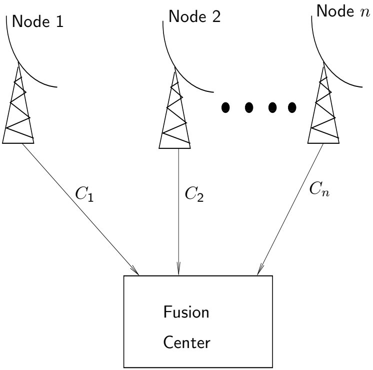
Distributed Statistical Inference using Type Based Random Access over Multi-access Fading Channels

Animashree Anandkumar and Lang Tong

Adaptive Communication and Signal Processing Group
School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853.

Supported by the MURI under Office of Naval Research and Army Research Laboratory CTA
03/22/2006.

Classical Distributed Inference

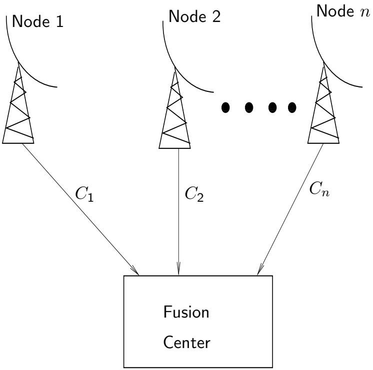


- Sensors : Sense physical phenomenon and transmit their local decisions.
- Fusion Center: Make inference on the phenomenon.
- Sensor-Fusion Center Communication Perfect (Error - free) with rate constraints.
- Typically in Radar communication.

Key Issues

- Quantization @ sensors.
- Inference @ fusion center.

Classical Distributed Inference

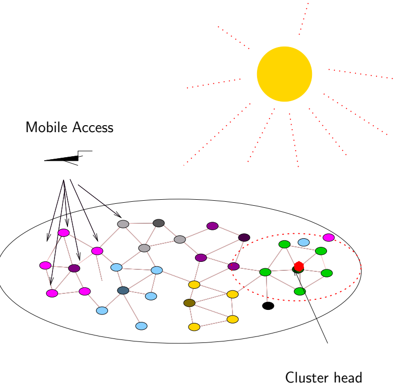


- Sensors : Sense physical phenomenon and transmit their local decisions.
- Fusion Center: Make inference on the phenomenon.
- Sensor-Fusion Center Communication
  Perfect (Error - free) with rate constraints.
- Typically in Radar communication.

Key Issues

- Quantization @ sensors.
- Inference @ fusion center.

Inference in Large Wireless Sensor Networks

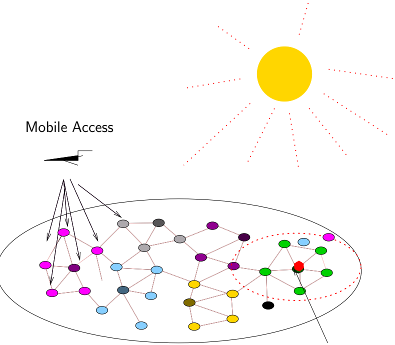


Characteristics

- Low Power and Low Rate Transmissions.
- Bandwidth Allocation.
- Multiaccess Channel with Fading.
- Energy Efficiency to prolong network life-time.
- Faulty, sleeping or poorly placed sensors.
- Deterministic scheduling (TDMA) may not be appropriate.

Medium Access Design is a key component.

Inference in Large Wireless Sensor Networks



Characteristics

- Low Power and Low Rate Transmissions.
- Bandwidth Allocation.
- Multiaccess Channel with Fading.
- Energy Efficiency to prolong network life-time.
- Faulty, sleeping or poorly placed sensors.
- Deterministic scheduling (TDMA) may not be appropriate.

Medium Access Design is a key component.

Random Access

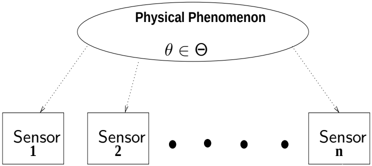
- Model : Random Number of Sensors in a data collection.
- Probabilistic Wake-up : Transmit based on a coin-flip.
- Transmit only Significant Data.
- Fusion center is a Mobile Access Point : collects data from different geographic regions.

