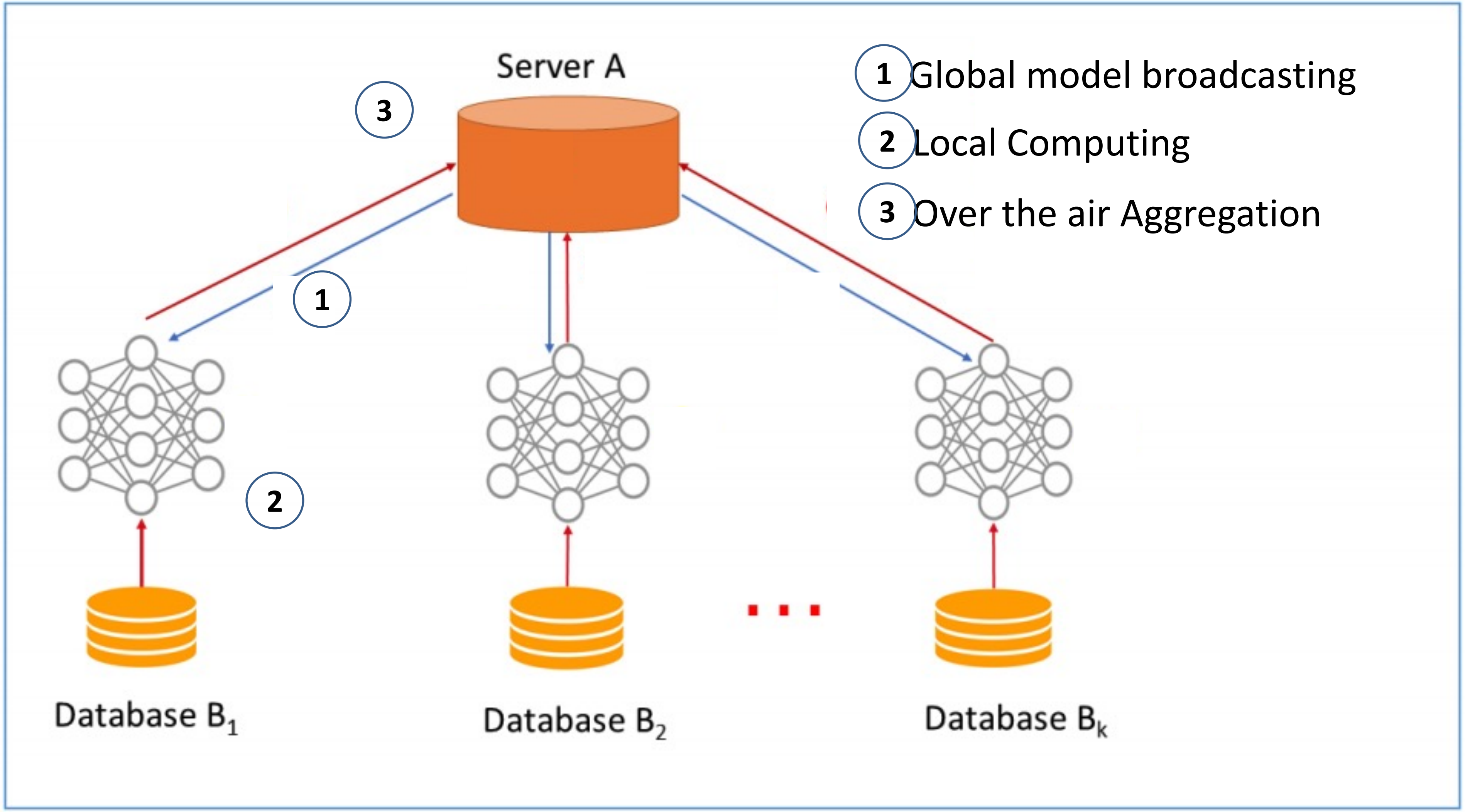
# Motivation – Over-the-air Federated Learning




**Communication Bottleneck:**

limited radio resources

# Motivation – Over-the-air Federated Learning




**Aggregate the gradients over the air can dramatically reduce the required resources**

# Outline

- Motivation
- AirComp and Federated Learning
- Computational-optimal AirComp
- Scaling Performance and conclusion

# AirComp and FL – Communication model

## OFDM modulation

- Bandwidth B into M orthogonal sub-channels
- Fixed digital constellation is employed by all the devices
- Map each gradient element to one digital symbol

## Channel Input Vector

Device index

The size of the model

$$\tilde{\mathbf{g}}_k = [\tilde{g}_{k,1}, \cdots, \tilde{g}_{k,q}]^T$$

Zhu, Guangxu, et al. "One-bit over-the-air aggregation for communication-efficient federated edge learning: Design and convergence analysis." *IEEE Transactions on Wireless Communications* 20.3 (2020): 2120-2135.

