An Information-Theoretic View of Integrated Sensing and Communication (ISAC)

Michèle Wigger

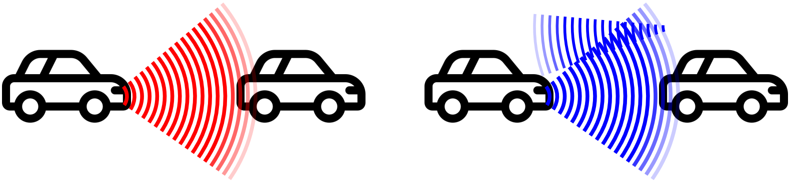
Tutorial, SPAWC 2024

Lucca, Italy, September 10, 2024

Traditional Sensing and Communications Separation

Communication

Sensing

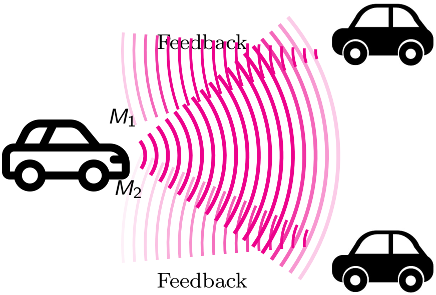


Conventional approach

- Individual hardware with own antenna and own RF chain for each of the two tasks
- Separate bandwidths for the two tasks

Integrated Sensing and Communication (ISAC)

Sensing and Communication



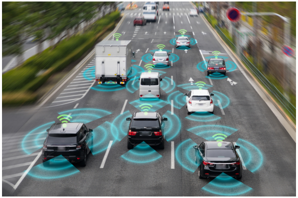
- Synergistic hardware, bandwidth, and waveform performing both tasks: Sensing and Communications

Motivation for Integrating Sensing and Communication

The most immediate benefits of ISAC:

- Cellular communication move up in frequencies, even to the THz regime
  $\rightarrow$ radar and cellular communication occupy similar bandwidths
- Integrating radar and communication will allow to free up precious bandwidth
- Savings in hardware costs, resources, and energy consumption

Important Use Cases



- But we can dream of much more...

Multi-Functional Networks in 6G

- Ubiquitous sensing capabilities for all terminals in the network: *Anywhere and anytime capabilities for all terminals*
- Network sensing (joint sensing capabilities for communicating terminals) can significantly improve local onboard sensing capabilities
- Precise positioning information allows for better communication performances
- Precise sensing of position, angles, speed, and structure of objects allows to obtain a reliable *digital twin*

Current Status of ISAC

- Predicted to be crucial building block of future 6G networks
- Heavily investigated in the communications and signal processing societies
- First prototypes available

Information-theoretic angle of attack

Determine the optimal performances of ISAC systems. And the inherent tradeoffs between sensing and communications

Outline of the Tutorial

1. Some Basics of Information Theory
2. Channel Capacity—Shannon's Channel Coding Theorem
3. Information-Theoretic Integrated Sensing and Communication (ISAC) with Distortion
4. Some Basics of Detection Theory
5. Information-Theoretic ISAC with Detection-Errors
6. Results on Network ISAC

Some Basics of Information Theory

Entropy

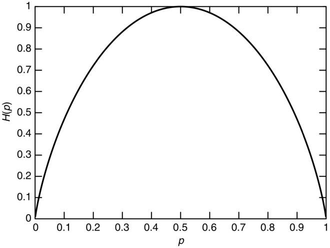
- Entropy measures randomness/uncertainty of a random variable

$$H(X) = \sum_{x \in \mathcal{X}} P_X(x) \log_2 \frac{1}{P_X(x)} = \mathbb{E}_{P_X}\left[\log_2 \frac{1}{P_X(X)}\right]$$

where $0 \log \frac{1}{0} := 0$

- Entropy of a binary random variable $X \sim \mathcal{B}(p)$:

$$H(X) = -p \log_2 p - (1-p) \log_2 (1-p).$$



Extreme Values of Entropy

- Deterministic random variable, $P_X(a) = 1$ for some $a \in \mathcal{X}$:

$$H(X) = \sum_{x \in \mathcal{X}} P_X(x) \log_2 \frac{1}{P_X(x)} = 0 \log \frac{1}{0} + 1 \cdot \log \frac{1}{1} = 0 + 1 \cdot 0 = 0$$

- Uniform random variable $P_X(x) = \frac{1}{|\mathcal{X}|}$, for all $x \in \mathcal{X}$:

$$H(X) = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} \log |\mathcal{X}| = |\mathcal{X}| \cdot \frac{1}{|\mathcal{X}|} \log |\mathcal{X}| = \log |\mathcal{X}|.$$

- Extreme values

$$0 \leq H(X) \leq \log |\mathcal{X}|;$$

lower bound tight iff X determ. and upper bound iff X uniform.

Conditional Entropy

- How the observation of a related random variable changes entropy

- Conditional entropy:

$$H(X|Y) = \sum_{y \in \mathcal{Y}} P_Y(y) \underbrace{\sum_{x \in \mathcal{X}} P_{X|Y}(x|y) \log_2 \frac{1}{P_{X|Y}(x|y)}}_{H(X|Y=y)}$$

$$= \mathbb{E}_{P_{XY}} \left[\log_2 \frac{1}{P_{Y|X}(Y|X)} \right]$$

- Conditioning reduces entropy: $0 \leq H(X|Y) \leq H(X)$

- $H(X|Y) = 0 \quad \Leftrightarrow \quad X = f(Y)$

- $H(X|Y) = H(X) \quad \Leftrightarrow \quad X$ independent of Y

Joint Entropy

- How much uncertainty is contained in a bunch of random variables
- Joint Entropy:

$$H(X,Y) = \sum_{x,y} P_{XY}(x,y) \log_2 \frac{1}{P_{XY}(x,y)} = \mathbb{E}_{P_{XY}}\left[\log_2 \frac{1}{P_{XY}(X,Y)}\right]$$

- Joint and conditional entropies extend to many r.v.:
  $H(X_1,\ldots,X_j|Y_1,\ldots,Y_i)$ and $H(X_1,\ldots,X_j)$

Chain rule of Entropy

- Entropy of X: $H(X) = \mathbb{E}\left[\log \frac{1}{P_X(X)}\right]$
- Conditional entropy of Y given X: $H(Y|X) = \mathbb{E}\left[\log \frac{1}{P_{Y|X}(Y|X)}\right]$
- Sum of the two: $(\log(ab) = \log a + \log b)$

$$\begin{aligned} H(X) + H(Y|X) &= \mathbb{E}\left[\log \frac{1}{P_X(X)P_{Y|X}(Y|X)}\right] \\ &= \mathbb{E}\left[\log \frac{1}{P_{XY}(X,Y)}\right] = H(X,Y) \end{aligned}$$

Chain rule of Entropy

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y).$$

Mutual Information

- Mutual Information is how much X tells about Y
- Reduction in Entropy:

$$I(X;Y) \triangleq H(X) - H(X|Y) = H(Y) - H(Y|X) \geq 0.$$

- X and Y are independent $\rightarrow \quad I(X;Y) = 0$
- $X = Y \rightarrow \quad I(X;Y) = H(X)$

Empirical Statistics or Types

- Empirical statistics (type) of a sequence $\boldsymbol{x} = (x_1, \dots, x_n)$

$$\boldsymbol{\pi}_{\boldsymbol{x}}(a) = \frac{|\{i : x_i = a\}|}{n}, \quad a \in \mathcal{X}.$$

- Example for the sequence $\boldsymbol{x} = (0, 1, 2, 1, 1, 2, 2, 0, 2, 1)$:

$$\boldsymbol{\pi}_{\boldsymbol{x}}(0) = \frac{2}{10} = \frac{1}{5}, \qquad \boldsymbol{\pi}_{\boldsymbol{x}}(1) = \frac{4}{10} = \frac{2}{5}, \qquad \boldsymbol{\pi}_{\boldsymbol{x}}(2) = \frac{4}{10} = \frac{2}{5}$$

Joint Empirical Statistics

- Joint empirical statistics of $\boldsymbol{x} = (x_1, \dots, x_n)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$:

$$\boldsymbol{\pi}_{\boldsymbol{xy}}(a, b) = \frac{|\{i : x_i = a \text{ and } y_i = b\}|}{n}, \quad a \in \mathcal{X},\ b \in \mathcal{Y}.$$

- Example for the sequences $\boldsymbol{x} = (0, 1, 2, 1, 1, 2, 2, 0, 2, 1)$ and $\boldsymbol{y} = (0, 3, 3, 0, 3, 0, 0, 0, 3, 0)$

$$\boldsymbol{\pi}_{\boldsymbol{xy}}(0, 0) = \frac{2}{10} \quad \boldsymbol{\pi}_{\boldsymbol{x},\boldsymbol{y}}(1, 0) = \frac{2}{10} \quad \boldsymbol{\pi}_{\boldsymbol{x},\boldsymbol{y}}(2, 0) = \frac{2}{10}$$

$$\boldsymbol{\pi}_{\boldsymbol{x},\boldsymbol{y}}(1, 3) = \frac{2}{10} \quad \boldsymbol{\pi}_{\boldsymbol{x},\boldsymbol{y}}(2, 3) = \frac{2}{10}$$