Random Access

- Model : Random Number of Sensors in a data collection.
- Probabilistic Wake-up : Transmit based on a coin-flip.
- Transmit only Significant Data.
- Fusion center is a Mobile Access Point : collects data from different geographic regions.

Distributed Inference over Multi-Access Channels

Detection (Binary Hypothesis) and Estimation.



Random number of sensors per collection
N_i are IID with mean λ.

Sensor Quantization: $X_{i,j}$ quantized to M levels and Conditionally IID given θ

$$X_{ij} \sim \mathbf{p}_\theta = (p_\theta(1), \cdots, p_\theta(M))$$

Multi-access model

- Flat IID fading: $H_{i,j}$
- AWGN $W(t)$ with PSD $=\sigma^2$.

Inference at Fusion Center

- Neyman Pearson or Bayesian Detection.
- Maximum Likelihood Estimation.

Multiple collections: i–time index, j–sensor index.

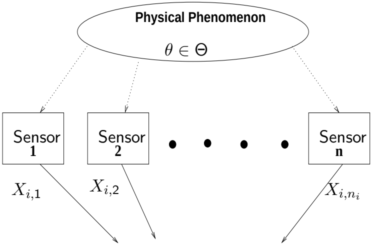
Distributed Inference over Multi-Access Channels

Detection (Binary Hypothesis) and Estimation.

Random number of sensors per collection
N_i are IID with mean λ.

Sensor Quantization: $X_{i,j}$ quantized to M levels and Conditionally IID given θ

$$X_{ij} \sim \mathbf{p}_\theta = (p_\theta(1), \cdots, p_\theta(M))$$



Multi-access model

- Flat IID fading: $H_{i,j}$
- AWGN $W(t)$ with PSD $=\sigma^2$.

Inference at Fusion Center

- Neyman Pearson or Bayesian Detection.
- Maximum Likelihood Estimation.

Multiple collections: i–time index, j–sensor index.

Distributed Inference over Multi-Access Channels

Detection (Binary Hypothesis) and Estimation.

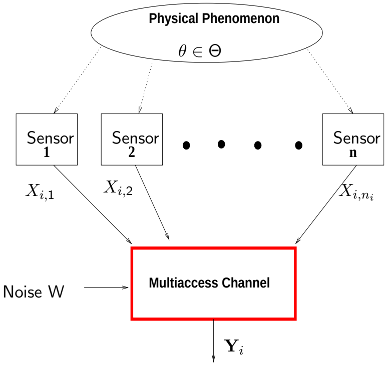
Random number of sensors per collection
N_i are IID with mean λ.

Sensor Quantization: $X_{i,j}$ quantized to M levels and Conditionally IID given θ

$$X_{ij} \sim \mathbf{p}_\theta = (p_\theta(1), \cdots, p_\theta(M))$$

Multi-access model

- Flat IID fading: $H_{i,j}$
- AWGN $W(t)$ with PSD $=\sigma^2$.



Inference at Fusion Center

- Neyman Pearson or Bayesian Detection.
- Maximum Likelihood Estimation.

Multiple collections: i–time index, j–sensor index.

Distributed Inference over Multi-Access Channels

Detection (Binary Hypothesis) and Estimation.

Random number of sensors per collection
N_i are IID with mean λ.

Sensor Quantization: $X_{i,j}$ quantized to M levels and Conditionally IID given θ

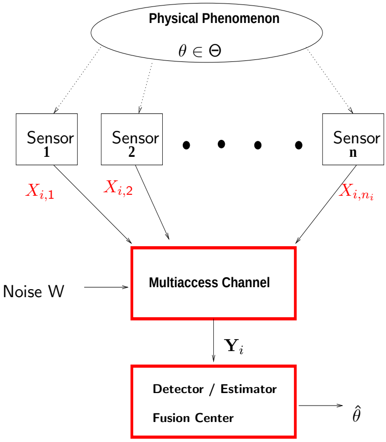
$$X_{ij} \sim \mathbf{p}_\theta = (p_\theta(1), \cdots, p_\theta(M))$$

Multi-access model

- Flat IID fading: $H_{i,j}$
- AWGN $W(t)$ with PSD $=\sigma^2$.

