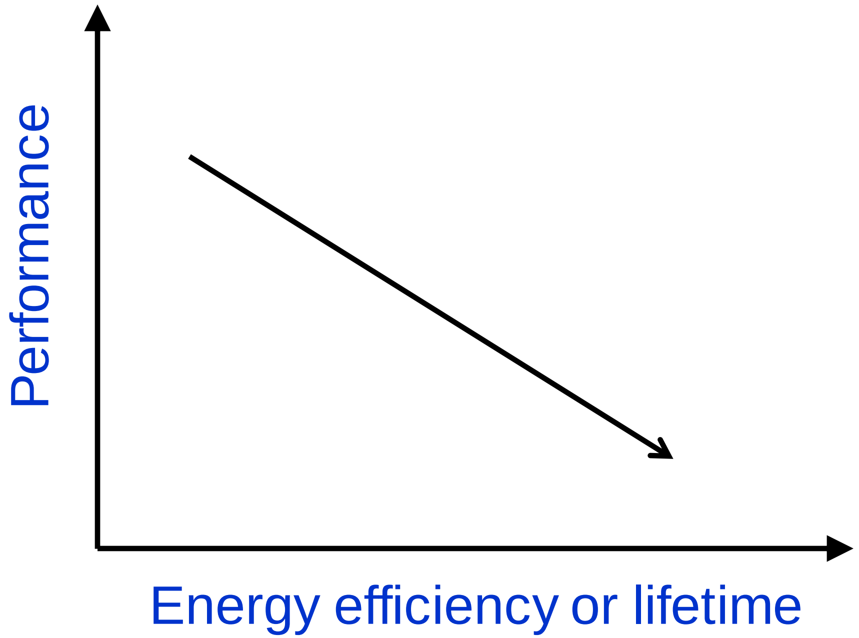
New Twists on the Ordered Transmissions Scheme for Energy-Efficient Detection in Wireless Sensor Networks

Neelesh B. Mehta

[Joint work with Sayan Sen Gupta]

Wireless Sensor Networks Design Challenge

- Minimize energy consumption ⇒ Maximize lifetime
  - Example: Make nodes sleep as much as possible, censor node transmissions
- But: More energy efficiency ⇒ Worse performance
  - Example: Increase in mean square error, latency



Outline

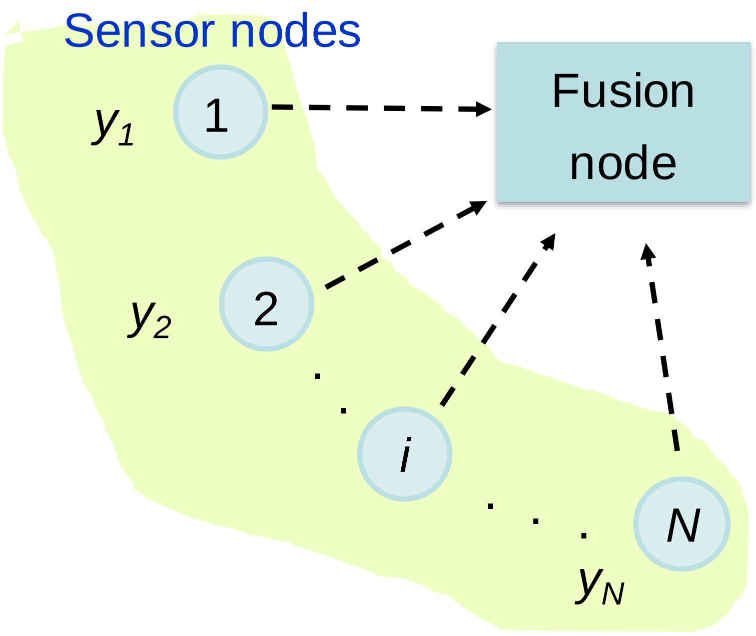
- Ordered transmission scheme (OTS)
- Two twists
  - Correlated observations
  - Energy harvesting WSNs
- Key takeaways

Binary Hypotheses Testing

- Two hypotheses
  - H_1: Intruder present, pollution present, primary on, cancer present
  - H_0: Intruder absent, pollution absent, primary off, cancer absent

Diverse applications

- Spectrum sensing
- Environmental monitoring
- Military surveillance
- Healthcare



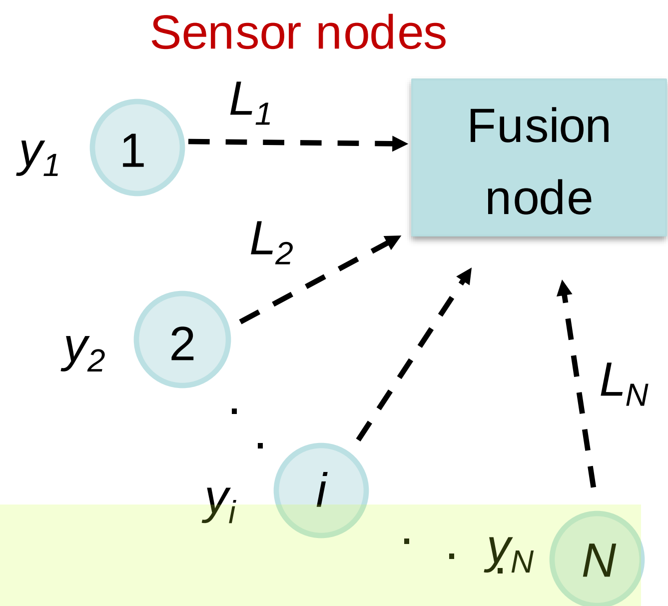
Background: Optimal Detection Rule

- Error-optimal decision rule

$$L_1 + L_2 + \cdots + L_N \underset{H_0}{\overset{H_1}{\gtrless}} \beta$$

- Log-likelihood ratio (LLR)

$$L_i = \log \frac{p(y_i|H_1)}{p(y_i|H_0)}$$



Example

$$H_1 : y_i = s + n_i \sim N(0, \sigma_1^2)$$

$$H_0 : y_i = n_i \sim N(0, \sigma_0^2)$$

} Gaussian signals with different variances

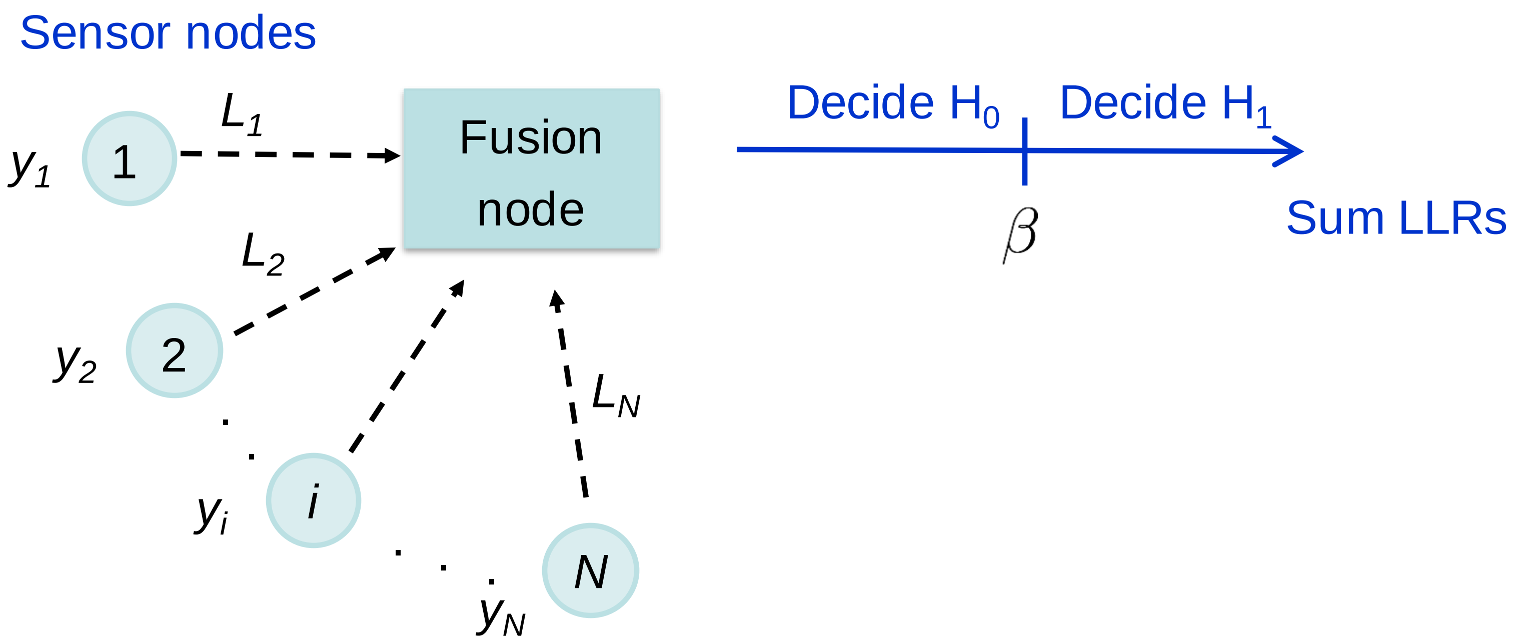
$$\text{LLR} : L_i = \log\left(\frac{\sigma_1^2}{\sigma_0^2}\right) + y_i^2\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)$$

$$L_i \propto y_i^2$$

$\rightarrow$ LLR is just the measured energy

To Make Distributed Detection Work

- Every node sends its LLR to fusion node
- Fusion node sums up all LLRs and decides
  - Needs *N* transmissions with *N* sensors



Ordered Transmission Scheme

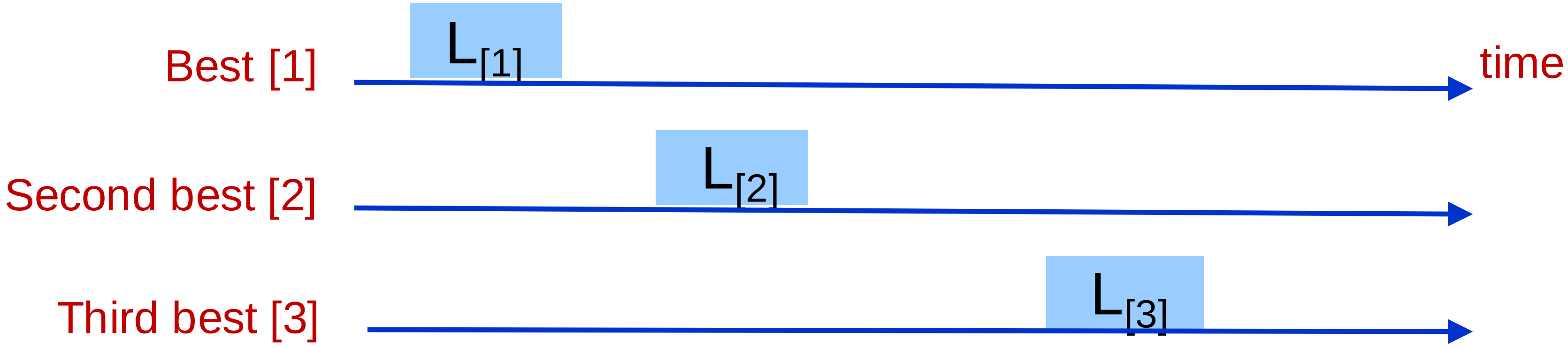
Ordered Transmissions [Blum & Sadler 2008]