Probability of having a Joint Empirical Statistic

- Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ be jointly drawn i.i.d. from P_{XY}.

$$\Rightarrow \quad \mathbb{P}[(\boldsymbol{X}, \boldsymbol{Y}) \text{ have joint empirical statistics } \approx P_{XY}] \to 1 \text{ as } n \to \infty.$$

(weak law of large numbers)

- Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be and $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ be independently drawn i.i.d. from P_X and P_Y.

$$\mathbb{P}[(\boldsymbol{X}, \boldsymbol{Y}) \text{ have joint empirical statistics } \approx P_{XY}] \approx 2^{-nI(X;Y)}$$

Recap from Basic Information Theory Part

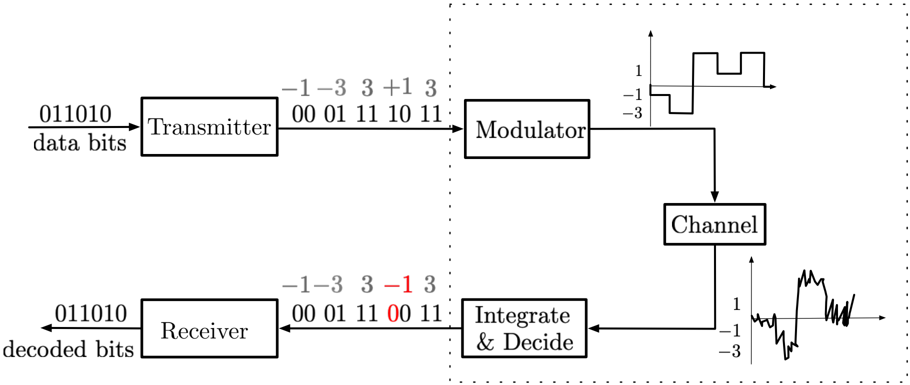
- Entropy and Conditional Entropy
- Mutual Information (Reduction in Entropy)
- Joint Empirical Statistics
- Probability that independently drawn sequences have joint empirical statistics P_{XY} is $\approx 2^{-nI(X;Y)}$

Channel Capacity

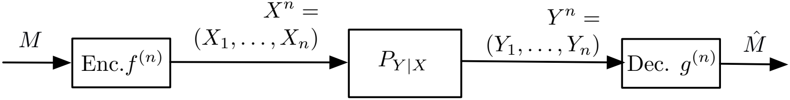
—

Shannon's Channel Coding Theorem

Data Transmission over a Noisy Channel



The Discrete Memoryless Channel (DMC)



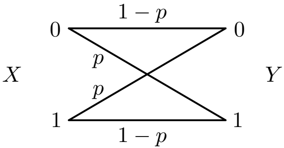
- Discrete-time and stationary memoryless channel law:

$$\mathbb{P}\left[Y_t = y \middle| X^t = x^t, Y_{t-1} = y^{t-1}\right] = P_{Y|X}(y|x_t)$$

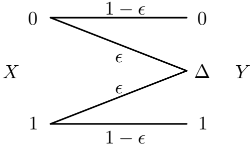
- Finite input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$

Examples of Discrete Memoryless Channel

- Binary Symmetric Channel (BSC):
  Each input, independently flipped with probability p



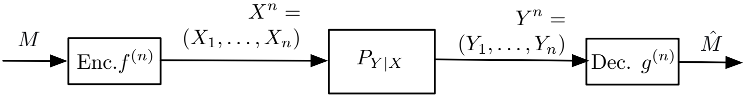
- Binary Erasure Channel (BEC):
  Each input, independently erased with probability ϵ



- Fast-fading channel $Y_t = S_t X_t + Z_t$ for i.i.d. $\{S_t\}$ and $\{Z_t\}$

- (Imperfect) receiver channel state information: $Y_t = \begin{pmatrix} S_t X_t + Z_t \\ \tilde{S}_t \end{pmatrix}$

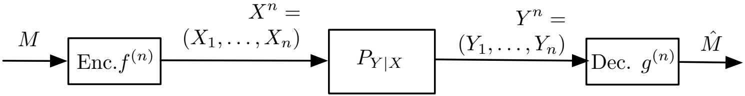
Capacity of Discrete Memoryless Channels



- M consists of nR random (i.i.d. Bernoulli-$1/2$) bits

C. E. Shannon, "A mathematical theory of communication," Bell System Journal, October 1948.

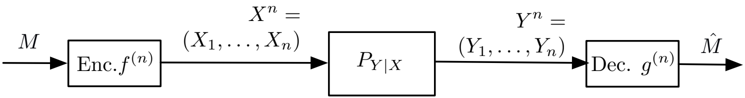
Capacity of Discrete Memoryless Channels



- M consists of nR random (i.i.d. Bernoulli-1/2) bits
- A rate $R > 0$ is achievable if $\exists$ sequence of encodings $f^{(n)}$ and decodings $g^{(n)}$ such that $\mathbb{P}\left[\hat{M} \neq M\right] \to 0$ as $n \to \infty$.

C. E. Shannon, "A mathematical theory of communication," Bell System Journal, October 1948.

Capacity of Discrete Memoryless Channels



- M consists of nR random (i.i.d. Bernoulli-$1/2$) bits
- A rate $R > 0$ is achievable if $\exists$ sequence of encodings $f^{(n)}$ and decodings $g^{(n)}$ such that $\mathbb{P}\left[\hat{M} \neq M\right] \to 0$ as $n \to \infty$.

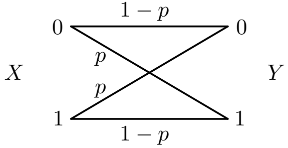
Theorem (Shannon's Channel Coding Theorem)

All rates $R < \mathsf{C} := \max_{P_X} I(X;Y)$ *are achievable.*
All rates $R > \mathsf{C}$ *are not achievable.*

C. E. Shannon, "A mathematical theory of communication," Bell System Journal, October 1948.

Capacity of the Binary Symmetric Channel

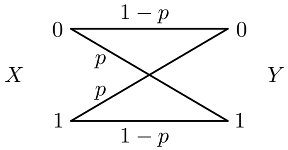
- Binary Symmetric Channel



$$
\begin{aligned}
C &= \max_{P_X} I(X;Y) = \max_{P_X} [H(Y) - H(Y|X)] \\
&= \max_{P_X} \left[H(Y) - \sum_{x \in \{0,1\}} P_X(x) H(Y|X=x) \right] \\
&= \max_{P_X} [H(Y)] - H_b(p) \\
&= 1 - H_b(p)
\end{aligned}
$$

Capacity of the Binary Symmetric Channel

- Binary Symmetric Channel

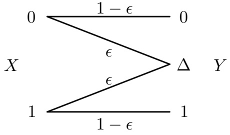


$$
\begin{aligned}
C &= \max_{P_X} I(X;Y) = \max_{P_X} [H(Y) - H(Y|X)] \\
&= \max_{P_X} \left[H(Y) - \sum_{x \in \{0,1\}} P_X(x) H(Y|X=x) \right] \\
&= \max_{P_X} [H(Y)] - H_b(p) \\
&= 1 - H_b(p)
\end{aligned}
$$

- Capacity achieved for P_X Bernoulli-1/2 $\rightarrow$ P_Y also Bernoulli-1/2

Capacity of Binary Erasure Channel

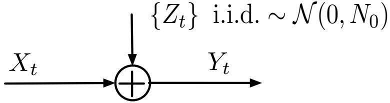
- Binary Erasure Channel



$$C = \max_{P_X} I(X;Y) = 1 - \epsilon.$$

Capacity again achieved with Bernoulli-1/2 input X

Capacity of the White Gaussian Noise Channel

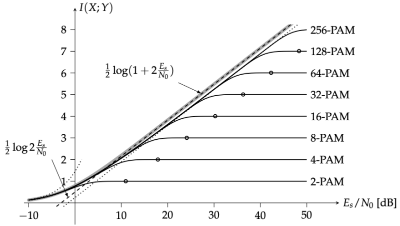


- Real inputs with blockpower constraint $\sum_{t=1}^{n} \mathbb{E}[X_i^2] \leq nE_s$:

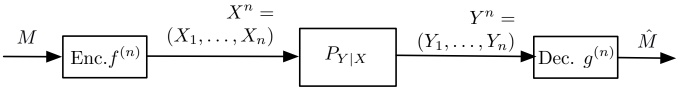
$$C = \frac{1}{2} \log \left(1 + \frac{E_s}{N_0}\right)$$

  Capacity achieved with Gaussian input $P_X \sim \mathcal{N}(0, N_0)$

- With M-PAM inputs: $I(X;Y)$ for uniform P_X



What is this Input Distribution P_X?



