Distributed Hypothesis Testing over Networks

Michèle Wigger
Telecom ParisTech

Joint work with R. Timo, S. Salehkalaibar, and L. Wang

Horizon Maths 2017—Mathematics and Networks
November 30, 2017

Example: Distributed Control-System for Smart Cars

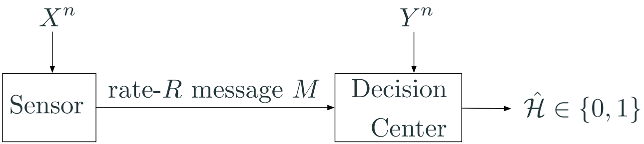
- Smart cars measuring speed, distance, road conditions
- Fixed road-side sensors measuring same parameters

- Intact car system: measurements highly correlated
- Erroneous car system: measurements independent

Task of Distributed Control-System

Decide on joint distribution underlying the observations

Distributed Hypothesis Testing



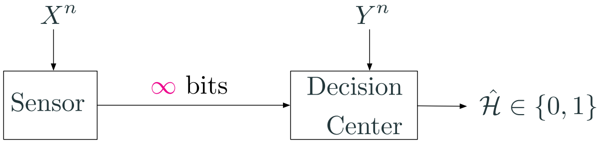
- "Normal situation" $\mathcal{H} = 0$: $(X^n, Y^n) \sim$ i.i.d. P_{XY}
- "Hazardous event" $\mathcal{H} = 1$: $(X^n, Y^n) \sim$ i.i.d. Q_{XY}

- Probability of false alarm: $\alpha_n = \mathbb{P}\big[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0\big] < \epsilon$
- Probability of miss detection: $\beta_n = \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\big] < 2^{-n\theta}$

Rate-Exponent Tradeoff $\theta^\star(R)$

Given $R > 0$, largest exponent θ that is achievable $\forall \epsilon > 0$

Local Hypothesis Testing

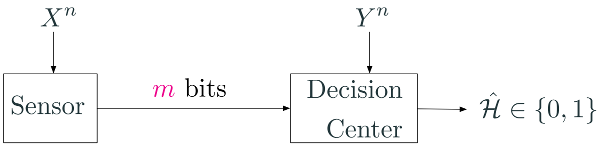


- Rate R is so large that sensor can send *all* X^n to decision center
- Decision center applies likelihood ratio test to *both* (X^n, Y^n)

$$\theta^\star(R=\infty) = D(P_{XY}||Q_{XY})$$

- Alternative: Decision center raises alarm if (X^n, Y^n) are typical (have good statistics) according to P_{XY}

Distributed Hypothesis Testing with $R = 0$ (Han'87)



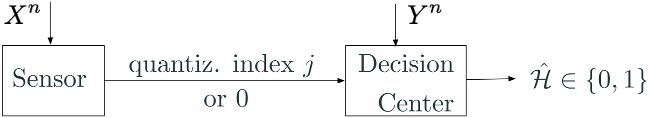
- $m = 1$ bit suffices
- Sensor: If X^n typical $\sim P_X \to$ send $M = 0$, otherwise $M = 1$
- Decision center raises alarm unless

$$X^n \text{ typical } \sim P_X \quad \text{and} \quad Y^n \text{ typical} \sim P_Y$$

- Optimal exponent:

$$\theta^\star(R = 0) = \min_{\substack{\pi_{XY}: \\ \pi_X = P_X \\ \pi_Y = P_Y}} D(\pi_{XY} || Q_{XY})$$

Distributed Hypothesis Testing with $R > 0$ (Han'87)



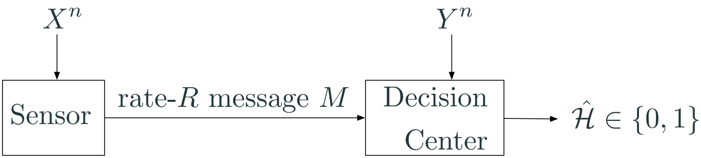
- Quantize X^n to $S^n(j)$
- If $(S^n(j), X^n)$ typical $\sim P_{S,X}$ send $M = j$, otherwise $M = 0$
- Decision center raises alarm $\hat{\mathcal{H}} = 1$ unless

  $(S^n(M), X^n)$ typical and $(S^n(M), Y^n)$ typical $\sim P_{SXY}$.