# AirComp and FL – Communication model

## Channel Input Vector

$$\tilde{\mathbf{g}}_k = [\tilde{g}_{k,1}, \ldots, \tilde{g}_{k,q}]^T$$

We require

$N_s = \frac{q}{M}$ OFDM symbols

## Over the Air Process

i-th aggregated gradient parameter, received at the m-th sub-carrier and t-th OFDM symbol is given by:

$$\tilde{g}_i = \sum_{k=1}^{K} h_k[t,m] p_k[t,m] \tilde{g}_{k,i} + z[t,m], \qquad \forall i,$$

Channel coefficients

Power Control policy

AWGN

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# AirComp and FL – Communication model

## Transmitter Design

one-bit quantization of local gradient estimates

$$\text{(One-bit quantization)} \qquad \tilde{g}_{k,i} = \mathsf{sign}(g_{k,i}), \quad \forall k, i.$$

Each gradient parameters is modulated into one binary phase shift keying (BPSK) symbol.

## Channel Inversion

We adopt the power control to invert the sub-channels so that gradient transmitted by different devices are received with identical amplitudes

$$\tilde{g}_i = \sum_{k=1}^{K} h_k[t,m] p_k[t,m] \tilde{g}_{k,i} + z[t,m], \qquad \forall i,$$

ANTS LAB



UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# AirComp and FL – Communication model

## Receiver Design

$$\text{(Over-the-air aggregation)}\ \tilde{\mathbf{g}}[t] = \sum_{k=1}^{K} \sqrt{\rho_0}\,\tilde{\mathbf{g}}_k^{(\mathsf{Tr})}[t] + \mathbf{z}[t]$$

Vector with M elements

Truncated gradients

$$\tilde{\mathbf{g}} = \left[\tilde{\mathbf{g}}[1]^T, \tilde{\mathbf{g}}[2]^T, \ldots, \tilde{\mathbf{g}}[N_s]^T\right]^T$$

$$\text{(Majority-vote based decoder)} \qquad \mathbf{v} = \mathrm{sign}(\tilde{\mathbf{g}})$$

# Outline

- Motivation
- AirComp and Federated Learning
- Computational-optimal AirComp
- Scaling Performance and conclusion

ANTS LAB

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# Computational Optimal AirComp