- The capacity formula $C = \max_{P_X} I(X;Y)$
- P_X describes the probability distribution of the input X_t averaged over all times t
- Since data bits are uniform, so are the codewords $x^n(M = 000\cdots000)$, $x^n(M = 000\cdots001)$, $x^n(M = 000\cdots010)$, $x^n(M = 000\cdots011)$, $\ldots$,

  $\Rightarrow P_X(0)$ simply indicates the frequency (empirical statistics) of the 0-symbol among all codewords

A Refined Shannon Theorem

What Shannon says:

- There is no family of encodings/decodings $\{(f^{(n)}, g^{(n)})\}_{n=1}^{\infty}$ of rate $R > C$
  such that $P_e^{(n)} \to 0$ as $n \to \infty$
- For any rate $R < C$ there does exist a family of encodings/decodings $\{f^{(n)}, g^{(n)})\}_{n=1}^{\infty}$
  s.t. $P_e^{(n)} \to 0$ as $n \to \infty$

A Refined Shannon Theorem

A stronger version:

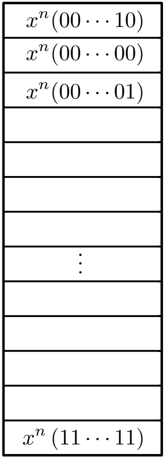
For any distribution P_X over $\mathcal{X}$:

- There is no family of encodings/decodings $\{(f^{(n)}, g^{(n)})\}_{n=1}^{\infty}$ of rate $R > I(X;Y)$ and with codebook statistics P_X
  such that $P_e^{(n)} \to 0$ as $n \to \infty$

- For any rate $R < I(X;Y)$ there does exist a family of encodings/decodings $\{f^{(n)}, g^{(n)})\}_{n=1}^{\infty}$ with codebook statistics P_X
  s.t. $P_e^{(n)} \to 0$ as $n \to \infty$

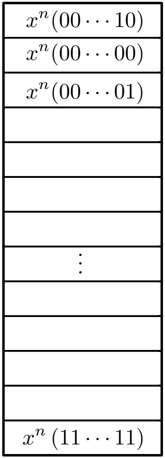
Achievability Proof (if-direction) for Capacity Theorem

- *Encoder: Send* $x^n(M)$



empirical statistics of
each codeword $\approx P_X$

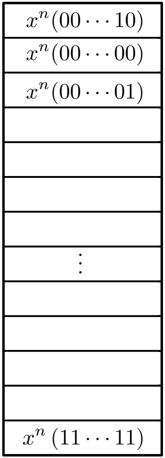
Achievability Proof (if-direction) for Capacity Theorem



empirical statistics of each codeword $\approx P_X$

- *Encoder: Send* $x^n(M)$
- *Decoder: Declare the unique* $\hat{M}$ *s.t.* $(x^n(\hat{M}), Y^n)$ *have joint empirical statistics* $\approx P_X P_{Y|X}$

Achievability Proof (if-direction) for Capacity Theorem



empirical statistics of each codeword $\approx P_X$

- *Encoder: Send* $x^n(M)$
- *Decoder: Declare the unique* $\hat{M}$ *s.t.* $(x^n(\hat{M}), Y^n)$ *have joint empirical statistics* $\approx P_X P_{Y|X}$
- *Analysis:*
  - $\mathbb{P}\big[(x^n(M), Y^n)$ of joint stat. $\approx P_X P_{Y|X}\big] \approx 1$
  - $\forall j \neq M$:
    $\mathbb{P}[(x^n(j), Y^n)$ of joint stat. $\approx P_X P_{Y|X}] \approx 2^{-nI(X;Y)}$
  - By union bound:

    $$\mathbb{P}[\exists j \neq M\colon (x^n(j), Y^n) \text{ of joint st.} \approx P_X P_{Y|X}] \approx 2^{nR} \cdot 2^{-nI(X;Y)}$$

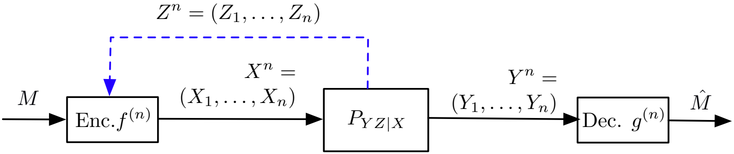
    which vanishes as $n \to \infty$ if $R < I(X;Y)$.

Converse Proof (only if-direction) for Capacity Theorem

- Fix encodings/decodings $(f^{(n)}, g^{(n)})$ so that $P_e^{(n)} \to 0$ as $n \to \infty$

$$
\begin{aligned}
nR &= H(M) = I(M;Y^n) - \underbrace{H(M|Y^n)}_{\to 0 \text{ as } n\to\infty, P_e\to 0} = I(M;Y^n) - \delta(P_e^{(n)}, n) \\
&= \sum_{t=1}^{n} I(M;Y_t|Y^{t-1}) - \delta(P_e^{(n)}, n) \\
&= \sum_{t=1}^{n} \left[H(Y_t|Y^{t-1}) - H(Y_t|Y^{t-1}, M)\right] - \delta(P_e^{(n)}, n) \\
&\leq \sum_{t=1}^{n} \left[H(Y_t) - H(Y_t|Y^{t-1}, M, X_t)\right] - \delta(P_e^{(n)}, n) \\
&= \sum_{t=1}^{n} [H(Y_t) - H(Y_t|X_t)] - \delta(P_e^{(n)}, n) \\
&= \sum_{t=1}^{n} I(Y_t;X_t) - \delta(P_e^{(n)}, n) \leq nC - \delta(P_e^{(n)}, n)
\end{aligned}
$$

What about Feedback?

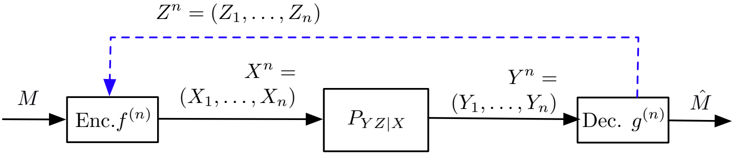


- Transmitter observes backscatterers or feedback

Theorem (Shannon's Capacity-Theorem for Feedback Channels)

Capacity with feedback and without feedback is the same for Discrete Memoryless Channels.

What about Feedback?



- Transmitter observes backscatterers or feedback

Theorem (Shannon's Capacity-Theorem for Feedback Channels)

Capacity with feedback and without feedback is the same for Discrete Memoryless Channels.

Holds also with active feedback.

Active feedback cannot be better than perfect feedback $Z = Y$

Capacity of Channels with Memory

- *Arbitrary* channel law $P_{Y^n|X^n}$ (for each n)
- Capacity formula

$$C := \sup_{\{P_{X^n}\}_{n=1}^{\infty}} p\text{-}\underline{\lim}_{n\to\infty} \frac{1}{n} i(X^n; Y^n)$$

where the *information density* is $i(X^n; Y^n) = \log \frac{P_{X^n Y^n}(X^n, Y^n)}{P_X^n(X^n) \cdot P_{Y^n}(Y^n)}$

- Supremum over distributions on tuples of inputs X^n
- For DMCs and i.i.d. inputs we obtain by the weak law of large numbers: $p\text{-}\underline{\lim}_{n\to\infty} \frac{1}{n} i(X^n; Y^n) = I(X; Y)$

S. Verdu and T.S. Han, "A general formula for channel capacity," in IEEE Trans. IT, July 1994.

Information-Theoretic Finite-Blocklength Bounds

- Information density: $i(X;Y) := \log \frac{P_{Y|X}(y|x)}{P_Y(y)}$

Theorem

Given a blocklength n. *Rate* R *is achievable with error probability* ϵ *if* $\exists$ P_X *and* $\mathsf{K} > 0$ *s.t.:*

$$\mathsf{R} \leq I(X;Y) - \sqrt{\frac{\mathsf{V}}{n}}\mathbb{Q}^{-1}(\epsilon - \beta_u) - \mathsf{K}\frac{\log(n)}{n} \quad (1)$$

with $\beta_u := \frac{1}{n^{\mathsf{K}}} + \frac{0.7975\mathsf{T}}{\sqrt{n\mathsf{V}^3}}$ *and* V / T *the 2nd / 3rd cent. mom. of* $i(X;Y)$.

