Information-Theoretic Tradeoffs in Distributed Hypothesis Testing

Michèle Wigger
Telecom Paris, IP Paris

Joint work with M. Hamad, S. Salehkalaibar, and Mireille Sarkiss

Swiss Seminar on Coding and Crypto
24 November, 2021

Example: Distributed Control-System for Smart Cars

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

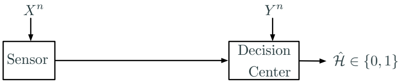
- Intact car system: measurements highly correlated
- Erroneous car system: measurements independent

Task of Distributed Control-System

Decide on joint distribution underlying the observations

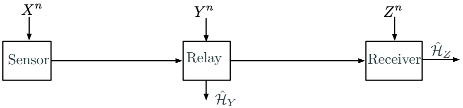
Outline of the Talk

- Simple single-sensor system



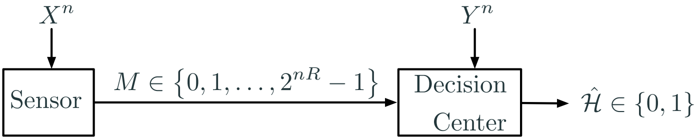
  - Infinitely many communication bits
  - Single communication bit
  - *nR* communication bits

- Multihop System



- Restriction only on expected communication rate
  - Simple single-sensor system
  - Multi-hop system

Distributed Hypothesis Testing



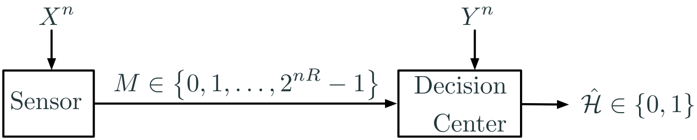
- "Normal situation" $\mathcal{H} = 0$: $(X^n, Y^n) \sim$ i.i.d. P_{XY}
- "Hazardous event" $\mathcal{H} = 1$: $(X^n, Y^n) \sim$ i.i.d. Q_{XY}

- Constraints on type-I and type-II error probabilities:

$$\overline{\lim_{n\to\infty}}\, \alpha_n = \overline{\lim_{n\to\infty}}\, \mathbb{P}\big[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0\big] \leq \epsilon$$

$$-\underline{\overline{\lim_{n\to\infty}}}\, \frac{1}{n} \log \beta_n = -\overline{\lim_{n\to\infty}}\, \frac{1}{n} \log \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\big] \geq \theta$$

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}

- Constraints on type-I and type-II error probabilities:

$$\overline{\lim_{n\to\infty}} \alpha_n = \overline{\lim_{n\to\infty}} \mathbb{P}\big[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0\big] \leq \epsilon$$

$$-\overline{\lim_{n\to\infty}} \frac{1}{n} \log \beta_n = -\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\big] \geq \theta$$

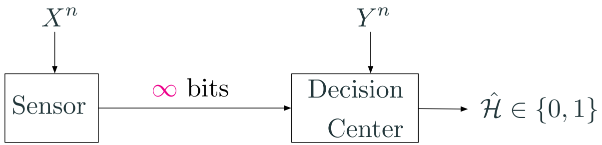
Rate-Exponent Tradeoff $\theta_\epsilon^*(R)$

Given $R > 0$, largest exponent θ that is ϵ-achievable.

- Ahlswede & Csiszár'86: $\theta_\epsilon^*(R)$ does not depend on $\epsilon \in [0, 1/2]$

Centralized Hypothesis Testing — The Ideal Case

Centralized Hypothesis Testing



- Rate R is so large that sensor can send *all* X^n to decision center

- Optimal decision: raise alarm unless statistics $\text{tp}\,(X^n, Y^n) \approx P_{XY}$

- Intuition: If system perfectly fits $\mathcal{H} = 0$, decide on $\hat{\mathcal{H}} = 0$, otherwise decide on $\hat{\mathcal{H}} = 1$

Example of Doubly-Symmetric Binary Sources

$$P_{XY}(x,y) = \begin{cases} 1/3 & x = y \\ 1/6 & x \neq y \end{cases} \qquad \text{and} \qquad Q_{XY}(x,y) = \frac{1}{2} \cdot \frac{1}{2}$$

- Ex. 1: if observations

$$\begin{aligned} X^n &= (0,1,1,0,1,0,1,0,1,1,0,1,0,0,1,1,0,0) \\ Y^n &= (0,0,1,0,1,0,1,0,1,0,1,1,1,0,1,0,1,0) \end{aligned}$$

$\text{tp}\left(X^n, Y^n\right) = P_{XY} \rightarrow$ decision center decides on $\hat{\mathcal{H}} = 0$

Example of Doubly-Symmetric Binary Sources

$$P_{XY}(x,y) = \begin{cases} 1/3 & x = y \\ 1/6 & x \neq y \end{cases} \qquad \text{and} \qquad Q_{XY}(x,y) = \frac{1}{2} \cdot \frac{1}{2}$$

- Ex. 1: if observations

$$\begin{aligned} X^n &= (0,1,1,0,1,0,1,0,1,1,0,1,0,0,1,1,0,0) \\ Y^n &= (0,0,1,0,1,0,1,0,1,0,1,1,1,0,1,0,1,0) \end{aligned}$$

tp $(X^n, Y^n) = P_{XY} \rightarrow$ decision center decides on $\hat{\mathcal{H}} = 0$