- Note: Decision statistic is separable in local LLRs
- Idea: Nodes transmit in descending order of |LLR|

$$|L_{[1]}| > |L_{[2]}| > \cdots > |L_{[i]}| > \cdots > |L_{[N]}|$$

[i]: Index of *i*th best node (Very useful, compact notation)

- Node conveys its LLR in its packet when it transmits
- After each transmission, fusion node can terminate round



OTS is Energy-Efficient and Error Rate Optimal

Provably reduces average number of transmissions by at least 50% with same error rate as optimal scheme

- Intuition: Two reasons why OTS succeeds
  - Most informative measurements come earlier
  - Ordering provides information about measurements that are yet to be received

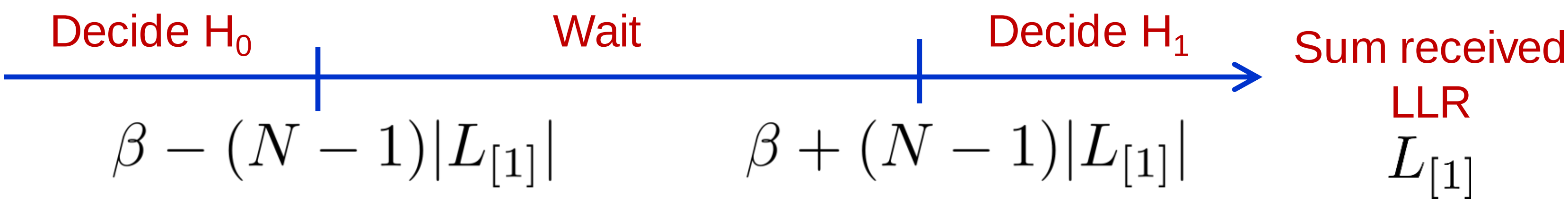
$$-|L_{[k]}| < L_{[j]} < |L_{[k]}|, \text{ for } j > k$$

Yet to be received LLR

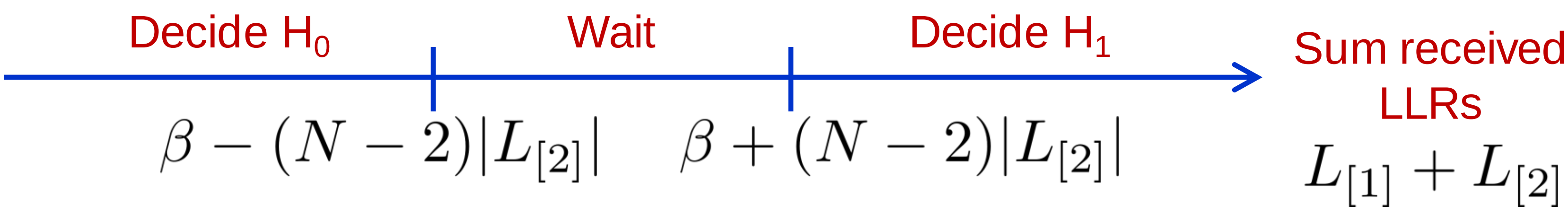
Last received LLR

Decision Rules: Illustration

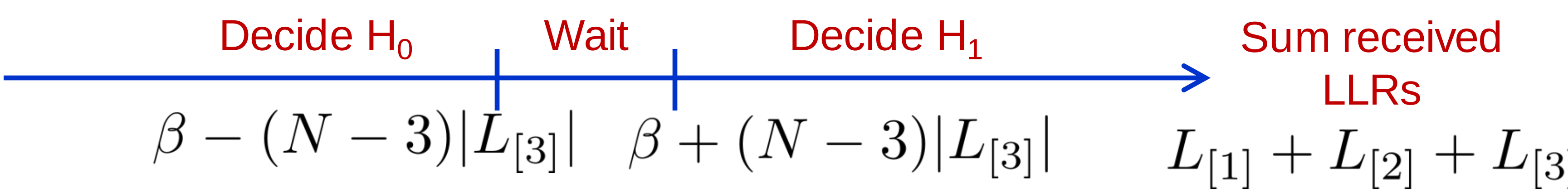
1. After receiving node [1] LLR:



2. After receiving nodes [1] and [2] LLRs:



3. After receiving nodes [1], [2], and [3] LLRs:



Decision Rules Derivation: Basic Idea

- Given: *k* best nodes transmitted.
- Idea: Bound yet-to-be-received LLRs

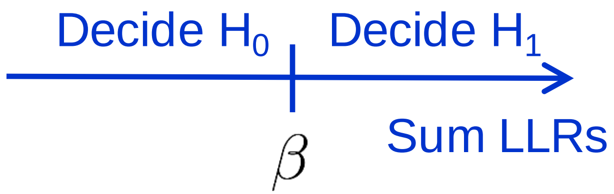
$$-|L_{[k]}| < L_{[j]} < |L_{[k]}|, \text{ for } j > k$$

- Implies bound on sum LLR $\sum_{j=1}^{k} L_{[j]} + \sum_{j=k+1}^{N} L_{[j]}$

$$< \sum_{j=1}^{k} L_{[j]} + (N-k)|L_{[k]}| \quad \text{(Upper bound)}$$

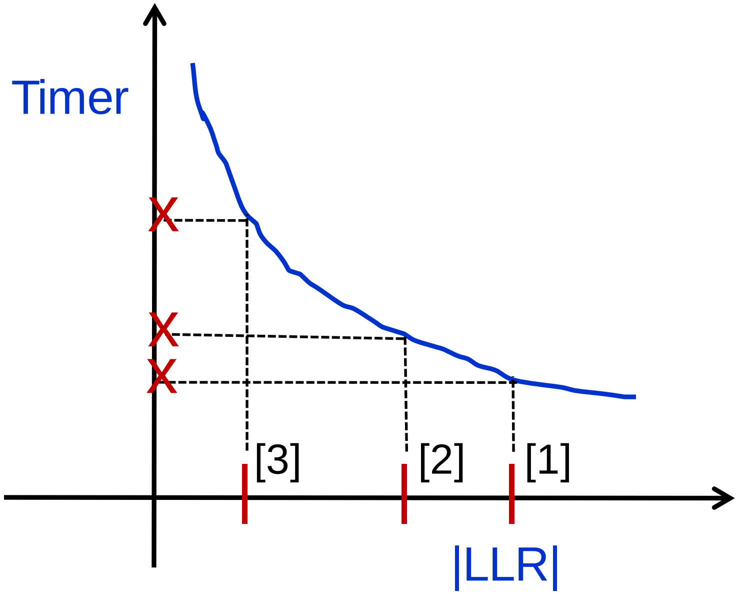
$$> \sum_{j=1}^{k} L_{[j]} - (N-k)|L_{[k]}| \quad \text{(Lower bound)}$$

- If upper bound < β, decide H_0
- If lower bound > β, decide H_1



OTS is Practically Implementable

1. Each node sets a timer depending on its |LLR|
2. Node transmits a packet when its timer expires



Key idea: Mapping is monotonically non-increasing

⇨ First timer to expire will be of the best node

[Bletsas et. al., TWC 2006][Shah, Mehta, Yim, TCom 2010]

Twist 1: Spatially Correlated Measurements

A More General Setting

Independent observations

- Node measurements are independent conditioned on hypothesis

$$H_1 : \quad \mathbf{y} \sim N(0, \sigma_1^2 \mathbf{I})$$
$$H_0 : \quad \mathbf{y} \sim N(0, \sigma_0^2 \mathbf{I})$$

Correlated observations

- Node observations are correlated conditioned on hypothesis
  - Ex.: Dense deployments of sensor nodes

$$H_1 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_1, \mathbf{R}_1)$$
$$H_0 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_0, \mathbf{R}_0)$$

$$\mathbf{R}_i = \gamma \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ & & & \\ \rho & \rho & \dots & 1 \end{bmatrix} \qquad \mathbf{R}_i = \gamma \begin{bmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ & & & \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{bmatrix}$$

Example: Optimal Decision Rule

- Optimal decision rule

$$H_1: \quad \mathbf{y} \sim N(\mathbf{0}, \mathbf{R})$$
$$H_0: \quad \mathbf{y} \sim N(\mathbf{0}, \sigma_0^2 \mathbf{I})$$

$$\log\left(\frac{f_{\mathbf{y}}(\mathbf{y}|H_1)}{f_{\mathbf{y}}(\mathbf{y}|H_0)}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \beta$$

- For Gaussian statistics, decision rule reduces to

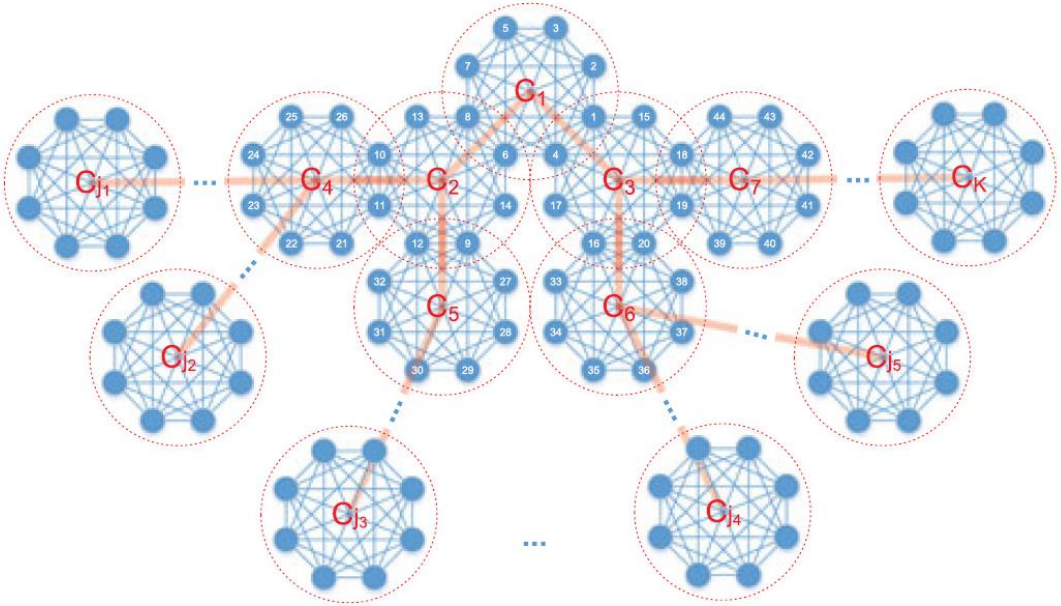
$$\mathbf{y}^T \underbrace{\left(\frac{1}{\sigma_0^2}\mathbf{I} - \mathbf{R}^{-1}\right)}_{\mathbf{G}} \mathbf{y} \underset{H_0}{\overset{H_1}{\gtrless}} \alpha$$

$$\sum_{i=1}^{N} \mathbf{G}_{ii} y_i^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{G}_{ij} y_i y_j \underset{H_0}{\overset{H_1}{\gtrless}} \alpha$$