Rate R *is not achievable with error probability* $\epsilon > 0$, *if* $\forall\delta > 0$ *and pmfs* P_X:

$$\mathsf{R} \geq I(X;Y) - \sqrt{\frac{\mathsf{V}}{n}}\mathbb{Q}^{-1}(\epsilon + \beta_l) + \frac{\log(n)}{2n} - \frac{\log\delta}{n}, \quad (2)$$

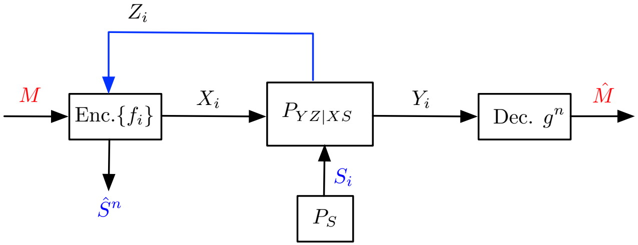
where $\beta_l := \frac{6\mathsf{T}}{\sqrt{n\mathsf{V}^3}} + \frac{\delta}{\sqrt{n}}$.

Take-Away Messages from the Capacity Part

- Capacity denotes the highest rates of reliable communication (error probability tending to 0)
- Capacity formula: $C = \max_{P_X} I(X;Y)$
- Capacity formula holds also with feedback
- Maximimaztion argument P_X refers to the codebook statistics
- For given codebook statistics P_X largest reliable rate is $I(X;Y)$

Information-Theoretic Integrated Sensing and Communication (ISAC) with Distortion

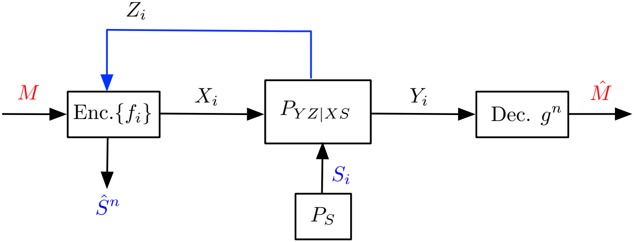
Information Theoretic Model for ISAC Kobayashi et al.



- state sequence $S^n = (S_1, \ldots, S_n)$ i.i.d. $\sim P_S$
- Behaviour of the channel depends on the state S^n (for example the acceleration of an object)
- Sensing Performance measured by Average Block-Distortion:

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[d(S_i, \hat{S}_i)] \leq D.$$

Distortion as a Sensing Performance

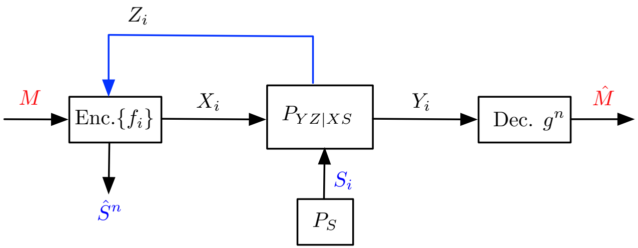


- Sensing Performance Measured by Average Block-Distortion:

$$\overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[d(S_i, \hat{S}_i)] \le D.$$

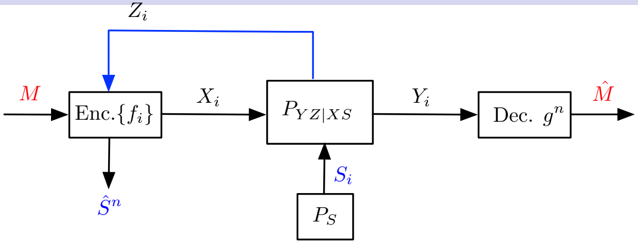
- Examples of distortion measures
  - Mean-Squared Error $d(s, \hat{s}) = (s - \hat{s})^2$
  - Hamming weight $d(s, \hat{s}) = \mathbb{1}\{s \neq \hat{s}\}$
  - Distortion on a function of the state $d(s, \hat{s}) = d'(f(s), \hat{s})$

Recall the Generality of our Model



- Arbitrary forward and backward channels $P_{Y|XS}$ and $P_{Z|YXS}$
- Examples:
  - Memoryless fading channels $Y = SX + N$
  - Backscatterer can be $Z = SX + N$ or $Z = Y + N$ or ...
  - Receiver CSI: Y_i can include S_i or imperfect versions of S_i

Information-Theoretic Fundamental Limit



Definition

Capacity-distortion tradeoff $C(D)$ is largest rate R such that there exist encoders, decoders and estimators with

$$\Pr\left(\hat{M} \neq M\right) \to 0 \qquad \text{as} \qquad n \to \infty$$

and

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[d(S_i, \hat{S}_i)] \leq D$$

Optimal Sensing under Distortion Constraints

- By memoryless assump.: Markov chain $(X^n, Z^n) \to (X_i, Z_i) \to S_i$

Lemma

The optimal estimator operates symbolwise on (X^n, Z^n)

$$\hat{s}^n(x^n, z^n) := (\hat{s}^*(x_1, z_1), \hat{s}^*(x_2, z_2), \ldots, \hat{s}^*(x_n, z_n)),$$

where the optimal per-symbol estimator is

$$\hat{s}^*(x, z) := \arg\min_{s' \in \hat{\mathcal{S}}} \sum_{s \in \mathcal{S}} P_{S|XZ}(s|x, z) d(s, s')$$

- Minimum achievable distortion $D_{\min} = \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[d(S_t, \hat{s}^*(X_t, Z_t))]$
- So, it depends only on the empirical statistics of the codebook

Capacity-Distortion Tradeoff $C(D)$

Theorem (Kobayashi et al.)

Capacity-distortion tradeoff

$$C(D) := \max I(X; Y)$$

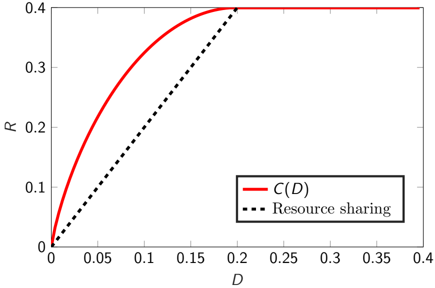
where maximum is over P_X *satisfying*

$$\mathbb{E}[d(S, \hat{s}^*(X, Z))] \leq D.$$

- Tradeoff between communication and sensing stems from P_X
- Generalized feedback not used for coding. Simple point-to-point codes are sufficient. It suffices to adjust input pmf P_X to desired sensing performance.

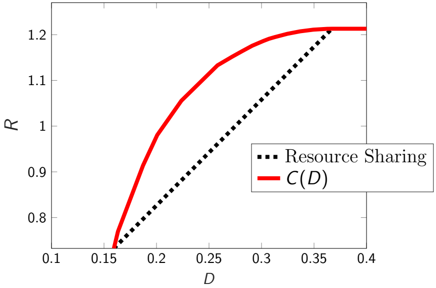
Ex. 1: Binary Multiplicative-State Channel

- $S \sim \mathcal{B}(q)$
- $Z = Y' = SX$ and $Y = (Y', S)$
- Hamming Distortion $d(s, \hat{s}) = s \oplus \hat{s}$.
- Minimize distortion: $X = 1 \ \rightarrow \ D = 0$ and $R = 0$
- Maximize rate: $X \sim \mathcal{B}(1/2) \ \rightarrow \ D = 1/2 \cdot \min\{q, 1-q\}$ and $R = q$



Ex. 2: Rayleigh Fading Channel

- Standard Gaussian state and noises S, N, N_{fb}
- Rayleigh fading channel $Y' = SX + N$
- Rx observes $Y = (Y', S)$ and Tx $Z = Y' + N_{\text{fb}}$
- Input power constraint $P = 10dB$
- Quadratic distortion $d(s, \hat{s}) = (s - \hat{s})^2$.



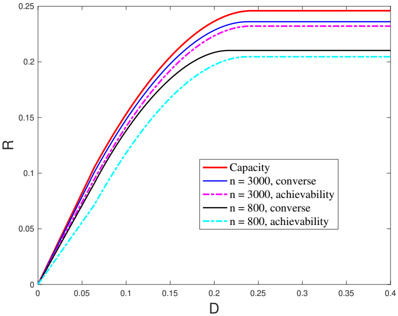
- $X \sim \mathcal{N}(0, P)$ achieves capacity
- $X \pm \sqrt{P}$ optimal for sensing

The Finite Blocklength Regime

- Given blocklength n, triple (R, D, ϵ) is called achievable if $\exists$ encoder, decoder, and estimator with

$$\Pr\left(\hat{M} \neq M\right) \leq \epsilon \qquad \text{and} \qquad \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[d(S_i, \hat{S}_i)] \leq D$$

- Can reuse the optimal estimator $s^\star(x, z)$ from the capacity-problem!
- Ex.: $Z = Y = XS$ and $S \sim \mathbb{B}(0.4)$ and $\epsilon = 10^{-3}$



[3] H. Nikbakht et al., "Integrated Sensing and Communication in the Finite Blocklength Regime", ISIT 2024.