- Ex. 2: if observations

$$\begin{aligned} X^n &= (0,1,0,1,1,1,0,1,0,1,1,0,1,0,0,1,1,0,0) \\ Y^n &= (1,0,0,1,0,1,0,0,0,1,0,1,1,1,1,1,0,1,0) \end{aligned}$$

tp (X^n, Y^n) far from $P_{XY} \rightarrow$ decision center decides on $\hat{\mathcal{H}} = 1$

Analysis of Proposed Centralized Scheme

- Type-I error probability (by the weak law of large numbers)

$$\alpha_n = 1 - \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 0\big] = 1 - P_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) \approx P_{XY}\Big)$$
$$\to 0 \quad \text{as} \quad n \to \infty$$

Analysis of Proposed Centralized Scheme

- Type-I error probability (by the weak law of large numbers)

$$\alpha_n = 1 - \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 0\big] = 1 - P_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) \approx P_{XY}\Big)$$
$$\to 0 \quad \text{as} \quad n \to \infty$$

- Probability of miss detection:

$$\begin{aligned} \beta_n &= Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) \approx P_{XY}\Big) \\ &= \dots \end{aligned}$$

Mathematical Preliminaries: Sanov's Theorem

- sequences $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ of type $\pi_{\mathbf{xy}}$

$$Q_{XY}^{\otimes n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} Q_{XY}(x_i, y_i) = \prod_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(Q_{XY}(x, y)\big)^{n\cdot\pi_{\mathbf{xy}}(x,y)}$$

Theorem (Sanov's theorem)

$$Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi\Big) \approx \exp\big(-n \cdot D(\pi||Q_{XY})\big)$$

Mathematical Preliminaries: Sanov's Theorem

- sequences $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ of type $\pi_{\mathbf{xy}}$

$$Q_{XY}^{\otimes n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} Q_{XY}(x_i, y_i) = \prod_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(Q_{XY}(x, y)\big)^{n\cdot\pi_{\mathbf{xy}}(x,y)}$$

$$= \quad \exp\Big(n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x, y) \cdot \log Q_{XY}(x, y)\Big)$$

Theorem (Sanov's theorem)

$$Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi\Big) \approx \exp\big(-n \cdot D(\pi || Q_{XY})\big)$$

Mathematical Preliminaries: Sanov's Theorem

- sequences $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ of type $\pi_{\mathbf{xy}}$

$$Q_{XY}^{\otimes n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} Q_{XY}(x_i, y_i) = \prod_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(Q_{XY}(x, y)\big)^{n\cdot\pi_{\mathbf{xy}}(x,y)}$$

$$= \quad \exp\Big(n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x, y) \cdot \log Q_{XY}(x, y)\Big)$$

$$= \quad \exp\Big(-n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x, y)\Big(\log \frac{\pi_{\mathbf{xy}}(x, y)}{Q_{XY}(x, y)} + \log \frac{1}{\pi_{\mathbf{xy}}(x, y)}\Big)\Big)$$

Theorem (Sanov's theorem)

$$Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi\Big) \approx \exp\big(-n \cdot D(\pi||Q_{XY})\big)$$

Mathematical Preliminaries: Sanov's Theorem

- sequences $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ of type $\pi_{\mathbf{xy}}$

$$
\begin{aligned}
Q_{XY}^{\otimes n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} Q_{XY}(x_i, y_i) &= \prod_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(Q_{XY}(x,y)\big)^{n\cdot\pi_{\mathbf{xy}}(x,y)} \\
&= \exp\Big(n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x,y) \cdot \log Q_{XY}(x,y)\Big) \\
&= \exp\Big(-n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x,y)\Big(\log \frac{\pi_{\mathbf{xy}}(x,y)}{Q_{XY}(x,y)} + \log \frac{1}{\pi_{\mathbf{xy}}(x,y)}\Big)\Big) \\
&= \exp(-n \cdot (D(\pi_{\mathbf{xy}}||Q_{XY}) + H(\pi_{\mathbf{xy}})))
\end{aligned}
$$

Theorem (Sanov's theorem)

$$
Q_{XY}^{\otimes n}\Big(\operatorname{tp}(X^n, Y^n) = \pi\Big) \approx \exp\big(-n \cdot D(\pi||Q_{XY})\big)
$$

- sequences $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ of type $\pi_{\mathbf{xy}}$

$$Q_{XY}^{\otimes n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} Q_{XY}(x_i, y_i) = \prod_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(Q_{XY}(x,y)\big)^{n\cdot\pi_{\mathbf{xy}}(x,y)}$$

$$= \exp\Big(n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x,y) \cdot \log Q_{XY}(x,y)\Big)$$

$$= \exp\Big(-n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x,y)\Big(\log \frac{\pi_{\mathbf{xy}}(x,y)}{Q_{XY}(x,y)} + \log \frac{1}{\pi_{\mathbf{xy}}(x,y)}\Big)\Big)$$

$$= \exp(-n \cdot (D(\pi_{\mathbf{xy}}||Q_{XY}) + H(\pi_{\mathbf{xy}})))$$

Theorem (Sanov's theorem)

$$Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi\Big) \approx \exp\big(-n \cdot D(\pi||Q_{XY})\big)$$

