Sensor-Fusion Center Communication Over Multiaccess Fading Channels

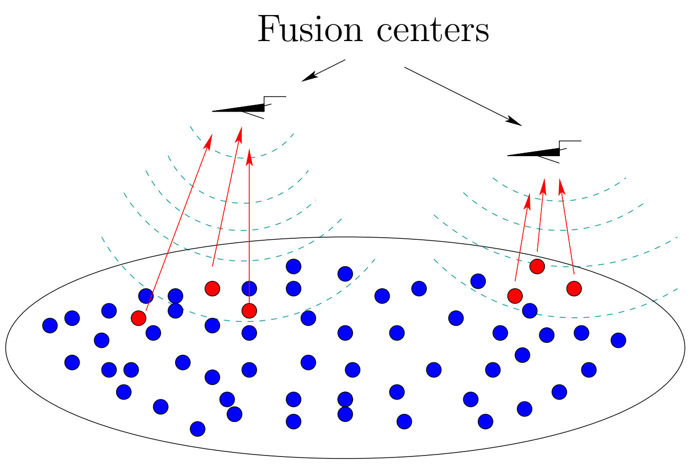
Gökhan Mergen, V. Naware* and Lang Tong*

School of Electrical and Computer Engineering

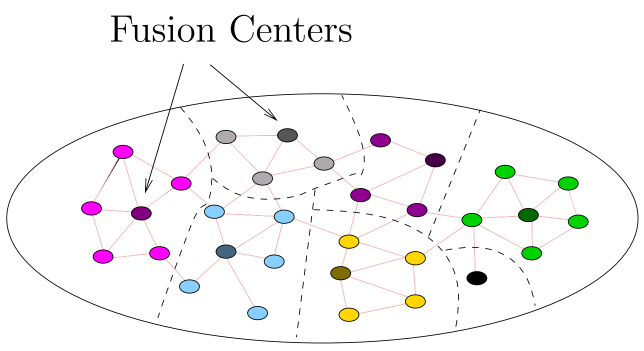
Cornell University, Ithaca, NY

* Now, with Qualcomm Inc., Campbell CA.

Motivation: Sensor Networks



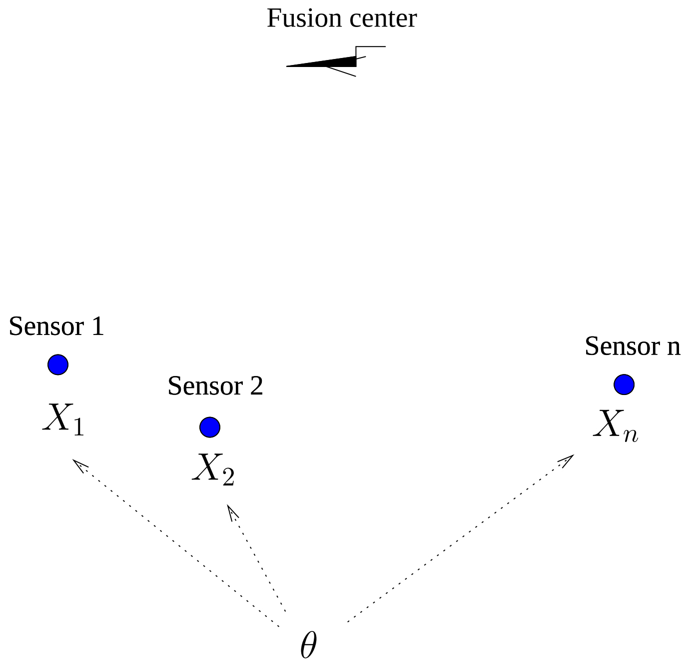
i) Sensor network with mobile access.



ii) Multi-hop ad-hoc network with local data fusion.

- Sense a physical phenomena, transmit it to a fusion center.
- Fusion centers estimate/detect the field parameters, and deliver the estimate.