Information-Theoretic Finite-Blocklength Bounds

Theorem

Given n*. Triple* $(\mathsf{R}, D, \epsilon)$ *is achievable if* $\exists\, P_X$ *and* $\mathsf{K} > 0$ *s.t.:*

$$\mathsf{R} \leq I(X;Y) - \sqrt{\frac{\mathsf{V}}{n}}\mathbb{Q}^{-1}(\epsilon - \beta_u) - \mathsf{K}\frac{\log(n)}{n}, \quad (3)$$

$$D \geq \mathbb{E}[d(S, \hat{s}^\star(X, Z))] \quad (4)$$

with $\beta_u := \frac{1}{n^{\mathsf{K}}} + \frac{0.7975\mathsf{T}}{\sqrt{n\mathsf{V}^3}}$ *and* V / T *the 2nd / 3rd cent. mom. of* $i(X;Y)$.
Triple $(\mathsf{R}, D, \epsilon)$ *not achievable if* $\forall \delta > 0$ *and pmfs* P *satisfying (4):*

$$\mathsf{R} \geq I(X;Y) - \sqrt{\frac{\mathsf{V}}{n}}\mathbb{Q}^{-1}(\epsilon + \beta_l) + \frac{\log(n)}{2n} - \frac{\log\delta}{n}, \quad (5)$$

where $\beta_l := \frac{6\mathsf{T}}{\sqrt{n\mathsf{V}^3}} + \frac{\delta}{\sqrt{n}}$.

[3] H. Nikbakht et al., "Integrated Sensing and Communication in the Finite Blocklength Regime", ISIT 2024.

Beyond the Memoryless Assumption

- *arbitrary* state sequence $S^n = (S_1, \ldots, S_n)$ (for each n)
- No feedback coding $X^n = f^n(M)$
- *arbitrary* channel law $P_{Z^n Y^n | X^n S^n}$ (for each n)
- *general* distortion constraint $\overline{\lim}_{n\to\infty} \frac{1}{n}\mathbb{E}[d(S^n, \hat{S}^n)] \leq D$

Theorem (Capacity-distortion tradeoff)

$$C(D) := \sup_{\{P_{X^n}\}_n} p - \underline{\lim}_{n\to\infty} \frac{1}{n} i(X^n; Y^n)$$

where supremum over all $\{P_{X^n}\}$ *s.t.* $\overline{\lim}_{n\to\infty} \frac{1}{n}\mathbb{E}[d(S^n, \hat{S}^n(X^n, Z^n))] \leq D$

[4] Chen et al. "On general capacity-distortion formulas of integrated sensing and communication," Arxiv 2023.

Take-Away for Distortion-Based Single-User ISAC

- Information-theoretic model with average distortion
- Symbol-by-symbol estimator optimal in memoryless case $\Rightarrow$ sensing performance depends only on empirical statistics of X^n

  $\Rightarrow$ Use optimal data communication scheme under restriction on empirical statistics of x^n
  $\Rightarrow$ No need to use backscatterers for coding
- Tradeoff between sensing and communication
- Resource-sharing schemes highly suboptimal

Beyond the Distortion Model: Estimation-Theoretic Sensing

- Gaussian channels:

$$\begin{aligned} \boldsymbol{Y}_c &= \boldsymbol{H}_c\boldsymbol{X} + \boldsymbol{Z}_c \\ \boldsymbol{Y}_s &= \boldsymbol{H}_s\boldsymbol{X} + \boldsymbol{Z}_s \end{aligned}$$

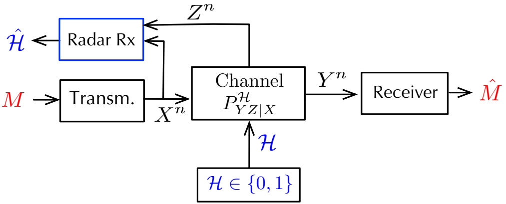
- Sensing parameter $\boldsymbol{\eta}$ s.t. $\boldsymbol{H}_s = \boldsymbol{g}(\boldsymbol{\eta})$ and metric (which lower-bounds MMSE)

$$\mathbb{E}\left[\mathrm{tr}\left(\mathsf{J}^{-1}_{\boldsymbol{\eta}|\boldsymbol{X}}\right)\right]$$

- Characterized/approximated extreme points of *Rate-Cramer-Rao-Bound* Tradeoff (max rate and min CRB)

[5] Y. Xiong et al., "On the fundamental tradeoff of ISAC under Gaussian channels," Trans. IT Sep. 2023.

Beyond the Distortion Model: Detection of Events



- Channel governed by single parameter $\mathcal{H}$ for entire transmission
- Radar Rx wishes to detect finite $\mathcal{H}$
- Information-theoretic limits derived under various sensing criteria: Stein exponent, minimum error exponent, exponents region

$$\overline{\lim_{n\to\infty}} -1/n \log \Pr[\hat{\mathcal{H}} \neq \mathcal{H} | \mathcal{H} = i], \quad i \in \{0,1\}$$

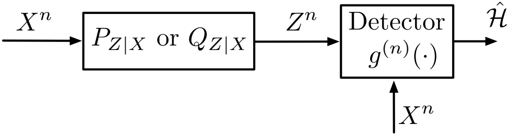
[6] H. Joudeh & F. Willems, "Joint communication and binary state detection," *JSAIT* 2022.

[7] M.-C. Chang, et al. "Rate and detection-error exponent tradeoff for joint communication and sensing of fixed channel states," *JSAIT 2023*.

[8] M. Ahmadipour, et al. "Strong Converse for Bi-Static ISAC with Two Detection-Error Exponents," *IZS 2024*.

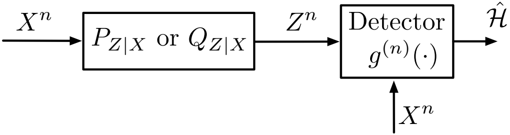
Some Basics of Detection Theory

Non-Adaptive Active Detection / Hypothesis Testing



- Hypothesis $\mathcal{H} \in \{0, 1\}$
  - Non-adaptive input sequence x^n (deterministic or random)
  - under $\mathcal{H} = 0$: outputs $Z^n \sim P_{Z|X}^{\otimes n}(\cdot|x^n)$
  - under $\mathcal{H} = 1$: outputs $Z^n \sim Q_{Z|X}^{\otimes n}(\cdot|x^n)$
  - Decision $\hat{\mathcal{H}} = g^{(n)}(Z^n, X^n)$

Non-Adaptive Active Detection / Hypothesis Testing



- Hypothesis $\mathcal{H} \in \{0, 1\}$
- Probability of type-I: $\alpha_n = \mathbb{P}\big[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0\big]$
- Probability of type-II: $\beta_n = \mathbb{P}\big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\big]$

Example on how to Detect

- Assume $\mathcal{Z} = \mathcal{X} = \{0, 1\}$ and for $x \neq z$: $P_{Z|X}(z|x) = 0.25$ and $Q_{Z|X}(z|x) = 0.7$
- Assume the input sequence is $X^n = (0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1)$
- How to decide on the following output sequences:

$$Z^n = (0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1) \qquad \rightarrow \hat{\mathcal{H}} = 0$$

$$Z^n = (0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1) \qquad \rightarrow \hat{\mathcal{H}} = 1$$

$$Z^n = (0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1) \qquad \rightarrow \hat{\mathcal{H}} = ?$$

Neyman-Pearson Test

For a fixed blocklength n and fixed input sequence x^n

- Log-likelihood ratio (LLR) $= \sum_{i=1}^{n} \log \frac{P_{Z|X}(Z_i|x_i)}{Q_{Z|X}(Z_i|x_i)}$

- Neyman-Pearson Test:

$$\hat{\mathcal{H}} = 0 \quad \text{if} \quad \text{LLR} \geq \gamma \qquad \text{and} \qquad \hat{\mathcal{H}} = 1 \qquad \text{otherwise}$$

- Varying γ over $\mathbb{R}$ we obtain the optimal tradeoff-curve between the type-I and type-II error probabilities

- Our interest will be in the behaviour of the error probabilities in function of $n \to \infty$:

  $\to$ error probabilities can decay exponentially!