Mathematical Preliminaries: Sanov's Theorem

- sequences $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ of type $\pi_{\mathbf{xy}}$

$$
\begin{aligned}
Q_{XY}^{\otimes n}(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^{n} Q_{XY}(x_i, y_i) &= \prod_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \big(Q_{XY}(x, y)\big)^{n\cdot\pi_{\mathbf{xy}}(x,y)} \\
&= \exp\Big(n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x, y) \cdot \log Q_{XY}(x, y)\Big) \\
&= \exp\Big(-n \cdot \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \pi_{\mathbf{xy}}(x, y)\Big(\log \frac{\pi_{\mathbf{xy}}(x, y)}{Q_{XY}(x, y)} + \log \frac{1}{\pi_{\mathbf{xy}}(x, y)}\Big)\Big) \\
&= \exp(-n \cdot (D(\pi_{\mathbf{xy}}||Q_{XY}) + H(\pi_{\mathbf{xy}})))
\end{aligned}
$$

Theorem (Sanov's theorem)

$$
Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi\Big) \approx \exp\big(-n \cdot D(\pi||Q_{XY})\big)
$$

- There are $\approx 2^{nH(\pi)}$ sequences of a given type π

Use Sanov's Theorem to Finalize the Analysis

- Type-I error probability (by the weak law of large numbers)

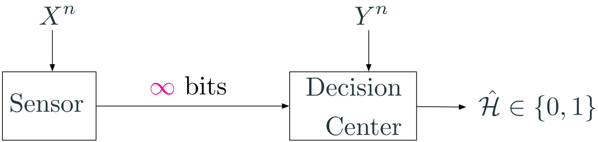
$$\alpha_n = 1 - P_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) \approx P_{XY}\Big) \to 0 \quad \text{as} \quad n \to \infty$$

- Type-II error probability

$$\begin{aligned} \beta_n &= Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) \approx P_{XY}\Big) \\ &= \sum_{\pi \approx P_{XY}} Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi\Big) \\ &\approx \exp\big(-n \cdot D(P_{XY} \| Q_{XY})\big) \end{aligned}$$

- N.B. The number of types π is polynomial in n

Optimal Exponent for Centralized Hypothesis Testing



- Rate R is so large that sensor can send *all* X^n to decision center

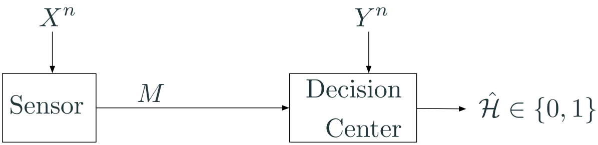
Theorem (Stein's exponent)

Largest achievable error exponent is:

$$\theta^\star(R=\infty) = D(P_{XY} \| Q_{XY}) := \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{Q_{XY}(x,y)}$$

Distributed Hypothesis Testing with Zero Communication Rate $R = 0$

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



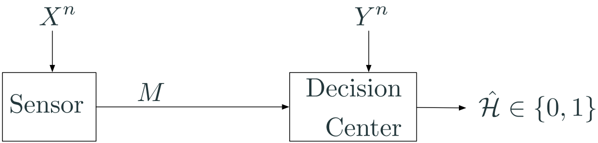
- Send message *M*, where number of bits representing *M* is sublinear in *n*

Theorem (Han'87)

Largest achievable error exponent is:

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

Scheme for $\underline{R = 0}$ (Han'87)



- $M \in \{0, 1\}$ (1 bit) suffices
- Sensor sends its own local decision:
  If $\text{tp}(X^n) \approx P_X \to$ send $M = 0$, otherwise send $M = 1$
- Decision center:
  1. Local decision: $L = 0$ if $\text{tp}\left(Y^n\right) \approx P_Y$ and $L = 1$ else
  2. Final decision: $\hat{\mathcal{H}} = 0$ if $M = L = 0$, and $\hat{\mathcal{H}} = 1$ else.

Example of Doubly-Symmetric Binary Sources

$$P_{XY}(x,y) = \begin{cases} 1/6 & x = y \\ 1/3 & x \neq y \end{cases} \quad \text{and } Q_{XY}(x,y) = \frac{1}{2} \cdot \frac{1}{2}$$

Ex.:

$$\begin{aligned} X^n &= (0,1,0,1,1,1,0,1,0,1,1,0,1,0,0,1,1,0,0) \\ Y^n &= (1,0,0,1,0,1,0,0,0,1,0,1,1,1,1,1,0,1,0) \end{aligned}$$

$\text{tp}\,(X^n) = P_X$ and $\text{tp}\,(Y^n) = P_Y \to$ decision center decides $\hat{\mathcal{H}} = 0$

For centralized setting ($R = \infty$), decision center decides $\hat{\mathcal{H}} = 1$!