- Separability exploited by OTS is lost due to cross terms

Limited Literature on Correlated Case

- Work by Blum's group
  - Decomposable Gaussian graphical models
  - Can be decomposed into maximal cliques with common nodes
  - Maximal clique $\rightarrow$ Cluster, Any node in clique $\rightarrow$ Cluster head
  - Decision statistic is separable across cluster heads



[Figure source: Zhang et al. TSP 2019]

Issues

- Applies only to decomposable graphical models
- Graph should consist of several smaller maximal cliques
  - Requires a sparse inverse correlation matrix to be effective
  - Not true in general
- Handles shift-in-mean or shift-in-covariance, not both
- Savings are only in the cluster head transmissions
  - Other nodes always transmit to their cluster heads

Problem Development

General case

$$H_1 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_1, \mathbf{R}_1)$$
$$H_0 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_0, \mathbf{R}_0)$$

Special cases

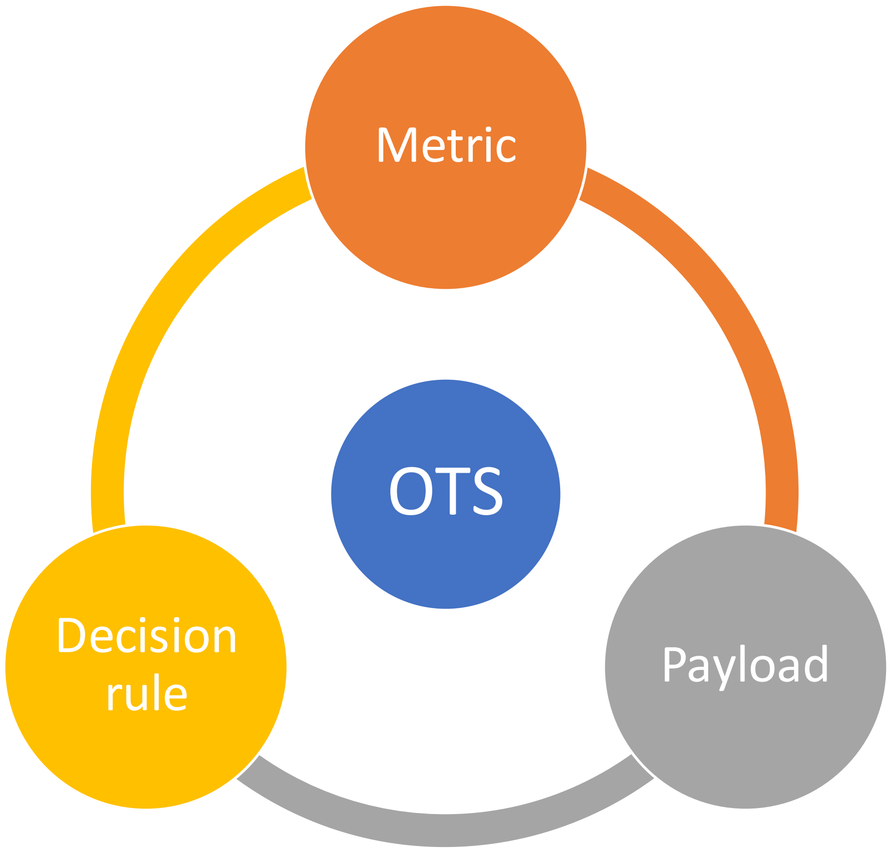
- Shift-in-covariance

$$H_1 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{R}_1)$$
$$H_0 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{R}_0)$$

- Shift-in-mean

$$H_1 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_1, \mathbf{R})$$
$$H_0 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_0, \mathbf{R})$$

OTS: Three Pillars and Mathematical Principle



Mathematical principle

- Previously: Decision statistic is separable
- Now: Use decision statistic bounds that are separable

Shift-in-Covariance Detection

- Optimal decision rule:

$$\mathbf{z}^T \mathbf{G} \mathbf{z} \underset{H_0}{\overset{H_1}{\gtrless}} \beta'$$

$$H_1 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{R}_1)$$
$$H_0 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{R}_0)$$

$z_i = y_i - \mu_i$ (Mean-shifted measurement)

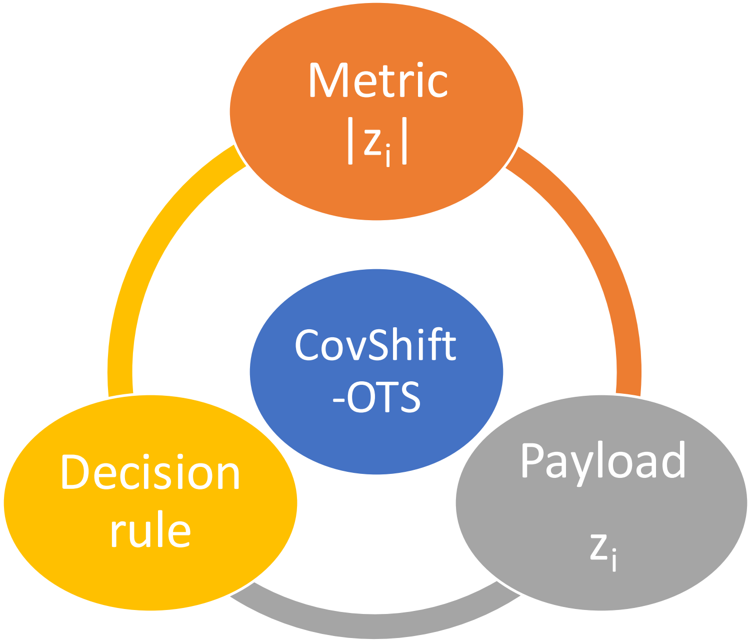
$$\mathbf{G} = \mathbf{R}_0^{-1} - \mathbf{R}_1^{-1}$$

CovShift-OTS

- Metric: $|z_i|$

$$|z_{[1]}| > |z_{[2]}| > \cdots > |z_{[N]}|$$

- Payload: z_i



Decision Rules: Eigenvalue Based Approach

- Lower and upper bounds on decision statistic

$$\lambda_{\min}(\mathbf{G})\mathbf{z}^T\mathbf{z} \leq \mathbf{z}^T\mathbf{G}\mathbf{z} \leq \lambda_{\max}(\mathbf{G})\mathbf{z}^T\mathbf{z}$$

- In terms of received and yet-to-be-received measurements

$$\mathbf{z}^T\mathbf{z} = \sum_{i=1}^{k} z_i^2 + \sum_{j=k+1}^{N} z_j^2$$

- Ordering gives information about yet-to-be-received measurements

$$0 \leq z_{[j]}^2 < z_{[k]}^2, \quad 1 \leq k < j < N$$

- Can derive lower and upper bounds on decision statistic in terms of received payloads

Refined Approach: Bound Sub-Matrices

- Write the decision statistic in terms of received and yet-to-be-received measurements and their cross-terms
- Bound only unknown terms

$$\mathbf{d} = \mathbf{z}^T \mathbf{G} \mathbf{z}$$

$$= \mathbf{r}_k^T \mathbf{P}_k \mathbf{r}_k + \mathbf{u}_{N-k}^T \mathbf{Q}_{N-k} \mathbf{u}_{N-k} + \mathbf{v}_k^T \mathbf{u}_{N-k}$$

$$\mathbf{r}_k = \begin{bmatrix} z_{[1]} & z_{[2]} & \cdots & z_{[k]} \end{bmatrix}$$

Received measurements

$$\mathbf{u}_{N-k} = \begin{bmatrix} z_{[k+1]} & z_{[k+2]} & \cdots & z_{[N]} \end{bmatrix}$$

Yet-to-be-received measurements [Unknown]

$$\mathbf{G} = \begin{bmatrix} \mathbf{P}_k & ? \\ ? & \mathbf{Q}_{N-k} \end{bmatrix}$$

Shift-in-Mean Detection

- Optimal decision rule:

$$\sum_{i=1}^{N} s_i y_i \underset{H_0}{\overset{H_1}{\gtrless}} \beta'$$

$$\mathbf{s} = 2\mathbf{R}^{-1}(\boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(0)})$$

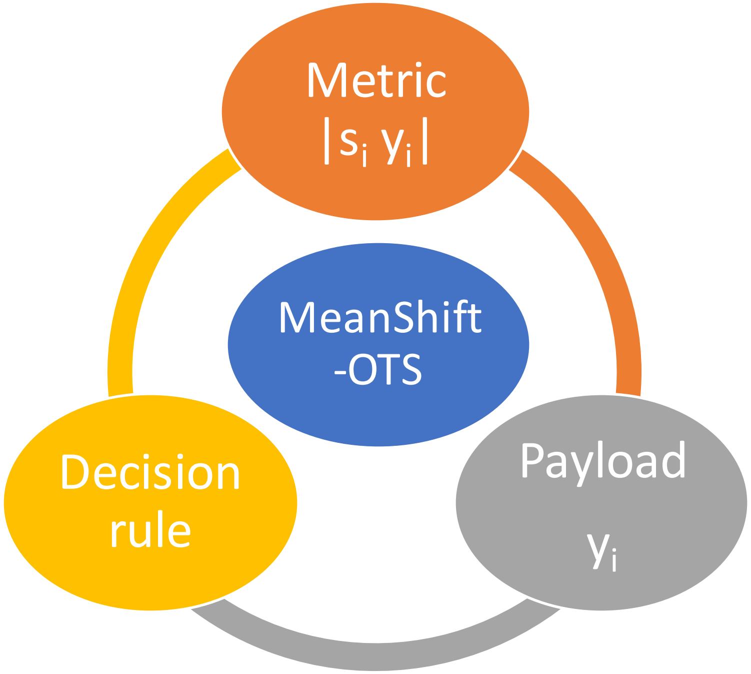
$$H_1: \quad \mathbf{y} \sim N(\boldsymbol{\mu}_1, \mathbf{R})$$
$$H_0: \quad \mathbf{y} \sim N(\boldsymbol{\mu}_0, \mathbf{R})$$