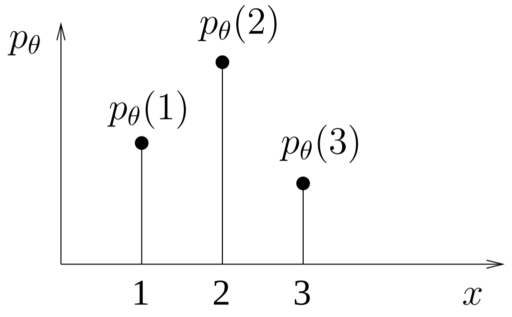
Problem Formulation



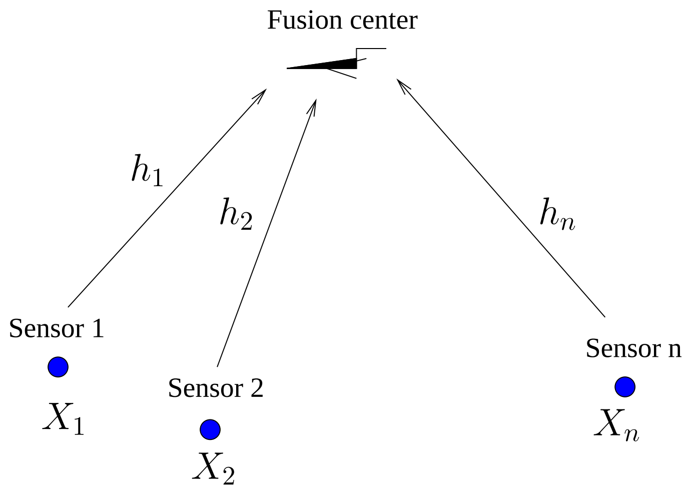
- Estimate $\theta \in \mathbb{R}$.
- Observation $X_1, \cdots, X_n$ are i.i.d. conditioned on θ. These are quantized observations.

Assumptions:

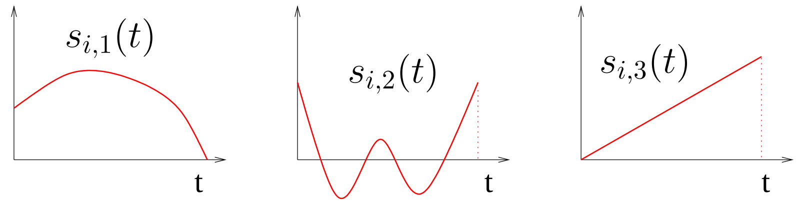
- X_i takes values in $\{1, \cdots, k\}$.
- $X_i \sim p_\theta$, where p_θ is a probability mass function.



Problem Formulation Cont'd



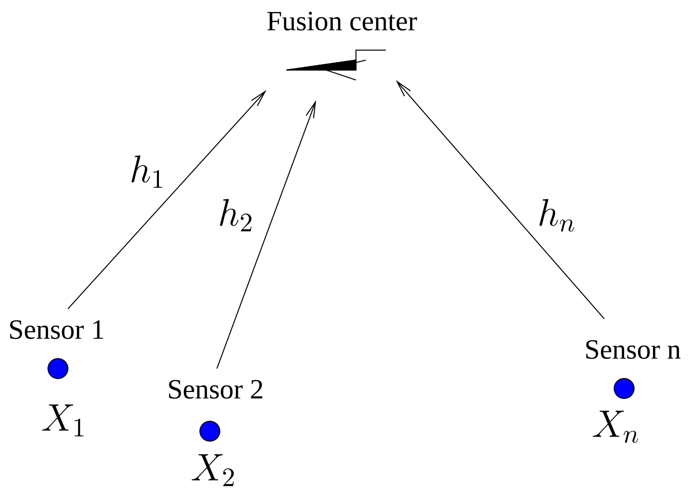
- Sensor i has a set of k channel waveforms $s_{i,1}, \cdots, s_{i,k}$.
- Upon observing X_i, it transmits s_{i,X_i}.



Received signal: $z = \sum_{i=1}^{n} h_i s_{i,X_i} + w$.

- Channel gains $h_1, \cdots, h_n \in \mathbb{R}$ are i.i.d.
- The noise is white $\mathcal{N}(0, \sigma^2)$.
- Energy constraint $||s_{i,j}||^2 \leq E$.

Problem Formulation Cont'd



Problem:

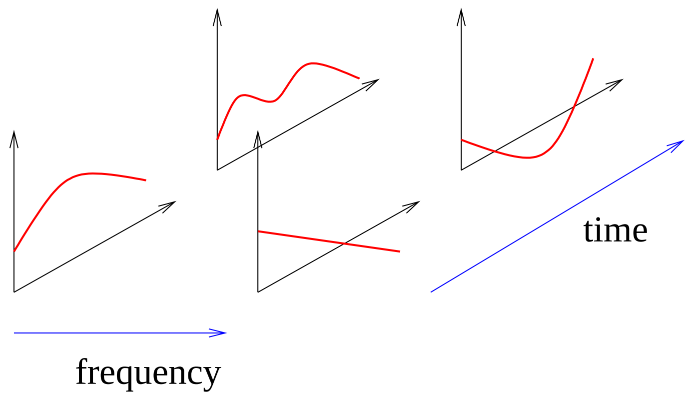
- Design the channel waveforms to facilitate the estimation or detection of parameter θ based on the received signal z.