Inference at Fusion Center

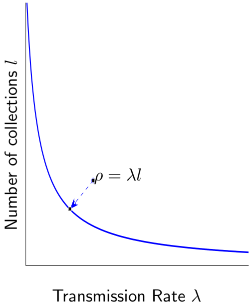
- Neyman Pearson or Bayesian Detection.
- Maximum Likelihood Estimation.



Multiple collections: i–time index, j–sensor index.

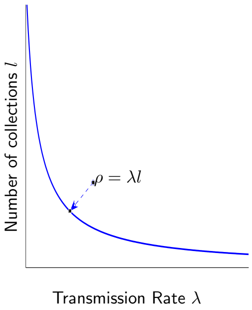
Spatio-Temporal Tradeoff

- Mean Transmitting Rate λ and Number of Data Collections l.
- Suppose we fix mean number of transmissions is $\rho \stackrel{\Delta}{=} \lambda l$, (proportional to energy budget).
- Should energy be allocated to simultaneous transmissions : large λ ?
- Or should we collect more data : large l ?
- Small λ : not enough sensors transmit , cannot counter noise.
- But, large λ : Less Observations as l is small.
- Role of multi access channel : Coherence or Cancelation ?



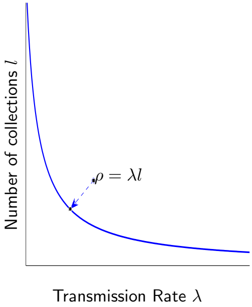
Spatio-Temporal Tradeoff

- Mean Transmitting Rate λ and Number of Data Collections l.
- Suppose we fix mean number of transmissions is $\rho \overset{\Delta}{=} \lambda l$, (proportional to energy budget).
- Should energy be allocated to simultaneous transmissions : large λ ?
- Or should we collect more data : large l ?
- Small λ : not enough sensors transmit , cannot counter noise.
- But, large λ : Less Observations as l is small.
- Role of multi access channel : Coherence or Cancelation ?



Spatio-Temporal Tradeoff

- Mean Transmitting Rate λ and Number of Data Collections l.
- Suppose we fix mean number of transmissions is $\rho \overset{\Delta}{=} \lambda l$, (proportional to energy budget).
- Should energy be allocated to simultaneous transmissions : large λ ?
- Or should we collect more data : large l ?
- Small λ : not enough sensors transmit , cannot counter noise.
- But, large λ : Less Observations as l is small.
- Role of multi access channel : Coherence or Cancelation ?



Fading Coherence Index γ

- Define Fading Coherence index as

$$\gamma = \frac{|\mathbb{E}(H)|^2}{\mathsf{Var}(H)}.$$

- Non Coherent ($\gamma = 0$) : Uniform phase uncertainty (*e.g.*,Rayleigh.)
- Perfectly Coherent ($\gamma = \infty$) : Deterministic Channel or no fading.

Outline

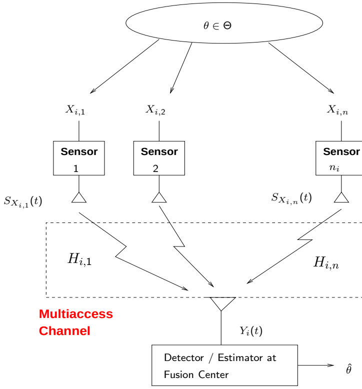
Outline

Type Based Random Access

- Signal Waveform
  $S_1(t), \ldots, S_M(t)$—a pre-determined set of M orthogonal waveforms with energy constraint $\mathcal{E}$.

- Sensor Encoding
  Quantized Data $X_{i,j} = x$ is encoded to waveform $S_x(t)$

$$S_{i,j}(t; x) = S_x(t)$$



- Observation @ FC:

$$Y_i(t) = \sum_{j=1}^{N_i} H_{i,j} \sqrt{\mathcal{E}} S_{X_{i,j}}(t - \tau_{i,j}) + W_i(t).$$

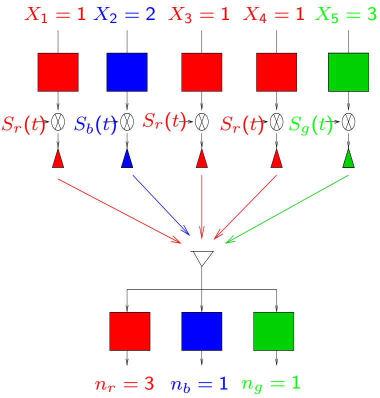
- Narrow band signal assumption :
  $S_{i,j}(t - \tau_{i,j}) \approx S_{i,j}(t)$.