Kullback Leibler Divergence

- Given two probability distributions P and Q on same alphabet $\mathcal{X}$
- Kullback Leibler (KL) Divergence:

$$D(P\|Q) \triangleq \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)}.$$

- KL divergence nonnegative: $D(P\|Q) \geq 0$ with equality iff $P(x) = Q(x)$ for all $x \in \mathcal{X}$
- KL divergence is non-symmetric $D(P\|Q) \neq D(Q\|P)$

  $\rightarrow$ not a real distance but still measures closeness
- $D(P_{XY}\|P_X P_Y) = I(X;Y)$

Probability of Having a Joint Statistics II

- Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be and $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ be jointly drawn i.i.d. from Q_{XY}. For any $P_{XY} \neq Q_{XY}$:

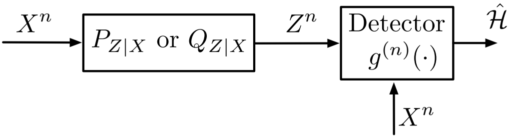
$$\Rightarrow \quad \mathbb{P}[(\boldsymbol{X}, \boldsymbol{Y}) \text{ have joint empirical statistics } \approx P_{XY}] \approx 2^{-nD(P_{XY} \| Q_{XY})}$$

- Let $\mathcal{S}$ be a set of joint empirical statistics:

$$\mathbb{P}[(\boldsymbol{X}, \boldsymbol{Y}) \text{ have joint empirical statistics in set } \mathcal{S}] \approx \sum_{P \in \mathcal{S}} 2^{-nD(P \| Q_{XY})} \approx \max_{P \in \mathcal{S}} 2^{-nD(P \| Q_{XY})} \approx 2^{-n \min_{P \in \mathcal{S}} D(P \| Q_{XY})}$$

- Notice that the number of empirical statistics is polynomial in n!

Stein-Exponent



- $\mathcal{H} = 1$ corresponds to an extreme alert situation! (Tsunami, fire, earthquake)

- Asymmetric requirements on the error probabilities

$$\overline{\lim_{n\to\infty}}\ \alpha_n \leq \epsilon$$

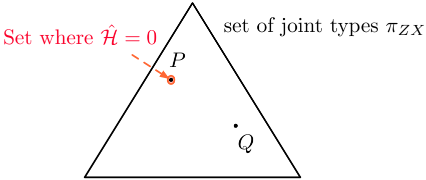
$$\underline{\lim_{n\to\infty}}\ -\frac{1}{n}\log\beta_n \geq \theta \qquad \text{“}\Longleftrightarrow\text{”}\ \beta_n \leq 2^{-n\theta}$$

Theorem (Stein Exponent)

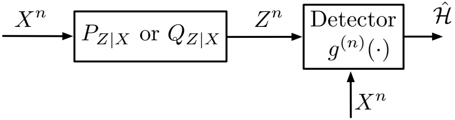
Largest possible exponential decay is $\theta_{\max} = \sum_x \pi(x) D(P_{Z|X=x} \| Q_{Z|X=x})$ *where* $\pi(x)$ *denotes the statistics of the input sequence* x^n.

Relation Stein-Exponent to Empirical Statistics

- Fix input sequence x^n of empirical statistics π_x
- Define $P = \pi_x P_{Z|X}$ and $Q = \pi_x Q_{Z|X}$
- Decide $\hat{\mathcal{H}} = g^{(n)}(z^n, x^n) = 0$ only if emp. statistics $\pi_{x^n z^n} \approx \pi_x P_{Z|X}$. Else declare $\hat{\mathcal{H}} = 1$
- $\mathbb{P}$[emp. statistics of $(x^n, Z^n) \approx P | \mathcal{H} = 1] \approx 2^{-nD(\pi_x P_{Z|X} \| \pi_x Q_{Z|X})}$



Exponents Region



- Exponential decay of both error exponents:

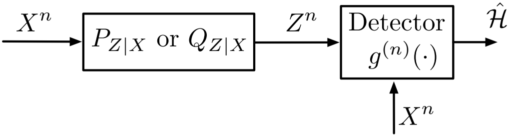
$$\begin{aligned} \varliminf_{n\to\infty} -\frac{1}{n}\log\alpha_n &\geq r \qquad \text{“} \Longleftrightarrow \text{”}\ \alpha_n \leq 2^{-nr} \\ \varliminf_{n\to\infty} -\frac{1}{n}\log\beta_n &\geq \nu \qquad \text{“} \Longleftrightarrow \text{”}\ \beta_n \leq 2^{-n\nu} \end{aligned}$$

Theorem (Exponentsregion)

For fixed type-I error exponent r the largest type-II error exponent is

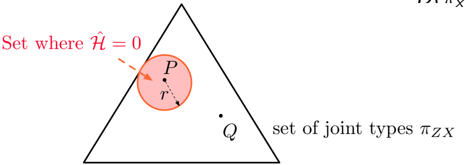
$$\nu^\star(r) = \min_{\substack{\pi_{Z|X}:\\ D(\pi_x\pi_{Z|X}\|\pi_x P_{Z|X})\leq r}} D(\pi_x\pi_{Z|X}\|\pi_x Q_{Z|X})$$

Exponents region to Empirical Statistics

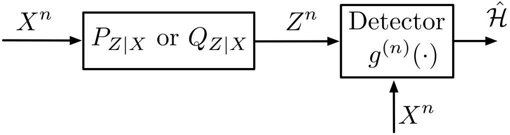


- Fix input sequence x^n of empirical statistics π_x
- Decide $\hat{\mathcal{H}} = g^{(n)}(z^n, x^n) = 0$ only if emp. statistics $\pi_{x^n z^n}$ is such that $D(\pi_{x^n z^n} \| \pi_x P_{Z|X}) \leq r$. Else declare $\hat{\mathcal{H}} = 1$
-

$$\mathbb{P}[D(\pi_{x^n Z^n} \| \pi_x P_{Z|X}) \leq r | \mathcal{H} = 1] \approx \sum_{\substack{\pi_{Z|X}: \\ D(\pi_x \pi_{Z|X} \| \pi_x P_{Z|X}) \leq r}} 2^{-nD(\pi_x \pi_{Z|X} \| \pi_x Q_{Z|X})}$$



Chernoff Exponent



- Interested in smallest exponent for input x^n

$$\min\left\{ \underline{\lim}_{n\to\infty} -\frac{1}{n}\log\alpha_n,\ \underline{\lim}_{n\to\infty} -\frac{1}{n}\log\beta_n \right\} \geq E$$

Theorem (Symmetric Error Exponent)

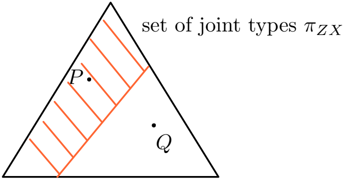
$$\begin{aligned}\nu^\star(r) &= \min_{\substack{\pi_{Z|X}:\\ D(\pi_x\pi_{Z|X}\|\pi_x P_{Z|X})\leq D(\pi_x\pi_{Z|X}\|\pi_x Q_{Z|X})}} D(\pi_x\pi_{Z|X}\|\pi_x Q_{Z|X}) \\ &= \max_{l\in[0,1]} -\sum_x \pi_{x^n}(x)\log\left(\sum_z P_{Z|X}(z|x)^l Q_{Z|X}(z|x,1)^{1-l}\right)\end{aligned}$$

- E is also the exponent of the sum probability $\frac{1}{2}\alpha_n + \frac{1}{2}\beta_n$
- Extends to multiple hypotheses!

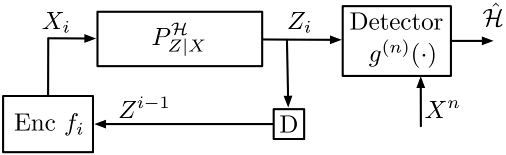
Chernoff Exponent and Empirical Statistics

- Fix input sequence x^n of empirical statistics π_x
- Define $P = \pi_x P_{Z|X}$ and $Q = \pi_x Q_{Z|X}$
- Decide $\hat{\mathcal{H}} = g^{(n)}(z^n, x^n) = 0$ only if emp. statistics $\pi_{x^n z^n}$ has smaller KL divergence with $\pi_x P_{Z|X}$ than with $\pi_x Q_{Z|X}$. Else $\hat{\mathcal{H}} = 1$
-

$$\mathbb{P}[D(\pi_{x^n Z^n} \| \pi_x P_{Z|X}) \leq D(\pi_{x^n Z^n} \| \pi_x Q_{Z|X}) | \mathcal{H} = 1]$$
$$\approx \sum_{\substack{\pi_{Z|X}: \\ D(\pi_x \pi_{Z|X} \| \pi_x P_{Z|X}) \\ \leq D(\pi_x \pi_{Z|X} \| \pi_x Q_{Z|X})}} 2^{-nD(\pi_x \pi_{Z|X} \| \pi_x Q_{Z|X})}$$



Active Adaptive Detection



- M-ary hypotheses: $\mathcal{H} \in \{0, 1, \ldots, M-1\}$
- Inputs can be chosen adaptively $x_i = f_i(Z^{i-1})$
- Decision $\hat{\mathcal{H}} = g^{(n)}(Z^n, X^n)$
- For more than two hypotheses $M \geq 2$, adaptive strategies allow to discriminate between different hypotheses!