Performance Metrics:

i) Minimize the Mean Square Error (MSE) $\mathbb{E}\{(\hat{\theta} - \theta)^2\}$ between the observation and the true parameters.

ii) Minimize the probability of error in detection.

Classical Approach



- Collision among users is "bad!"
- Solution: Orthogonalize different users.
- Can be done by time/frequency/code division.

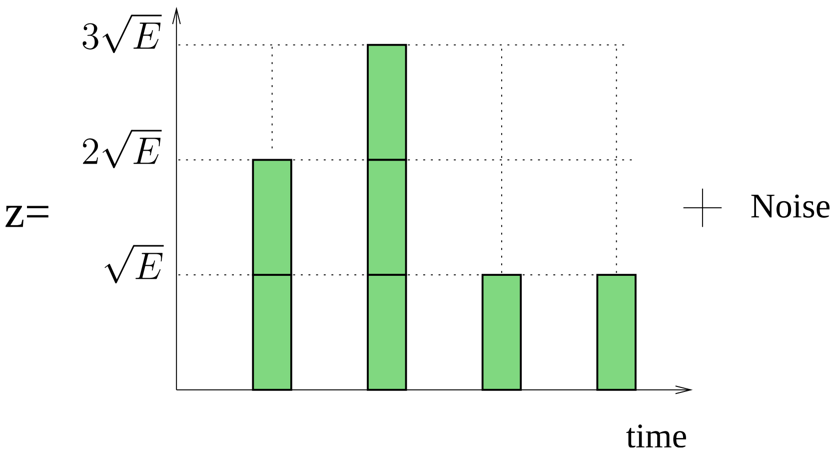
Advantages:

- Allows us to use the standard layered approach.
- Well understood. Rather easy to implement.

Caveat:

- The bandwidth requirement is significant for large n.
- Neglects the dependency among sensor data.

Proposed Approach: Type-Based Multiple Access (TBMA)



- Idea: Orthogonalize with respect to observations, not with respect to users.
- $s_{i,X_i} = \sqrt{E}\delta_{X_i}$, where $\delta_1, \cdots, \delta_k$ are orthonormal waveforms.
- When all $h_i = 1$,

$$y := \frac{1}{\sqrt{E}n}(\langle z, \delta_1 \rangle, \cdots, \langle z, \delta_k \rangle)$$
$$= \text{ empirical measure + noise.}$$

Advantages:

- Uses much less bandwidth/time than orthogonal approaches if k is small compared to n (*i.e.*, if there are many sensors).
- Delivers a noisy version of a sufficient (the empirical measure). Also, the noise power is attenuated by $1/n^2$.

Previous Work

Estimation Setup:

- G. Mergen and L. Tong, "Estimation over deterministic multiaccess channels," in *42nd Annual Allerton Conf. on Commun., Control and Comp.*, 2004.
- G. Mergen and L. Tong, "Type-based estimation over multiaccess channels," to appear in *IEEE Trans. on Signal Processing*.

Detection Setup:

- Ke Liu and A. M. Sayeed, "Optimal distributed detection strategies for wireless sensor networks," in *42nd Annual Allerton Conf. on Commun., Control and Comp.*, Sep.29-Oct.1 2004.

Outline

- Introduction
- Performance analysis
  - ▶ Fundamental limits of Estimation
  - – Asymptotic Estimation Performance of TBMA
  - – Comparison with Other Orthogonal Allocation Methods
  - – Asymptotic Detection Performance of TBMA
- Conclusions

Fundamental Limits

Cramer-Rao Bound:

Let $\hat{\theta}$ be an *unbiased* estimator based on $X_1, \cdots, X_n$. Then,

$$\mathbb{E}\{(\hat{\theta}-\theta)^2\} \geq \frac{1}{nI(\theta)}, \tag{1}$$

where $I(\theta) = \sum_{i=1}^{k} \frac{(dp_\theta(i)/d\theta)^2}{p_\theta(i)}$ is the *Fisher information* in X_i.