MeanShift-OTS

- Metric: $w_i = |s_i \, y_i|$

$$w_{[1]} > w_{[2]} > \cdots > w_{[N]}$$

- Payload: y_i



General Case: Correlation-Aware OTS

- Bounds are more involved, but can be worked out

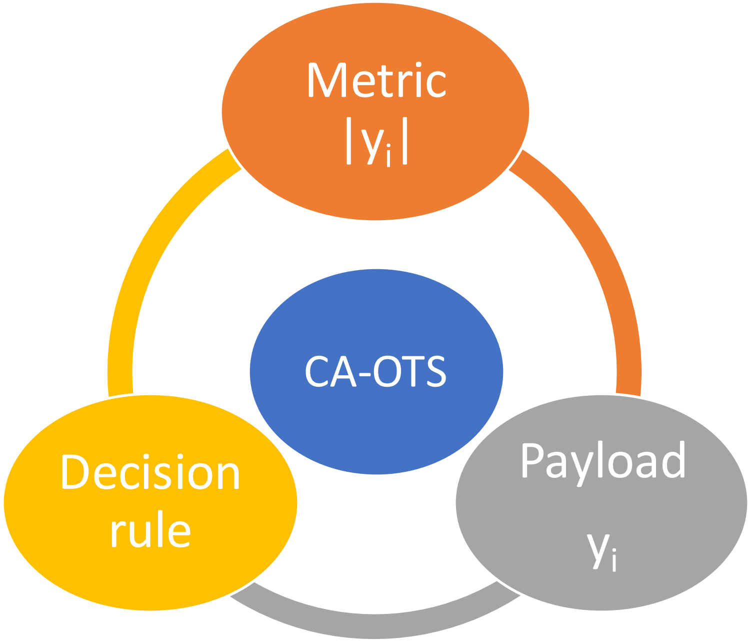
$$H_1 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_1, \mathbf{R}_1)$$
$$H_0 : \quad \mathbf{y} \sim N(\boldsymbol{\mu}_0, \mathbf{R}_0)$$

CA-OTS

- Metric: $|y_i|$

$$|y_{[1]}| > |y_{[2]}| > \cdots > |y_{[N]}|$$

- Payload: y_i



Product Correlation Model [Beaulieu]

- Different pairs of nodes have different correlation coefficients

$$\mathbf{R} = \gamma \begin{bmatrix} 1 & \theta_1\theta_2 & \dots & \theta_1\theta_N \\ \theta_2\theta_1 & 1 & \dots & \theta_2\theta_N \\ & & & \\ \theta_N\theta_1 & \theta_N\theta_2 & \dots & 1 \end{bmatrix}$$

- Subsumes uniform correlation model [Mallik]

$$\mathbf{R} = \gamma \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ & & & \\ \rho & \rho & \dots & 1 \end{bmatrix}$$

Upper Bound: Ave. Number of Transmissions

- Shift-in-covariance: CovShift-OTS with

$$H_1 : \; R_1(i,j) = \gamma_1 \theta_i \theta_j$$ (Product correlation model)

$$H_0 : \; \mathbf{R}_0 = \gamma_0 \mathbf{I}$$

  – Average number of transmissions $\rightarrow$ 1 as SNR $(\gamma_1/\gamma_0) \rightarrow \infty$

- Shift-in-mean: MeanShift-OTS with

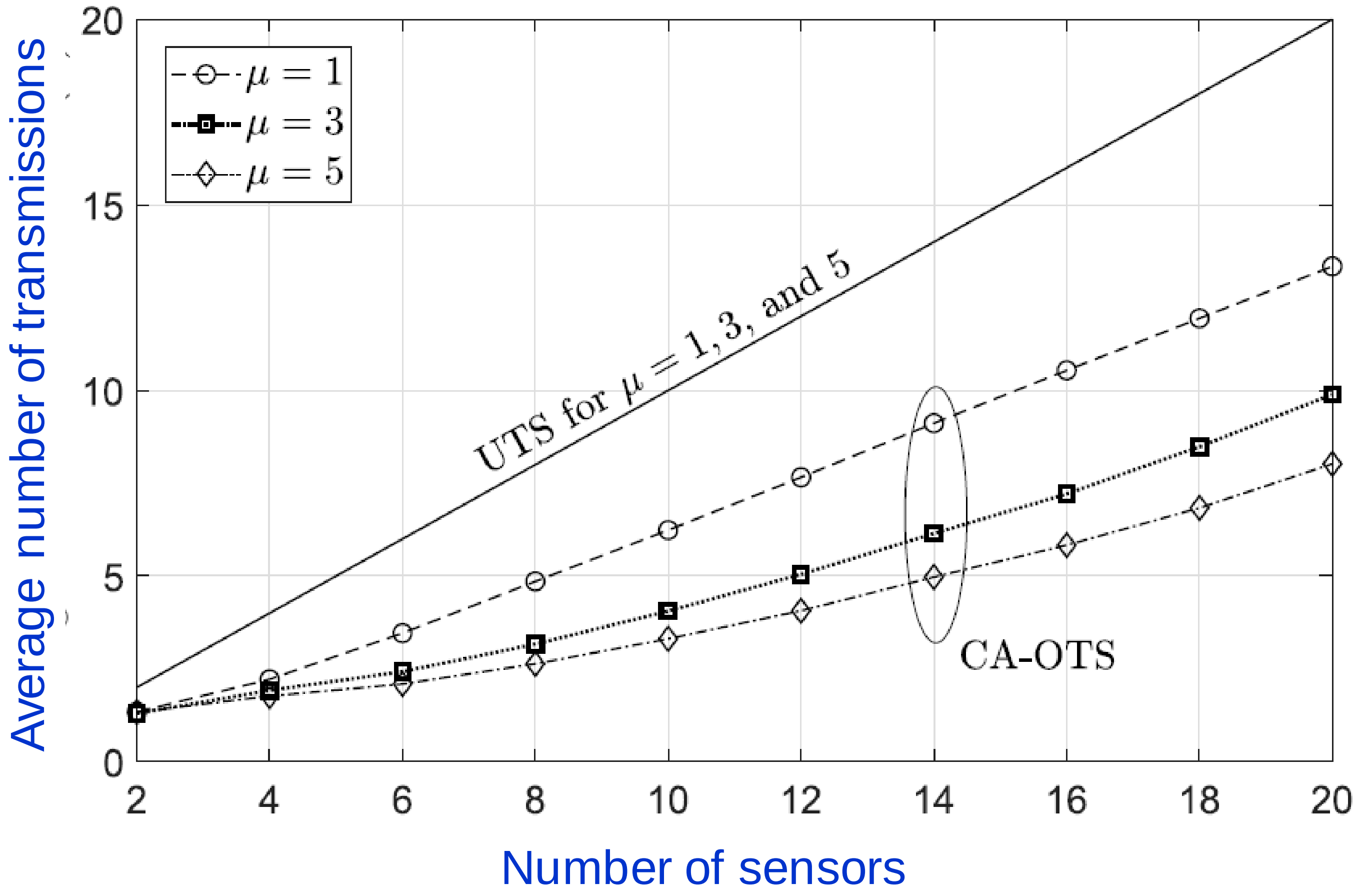
$$\mu_1^{(1)} = \cdots = \mu_N^{(1)} = \mu$$

$$\mu_1^{(0)} = \cdots = \mu_N^{(0)} = 0$$

Uniform correlation model with equal priors

  – Average number of transmissions $\rightarrow$ N/2 as $\mu \rightarrow \infty$

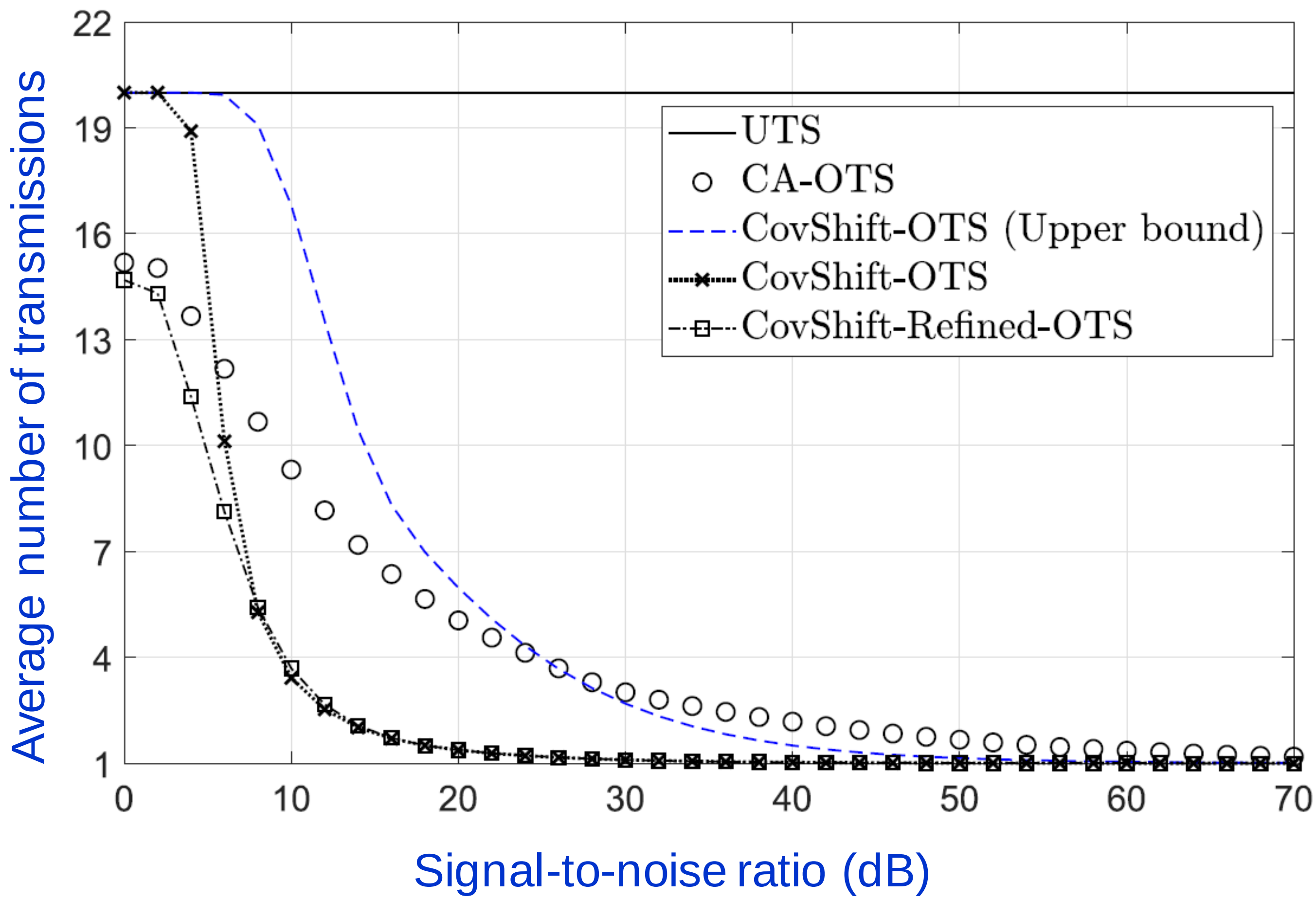
Performance Benchmarking: General Case



- Markedly fewer transmissions than OTS

[Setting: H_1: Product correlation model with non-identical sensors with mean $\mu[1\ 1\ \ldots\ 1]$, H_0: Uniform correlation model with mean $\mathbf{0}$]