- Data-centric in contrast to user-centric schemes (TDMA, FDMA or CDMA).

Key: MAC adds transmissions with same data level

Matched Filter Output

Schematic of TBRA



Ideal Conditions with $\lambda = 5$

Matched Filtering

$$\begin{aligned} \mathbf{Y}_i &\triangleq \frac{1}{\sqrt{\mathcal{E}}}\Big[\langle Y_i(t), S_1(t)\rangle, \cdots, \langle Y_i(t), S_M(t)\rangle\Big] \\ &= \sum_{j=1}^{N_i} H_{i,j}\mathbf{e}_{X_{i,j}} + \mathbf{W}_i, \quad \mathbf{W}_i \overset{iid}{\sim} \mathcal{N}(0, \frac{\sigma^2}{\mathcal{E}}\mathbf{I}) \end{aligned}$$

where $\mathbf{e}_{X_1}, \cdots, \mathbf{e}_{X_M}$ are basis vectors.

Ideal Conditions (Deterministic $N_i \equiv \lambda$, $H_{i,j} \equiv 1$ or $\gamma = \infty$ and $\mathbf{W}_i \equiv 0$)

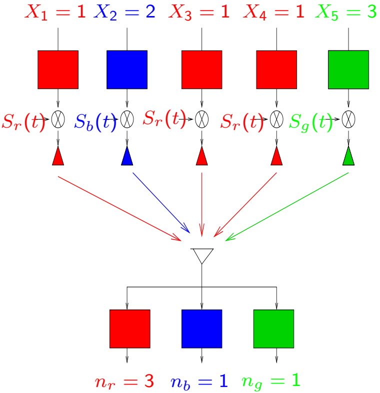
$$\mathbf{Y}_i = \sum_{j=1}^{\lambda} \mathbf{e}_{X_{i,j}}.$$

jth entry of $\mathbf{Y}_i$: no. of sensors quantizing to level j.

$\frac{\mathbf{Y}_i}{\lambda}$ gives Type or Empirical Distribution of $X_{i,j}$.

Matched Filter Output

Schematic of TBRA



Ideal Conditions with $\lambda = 5$

Matched Filtering

$$\mathbf{Y}_i \triangleq \frac{1}{\sqrt{\mathcal{E}}}\Big[\langle Y_i(t), S_1(t)\rangle, \cdots, \langle Y_i(t), S_M(t)\rangle\Big]$$

$$= \sum_{j=1}^{N_i} H_{i,j}\mathbf{e}_{X_{i,j}} + \mathbf{W}_i, \quad \mathbf{W}_i \overset{iid}{\sim} \mathcal{N}(0, \frac{\sigma^2}{\mathcal{E}}\mathbf{I})$$

where $\mathbf{e}_{X_1}, \cdots, \mathbf{e}_{X_M}$ are basis vectors.

Ideal Conditions (Deterministic $N_i \equiv \lambda$, $H_{i,j} \equiv 1$ or $\gamma = \infty$ and $\mathbf{W}_i \equiv 0$)

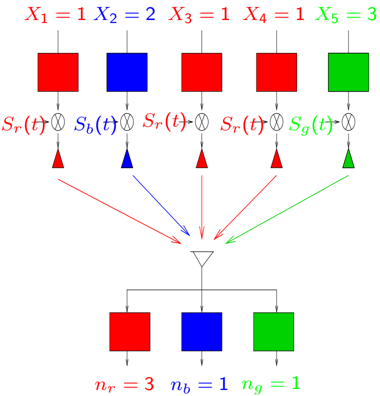
$$\mathbf{Y}_i = \sum_{j=1}^{\lambda} \mathbf{e}_{X_{i,j}}.$$

jth entry of $\mathbf{Y}_i$: no. of sensors quantizing to level j.

$\frac{\mathbf{Y}_i}{\lambda}$ gives Type or Empirical Distribution of $X_{i,j}$.