Asymptotic Efficiency:

There exists a class of estimators based on $X_1, \cdots, X_n$ satisfying

$$\hat{\theta} \simeq \mathcal{N}(\theta, \frac{1}{nI(\theta)}), \qquad \text{for large } n. \tag{2}$$

Notation "$\simeq$" means $\hat{\theta} \xrightarrow{\text{p}} \theta$ and $\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{\text{d}} \mathcal{N}(0, \frac{1}{I(\theta)})$ as $n \to \infty$.

Estimation with TBMA

- Assume the channel gain h_i has *non-zero* mean $h := \mathbb{E}(h_i)$, and variance $\sigma_h^2 := \mathrm{Var}(h_i)$.

Lemma 1 (Asymptotic distribution of received signal):

Let $y = \frac{1}{\sqrt{E}hn}(\langle z, \delta_1\rangle, \cdots, \ \langle z, \delta_k\rangle)$. Then,

$$y \simeq \mathcal{N}(p_\theta, \frac{1}{n}\Sigma), \quad \text{where} \quad \Sigma = (1 + \frac{\sigma_h^2}{h^2})\text{Diag}(p_\theta) - p_\theta p_\theta^T.$$

Definition (Asymptotic ML Estimator):

If y were $\mathcal{N}(p_\theta, \frac{1}{n}\Sigma)$, then its pdf would be

$$f(y) = \frac{1}{(2\pi/n)^{\frac{k}{2}}}\text{exp}\left(-\frac{n(y-p_\theta)^T\Sigma_\theta^{-1}(y-p_\theta) + \log|\Sigma_\theta|}{2}\right).$$

The first term in the exponent dominates for large n. Define the estimator $\hat{\theta}$ to be the $\theta \in \mathbb{R}$ that minimizes $M(\theta) := (y-p_\theta)^T\Sigma_\theta^{-1}(y-p_\theta)$ with respect to θ.

Estimation with TBMA Cont'd

Theorem 1:

The proposed asymptotic ML estimator based on y satisfies

$$\hat{\theta} \doteq \mathcal{N}(\theta, \frac{1+\frac{\sigma_h^2}{h^2}}{nI(\theta)}), \qquad \text{for large } n.$$

Remarks:

- In a deterministic channel, the CRB is asymptotically achieved, *i.e.*, the asymptotic performance of TBMA is as if the fusion center has direct access to X_i's.
- This performance is optimal, *i.e.*, no unbiased $\hat{\theta}$, even the ones with direct access to X_i's, can do better than this.
- In case of fading, the MSE degrades by a factor $1+\sigma_h^2/h^2$.

Fading Channels with Zero Mean

Theorem 2:

Suppose that the channel mean is zero. Let $y = \frac{1}{\sqrt{E}\sqrt{n}}(\langle z, \delta_1 \rangle, \cdots, \langle z, \delta_k \rangle)$. Then,

$$y \xrightarrow{\mathrm{d}} \mathcal{N}(0, \text{Diag}(p_\theta)), \quad \text{as} \quad n \to \infty.$$

Remarks:

- The parameter information appears as a second order effect.
- The Fisher information (FI) in $\mathcal{N}(0, \text{Diag}(p_\theta))$ is finite. Therefore, the MSE of any unbiased estimator should be $\geq$ FI.
- Conclusion: The MSE can not go to zero as $n \to \infty$. The asymptotic performance of TBMA is not too promising in zero-mean channels. However, with the use of transmitter CSI this situation can be rectified; see our SP Transactions paper.

Other Orthogonal Allocation Methods (T/FDMA, etc.)

- Let $s_1, \cdots, s_k \in \mathbb{R}^m, \ ||s_i||^2 \leq 1,$ be constellation points.
- Signal received from the i'th sensor:

$$z^{(i)} = h_i \sqrt{E} s_{X_i} + v^{(i)}, \tag{3}$$

where $v^{(i)} \sim \mathcal{N}(0, \sigma^2 I)$.