**Related works**

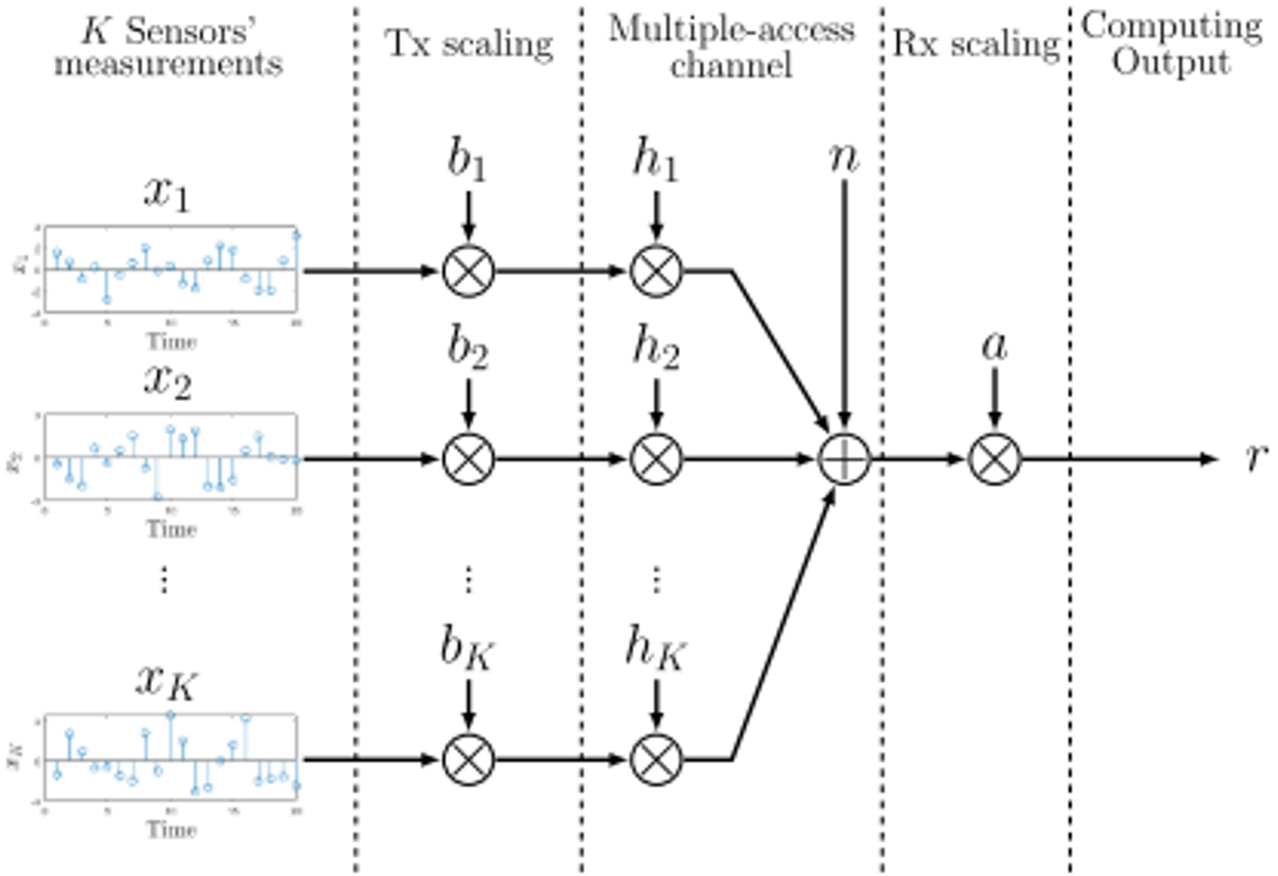
- Processing functions design

  Compute a given function, ideal channel

- Performance analysis on practical wireless MAC

  Compute the sum, non-ideal channel

# Computational Optimal AirComp – System Model




Function

$$r = a\left(\sum_{k=1}^{K} h_k b_k x_k + n\right),$$

Distortion

$$\text{MSE} = \sum_{k=1}^{K} |a h_k b_k - 1|^2 + \sigma^2 |a|^2.$$

Power consumption (bound)

$$\text{PW} \triangleq \sum_{k=1}^{K} |b_k|^2.$$

ANTS LAB



UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# Computational Optimal AirComp

*"Computational-optimal: minimize MSE under power constraint"*

Tx-scaling devotes power to improve SNR, while Rx-scaling (post-processing) amplifies noise without sacrificing power.

Assume the channel coefficients are known by both the transmitters and receiver.

Pre-processed signals are assumed to have zero mean and unit variance, and be independent.

ANTS LAB



UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# Computational Optimal AirComp

## Solve the problem

$$\min_{a,\{b_k\}} \quad \text{MSE}$$
$$\text{subject to } |b_k|^2 \leq P, \quad \forall k.$$

$$\text{MSE} = \sum_{k=1}^{K} |a h_k b_k - 1|^2 + \sigma^2 |a|^2.$$

Step 1: sort the channel coefficient ascendingly, and fix a

$$0 \triangleq h_0 < h_1 \leq h_2 \leq \cdots$$

$$\mathcal{S}_k \triangleq \begin{cases} \left(\frac{1}{h_1\sqrt{P}}, \infty\right), & k = 0, \\ \left(\frac{1}{h_{k+1}\sqrt{P}}, \frac{1}{h_k\sqrt{P}}\right], & k = 1, \cdots, K-1, \\ \left[0, \frac{1}{h_K\sqrt{P}}\right], & k = K. \end{cases}$$

$$b_k = \begin{cases} \sqrt{P}, & 1 \leq k \leq i \\ \frac{1}{a h_k}, & i < k \leq K. \end{cases}$$