Analysis of 1-bit Communication Scheme

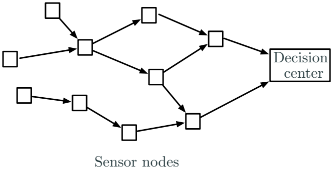
- Type-I error probability (by the weak law of large numbers)

$$\alpha_n = 1 - P_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n) \approx P_X \quad \text{and} \quad \mathrm{tp}\,(Y^n) \approx P_Y\Big) \to 0 \quad \text{as} \quad n \to \infty$$

- Type-II error probability:

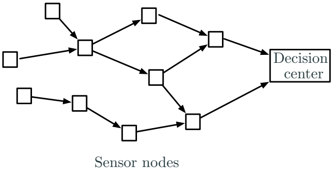
$$\begin{aligned}
\beta_n &= Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n) \approx P_X \quad \text{and} \quad \mathrm{tp}\,(Y^n) \approx P_Y\Big) \\
&= \sum_{\substack{\pi_{xy}:\ \pi_x \approx P_X \\ \pi_y \approx P_Y}} Q_{XY}^{\otimes n}\Big(\mathrm{tp}\,(X^n, Y^n) = \pi_{xy}\Big) \\
&\approx \exp\Big(- n \cdot \big(\min_{\substack{\pi_{xy}:\ \pi_x = P_X \\ \pi_y = P_Y}} D(\pi_{xy} \| Q_{XY}) \big) \Big)
\end{aligned}$$

Many-Sensors and One Detector



- “Normal situation” $\mathcal{H} = 0$: $(X_1^n, \dots, X_K^n, Y^n) \sim$ i.i.d. $P_{X_1 \cdots X_K Y}$
- “Hazardous event” $\mathcal{H} = 1$: $(X_1^n, \dots, X_K^n, Y^n) \sim$ i.i.d. $Q_{X_1 \cdots X_K Y}$

Many-Sensors and One Detector



- Each sensor k produces local decision
- Unanimous decision forwarding: sensor sends $M_k = 0$ only if all incoming messages 0 and local decision $\mathcal{H} = 0$
- Decision center: raises alarm unless all incoming messages 0 and local decision $\mathcal{H} = 0$
- Optimal exponent (independent of ϵ)

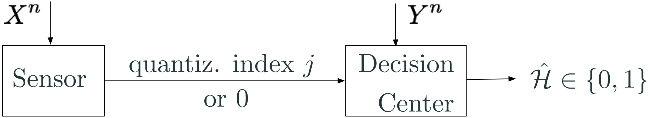
$$\theta_\epsilon^* = \min_{\substack{\pi_{X_1\cdots X_K Y}: \\ \pi_{X_k}=P_{X_k},\ \forall k \\ \pi_Y = P_Y}} D(\pi_{X_1\cdots X_K Y} || Q_{X_1\cdots X_K Y})$$

When is a single-bit transmission optimal for $R = 0$?

- Single-bit transmission (unanimous decision forward) optimal if
  - Two hypotheses $\mathcal{H} = 0$ and $\mathcal{H} = 1$
  - All decision centers interested in same exponent
  - Interactive communication allowed [Katz-Piantanida-Debbah-2016]
- Sending a single bit is not sufficient if
  - Exponents under both decisions need to be maximized
  - Different decision centers interested in different exponents
  - More than $K > 2$ hypotheses

Distributed Hypothesis Testing under Positive Rates $R > 0$

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



- Random codebook $\mathcal{C}_S = \{S^n(1), \ldots, S^n(2^{nR} - 1)\}$
- Quantization: Send $M = j$ if $\mathrm{tp}\,(S^n(j), X^n) \approx P_{SX}$; else $M = 0$
- Alarm $\hat{\mathcal{H}} = 1$ unless $M \geq 1$ and $\mathrm{tp}\,(S^n(M), Y^n) \approx P_{SY}$

- If $R \geq I(S;X)$ quantization succeeds and $\alpha_n \to 0$ as $n \to \infty$
- Achievable type-II error exponent

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

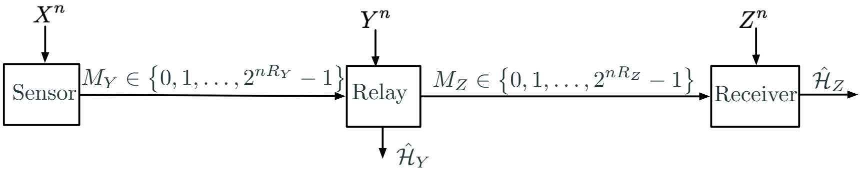
Testing Against Independence (Csiszar-Ahlswede '86)

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

Optimal Rate-Exponent Tradeoff (does not depend on ϵ)

$$\theta^*_\epsilon(R) = \max_{\substack{P_{S|X}: \\ R \geq I(S;X)}} I(S;Y) =: \eta_{XY}(R)$$

Testing Against Independence over Two Hops



- Two decision centers (relay and receiver)
- Markov chain $X \to Y \to Z$ under both hypotheses
- $\mathcal{H} = 0 : (X^n, Y^n, Z^n)$ i.i.d. $\sim P_X \cdot P_{Y|X} \cdot P_{Z|Y}$
- $\mathcal{H} = 1 : (X^n, Y^n, Z^n)$ i.i.d. $\sim P_X \cdot P_Y \cdot P_Z$