Asymptotic Performance:

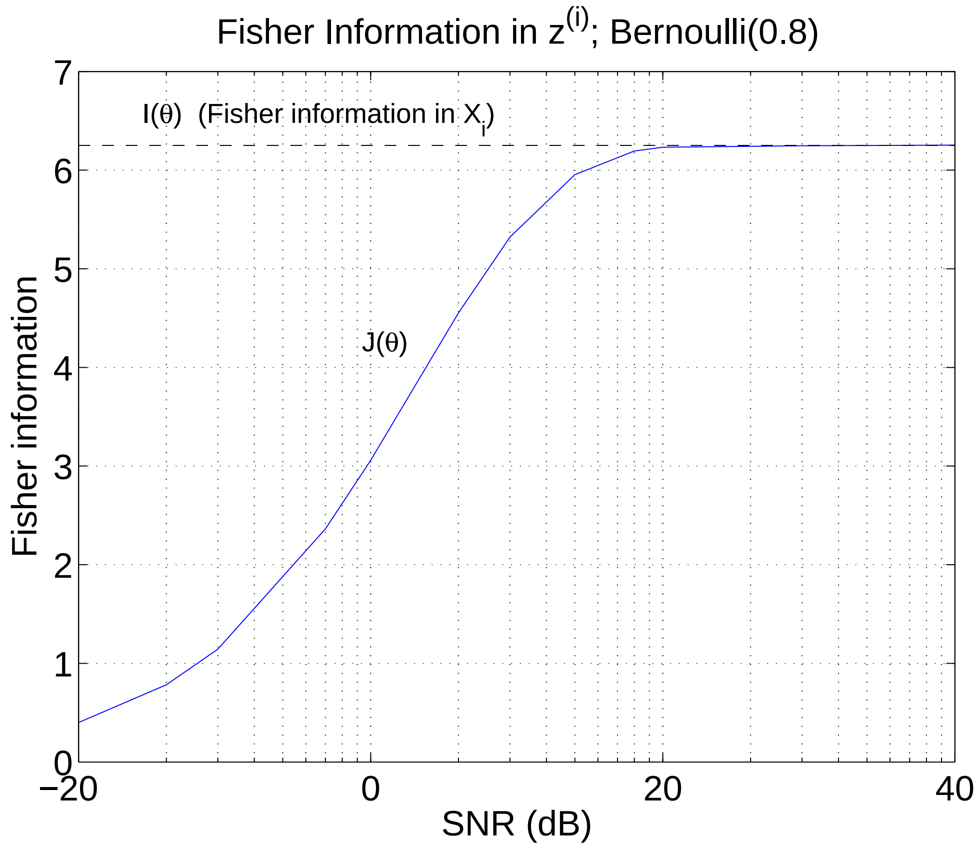
- For any unbiased $\hat{\theta}$,

$$\mathbb{E}\{(\hat{\theta} - \theta)^2\} \geq \frac{1}{nJ(\theta)}, \tag{4}$$

where $J(\theta) = \mathbb{E}_{z^{(i)}} \left[\left(\frac{d \log f(z^{(i)})}{d\theta} \right)^2 \right]$ is the *Fisher information* in $z^{(i)}$.