Shift-in-Covariance: Role of Metric



20 sensors

- Using refined metric reduces average number of transmissions even further

[Setting: H_1: Product correlation model, H_0: noise]

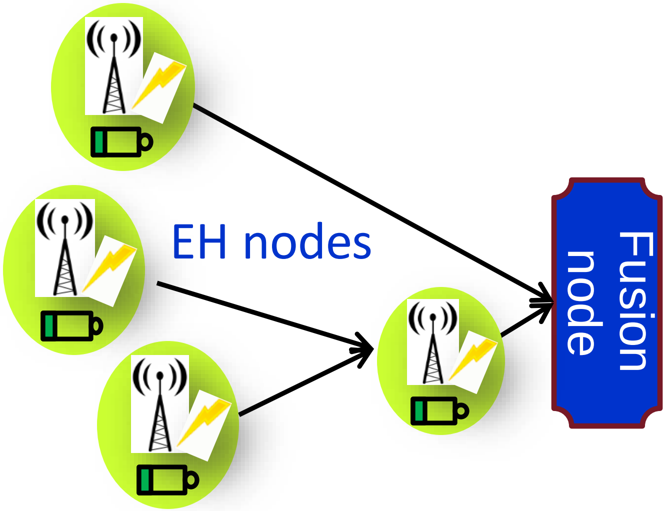
OTS is Not …

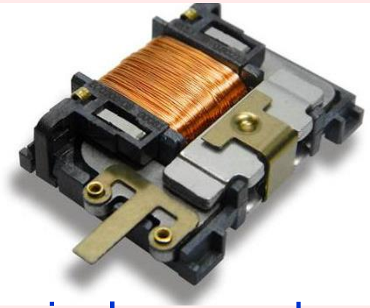
- Reduced complexity detection
  - No projection of observed vector **y** onto a basis possible since no node knows the entire **y**
- Sequential detection
  - Similarity: Compare accumulated decision statistic with two thresholds
  - But, nodes transmit in a random order
  - Often thresholds are designed assuming there are enough nodes
    - Wald’s rules

Twist 2: In Energy Harvesting Wireless Sensor Networks

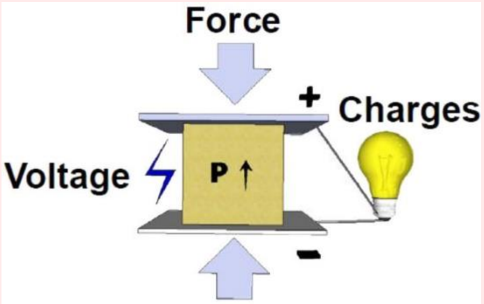
Energy Harvesting Wireless Sensor Networks

- Nodes harvest energy from the environment
  - Use renewable energy sources
- Can store harvested energy
- Use harvested energy for sensing, processing, and communication

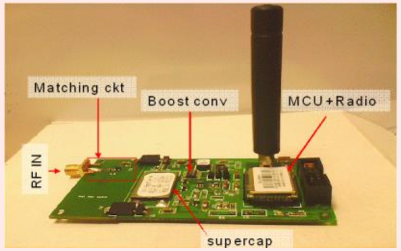




Mechanical energy harvester (Enocean)



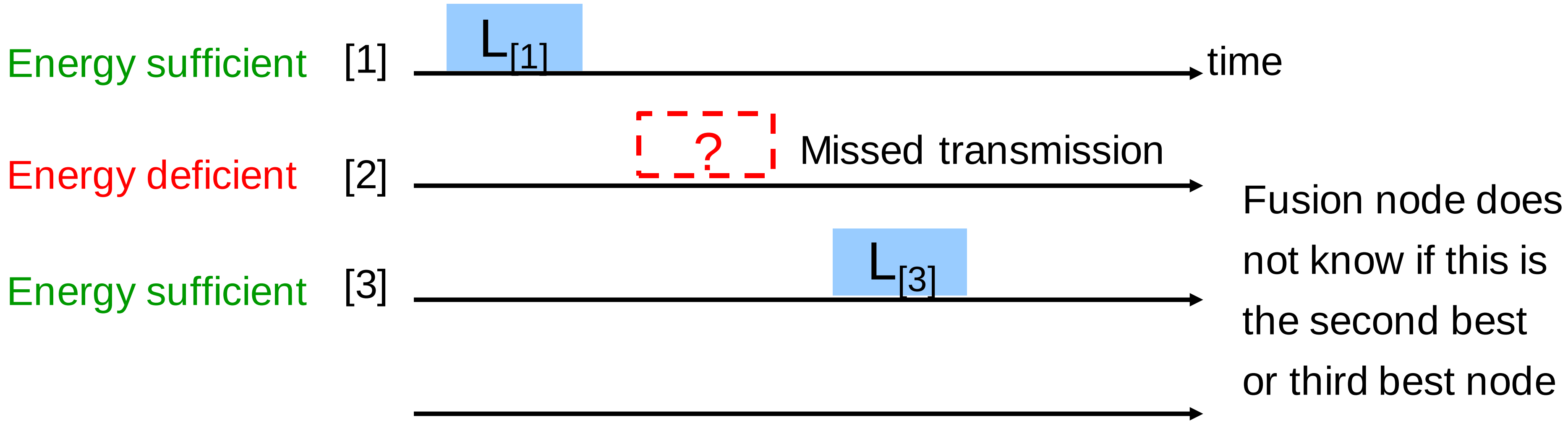
Piezoelectric effect harvester



RF energy harvester (ZEN lab @ IISc)

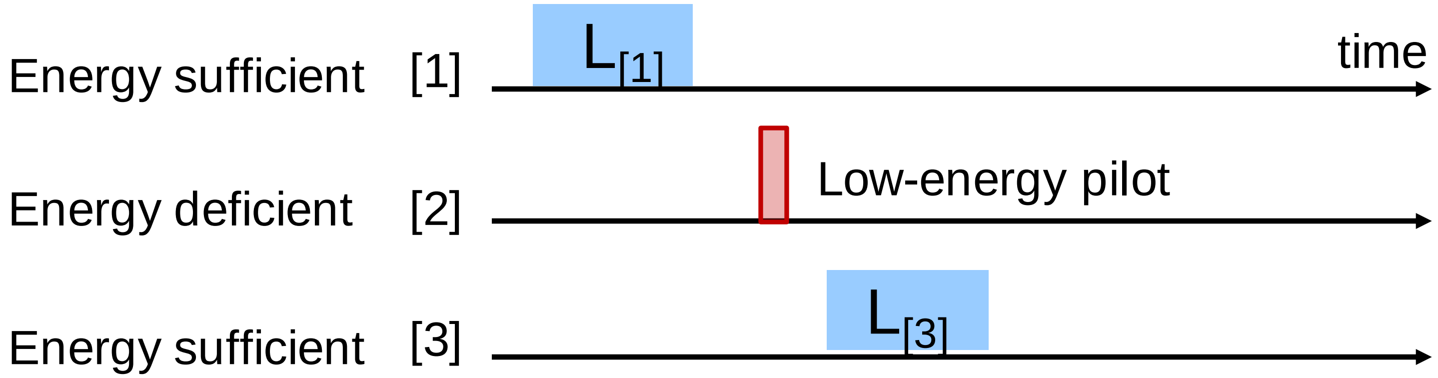
OTS Breaks Down in EH WSNs!

- Energy harvesting nodes occasionally run out of energy
  - Energy harvested is random
- Missing transmissions can mess up the sequence of ordered transmissions!



New Scheme for EH WSNs

- Fix: If low on energy, transmit a low-energy pilot/tone
  - Needs much less transmit energy than a data packet
  - Have a small, separate energy reserve for transmitting pilots
- Fusion node detects pilot and waits for next measurement

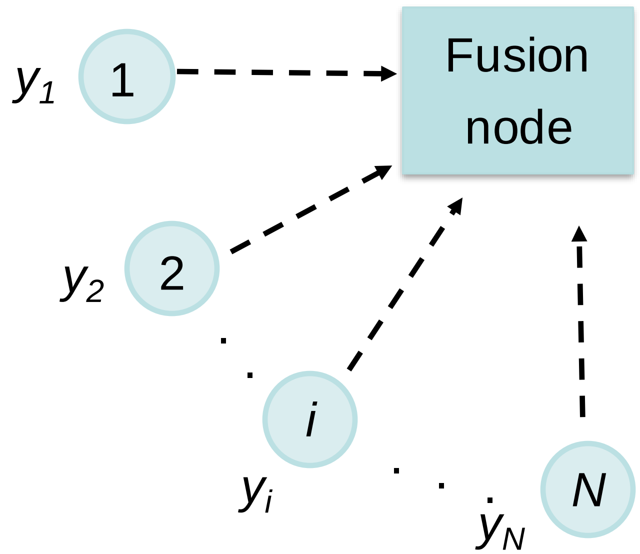


General System Model

- Measurement at sensor node *i*

$$y_i \sim \begin{cases} f(y_i|H_1), & \text{under } H_1 \\ f(y_i|H_0), & \text{under } H_0 \end{cases}$$

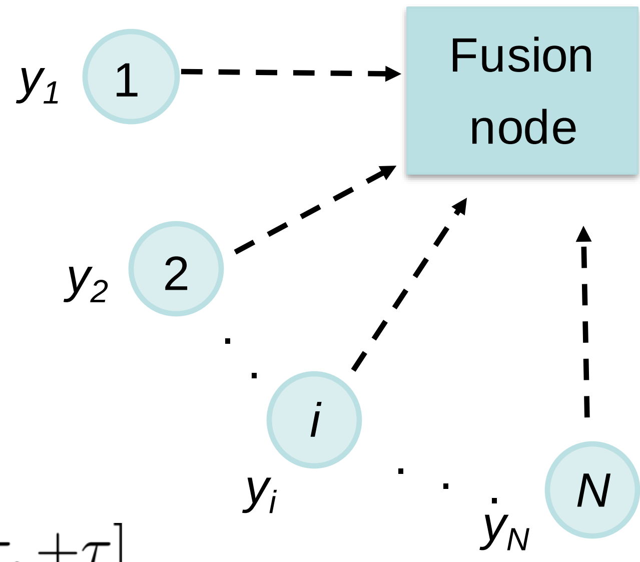
$$L_i = \log \frac{p(y_i|H_1)}{p(y_i|H_0)}$$



- Assumption: LLRs are bounded $|L_i| \leq \Psi$
  - Easily holds in practice
- Measurements conditioned on the hypotheses are mutually independent
  - Future work: Extension to spatially correlated model

