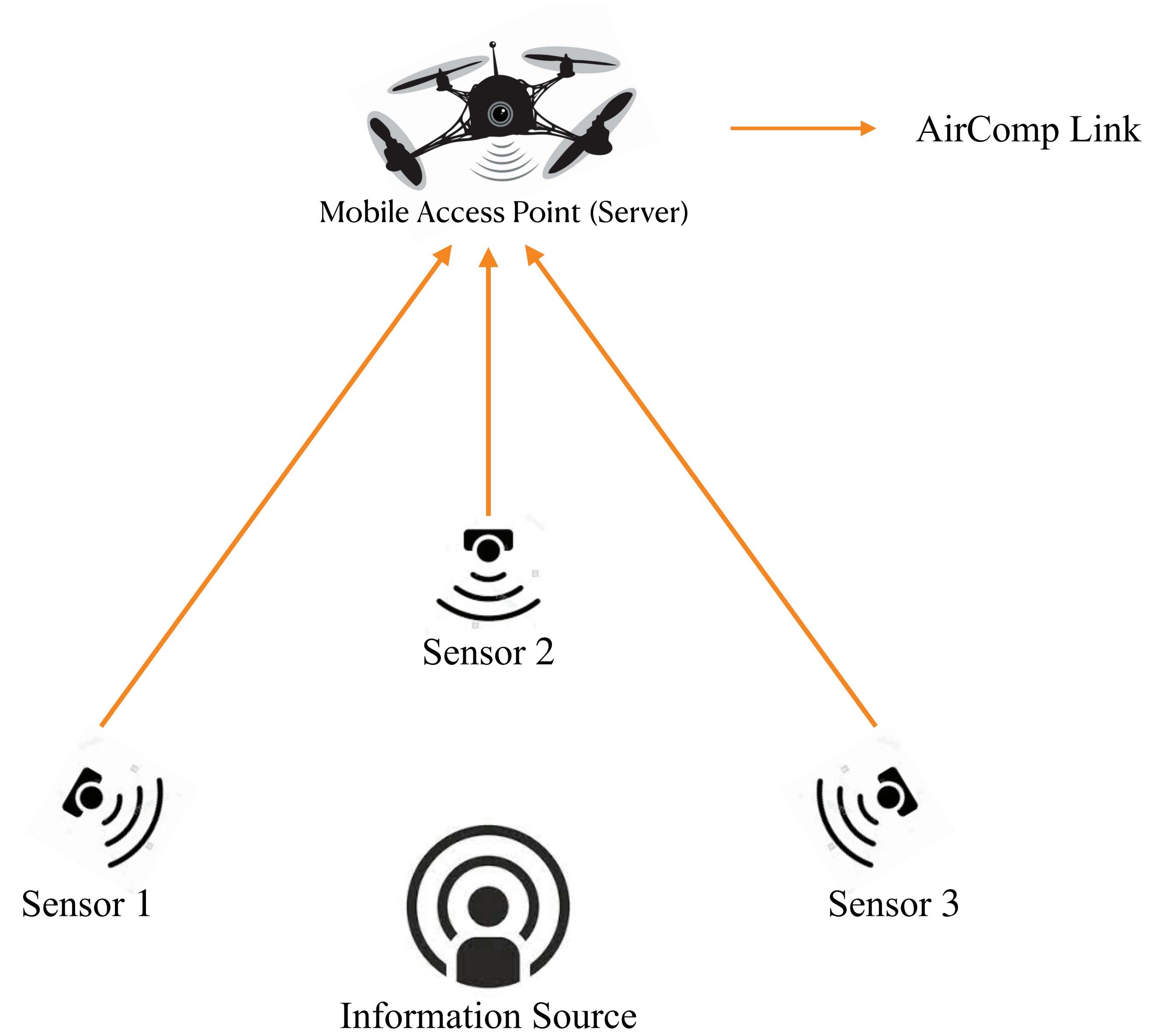
Background

- A Mobile Server
  - Real-time inference applications
  - Collect features from sensors
- Multiple Sensors
  - Sensing the same source
  - Each observation is a corrupted version of ground-true data [1],
- AirComp
  - Aggregate the features to average the corruptions.
- Metric: Discriminant Gain



[1] J.-J. Xiao, S. Cui, Z.-Q. Luo and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks", IEEE Trans. Signal Processing, vol. 54, no. 2, pp. 413-422, Feb. 2006.