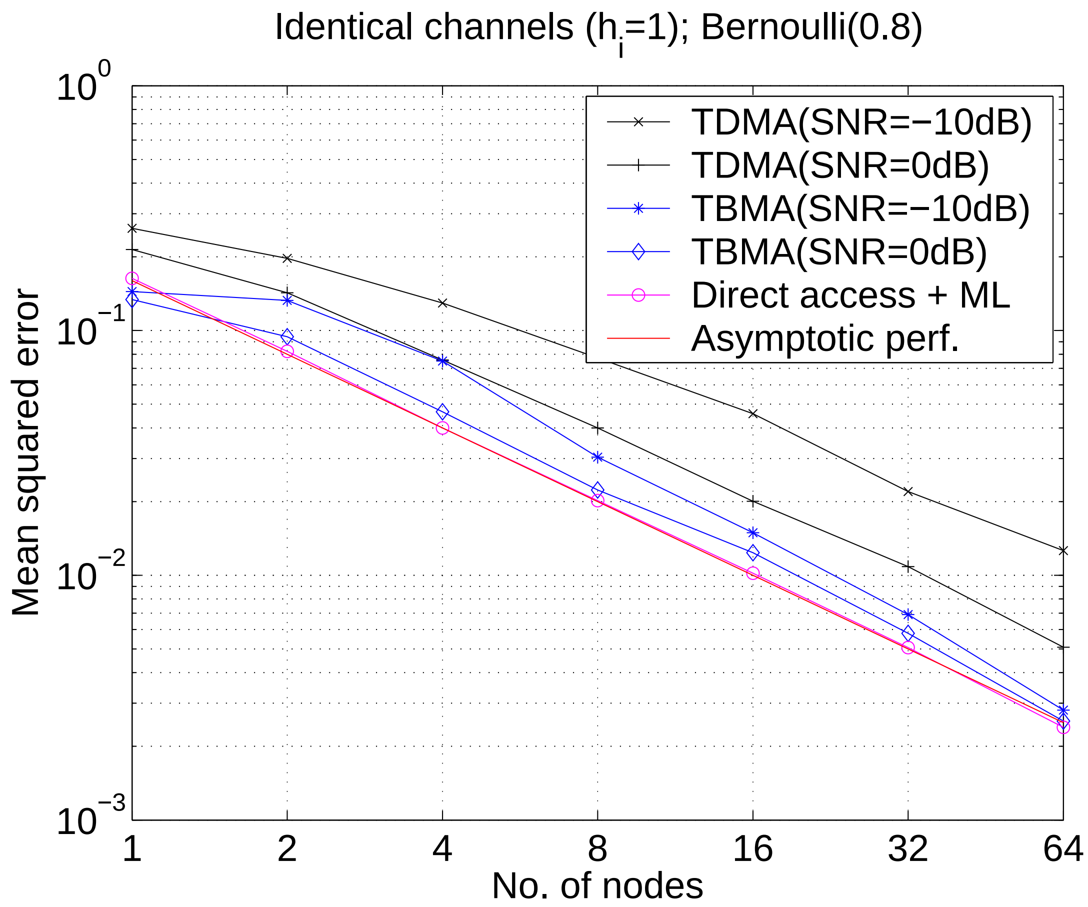
- The MLE based on $z^{(1)}, \cdots, z^{(n)}$ satisfies $\hat{\theta} \simeq \mathcal{N}(\theta, \frac{1}{nJ(\theta)})$.

Performance of Orthogonal Allocation Cont'd

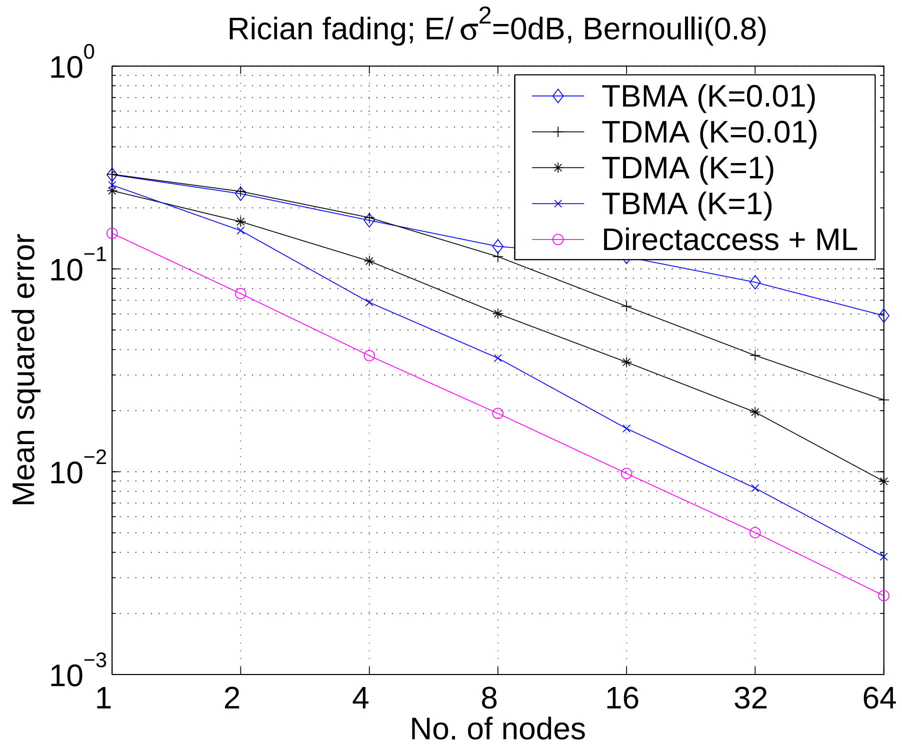


- For the best asymptotic performance $J(\theta)$ should be maximized with respect to the constellation.
- Theorem 3: For $h_i = 1$, the antipodal constellation maximizes $J(\theta)$ for $k = 2$.
- In general, the optimal constellation depends on the family $\{p_\theta : \theta \in \mathbb{R}\}$.