Matched Filter Output

Schematic of TBRA



Ideal Conditions with $\lambda = 5$

Matched Filtering

$$\mathbf{Y}_i \triangleq \frac{1}{\sqrt{\mathcal{E}}}\Big[\langle Y_i(t), S_1(t)\rangle, \cdots, \langle Y_i(t), S_M(t)\rangle\Big]$$

$$= \sum_{j=1}^{N_i} H_{i,j}\mathbf{e}_{X_{i,j}} + \mathbf{W}_i, \quad \mathbf{W}_i \overset{iid}{\sim} \mathcal{N}(0, \frac{\sigma^2}{\mathcal{E}}\mathbf{I})$$

where $\mathbf{e}_{X_1}, \cdots, \mathbf{e}_{X_M}$ are basis vectors.

Ideal Conditions (Deterministic $N_i \equiv \lambda$, $H_{i,j} \equiv 1$ or $\gamma = \infty$ and $\mathbf{W}_i \equiv 0$)

$$\mathbf{Y}_i = \sum_{j=1}^{\lambda} \mathbf{e}_{X_{i,j}}.$$

jth entry of $\mathbf{Y}_i$: no. of sensors quantizing to level j.

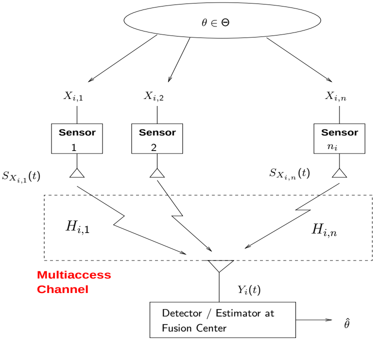
$\frac{\mathbf{Y}_i}{\lambda}$ gives Type or Empirical Distribution of $X_{i,j}$.

Role of Coherence Index γ for TBRA

- MAC adds transmissions with same data level
- Large γ : Better addition of signals in the mean since

$$\mathbb{E}\mathbf{Y}_i = \lambda \mathbb{E}(H)\mathbf{p}_\theta.$$

- Small γ : Effect of sensor data is only a second order effect (through the Channel Variance).



$$\gamma = \frac{|\mathbb{E}(H)|^2}{\mathrm{Var}(H)}.$$

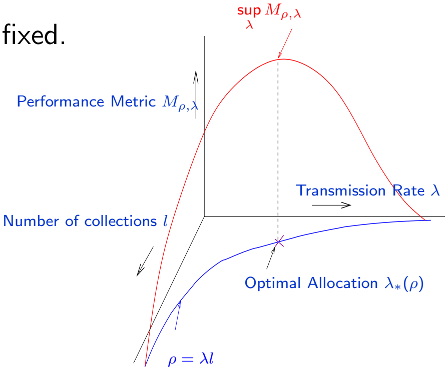
Outline

Spatio-Temporal Tradeoff

- Mean no. of transmissions $\rho = \lambda l$ fixed.
- Performance Metric $M_{\rho,\lambda}$.
- Optimal allocation in λ and l

$$\lambda_*(\rho) = \arg\sup_{\lambda} M_{\rho,\lambda}.$$

- For finite ρ, $M_{\rho,\lambda}$ intractable in our setup.



Asymptotic Performance Metric $M(\lambda)$

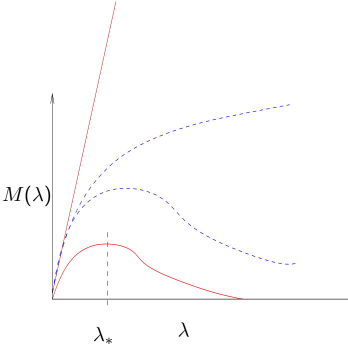
- Metric : Asymptotic (in ρ) Performance.

$$M(\lambda) \overset{\Delta}{=} \lim_{\rho \to \infty} M_{\rho,\lambda}.$$

- Goal : Existence of a optimal λ_* such that

$$\lambda_* = \arg\max_{\lambda} M(\lambda).$$

- λ_* is bounded : Cancelation, Avoid Interference.
- λ_* is unbounded : Coherence, Simultaneous Transmissions.