- Achievable exponent

$$\theta^\star(R) \geq \max_{\substack{P_{S|X}: \\ R \geq I(S;X)}} \min_{\substack{\pi_{SXY}: \\ \pi_{SX} = P_{SX} \\ \pi_{SY} = P_{SY}}} D(\pi_{SXY} || P_{S|X} Q_{XY})$$

Testing Against Independence (Csiszar-Ahlswede '86)



- $\mathcal{H} = 0$: $(X^n, Y^n) \sim$ i.i.d. P_{XY}
- $\mathcal{H} = 1$: $(X^n, Y^n) \sim$ i.i.d. $P_X P_Y$

Optimal Rate-Exponent Tradeoff

$$\theta^\star(R) = \max_{\substack{P_{S|X}: \\ R \geq I(S;X)}} I(S;Y)$$

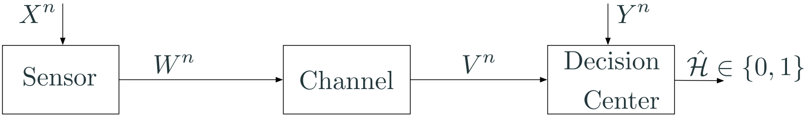
Using Wyner-Ziv Compression (Shimokawa, Han, Amari'94)

- Rx has side-info. Y^n about source X^n
- Wyner-Ziv coding: send a list of possible quantization indices $\rightarrow$ Rx decodes the correct index using Y^n
- Rx decodes with minimum empirical-entropy decoder (a universal capacity-achieving decoder)

$$\theta^\star(R) \geq \max_{\substack{P_{S|X}: \\ R \geq I(S;X|Y)}} \min \left\{ \min_{\substack{\pi_{SXY}: \\ \pi_{SX}=P_{SX} \\ \pi_{SY}=P_{SY}}} D(\pi_{SXY} \| P_{S|X} Q_{XY}), \right.$$
$$\left. \min_{\substack{\pi_{SXY}: \\ \pi_{SX}=P_{SX} \\ \pi_Y=P_Y \\ H(S|Y) \leq H_{\pi_{SY}}(S|Y)}} D(\pi_{SXY} \| P_{S|X} Q_{XY}) + R - I(S;X|Y) \right\}$$

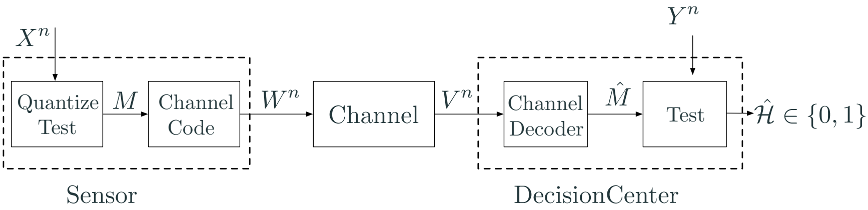
Testing over Noisy Channels

Distributed Testing over Noisy Channels



- Discrete memoryless channel $P_{V|W}$
  (can model fast fading, additive noise, etc.)

Our Coding and Testing Scheme for Noisy Channels



- "Quantize and test" with lists (Wyner-Ziv)
- *Unequal error protection* code protects $M = 0$ (Borade'08)
    - Send t^n if $M = 0$
    - Use codebook $\mathcal{C}_W = \{W^n(1), \ldots, W^n(2^{n(R)})\}$ if $M \neq 0$
- New error events related to erroneous decoding of 0-message