Take-Away from Detection Theory

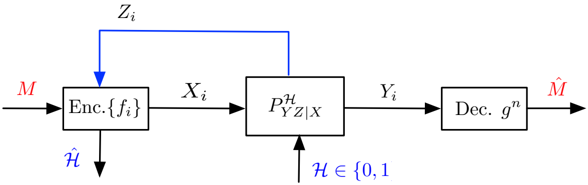
- $\mathcal{S}$ a set of joint statistics and $Z^n \sim Q_{Z|X}$ given x^n. Then:

$$\mathbb{P}[\pi_{x^n Z^n} \in \mathcal{S}] \approx 2^{-n \min_{\pi \in \mathcal{S}} D(\pi \| Q)}$$

- Non-adaptive active hypothesis testing:
  - Asymptotically-optimal detection rule based solely on the joint empirical statistics of (x^n, Z^n)
  - Detection performances only depend on statistics of input x^n!
  - Channel discrimination possible with exp.-decaying error prob.
- Adaptive strategies beneficial for active hypothesis testing if more than 2 hypotheses

Information-Theoretic Results on ISAC with Detection Errors

ISAC Models with Detection Exponents



- Single sensing parameter $\mathcal{H} \in \{0, 1\}$ constant for all times

- Sensing performance measured in detection-error exponents

  - Symmetric detection exponent:

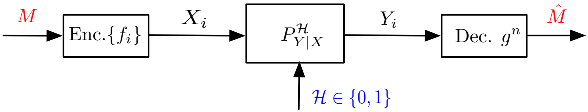
$$E_{\text{Sym}} := \lim_{n\to\infty} -\frac{1}{n} \log \left(\max \left\{ \mathbb{P}\left[\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\right], \ \mathbb{P}\left[\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\right] \right\} \right)$$

  - Stein's exponent requires

$$\lim_{n\to\infty} \mathbb{P}\left[\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\right] \leq \epsilon$$

$$\theta_{\text{Stein}} := \lim_{n\to\infty} -\frac{1}{n} \log \mathbb{P}\left[\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\right]$$

Relation to a Compound Channel

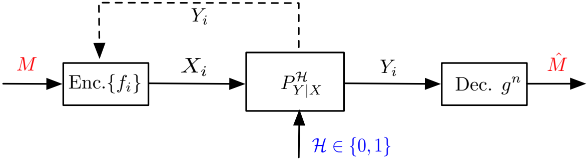


- Without the sensing it is a compound channel
- Compound capacity without feedback:

$$C_{\text{compound}} \leq \max_{P_X} \min_{\mathcal{H}} I(X;Y|\mathcal{H})$$

Cannot adjust codebook statistics P_X to the channel $\mathcal{H}$!

Relation to a Compound Channel



- Without the sensing it is a compound channel
- Compound capacity without feedback:

$$C_{\text{compound}} \leq \max_{P_X} \min_{\mathcal{H}} I(X;Y|\mathcal{H})$$

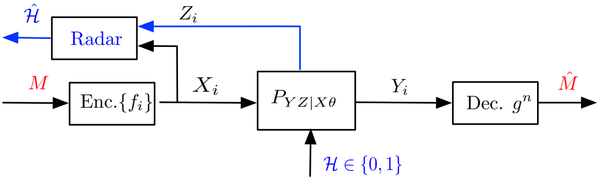
Cannot adjust codebook statistics P_X to the channel $\mathcal{H}$!

- Compound capacity with feedback:

$$C_{\text{compound,fb}} \leq \min_{\mathcal{H}} \max_{P_X} I(X;Y|\mathcal{H}) = \min_{\mathcal{H}} C(P^{\mathcal{H}}_{Y|X}).$$

Can learn channel and adapt codebook statistics!

Bi-Static ISAC with Chernoff Exponent



- Only non-adaptive strategies are possible

Theorem (Joudeh et al. and Chang et al.)

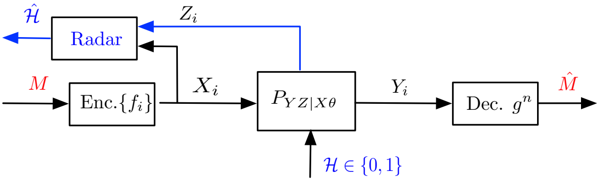
(R, E_{Sym}) *pairs are achievable iff for some* P_X*:*

$$
\begin{aligned}
R &\leq \min_{\mathcal{H}} I(X;Y|\mathcal{H}),\\
E_{\text{Sym}} &\leq \max_{l\in[0,1]} -\sum_x P_X(x) \log\left(\sum_z P^{\mathcal{H}}_{Z|X}(z|x)^l P^{\mathcal{H}}_{Z|X}(z|x)^{1-l}\right)
\end{aligned}
$$

- Tradeoff between sensing and communication due to common P_X! (input statistics)

[6] H. Joudeh & F. Willems, "Joint communication and binary state detection,"

ISAC Model with Stein Error Exponent



Theorem (Ahmadipour et al.)

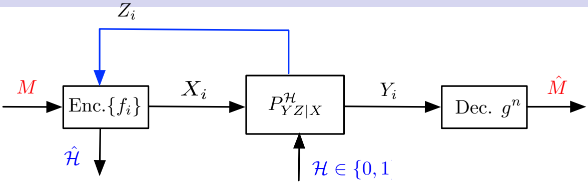
(R, E_{Stein}) *pairs are achievable iff for some* P_X*:*

$$\begin{aligned} R &\leq \min_{\mathcal{H}} I(X;Y|\mathcal{H}), \\ E_{\text{Stein}} &\leq \sum_x P_X(x) D(P_{Z|X}(\cdot|x) \| Q_{Z|X}(\cdot|x)) \end{aligned}$$

- Tradeoff between sensing and communication due to common P_X! (input statistics)

[8] M. Ahmadipour, M. Kobayashi, M. W. and G. Caire, "An Information-Theoretic Approach to Joint Sensing and Communication," *Trans. IT*, 2022.

Adaptive Channel Coding/Sensing



- Feedback allows the encoder to "learn" the channel parameter $\mathcal{H}$ and to adapt its coding to the correct channel
- Sensing (detection) problem is still open when $\mathcal{H}$ non-binary $\rightarrow$ adaptive inputs also improve detection performance
- Chang et al. propose joint sensing and communication schemes
- Problem seems difficult and is open!

[[7] M.-C. Chang, et al. "Rate and detection-error exponent tradeoff for joint communication and sensing of fixed channel states," *JSAIT 2023*.

Take-Aways from ISAC with Detection-Errors

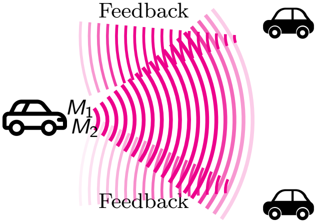
- Non-Adaptive case:
  - Complete information-theoretic results for various Detection Exponents
  - Optimal asymptotic performance only depends on x^n through its empirical statistics.
  - Sensing detector decides based on joint type of x^n and backscatterer depending on the exponent to maximize
  - Again a tradeoff between sensing and communication depending on input statistics

Take-Aways from ISAC with Detection-Errors

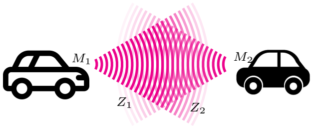
- Adaptive case:
  - Much more involved, communication problem is closed but not the optimal sensing $\rightarrow$ ISAC also widely open
  - Adaptive inputs can be used because the transmitter learns the channel and to better discriminate between different pairs of hypotheses.

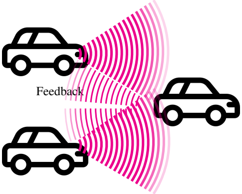
Network ISAC

Network ISAC



- Data sent to both receivers
- Fundamental limits partly characterized

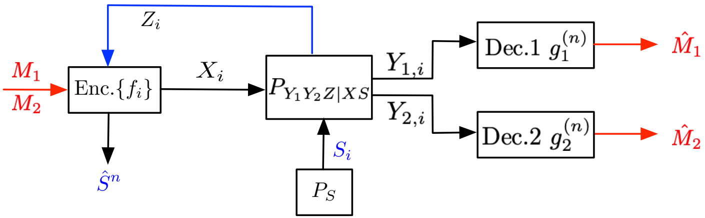




- Both Transmitters sense and send data
- Comm. path between Txs! → Collaborative comm. and sensing

Multi-Receiver ISAC

- State-dependent broadcast channel with backscatterer



- Arbitrary forward and backward channels $P_{Y_1Y_2|XS}$ and $P_{Z|Y_1Y_2XS}$
- Model includes as special cases Rx-CSI and two states $S = (S_1, S_2)$
- For most channels, feedback does increase capacity!!!

Fundamental Capacity-Distortion Region for BC

Definition

Capacity-distortion region is the set of triples (R_1, R_2, D) so that there exist encoder, decoders, and estimator with

$$\lim_{n\to\infty} \Pr(\hat{M}_k \neq M_k) = 0, \quad k \in \{1,2\}, \qquad \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[d(S_i, \hat{S}_i)] \leq D$$

Same per-symbol optimal estimator as for single-receiver!