Performance Metric $M(\lambda)$

Detection

- Performance Metric is Detection error exponent

$$M(\lambda) \triangleq -\lim_{\rho\to\infty} \frac{1}{\rho} \log P_e(\rho,\lambda),$$

where $P_e(\rho,\lambda)$ is detection error probability.

- Under Neyman-Pearson or Bayesian setting,

$$\begin{aligned} M^{\mathrm{NP}}(\lambda) &= \frac{1}{\lambda} D_\lambda(f_0 \| f_1), \\ M^{\mathrm{B}}(\lambda) &= \frac{1}{\lambda} C_\lambda(f_0, f_1), \end{aligned}$$

$D_\lambda(f_0\|f_1)$: Kullback-Leibler distance.
$C_\lambda(f_0, f_1)$: Chernoff information.

Estimation

- Define Performance Metric

$$M(\lambda) \triangleq \frac{I_\lambda(\theta)}{\lambda},$$

$I_\lambda(\theta)$: Fisher Information of $\mathbf{Y}_i$.

- Cramer Rao Bound for any unbiased estimator $\hat\theta$ based on ρ mean no. of transmissions and transmission rate λ,

$$\mathrm{Var}(\hat\theta - \theta) \geq \frac{\lambda}{\rho I_\lambda(\theta)}.$$

- Asymptotic Efficiency of ML estimator

$$\sqrt{\rho}(\hat\theta_{\mathrm{MLE}} - \theta) \xrightarrow{d} \mathcal{N}(0, \frac{\lambda}{I_\lambda(\theta)}).$$

Performance Metric $M(\lambda)$

Detection

- Performance Metric is Detection error exponent

$$M(\lambda) \triangleq -\lim_{\rho \to \infty} \frac{1}{\rho} \log P_e(\rho, \lambda),$$

where $P_e(\rho, \lambda)$ is detection error probability.

- Under Neyman-Pearson or Bayesian setting,

$$\begin{aligned} M^{\mathsf{NP}}(\lambda) &= \frac{1}{\lambda} D_\lambda(f_0 || f_1), \\ M^{\mathsf{B}}(\lambda) &= \frac{1}{\lambda} C_\lambda(f_0, f_1), \end{aligned}$$

$D_\lambda(f_0||f_1)$: Kullback-Leibler distance.
$C_\lambda(f_0, f_1)$: Chernoff information.

Estimation

- Define Performance Metric

$$M(\lambda) \triangleq \frac{I_\lambda(\theta)}{\lambda},$$

$I_\lambda(\theta)$: Fisher Information of $\mathbf{Y}_i$.

- Cramer Rao Bound for any unbiased estimator $\hat{\theta}$ based on ρ mean no. of transmissions and transmission rate λ,

$$\mathsf{Var}(\hat{\theta} - \theta) \geq \frac{\lambda}{\rho I_\lambda(\theta)}.$$

- Asymptotic Efficiency of ML estimator

$$\sqrt{\rho}(\hat{\theta}_{\mathsf{MLE}} - \theta) \xrightarrow{d} \mathcal{N}(0, \frac{\lambda}{I_\lambda(\theta)}).$$

Outline

Type Based Random Access

Performance Metric

Optimal TBRA

Conclusion

Theorem on Existence of optimal λ

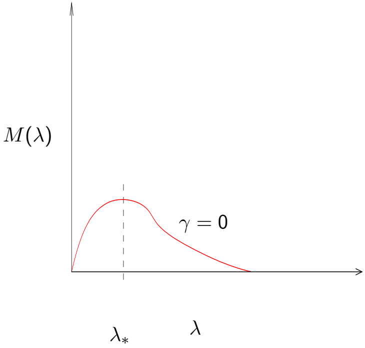
Under Regularity Conditions in the paper,