Result on Testing over Noisy Channels

Achievable Exponent (Salehkalaibar&W'2017)

$$\theta^\star \geq \max_{\substack{P_{S|X}, P_{TW}: \\ I(S;X|Y) \leq I(W;V|T)}} \min\left\{\theta_{\text{standard}}, \theta_{\text{wrong-dec.}}, \theta_{\text{missed-0}}\right\},$$

where

$$\theta_{\text{standard}} = \min_{\substack{\pi_{SXY}: \\ \pi_{SX}=P_{SX} \\ \pi_{SY}=P_{SY}}} D(\pi_{SXY}||Q_{XY}P_{S|X}),$$

$$\theta_{\text{wrong-dec.}} = \min_{\substack{\pi_{SXY}: \\ \pi_{SX}=P_{SX} \\ \pi_{Y}=P_{Y} \\ H(S|Y)\leq H_\pi(S|Y)}} D(\pi_{SXY}||P_{S|X}Q_{XY}) + I(W;V|T) - I(S;X|Y),$$

$$\theta_{\text{missed-0}} = D(P_Y||Q_Y) + I(W;V|T) - I(S;X|Y) + \mathrm{E}_T[D(P_{V|T}||P_{V|W=T})]$$

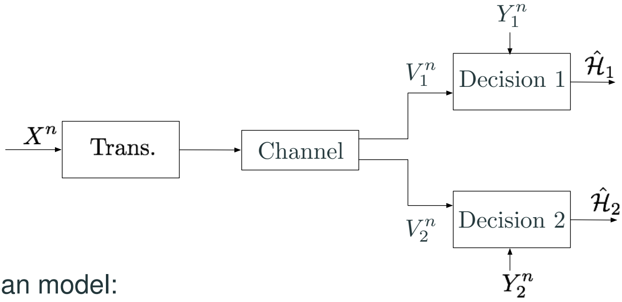
Quantization rate R limited by $I(W;V|T)$.

Penalty because of Noisy Channel

- Sometimes noisy channel only limits communication rate:
  - In case of large missed-0 exponent
  - When sensor cannot decide
- Noisy channel can severely limit error exponent
  - In case of small missed-0 exponent
  - Sensor decision very important

Testing at Multiple Centers

Two Simultaneous Hypothesis Tests



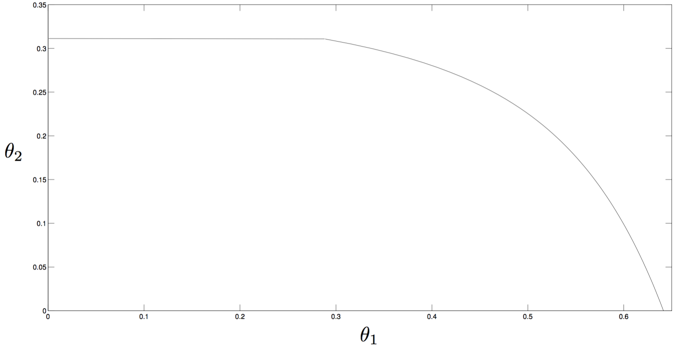
Setup can model:

- Two different decision centers
- Single decision center with uncertain $P_{XY} \in \{P_{XY_1}, P_{XY_2}\}$

Tension: Communication needs to serve both decisions!
E.g.: find quantization that is useful for both centers

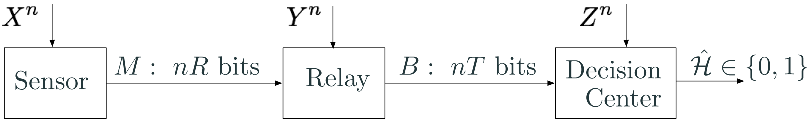
Tradeoff in Exponents Region (Salehkalaibar/W'/Timo'2017)

- Probabilities of false alarms: $\alpha_{i,n} = \mathbb{P}\big[\hat{\mathcal{H}}_i = 1 | \mathcal{H} = 0\big] < \epsilon$
- Probabilities of miss detections: $\beta_{i,n} = \mathbb{P}\big[\hat{\mathcal{H}}_i = 0 | \mathcal{H} = 0\big] < 2^{-n\theta_i}$
- Find optimal exponents region (θ_1, θ_2)
- For an example with Gaussian sources and channel