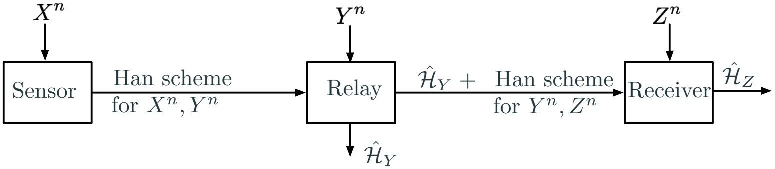
Definition of Exponents Region

- Error Constraints at the Relay
  - $\alpha_{Y,n} \triangleq \mathbb{P}[\hat{\mathcal{H}}_Y = 1|\mathcal{H} = 0] \leq \epsilon_Y$
  - $\beta_{Y,n} \triangleq \mathbb{P}[\hat{\mathcal{H}}_Y = 0|\mathcal{H} = 1] \leq 2^{-n\theta_Y}$
- Error Constraints at the Receiver
  - $\alpha_{Z,n} \triangleq \mathbb{P}[\hat{\mathcal{H}}_Z = 1|\mathcal{H} = 0] \leq \epsilon_Z$
  - $\beta_{Z,n} \triangleq \mathbb{P}[\hat{\mathcal{H}}_Z = 0|\mathcal{H} = 1] \leq 2^{-n\theta_Z}$

Definition

$\mathcal{E}^*_{\epsilon_Y,\epsilon_Z}(R_Y, R_Z)$ is the closure of the set of all (ϵ_Y, ϵ_Z)-achievable exponent pairs (θ_Y, θ_Z)

Optimal Scheme and Exponents Region



- Independent Han-scheme over each link
- Unanimous Decision-Forwarding
- $\hat{\mathcal{H}}_Z = 0$ only if $\hat{\mathcal{H}}_Y = 0$ *and* Han scheme for (Y^n, Z^n) indicates $\mathcal{H} = 0$

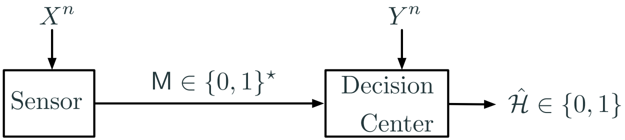
Theorem (Salehkalaibar, W', Wang 2017, Cao, Zhu, Tan'2019)

$\mathcal{E}_{\epsilon_Y,\epsilon_Z}(R_Y, R_Z)$ *is the set of all pairs* (θ_Y, θ_Z) *satisfying*

$$\theta_Y \leq \eta_{XY}(R_Y), \qquad \theta_Z \leq \eta_{XY}(R_Y) + \eta_{YZ}(R_Z).$$

Testing under Expected Rate Constraints

Point-to-Point Testing under Expected Rate Constraints



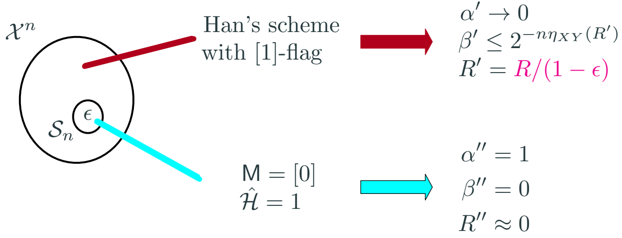
- Message M is a variable-length bit-string
- Rate constraint $\mathbb{E}[\mathrm{len}(\mathsf{M})] \leq nR$

Theorem (Salehkalaibar and Wigger'2020)

The largest ϵ-achievable exponent is

$$\theta^*_{\mathsf{VL},\epsilon}(R) = \eta_{XY}(R/(1-\epsilon)) \qquad \text{(depends on } \epsilon\text{)}$$

Optimal Point-to-Point Variable-Length Scheme



- Set $\mathcal{S}_n$ has probability ϵ
- Average performances

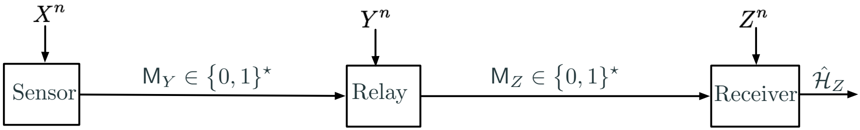
$$\alpha_n \leq \epsilon$$

$$\beta_n \leq (1-\epsilon)2^{-n \cdot \eta_{XY}(R/(1-\epsilon))}$$

and total rate is $(1-\epsilon)R' = R$

$\rightarrow$ Achievability of $\theta = \eta_{XY}(R/(1-\epsilon))$

Two-Hop Testing under Expected Rate Constraints

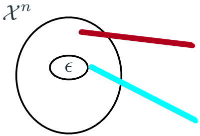


- Messages M_Y and M_Z are variable-length bit-strings
- Rate constraints $\mathbb{E}[\mathrm{len}(\mathsf{M}_Y)] \leq nR_Y$ and $\mathbb{E}[\mathrm{len}(\mathsf{M}_Z)] \leq nR_Z$

Definition

Exponents region $\mathcal{E}_{\mathsf{VL},\epsilon_Y,\epsilon_Z}(R_Y, R_Z)$ is the set of all (ϵ_Y, ϵ_Z)-achievable exponent pairs (θ_Y, θ_Z) under expected rate constraints

Achievability with Expected Rate Constraints for $\epsilon_Y = \epsilon_Z \triangleq \epsilon$