A Numerical Example



Numerical Example - 2



- The channel is Rician distributed, $i.e.,$

$$h_i = \sqrt{\frac{K}{K+1}} + \sqrt{\frac{1}{K+1}}\mathcal{CN}(0,1),$$

where $K > 0$ is a deterministic number ($K = 0 \;\Rightarrow$ Rayleigh).

Detection with TBMA

- Suppose that the channel gain h is non-zero. Let $D(X||Y)$ denote the relative entropy between random variables X and Y.

Theorem 4:

Let $I_h : \mathbb{R} \rightarrow \mathbb{R}_+ \cup \{\infty\}$ be the function defined as

$$I_h(r) = \inf_{\tilde{h}:\ \mathbb{E}(\tilde{h})=r} D(\tilde{h}||h_i), \quad r \in \mathbb{R}, \tag{5}$$

where the minimization is over real valued random variables $\tilde{h}$. Similarly, define $I : \mathbb{R}^k \rightarrow \mathbb{R}_+ \cup \{\infty\}$

$$I(x) = \inf_{\tilde{p}} \{D(\tilde{p}||p_\theta) + \sum_{j=1}^{k} \tilde{p}_j I_h(\frac{x_j}{\tilde{p}_j})\}, \quad x \in \mathbb{R}^k, \tag{6}$$

where the minimization is over all probability vectors $\tilde{p} \in \mathbb{R}^k$. The $y = \frac{1}{\sqrt{E}n}(\langle z, \delta_1 \rangle, \cdots, \ \langle z, \delta_k \rangle)$ satisfies the Large Deviations Principle (LDP) with the rate function I.

Detection with TBMA Cont'd

Interpretation of LDP:

- The probability that the y turns out to be in a close vicinity of x decays as $e^{-nI(x)}$.
- More precisely, let $B_\epsilon(x)$ be the open ball centered at x with radius ϵ. Then,

$$\lim_{n\to\infty} \frac{1}{n} \log \Pr(y \in B_\epsilon(x)) = I(x) + O(\epsilon) \tag{7}$$

Definition (Asymptotic ML Detector):

- Consider a hypothesis test between $\mathcal{H}_0: \ \theta = \theta_0$ vs. $\mathcal{H}_1: \ \theta = \theta_1$.
- Let I_i be the rate function under $\mathcal{H}_i$. According to above interpretation, $e^{-nI_i(x)}$ can be viewed as the likelihood under hypothesis $\mathcal{H}_i$.
- Define the detector θ^* with detection regions $\Gamma_0 = \{I_0(x) \leq I_1(x)\}$, $\Gamma_1 = \mathbb{R}^k \setminus \Gamma_0$

Detection with TBMA Cont'd

Theorem 5:

i) The proposed asymptotic ML detector θ^* gives the best error exponent in Bayesian hypothesis testing among all detectors which is a function of y.

ii) If the rate functions I_0 and I_1 are continuous at the boundary between Γ_0 and Γ_1, then the Type-I and Type-II error exponents of the detector θ^* are equal and given by

$$\lim_{n \to \infty} \frac{1}{n} \log \Pr(y \in \Gamma_1;\ \theta = \theta_0) = -\inf_{x \in \Gamma_1} I_0(x)$$

Corollary 1:

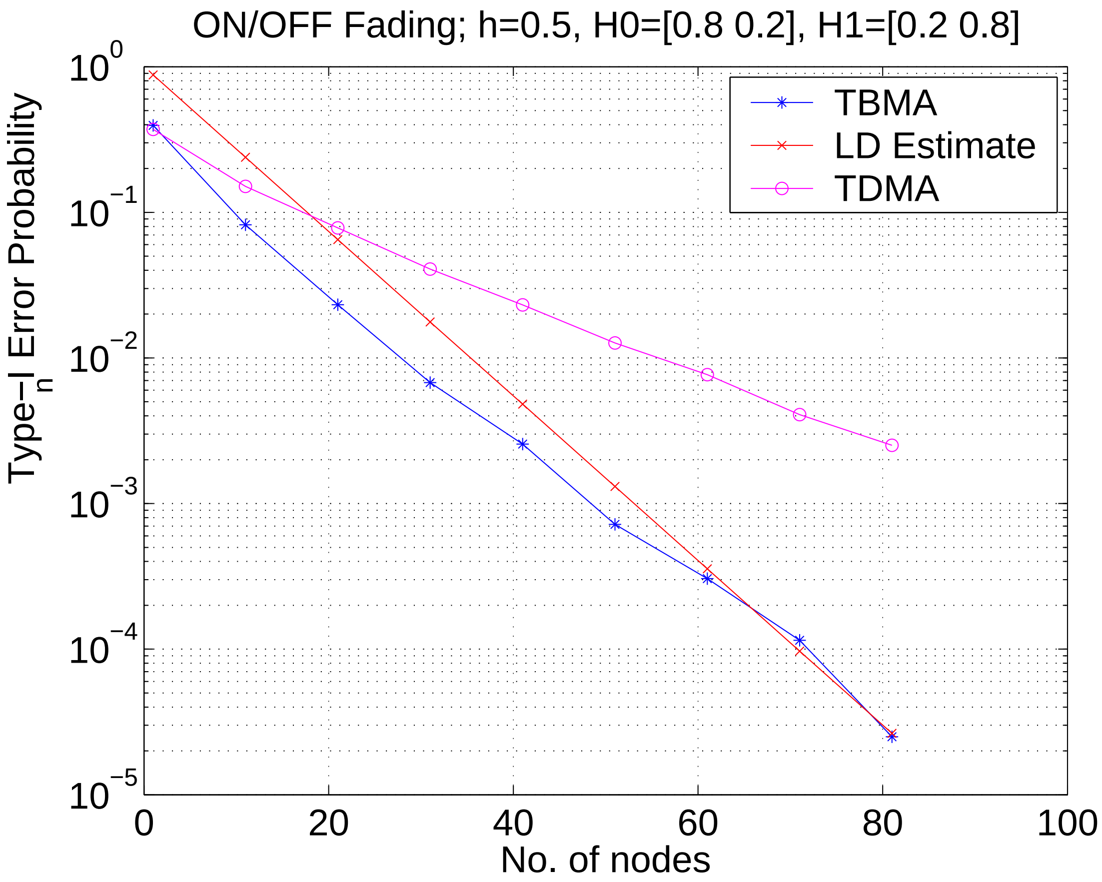
If the sensor channel gains are deterministic and identical, then $I_i(x) = D(x||p_{\theta_i})$ for all probability vectors $x \in \mathbb{R}^k$. This implies that the TBMA has the best Bayesian exponent (the so-called Chernoff information) achievable among all detectors which is a function of $X_1, \cdots, X_n$.

Detection with TBMA Cont'd

Remarks:

- Corollary 1 implies that the TBMA scheme with the proposed detector is asymptotically optimal in terms of error exponents.
- The same conclusion was obtained independently using different methods by Liu and Sayeed (Allerton, Oct.'04).
- The advantage of the Large Deviations framework, and in particular, Theorem 5 is that these results are applicable in the fading scenario as well.

Numerical Example: ON/OFF Channel



- The channel gain is $\{0, 1\}$ with probabilities $\{0.5, 0.5\}$. LD estimate refers to e^{-nE}, where E is the exponent given by Theorem 5.

Detection in Zero-mean Channels

Case 1 ($p_{\theta_i}(j) > 0, \ \forall i, j$):

- The detection probabilities with the ML detector do not go to zero as $n \to \infty$ (similar to what happened in estimation).

Case 2 ($p_{\theta_0}(j) = 0, \ p_{\theta_1}(j) > 0,$ for some j, or vice versa):

- The detection probabilities go to zero as $n \to \infty$, but not exponentially fast. The rate of decay is characterized in [1].

[1] G. Mergen, V. Naware and L. Tong, "Asymptotic Detection Performance of Type-Based Multiple Access Over Multiaccess Fading Channels," to be submitted.

Conclusions

- Considered the problem of communicating sensor readings over a Gaussian multiaccess channel.
- The TBMA scheme is proposed as an alternative to the conventional orthogonal allocation.
- The estimation and detection performance of TBMA is analyzed by using certain asymptotics of the received signal and the Large Deviations Theory.
- The TBMA requires less bandwidth than the traditional orthogonal allocation methods in a large network, and also has favorable MSE/Error Probability depending on channel conditions.
- The detection/estimation performance of TBMA is shown to be asymptotically optimal if there is no fading.
- The TBMA scheme needs improvement in channels with zero mean (Transmitter CSI is very useful in that scenario).