Testing over Multi-Hop Networks

Single-Relay Multi-Hop Channel

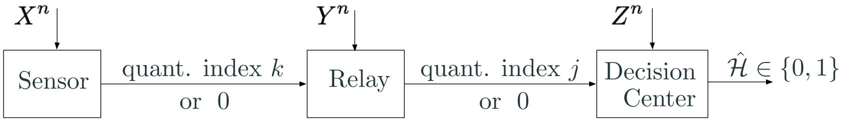


- $\mathcal{H} = 0$: $(X^n, Y^n, Z^n) \sim$ i.i.d. P_{XYZ}
- $\mathcal{H} = 1$: $(X^n, Y^n, Z^n) \sim$ i.i.d. Q_{XYZ}
- Probability of false alarm: $\alpha_n = \mathbb{P}\big[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0\big] < \epsilon$
- Probability of miss-detection: $\beta_n = \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\big] < 2^{-n\theta}$

Rate-Exponent Tradeoff $\theta^\star(R, T)$

Largest exponent θ achievable $\forall \epsilon > 0$ given rates $R, T \geq 0$

Markov chain $X^n \to Y^n \to Z^n$: Independent Tests



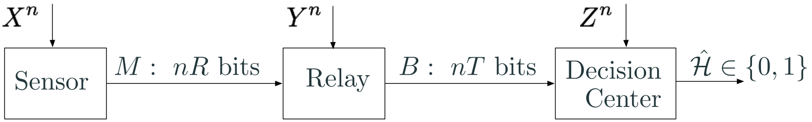
- Independent "Quantize and Test" at sensor and relay
- "Unanimous-Decision Forwarding": forward 0 if 0 received

Testing for Independence (Salehkalaibar/W'/Wang 2017)

$$\theta^\star(R,T) = \theta^\star_{\text{Sensor}\to\text{Relay}}(R) \quad + \quad \theta^\star_{\text{Relay}\to\text{Decision}}(T)$$

Accumulation of error exponents

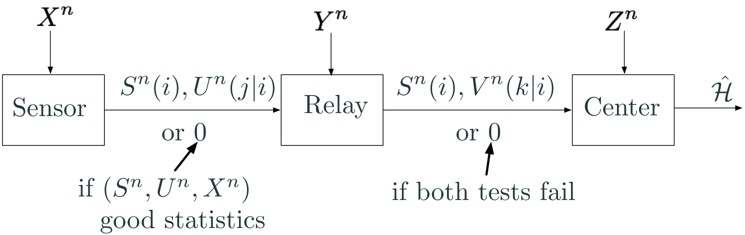
Challenges/Features of Solutions



- Each terminal will take a decision $\rightarrow$ alarm if one raises alarm
- Message sent from transmitter (sensor): tradeoff between serving decision center / providing useful information to relay
- Relay processing of Y^n and message from transmitter $\rightarrow$ reduce communication rate or send joint information to receiver

General Coding- and Testing-Scheme

- Coarse quantization of $X^n \rightarrow S^n(i)$
- Finer quantization of X^n given $S^n(i) \rightarrow U^n(j|i)$
- Joint quantization of $U^n(j|i)$, Y^n given $S^n(i) \rightarrow V^n(k|i)$



- Distributed quantization for cascade channels and unanimous-decision forwarding

An Achievable Exponent for Single-Relay Multi-Hop

Theorem

$$\theta^\star(R,T) \geq \max_{\substack{P_{US|X}, P_{V|SUY}: \\ R \geq I(US;X) \\ T \geq I(X;S) + I(V;YU|S)}} \min_{\substack{\pi_{SUVXYZ}: \\ \pi_{SUX} = P_{SUX} \\ \pi_{SVUY} = P_{SVUY} \\ \pi_{SVZ} = P_{SVZ}}} D(\pi_{SUVXYZ} \| P_{SU|X} P_{V|SUY} Q_{XYZ}).$$

- KL-divergence between auxiliary "π"- and "Q"-distribution
- max-constraints on R, T from applied source coding
- min-constraints from joint-typicality tests

Summary

- Hypothesis testing for multi-hop / multi-receiver networks and noisy channels
  (Extensions to multi-relay networks, parallel relay networks)

- Schemes based on distributed quantization, unanimous-decision forwarding, and unequal error protection
  - Accumulation of error exponents
  - Competition for network resources $\rightarrow$ tradeoff in exponents
  - Intermediate processing required for optimal communication

- Derived error exponents are optimal for some testing against conditional independence