Specialization: Gaussian Statistics

- Measurement at sensor node *i*



$$y_i = \begin{cases} x_i + n_i, & \text{under } H_1 \\ n_i, & \text{under } H_0 \end{cases}$$

$$x_i \sim \mathcal{N}(0, \sigma_x^2),\ n_i \sim \mathcal{N}(0, \sigma_n^2)$$

- Measurements are truncated $y_i \in [-\tau, +\tau]$
  - Reason: Readings of a sensor fall within a range
- Log-likelihood ratio at EH sensor node *i*

$$L_i = \ln\left(\frac{\sigma_n}{\sigma_1}\right) + \ln\left(\frac{1 - 2Q\left(\tau/\sigma_n\right)}{1 - 2Q\left(\tau/\sigma_1\right)}\right) + \frac{y_i^2}{2}\left(\frac{\sigma_x^2}{\sigma_1^2\sigma_n^2}\right)$$

$$\sigma_1 \triangleq \sqrt{\sigma_x^2 + \sigma_n^2}$$

Gaussian Model: Equivalent Hypothesis Test

- Bayesian hypothesis test $\sum_{i=1}^{N} L_i \underset{H_0}{\overset{H_1}{\gtrless}} \beta$ is equivalent to

$$\sum_{i=1}^{N} y_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

where

$$\lambda = \frac{2\sigma_n^2\sigma_1^2}{\sigma_x^2}\left[\beta + N\ln\left(\frac{\sigma_1}{\sigma_n}\right) + N\ln\left(\frac{1-2Q\left(\tau/\sigma_1\right)}{1-2Q\left(\tau/\sigma_n\right)}\right)\right]$$

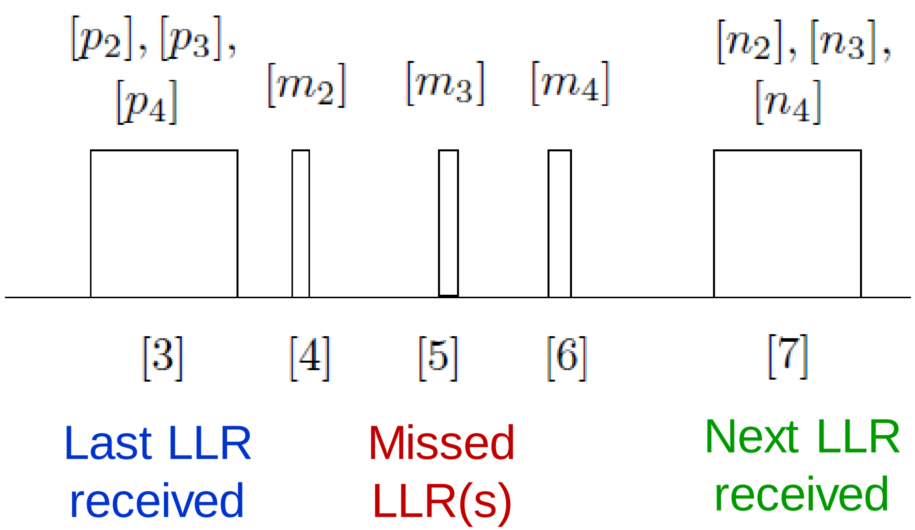
- New metric: $\Theta_i = y_i^2$
  - No need for taking absolute value of LLR
  - Advantage: Removes sign ambiguity
  - Bounded between 0 and τ^2

New Decision Rules for EH WSNs

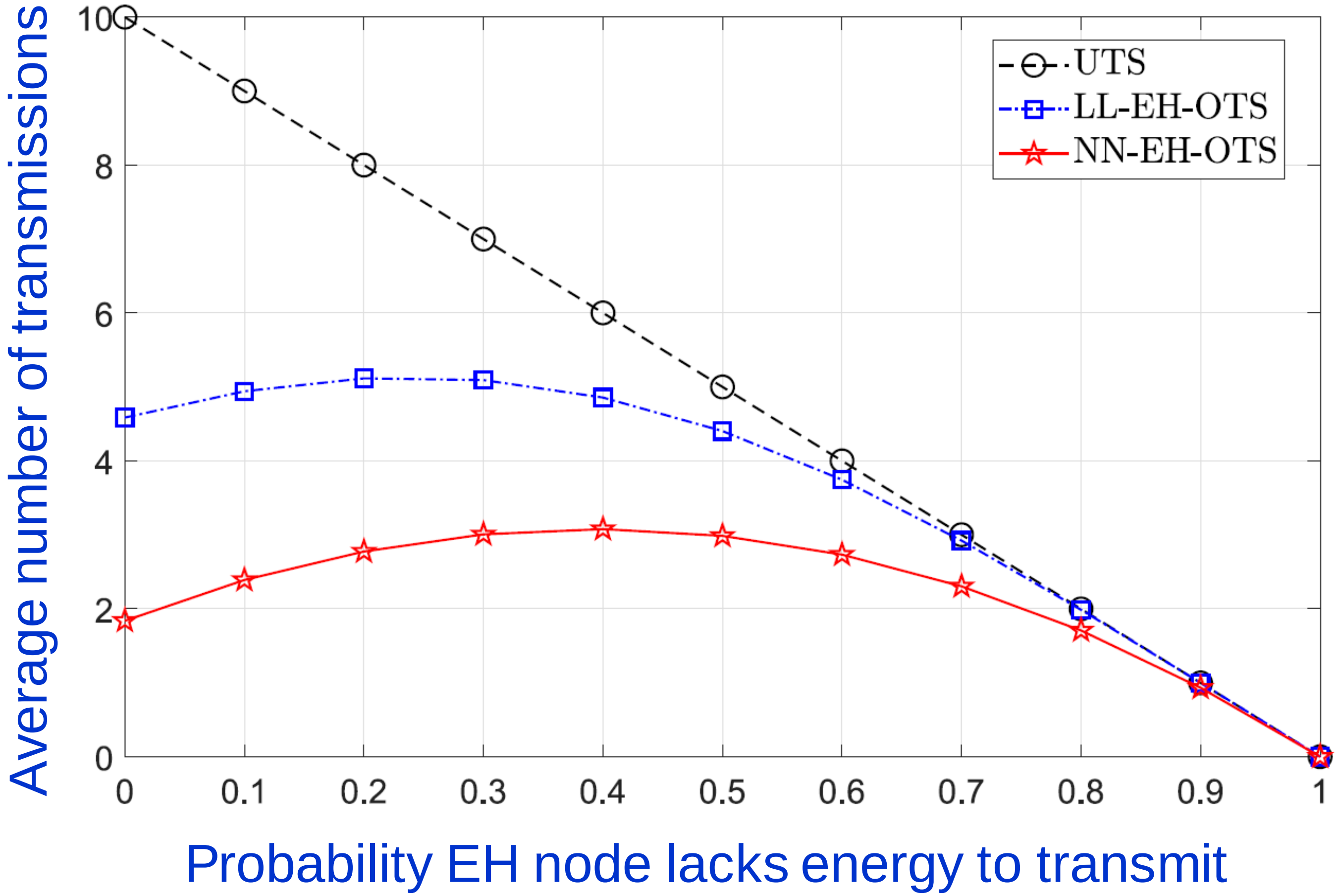
$$\text{Decide } H_1 \text{ if : } \sum_{\substack{i=1 \\ i \notin \{m_1,\dots,m_j\}}}^{k} \Theta_{[i]} > \lambda - \sum_{l=1}^{j} \Theta^2_{[n_l]}$$

$$\text{Decide } H_0 \text{ if : } \sum_{\substack{i=1 \\ i \notin \{m_1,\dots,m_j\}}}^{k} \Theta_{[i]} < \lambda - \sum_{l=1}^{j} \Theta_{[p_l]} - (N-k)\Theta_{[k]}$$

Else, wait for next measurement



Average Number of Transmissions Comparison

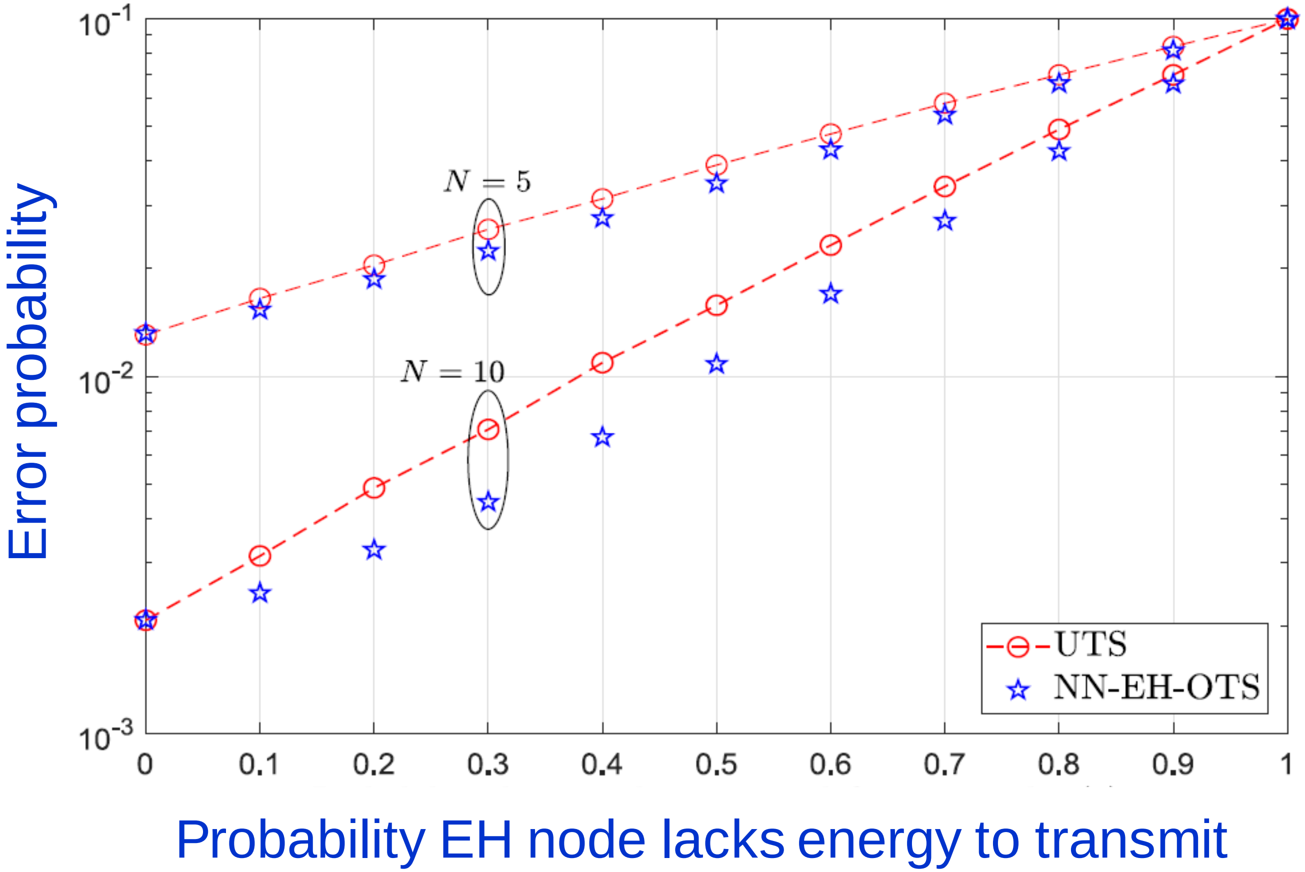


SNR = 10 dB

$Pr(H_1) = 0.1$

- NN-EH-OTS: Requires far fewer transmissions than LL-EH-OTS and unordered transmissions scheme (UTS)
  - Removing sign ambiguity in metric helps!