Two-Hop Scheme with [1]-flag

$$\alpha'_Y, \alpha'_Z \to 0$$
$$\beta'_Y \leq 2^{-n\eta_{XY}(R'_Y)}$$
$$\beta'_Z \leq 2^{-n(\eta_{XY}(R'_Y)+\eta_{YZ}(R'_Z))}$$
$$R'_Y = R_Y/(1-\epsilon)$$
$$R'_Z = R_Z/(1-\epsilon)$$

$\mathsf{M}_Y = \mathsf{M}_Z = [0]$
$\hat{\mathcal{H}}_Y = \hat{\mathcal{H}}_Z = 1$

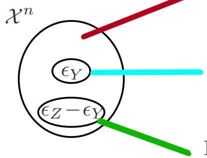
$$\alpha''_Y = \alpha''_Z = 1$$
$$\beta''_Y = \beta''_Z = 0$$
$$R''_Y = R''_Z = 0$$

Theorem (Hamad, Sarkiss, W'2021)

$\mathcal{E}_{\epsilon,\epsilon}(R_Y, R_Z)$ is the set of all pairs (θ_Y, θ_Z) satisfying

$$\theta_Y \leq \eta_{XY}(R_Y/(1-\epsilon)), \qquad \theta_Z \leq \eta_{XY}(R_Y/(1-\epsilon)) + \eta_{YZ}(R_Z/(1-\epsilon)).$$

Scheme for Expected Rate Constraints with $\epsilon_Y < \epsilon_Z$



Two-Hop Scheme with [1]-flag → $\alpha'_Y, \alpha'_Z \to 0$; $\beta'_Y \leq 2^{-n\eta_{XY}(R'_Y)}$; $\beta'_Z \leq 2^{-n(\eta_{XY}(R'_Y)+\eta_{YZ}(R'_Z))}$; $R'_Y, R'_Z = R_Z/(1-\epsilon_Z)$

$\mathsf{M}_Y = \mathsf{M}_Z = [0]$, $\hat{\mathcal{H}}_Y = \hat{\mathcal{H}}_Z = 1$ → $\alpha''_Y = \alpha''_Z = 1$; $\beta''_Y = \beta''_Z = 0$; $R''_Y = R''_Z = 0$

Han scheme on first link, $\mathsf{M}_Z = [0]$, $\hat{\mathcal{H}}_Z = 1$ → $\alpha'''_Y \to 0$; $\beta'''_Y \leq 2^{-n\eta_{XY}(R'''_Y)}$; $\alpha'''_Z = 1$, $\beta'''_Z = 0$; $R'''_Y, R'''_Z = 0$

- Average type-I error prob. $\alpha_Y \to \epsilon_Y$ and $\alpha_Z \to \epsilon_Z$ as $n \to \infty$
- Average type-II error prob. at Relay:
  $\beta_Y \leq (1-\epsilon_Z)2^{-n\eta_{XY}(R'_Y)} + (\epsilon_Z - \epsilon_Y)2^{-n\eta_{XY}(R'''_Y)})$
- Average rate of first link $(1-\epsilon_Z)R'_Y + (\epsilon_Z - \epsilon_Y)R'''_Y$

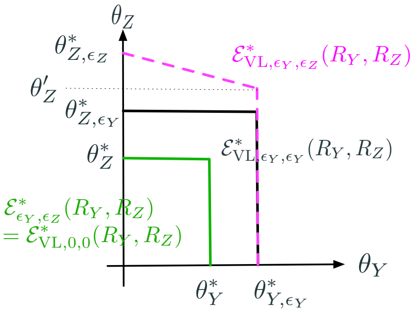
Optimal Exponents Region for $\epsilon_Y < \epsilon_Z$

Theorem (Hamad, W', Sarkiss '2021)

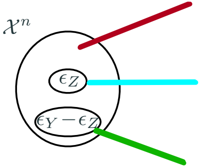
$\mathcal{E}^*_{\text{VL},\epsilon_Y,\epsilon_Z}(R_Y, R_Z)$ set of (θ_Y, θ_Z) pairs satisfying

$$\theta_Y \leq \min\{\eta_{XY}(R'_Y), \eta_{XY}(R'''_Y)\}$$
$$\theta_Z \leq \eta_{XY}(R'_Y) + \eta_{YZ}(R_Z/(1-\epsilon_Z)),$$

for some $R_{Y'}, R'''_Y > 0$ satisfying $R_Y \geq (1-\epsilon_Z)R'_Y + (\epsilon_Z - \epsilon_Y)R'''_Y$.



Scheme for Expected Rate Constraints with $\epsilon_Y > \epsilon_Z$



Two-Hop Scheme with [1,0]-flag → $\alpha'_Y, \alpha'_Z \to 0$, $\beta'_Y \leq 2^{-n\eta_{XY}(R'_Y)}$, $\beta'_Z \leq 2^{-n(\eta_{XY}(R'_Y)+\eta_{YZ}(R'_Z))}$, R'_Y, R'_Z

$\mathsf{M}_Y = \mathsf{M}_Z = [0]$, $\hat{\mathcal{H}}_Y = \hat{\mathcal{H}}_Z = 1$ → $\alpha''_Y = \alpha''_Z = 1$, $\beta''_Y = \beta''_Z = 0$, $R''_Y = R''_Z = 0$