Full power transfer

Channel inversion

# Computational Optimal AirComp

Step 2: solve optimal {a}, or equivalently, the critical number i

$$\text{MSE} = \sum_{k=1}^{i} \left| a h_k \sqrt{P} - 1 \right|^2 + \sigma^2 |a|^2, \quad a \in \mathcal{S}_i,$$

$$g_i \triangleq \begin{cases} 0, & i = 0 \\ \dfrac{\sqrt{P} \sum_{k=1}^{i} h_k}{\sigma^2 + P \sum_{k=1}^{i} h_k^2}, & i = 1, \cdots, K. \end{cases}$$

$$i^\star = \arg \max_{1 \le i \le K} g_i.$$

Minimal MSE and its power consumption

$$\text{MSE}^\star = \sum^{i^\star} \left( a_{i^\star} h_k \sqrt{P} - 1 \right)^2 + \sigma^2 (a_{i^\star})^2.$$

$$\text{PW} = \sum_{k=1}^{K} b_k^{\star 2} = P i^\star + \frac{1}{(a_{i^\star})^2} \sum_{k=i^\star+1}^{K} \frac{1}{h_k^2}.$$

Full power transfer

Channel inversion

# Outline

- Motivation
- AirComp and Federated Learning
- Computational-optimal AirComp
- Scaling Performance and Conclusion

ANTS LAB



UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# Scaling Performance – Benchmark Policies

Channel-Inversion Policy: set the critical number $i = 1$ (i.e., force to channel inversion)

Energy-Greedy Policy: set the critical number $i = K$ (i.e., force to full power)

# Scaling Performance – a Suboptimal Policy

Define:

Policy is computation-effective iff

$$\lim_{K\to\infty} \frac{\mathsf{E}[\mathsf{MSE}]}{K} = 0.$$

Policy is energy-efficient iff

$$\lim_{K\to\infty} \frac{\mathsf{E}[\mathsf{PW}]}{K} = 0.$$

*Definition 6 (First-$\imath$ Policy):* A Tx-Rx scaling policy of the AirComp system is a first-$\imath$ policy if it satisfies:

i) the critical number is determined by a function, i.e., $\imath = \imath(K)$, where $\imath : \mathbb{N} \to \mathbb{N}$ and $\imath(K) \leq K$,

ii) the Rx-scaling factor $a \in \mathcal{S}_\imath$, and

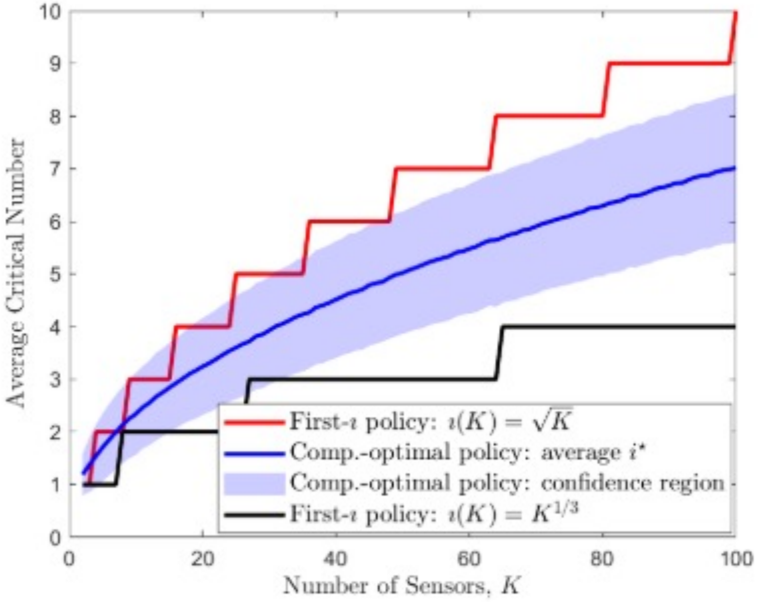
iii) the Tx-scaling factor $b_k$ is given by (8), $\forall k \in \{1, \cdots, K\}$.