Error Probability Comparison



SNR = 10 dB

$Pr(H_1) = 0.1$

- NN-EH-OTS even reduces the error probability compared to unordered transmissions (UTS)
- Surprising double benefit of ordering for EH WSNs

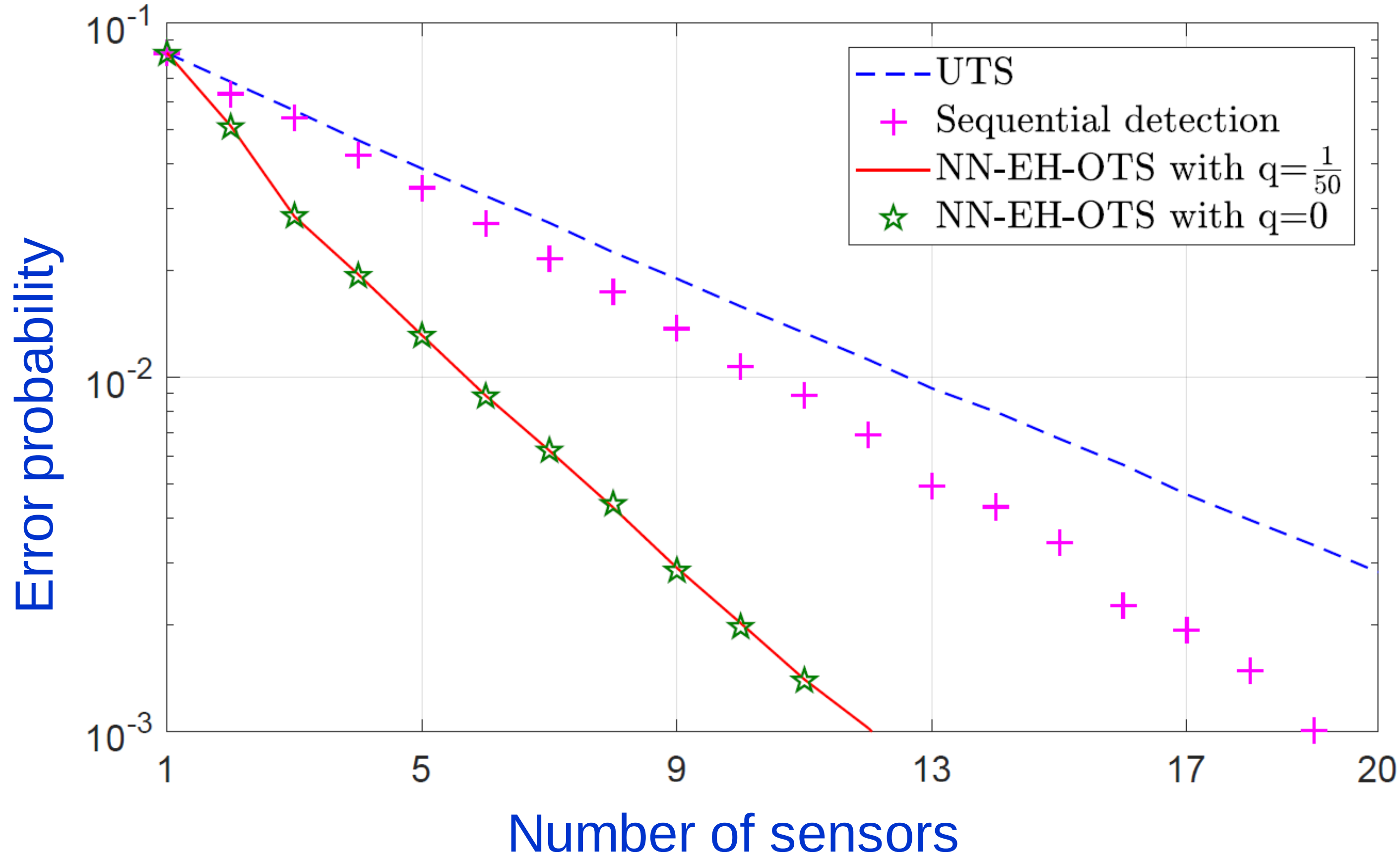
Capturing Time Dynamics

No longer pre-specify transmission miss probability

Physically realistic simulation that tracks battery evolution of each EH node and the coupling between their transmissions

- Energy harvesting model: Energy harvested in a round with probability *p*
- Energy storage model: Battery with a finite capacity
- Transmission model: E Joules required to transmit packet, qE required to transmit a pilot (q << 1)

NN-EH-OTS vs. Sequential Detection & UTS



- Error probability is now the key performance
  - Average number of transmissions and transmission miss probability are both scheme-dependent
- Much lower error probability than UTS, sequential detection

Conclusions

- OTS saves energy without compromising on error rate
  - Key ideas: Separability of decision statistic, and timer scheme
  - Three pillars: Metric, payload, and decision rule
- OTS needs to be redesigned in energy harvesting sensor networks due to missing transmissions
  - Reduces average number of transmissions and also error rate!
- OTS needs to be redesigned for correlated measurements
  - Key idea: Separability of bounds on the decision statistic
  - Reduces average number of transmissions substantially

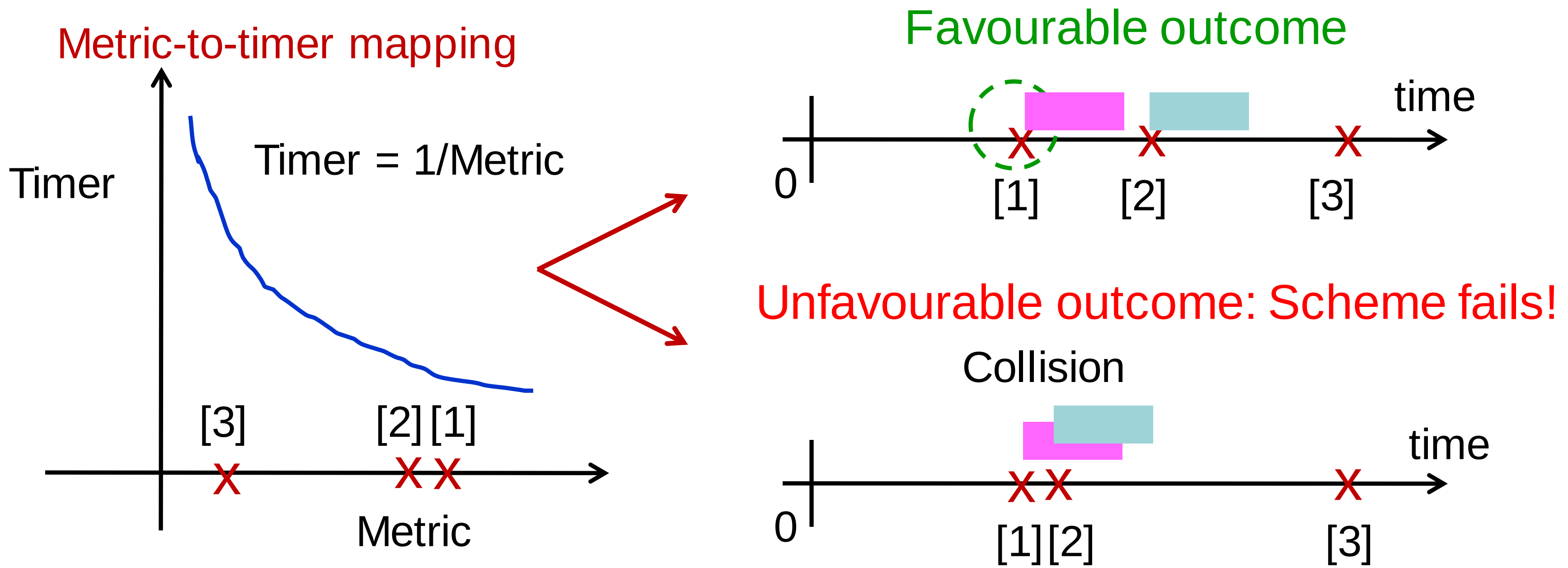
Ordering is a powerful MAC layer technique to improve performance of the physical layer

For More Details: ece.iisc.ac.in/~nextgenwrl

1. S. Sen Gupta, Neelesh B. Mehta, "Ordered Transmissions Schemes for Detection in Spatially Correlated Wireless Sensor Networks," *To appear in IEEE Trans. on Communications*, 2021.
2. S. Sen Gupta, S. Pallapothu, N. B. Mehta, "Ordered Transmissions for Energy-Efficient Detection in Energy Harvesting Wireless Sensor Networks," *IEEE Trans. on Communications*, Apr. 2020.
3. R. Talak, N. B. Mehta, "Feedback Overhead-Aware, Distributed, Fast, and Reliable Selection," *IEEE Trans. on Communications*, Nov. 2012.
4. V. Shah, N. B. Mehta, R. Yim, "Optimal Timer Based Selection Schemes," *IEEE Trans. on Communications*, Jun. 2010.
5. A. Anand, N. B. Mehta, "Quick, Decentralized, Energy-Efficient One-Shot Max Function Computation Using Timer-Based Selection," *IEEE Trans. on Communications*, Mar. 2015.

But, There is a Catch!