Two-Hop Scheme with [1,1]-flag except $\hat{\mathcal{H}}_Y = 1$ → $\alpha'''_Y = 1,\ \beta'''_Y = 0$, $\alpha'''_Z \to 0$, $\beta'''_Z \leq 2^{-n(\eta_{XY}(R'''_Y)+\eta_{YZ}(R'''_Z))}$, R'''_Y, R'''_Z

- Average type-I error prob. $\alpha_Y \to \epsilon_Y$ and $\alpha_Z \to \epsilon_Z$ as $n \to \infty$
- Average type-II error prob. at Receiver:
  $\beta_Y \leq (1-\epsilon_Y)2^{-n(\eta_{XY}(R'_Y)+\eta_{YZ}(R'_Z))} + (\epsilon_Y - \epsilon_Z)2^{-n(\eta_{XY}(R'''_Y)+\eta_{YZ}(R'''_Z))}$
- Average rate of first link $(1-\epsilon_Y)R'_Y + (\epsilon_Y - \epsilon_Z)R'''_Y$

Optimal Exponents Region for $\epsilon_Y > \epsilon_Z$

Theorem (Hamad, W', Sarkiss '2021)

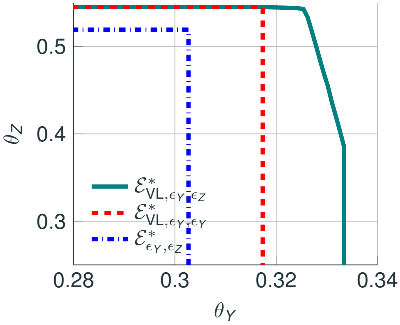
$\mathcal{E}^*_{\mathrm{VL},\epsilon_Y,\epsilon_Z}(R_Y, R_Z)$ *set of* (θ_Y, θ_Z) *pairs satisfying*

$$\theta_Y \leq \eta_{XY}(R'_Y)$$
$$\theta_Z \leq \min\{\eta_{XY}(R'_Y) + \eta_{YZ}(R'_Z), \eta_{XY}(R'''_Y) + \eta_{YZ}(R'''_Z)\}$$

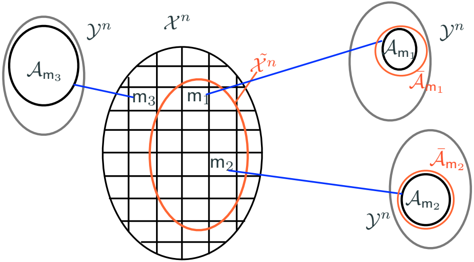
for some $R_{Y'}, R'''_Y, R'_Z, R''''_Z > 0$ *satisfying*
$R_Y \geq (1-\epsilon_Z)R'_Y + (\epsilon_Z - \epsilon_Y)R'''_Y$ *and* $R_Z \geq (1-\epsilon_Z)R'_Z + (\epsilon_Z - \epsilon_Y)R'''_Z$

Two-Hop Setup with $\epsilon_Y > \epsilon_Z$: Numerical Results

- Binary example
  - $X \sim$ Bernoulli(0.5)
  - $Y = X \oplus$ Bernoulli(0.9)
  - $Z = Y \oplus$ Bernoulli(0.8)
- Type-I error probabilities $\epsilon_Y = 0.1$ and $\epsilon_Z = 0.05$
- Rates $R_Y = R_Z = 0.5$



Idea of Converse Proof for Single-Sensor System



Original acceptance region:
$\mathcal{A} = \bigcup_m \{m\} \times \mathcal{A}_m$

New problem:
$\tilde{X}^n$ lives on $\tilde{\mathcal{X}}^n$

New acceptance region:
$\bar{\mathcal{A}} = \bigcup_m \{m\} \times \bar{\mathcal{A}}_m$

- Restrict to subset $\tilde{\mathcal{X}}^n \rightarrow$ change of measure
- Slightly-enlarge the acceptance region $\rightarrow$ blowing-up lemma

New problem has $\tilde{\alpha}_n \approx 0$ and $\mathbb{E}[\mathrm{len}(\tilde{M})] \leq \frac{\mathbb{E}[\mathrm{len}(M)]}{1-\epsilon}$ and $-\frac{1}{n}\log\tilde{\beta}_n \approx -\frac{1}{n}\log\beta_n$

Converse Proof: Bound on Rate

- Pick set $\mathcal{D}_n := \left\{ x^n \in \mathcal{T}_\epsilon^{(n)}(P_X) \colon \mathbb{P}[\hat{\mathcal{H}} = 0 | \mathcal{H} = 0, X^n = x^n] \geq \eta \right\}$ for small parameters $\eta, \epsilon > 0$ that will tend to 0
- Since $\mathbb{P}[\hat{\mathcal{H}} = 0 | \mathcal{H} = 0, X^n = x^n] \geq 1 - \epsilon \quad \rightarrow \quad \mathbb{P}[\mathcal{D}_n] \geq \frac{1-\epsilon-\eta}{1-\eta}$
- Change of measure:

$$(\tilde{X}^n, \tilde{Y}^n, \tilde{\mathsf{M}}) \sim \prod_{i=1}^{n} P_X(x_i) \frac{\mathbb{1}\{x^n \in \mathcal{D}_n\}}{\mathbb{P}[\mathcal{D}_n]} \prod_{i=1}^{n} P_{Y|X}(y_i|x_i) \mathbb{1}\{m = \text{enc}(x^n)\}$$