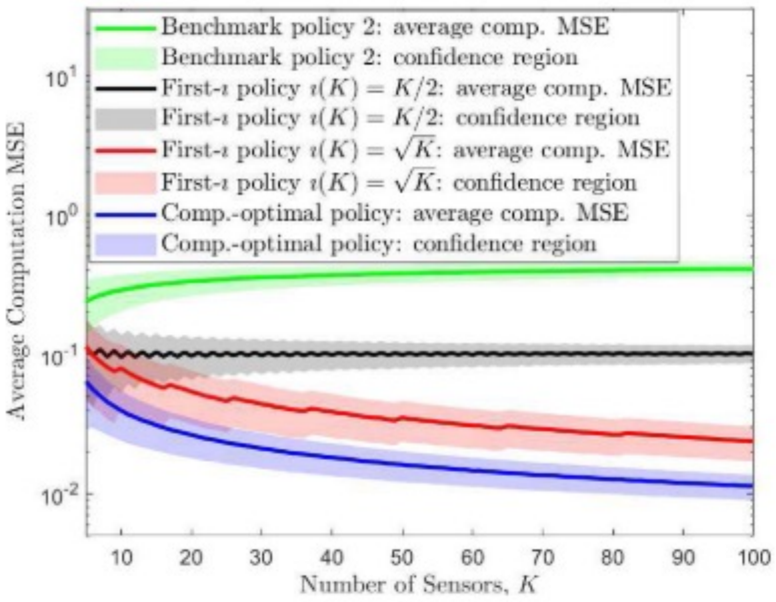
Does not depend on channel realizations!

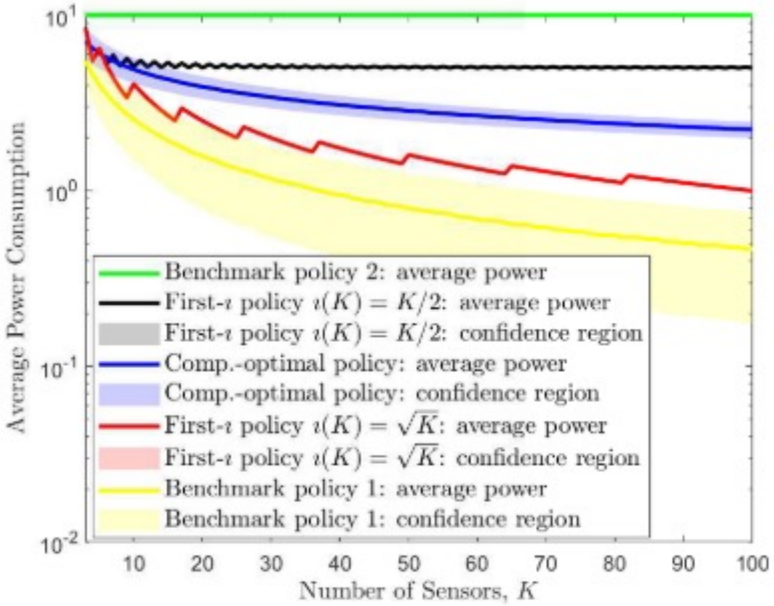
# Numerical Result

## Simulation parameters

Transmission power = 10
Receiving noise power = 1
Channel coefficients: i.i.d. Rayleigh fading with unit variance




ANTS LAB



UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

# Conclusion

COMPUTATION EFFECTIVENESS AND ENERGY EFFICIENCY OF AIRCOMP POLICIES

| | Computation-Effective Policy | Energy-Efficient Policy |
|---|---|---|
| Benchmark Policy 1 [12, 13] | ✗ | ✓ |
| Benchmark Policy 2 | ✗ | ✗ |
| Computation-Optimal Policy | ✓ | ✓ |
| First-$\imath$ Policy with $\imath(K) = \sqrt{K}$ | ✓ | ✓ |
| First-$\imath$ Policy with $\imath(K) = K/2$ | ✗ | ✗ |

Computation-optimal policy has a vanishing average MSE and a vanishing average power consumption with the increasing number of sensors

First-i policy (square root) reveals the tradeoff between computation effectiveness and energy efficiency, which is important in practical AirComp system design

# THANK YOU