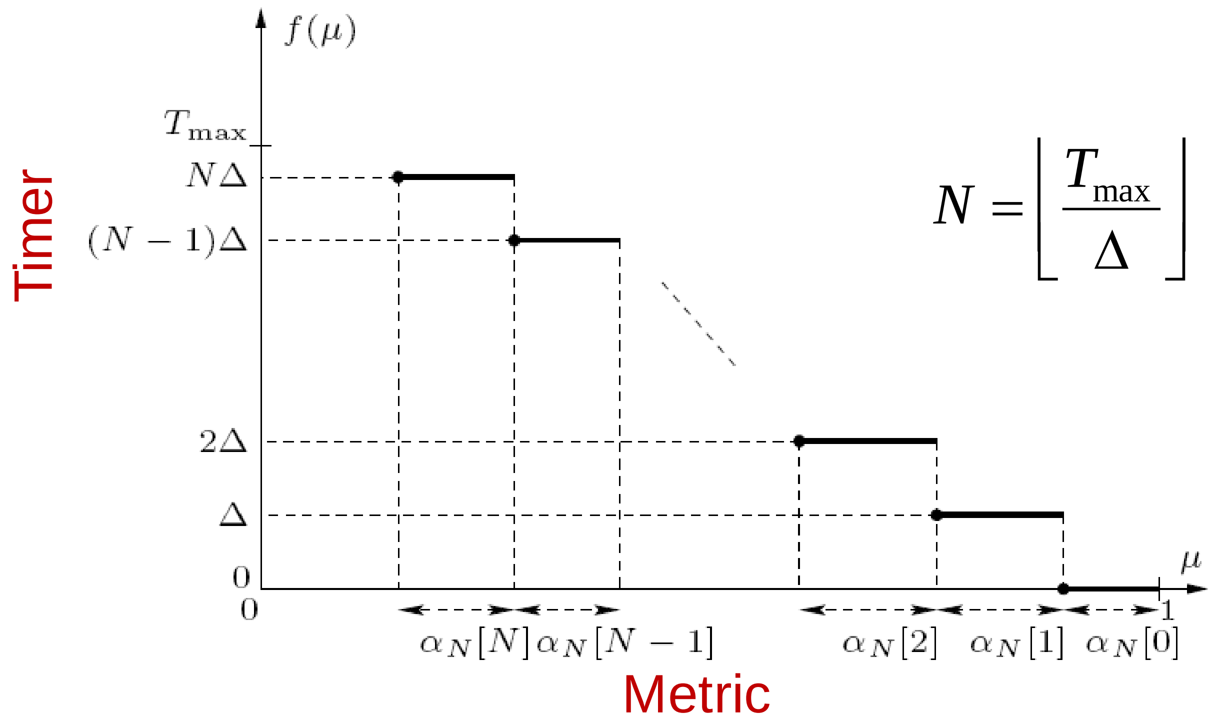
- Timer packets can collide in a wireless channel
  - Scheme can fail to select the best node or reveal proper order



- Success probability depends on metric-to-timer mapping

Research Question: Optimal Timer Mapping?

Question: Which timer mapping maximizes the probability of selecting the best node?

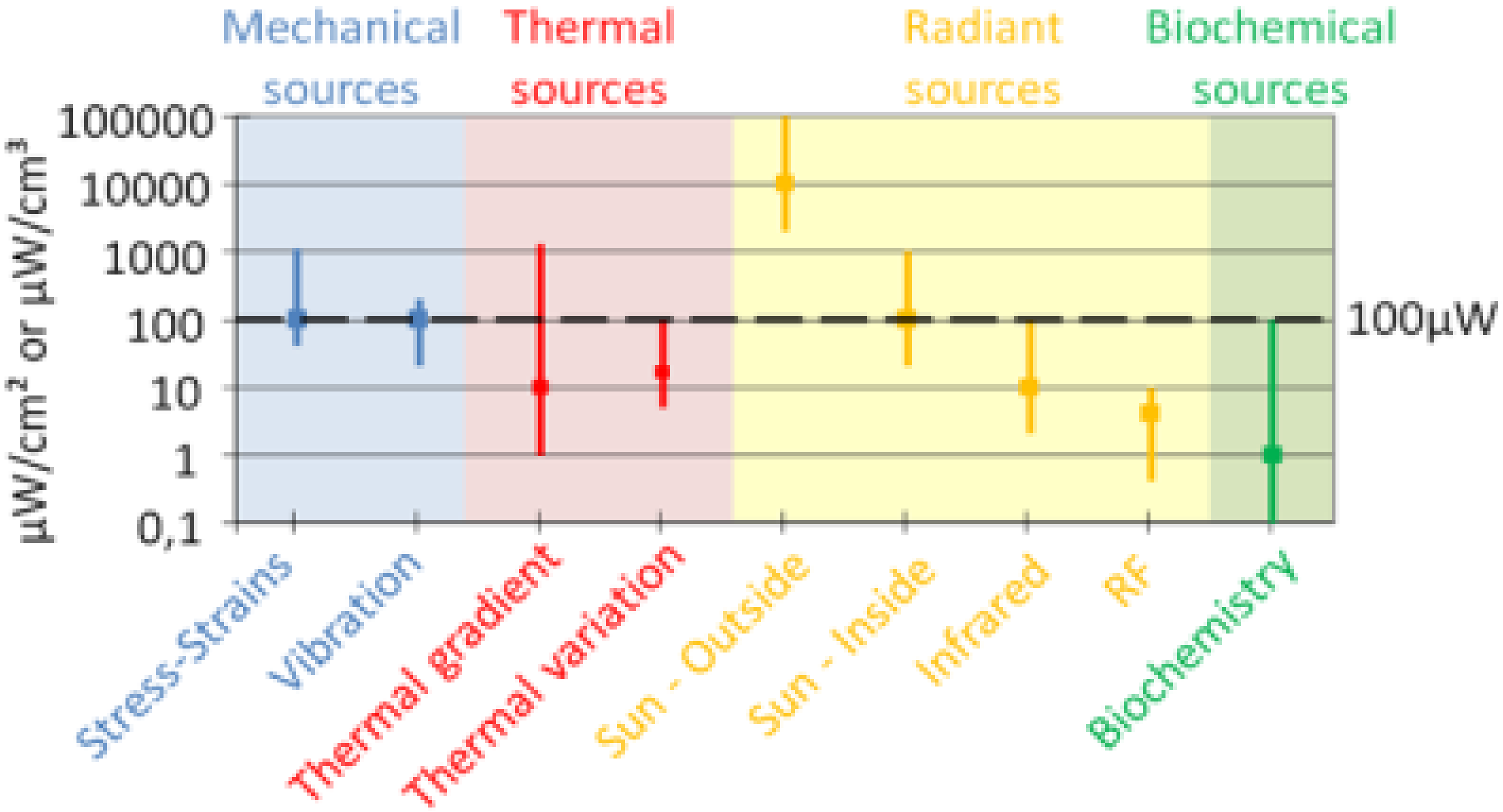


$$N = \left\lfloor \frac{T_{\max}}{\Delta} \right\rfloor$$

Timers can expire only at *0, Δ, 2Δ,…, NΔ*

(Mapping looks like a stair case with uneven lengths)

How Much Energy Can Be Harvested?



- Ballpark range of energy harvesting: 10-100 µW/cm^3

[Source: EE Times, Boisseau and Despesse, 2012]

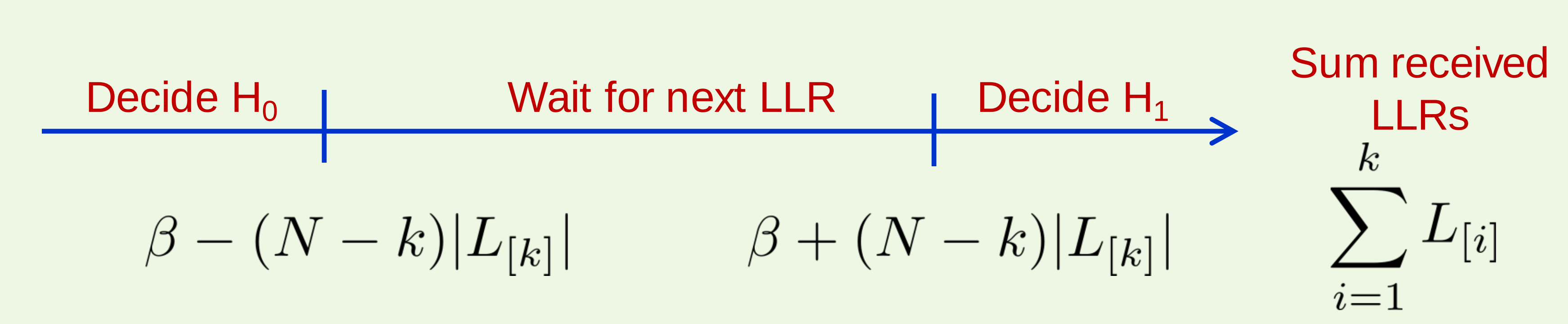
New Decision Rules: Formal Statement

- When LLRs from best *k* nodes *[1],…,[k]* received:

$$\text{Decide } H_1 \text{ if}: \sum_{i=1}^{k} L_{[i]} > \beta + (N-k)|L_{[k]}|$$

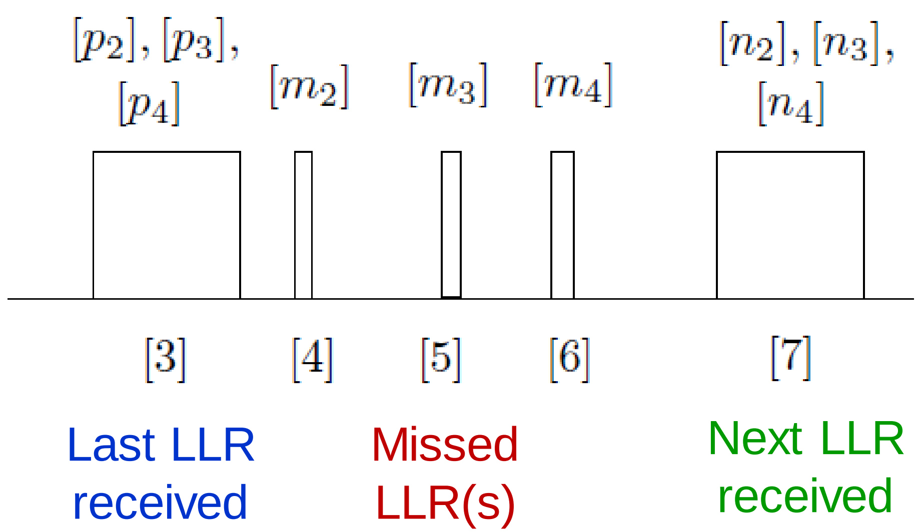
$$\text{Decide } H_0 \text{ if}: \sum_{i=1}^{k} L_{[i]} < \beta - (N-k)|L_{[k]}|$$

Else, wait for next measurement



Order Statistics: Track Missed Transmissions

- Missing transmissions: $[m_1], [m_2], \ldots, [m_j]$
  - i.e., m_1th best node, m_2th best node,…, m_jth best node lack energy
- For l^{th} missing transmission $[m_l]$
  - p_l: Last LLR received
  - n_l: Next LLR received



$$\Theta_{[p_l]} > \Theta_{[m_l]} > \Theta_{[n_l]}$$

Boundary cases:

$$\Theta_{[0]} = \tau^2, \ \Theta_{[N+1]} \triangleq 0$$

Result: New EH-OTS Decision Rules

- If EH node *[k]* last transmitted and *j* transmissions missed:

$$\text{declare } H_1 \text{ if}: \sum_{i=1, i \notin \{m_1, \ldots, m_j\}}^{k} \Theta_{[i]} > \lambda - \sum_{l=1}^{j} \Theta_{[n_l]}$$

$$\text{declare } H_0 \text{ if}: \sum_{i=1, i \notin \{m_1, \ldots, m_j\}}^{k} \Theta_{[i]} < \lambda - \sum_{l=1}^{j} \Theta_{[p_l]} - (N-k)\Theta_{[k]}$$

Achieve same error probability as having all nodes transmit (including the missing ones)!

Boundary cases:

- No more energy-sufficient nodes remain: Decide based on received metrics from energy-sufficient nodes (similar to UTS)
- No nodes transmit: Declare hypothesis with higher prior probability