- $R \geq \frac{1}{n}\mathbb{E}[\mathrm{len}(\mathsf{M})] \geq \frac{1}{n}\mathbb{P}[\mathcal{D}_n] \cdot \mathbb{E}[\mathrm{len}(\tilde{\mathsf{M}})]$

$$\mathbb{E}[\mathrm{len}(\tilde{\mathsf{M}})] \geq H(\tilde{\mathsf{M}} \big| \mathrm{len}(\tilde{\mathsf{M}})) = (H(\tilde{\mathsf{M}}) - H(\mathrm{len}(\tilde{\mathsf{M}}))) \approx H(\tilde{\mathsf{M}})$$

$$\geq I(\tilde{\mathsf{M}}; \tilde{X}^n) = \sum_{i=1}^{n} I(\tilde{\mathsf{M}}; \tilde{X}_i | \tilde{X}^{i-1})$$

$$= \sum_{i=1}^{n} I(\tilde{\mathsf{M}}, \tilde{X}^{i-1}; \tilde{X}_i) + \log(\mathbb{P}[\mathcal{D}_n]) = \sum_{i=1}^{n} I(U_i; \tilde{X}_i) + \log(\mathbb{P}[\mathcal{D}_n])$$

Converse Proof: Marton's Blowing Up Lemma

Lemma (Marton's blowing up lemma)

Let $S_1, S_2, \ldots$ be i.i.d. $\sim P_S$ and $\{\epsilon_n\} \downarrow 0$.
There exist sequences $\{\ell_n\}$ and $\{\zeta_n\}$ both $\downarrow 0$ s.t. for any set $\mathcal{B}_n$:

If $P_S^{\otimes n}(\mathcal{B}_n) \geq exp(-n\epsilon_n)$, then $P_S^{\otimes n}(\bar{\mathcal{B}}_n^{(\ell_n)}) \geq 1 - \zeta_n$.

- Blow-up acceptance regions
  $\bar{\mathcal{A}}_{\mathrm{m}} := \left\{ y^n \colon \exists \tilde{y}^n \ \text{ s.t. } \ d_{\mathrm{H}}(y^n, \tilde{y}^n) \leq \ell_n, \ \tilde{y}^n \in \mathcal{A}_{\mathrm{m}} \right\}$
- By the blowing-up lemma and the definition of $\mathcal{D}_n$

$$\mathbb{P}[(\tilde{\mathrm{M}}, \tilde{Y}^n) \in \bar{\mathcal{A}} \big| \mathcal{H} = 0] \geq 1 - \zeta_n$$

- Since the blowup is very small

$$\begin{aligned} -\frac{1}{n} \log \mathbb{P}[(\tilde{\mathrm{M}}, \tilde{Y}^n) \in \bar{\mathcal{A}} \big| \mathcal{H} = 1] \quad &\approx \quad -\frac{1}{n} \log \mathbb{P}[(\tilde{\mathrm{M}}, \tilde{Y}^n) \in \mathcal{A} \big| \mathcal{H} = 1] \\ &\approx \quad -\frac{1}{n} \log \mathbb{P}[(\mathrm{M}, Y^n) \in \mathcal{A} \big| \mathcal{H} = 1] \end{aligned}$$

Converse Proof: Bound on Exponent

$$-\frac{1}{n}\log \mathbb{P}[(\tilde{\mathrm{M}},\tilde{Y}^n)\in\bar{\mathcal{A}}|\mathcal{H}=1]\approx\frac{1}{1-\zeta_n}D\big(P_{(\tilde{\mathrm{M}},\tilde{Y}^n)|\mathcal{H}=0}\|P_{(\tilde{\mathrm{M}},\tilde{Y}^n)|\mathcal{H}=1}\big)$$

$$\approx\quad\frac{1}{1-\zeta_n}I(\tilde{\mathrm{M}};\tilde{Y}^n)=\frac{1}{1-\zeta_n}\sum_{i=1}^n I(\tilde{\mathrm{M}};\tilde{Y}_i|\tilde{Y}^{i-1})$$

$$\approx\quad\frac{1}{1-\zeta_n}\sum_{i=1}^n I(\tilde{\mathrm{M}},\tilde{Y}^{i-1};\tilde{Y}_i)$$

$$\geq\quad\frac{1}{1-\zeta_n}\sum_{i=1}^n I(\tilde{\mathrm{M}},|\tilde{X}^{i-1};\tilde{Y}_i)=\frac{1}{1-\zeta_n}\sum_{i=1}^n I(U_i;\tilde{Y}_i)$$

- We related the exponent to the rate via the same variables U_i
- Remaining steps by introducing a uniform random variable T for the time index and taking $n\to\infty,\epsilon,\eta\downarrow 0$

Summary

- Distributed hypothesis testing under zero-rate constraints $\rightarrow$ Local type-based decisions and unanimous decision-forwarding
- Distributed hypothesis testing under maximum rate-constraints $\rightarrow$ Quantization and unanimous-decision forwarding
- Distributed hypothesis testing under *expected* rate-constraints:
    - Combine different degenerate versions of optimal fixed-length schemes
    - Rate-boost on all rates
    - Tradeoff between different decisions