System Model

- Network Model
  - One mobile server (access point) with N_r receive antennas
  - K sensors with each having $N_t \leq N_r$ transmit antennas
- Sensing Data (Features)
  - The sensed data (feature) vector, denoted as $\mathbf{x}$, has $M = 2N_t$ dimensions.
  - For the k-th sensor, its observed feature vector is $\mathbf{x}_k$, given as

$$\mathbf{x}_k = \mathbf{x} + \mathbf{d}_k,$$

    - $\mathbf{x}$ is the ground-true data (feature).
    - $\mathbf{d}_k$ is the observation distortion with the following distribution,

$$\mathbf{d}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{D}_k\right).$$

      - $D_k = \text{diag}\{\delta_{k,1}^2, \delta_{k,2}^2, \ldots, \delta_{k,M}^2\}$ is the diagonal covariance matrix.

Discriminant Gain Model

- The ground-true feature $\mathbf{x}$ has the following distribution

$$\mathbf{x} \sim \sum_{\ell=1}^{L} \mathcal{N}\left(\boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell\right),$$

  - L is the total number of classes.
  - $\boldsymbol{\mu}_\ell = \left[\mu_{\ell,1}, \mu_{\ell,2}, \ldots, \mu_{\ell,M}\right]^T$, is the mean vector of the $\ell-$th class.
  - $\boldsymbol{\Sigma}_\ell = \mathrm{diag}\left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2\right\}$, is the covariance matrix.
    - $\boldsymbol{\Sigma}$ is a diagonal matrix, as PCA is used to pre-process the features and different feature elements are independent.
    - All classes are assumed to have the same variance.

Discriminant Gain Model

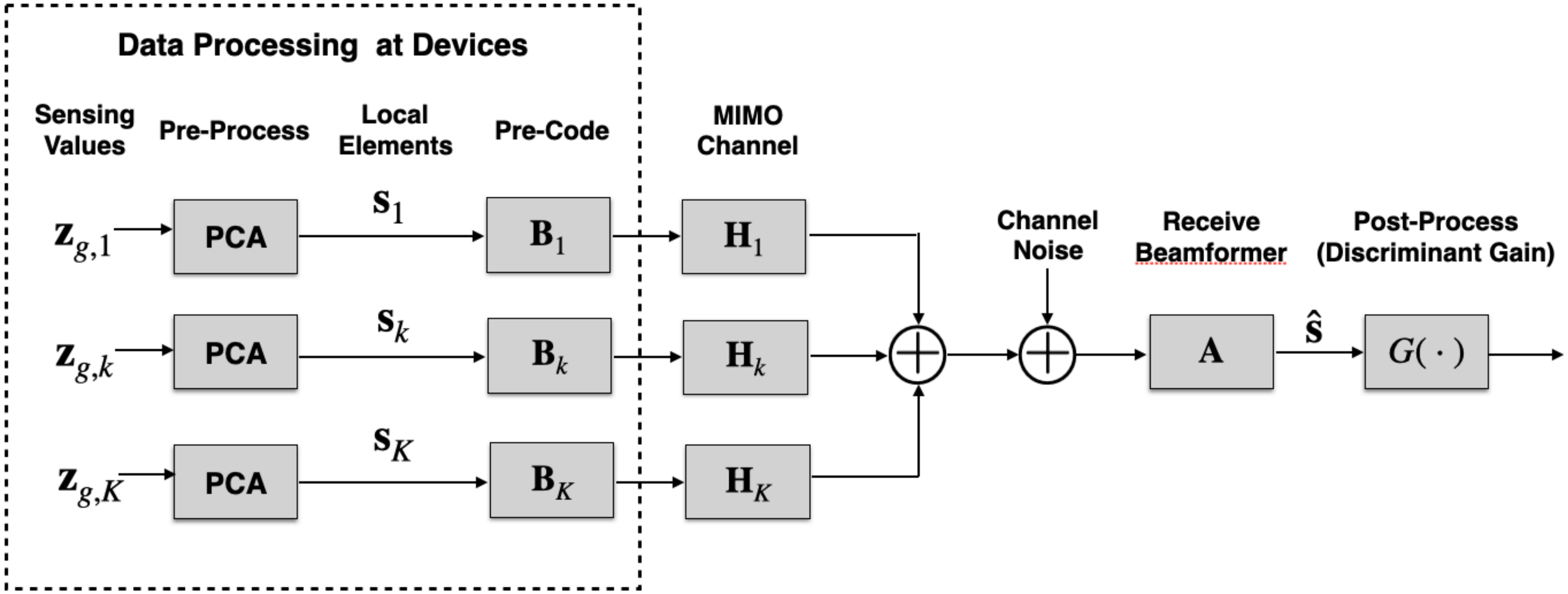
- Discriminant Gain
  - Consider the $m-$th feature dimension, its discriminant gain is