Non Coherent Channels ($\mathbb{E}(H) = 0$ or $\gamma = 0$) :
Existence of Bounded optimal λ_*

$$\lim_{\lambda \to 0} M(\lambda) = \lim_{\lambda \to \infty} M(\lambda) = 0,$$

which implies that there exists $0 < \lambda_* < \infty$ such that

$$\sup_{\lambda} M(\lambda) = \frac{1}{\lambda_*} I_{\lambda_*}(\theta).$$



Deterministic Channels ($\text{Var}(H) = 0$ or $\gamma = \infty$) : No Bounded optimizing λ_*,

$$M(\lambda) = \Theta(\lambda) \quad , \lambda \to \infty,$$

$$\gamma = \frac{|\mathbb{E}(H)|^2}{\text{Var}(H)}.$$

Theorem on Existence of optimal λ

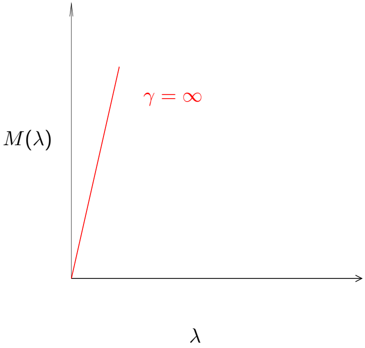
Under Regularity Conditions in the paper,

Non Coherent Channels ($\mathbb{E}(H) = 0$ or $\gamma = 0$) :
Existence of Bounded optimal λ_*

$$\lim_{\lambda \to 0} M(\lambda) = \lim_{\lambda \to \infty} M(\lambda) = 0,$$

which implies that there exists $0 < \lambda_* < \infty$ such that

$$\sup_{\lambda} M(\lambda) = \frac{1}{\lambda_*} I_{\lambda_*}(\theta).$$

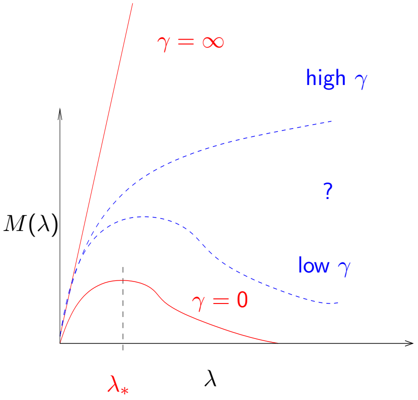


Deterministic Channels ($\mathrm{Var}(H) = 0$ or $\gamma = \infty$) : No Bounded optimizing λ_*,

$$M(\lambda) = \Theta(\lambda) \quad , \lambda \to \infty,$$

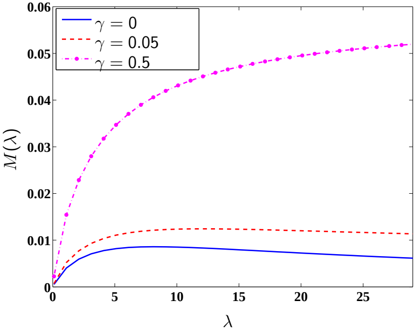
$$\gamma = \frac{|\mathbb{E}(H)|^2}{\mathrm{Var}(H)}.$$

$M(\lambda)$ for different Coherence Indices.

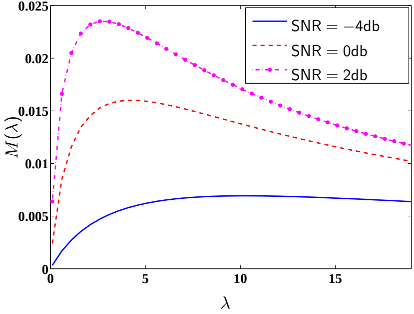


$$\gamma = \frac{|\mathbb{E}(H)|^2}{\mathrm{Var}(H)}.$$

Numerical Plots of Performance Metric



(SNR = 7db, $\sigma_H^2 = 0.1$)



($\gamma = 0$, $\sigma_H^2 = 1$)

Normalized Chernoff Information vs. Transmission Rate.

Asymptotic Normal Distribution

- To analyze asymptotic behavior of $M(\lambda)$: compute Performance metric for limiting distribtution $\tilde{M}(\lambda)$.
- Since by continuity $\lim_{\lambda\to\infty} M(\lambda) = \lim_{\lambda\to\infty} \tilde{M}(\lambda)$.
- CLT for Random Number of Summands :

$$\frac{\mathbf{Y} - \lambda\mathbb{E}(H)\mathbf{p}_\theta}{\sqrt{\lambda}} \xrightarrow{d} \mathcal{N}\Big(0, \mathrm{Var}(H)\mathrm{Diag}(\mathbf{p}_\theta)\Big) \quad \text{as} \quad \lambda \to \infty.$$

- Gaussian Metric $\tilde{M}(\lambda)$: closed form expressions.
- For Large λ, approximate actual $M(\lambda)$ by Gaussian $\tilde{M}(\lambda)$.

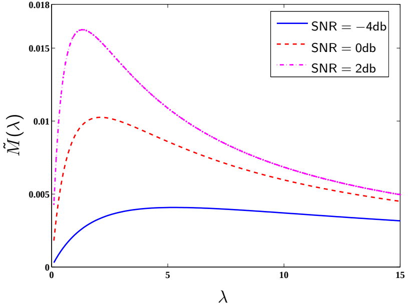
Asymptotic Normal Distribution

- To analyze asymptotic behavior of $M(\lambda)$: compute Performance metric for limiting distribtution $\tilde{M}(\lambda)$.
- Since by continuity $\lim_{\lambda\to\infty} M(\lambda) = \lim_{\lambda\to\infty} \tilde{M}(\lambda)$.
- CLT for Random Number of Summands :

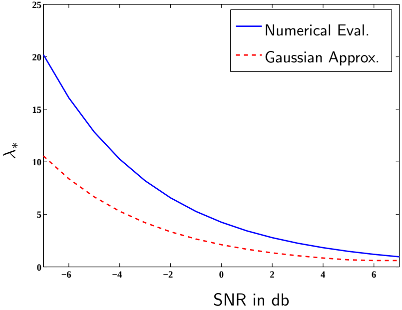
$$\frac{\mathbf{Y} - \lambda \mathbb{E}(H)\mathbf{p}_\theta}{\sqrt{\lambda}} \xrightarrow{d} \mathcal{N}\Big(0, \mathsf{Var}(H)\mathsf{Diag}(\mathbf{p}_\theta)\Big) \quad \text{as} \quad \lambda \to \infty.$$

- Gaussian Metric $\tilde{M}(\lambda)$: closed form expressions.
- For Large λ, approximate actual $M(\lambda)$ by Gaussian $\tilde{M}(\lambda)$.

Gaussian Approximation



Gaussian Metric vs. Transmission Rate. ($\gamma = 0$, $\sigma_H^2 = 1$)



Optimal λ_* vs. SNR in db. ($\gamma = 0$, $\sigma_H^2 = 1$)

Outline

Type Based Random Access

Performance Metric

Optimal TBRA

Conclusion

Summary

- Introduced TBRA : removes requirement of channel coherency and handles random number of sensors.
- Provided a general characterization of Performance metric of estimation and provided approximate solutions.
- Proved the existence of optimal spatio-temporal allocation scheme dependent on Channel Coherence Index.

Related Publication

- A. Anandkumar and L. Tong, A Large Deviation Analysis of Detection over Multi-Access Channels with Random Number of Sensors, *accepted to Proc. of ICASSP 06, Toulouse, France, May 2006.*

- A. Anandkumar and L. Tong, Type-Based Random Access for Distributed Detection over Multiaccess Fading Channels, *Submitted to IEEE Trans. Signal Proc., Dec. 2005.*

- A. Anandkumar, L. Tong and A. Swami, Large deviation analysis of Sequential distributed detection using Type based Random Access, *To be submitted to Proc. of EUSIPCO, Sep. 2006.*

References

- G. Mergen, V. Naware, and L. Tong, Asymptotic Detection Performance of Type-Based Multiple Access Over Multiaccess Fading Channels, submitted to IEEE Trans. on Signal Processing, May 2005.

- Ke Liu and A. M. Sayed, Optimal distributed detection strategies for wireless sensor networks,, in 42nd Annual Allerton Conf. on Commun., Control and Comp., Oct. 2004.

- J.-F. Chamberland and V. V. Veeravalli, Asymptotic results for decentralized detection in power constrained wireless sensor networks, IEEE JSAC Special Issue on Wireless Sensor Networks, 2004.

- S.A. Aldosari and J.M.F. Moura, Detection in decentralized sensor networks, in Proc. of ICASSP 04 Conf., Montreal, Canada.

- B. Chen, R. Jiang, T. Kasetkasem, and P.K. Varshney, Channel aware decision fusion in wireless sensor networks, IEEE Trans. on Signal Processing, vol. 52, Dec. 2004.

Thank You !