$$G(\mathbf{x}_m) = \frac{2}{L(L-1)} \sum_{\ell'=1}^{L} \sum_{\ell<\ell'} \frac{\left(\mu_{\ell,m} - \mu_{\ell',m}\right)^2}{\sigma_m^2},$$

  - The total discriminant gain can be written as

$$G(\mathbf{x}) = \sum_{m=1}^{M} G(\mathbf{x}_m).$$

AirComp Model



- Transmit Symbols $\{\mathbf{s}_k\}$

$$\mathbf{s}_k = [s_{k,1}, s_{k,2}, \ldots, s_{k,M}]^T,$$

where

$$s_{k,i} = x_{k,(2i-1)} + jx_{k,2i}, \quad 1 \leq i \leq N_t.$$

  - Note: The order of the elements can be re-scheduled.

AirComp Model

- At the server, the received signal can be written as

$$\hat{\mathbf{s}} = \mathbf{A} \sum_{k=1}^{K} \mathbf{H}_k \mathbf{B}_k \mathbf{s}_k + \mathbf{n},$$

  - $\mathbf{s}_k \in \mathbb{C}^{N_t}$ is the transmit symbol,
  - $\mathbf{B}_k \in \mathbb{C}^{N_t \times N_t}$ is the pre-coding matrix of sensor k,
  - $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix of sensor k,
  - $\mathbf{A} \in \mathbb{C}^{N_t \times N_r}$ is the receive beam-forming matrix at the receiver,
  - $\mathbf{n} \in \mathbb{C}^{N_r}$ is the noise with the following distribution,

$$\mathbf{n} \sim \mathcal{N}\left(\mathbf{0}, \delta_0^2 \mathbf{I}\right),$$

    - δ_0^2 is the noise variance.

Problem Formulation

- Received Feature Vector

$$\hat{\mathbf{x}} = \left[\mathrm{Re}(\hat{\mathbf{s}})^T, \mathrm{Im}(\hat{\mathbf{s}})^T\right]^T.$$

- Objective: Maximize the Receive Signal's Discriminant Gain

$$\max_{\mathbf{A},\{\mathbf{B}_k\}} \quad G(\hat{\mathbf{x}})$$

- Power Constraint of Each Sensor

$$\mathrm{tr}\left(\mathbf{B}_k\mathbf{B}_k^H\right) \leq P_k,$$

  - P_k is the maximum transmit power of sensor k.

Objective Simplification

✦ Zero-Forcing (ZF) Pre-Coding

$$\mathbf{A}\mathbf{H}_k\mathbf{B}_k\mathbf{x}_k = \mathbf{C}_k,$$

- $\mathbf{C}_k = \mathrm{diag}\{c_{k,1}, c_{k,2}, \ldots, c_{k,N_t}\},$
- $c_{k,i} \in \mathbb{R}$ is non-negative.

★ Based on ZF precodes, simplify the objective, i.e., discriminant gain.

★ Note: One symbol for two elements may be too difficult. Could revise the assumption that using one symbol for transmitting one element.