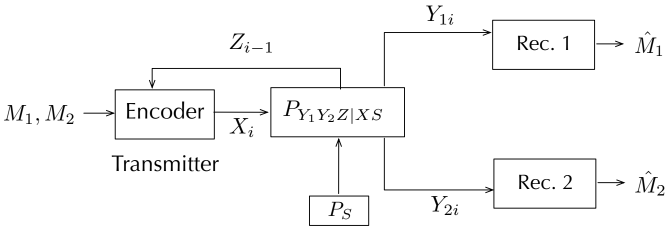
Optimal estimator: $\quad s^n = (\hat{s}^*(x_1, z_1), \hat{s}^*(x_2, z_2), \ldots, \hat{s}^*(x_n, z_n)),$

with

$$\hat{s}^*(x,z) := \arg\min_{s' \in \hat{\mathcal{S}}} \sum_{s \in \mathcal{S}} P_{S|XZ}(s|x,z) d(s,s').$$

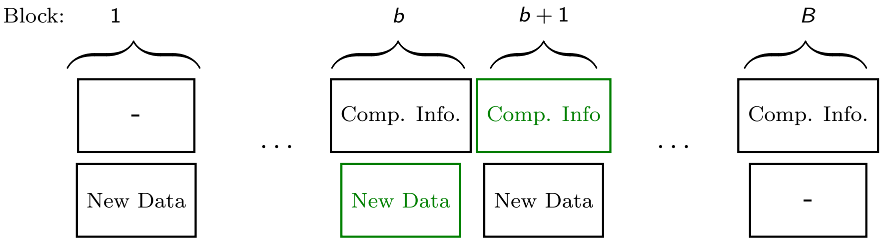
- Sensing performance depends only on statistics of x^n
- Find optimal feedback BC code under constrained x^n-statistics.

State-Dependent BCs with Generalized Feedback



- Feedback does not increase capacity of degraded BCs (El Gamal,'79)
- Achievable scheme for general BCs (Shayevitz et al'12, Venkataramanan et al'13)
- Capacity of several BCs with full Receiver-CSI (Kim et al'16)

Intuition about the Shayevitz-Wigger BC Scheme



1. Block-Markov strategy:
   - Compression info sent in block $b+1$: info about channel in block b learned via feedback
   - Block-b outputs improved with compression info sent in block $b+1$
2. New data and compression info sent with Marton's BC scheme (without feedback)

Results on Capacity-Distortion Region of BCs

Degraded Broadcast Channels $X \to Y_1 \to Y_2$

Capacity-distortion region: all (R_1, R_2, D) that for some P_{UX} satisfy

$$\begin{aligned} R_1 &\leq I(X; Y_1 \mid U) \\ R_2 &\leq I(U; Y_2), \\ \mathbb{E}[d(S, \hat{s}^*(X, Z))] &\leq D. \end{aligned}$$

- Tradeoff between communication and sensing from P_X.
- No-feedback codes with appropriate P_X.

Results on Capacity-Distortion Region of BCs

Degraded Broadcast Channels $X \to Y_1 \to Y_2$

Capacity-distortion region: all (R_1, R_2, D) that for some P_{UX} satisfy

$$\begin{aligned} R_1 &\leq I(X; Y_1 \mid U) \\ R_2 &\leq I(U; Y_2), \\ \mathbb{E}[d(S, \hat{s}^*(X, Z))] &\leq D. \end{aligned}$$

- Tradeoff between communication and sensing from P_X.
- No-feedback codes with appropriate P_X.

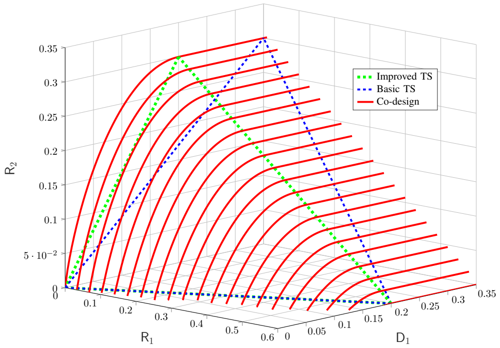
General Broadcast Channels

Inner and outer bounds (feasible and infeasible regions) based on Shayevitz-W. scheme and genie-aided bound

- Bounds in general case tight only in special cases.

Binary Fading Example: Capacity-Distortion Region

- Double-State $S = (S_1, S_2)$ with corr. components, known at Rxs!
- Fading outputs $Y_k = S_k X$, for $k = 1, 2$ (without noise)
- Perfect Rx CSI and both outputs fed back $Z = (Y_1, Y_2)$
- When $X = 1$ Tx learns S_1, S_2; when $X = 0$ it learns nothing



- Resource/time-sharing approaches sub-optimal
- Tradeoff betwen optimal sensing and comm. performances

Take-Away Messages for BC

- Information-theoretic model based on generalized feedback, memoryless state sequence, average distortion

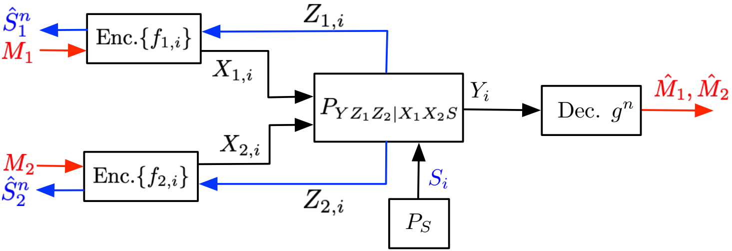
- Symbol-by-symbol estimator optimal; sensing performance depends only on empirical statistics of x^n

- Use optimal data communication scheme under restriction on empirical statistics of x^n
  $\rightarrow$ generalized feedback used for data communication

- 3-dimensional tradeoff between 2 rates and distortion

- Resource-sharing schemes highly suboptimal

Multi-Access ISAC



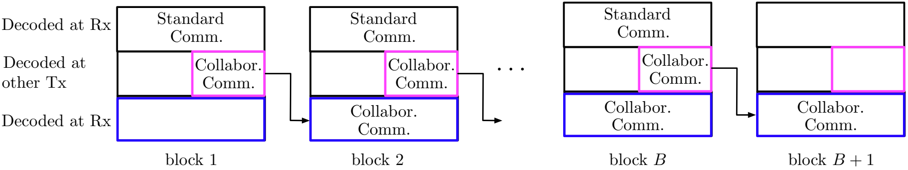
- Both Txs sense the state and send data
- Communication path between Txs!!

- Symbol-wise estimator at Tx k based on $(X_{k,i}, Z_{k,i})$ is suboptimal!
- Collaborative coding and sensing through Tx-Tx- paths!

Idea of Willems' Collaborative Comm Scheme

[Willems'83] scheme for the MAC with generalized feedback

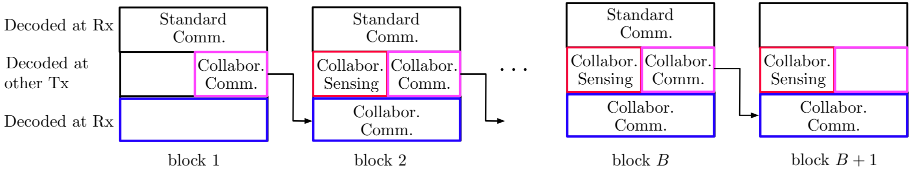
- Block-Markov coding and backwards decoding



- Txs exchange message parts over Tx-Tx paths
- Exchanged message parts are collaboratively re-transmitted in the next block
- Collaboratively transmitted message parts are easier to decode

Our Joint Collaborative Comm/Sensing Scheme

- Reuse Willem's multilayer block-Markov coding scheme



- After each block, each Tx extracts sensing info of interest to the other Rx

- This sensing info is transmitted to the other Tx during the next block

- Each Tx i estimates the state S_i^n based on its inputs/outputs, decoded codewords, and sensing info from the other Tx.

Binary MAC Example

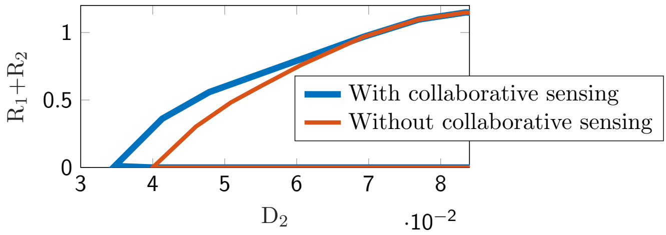
- S_1, S_2 i.i.d. Bernouilli-0.9, noises B_0, B_1, B_2 ind. Bernoulli, and

$$\begin{aligned} Y' &= S_1X_1 + S_2X_2 + B_0, && Y = (Y', S_1, S_2), \\ Z_k &= S_1X_1 + S_2X_2 + B_k, && \forall k \in \{1,2\}. \end{aligned}$$

- Hamming distortion $d(s, \hat{s}) = s \oplus \hat{s}$

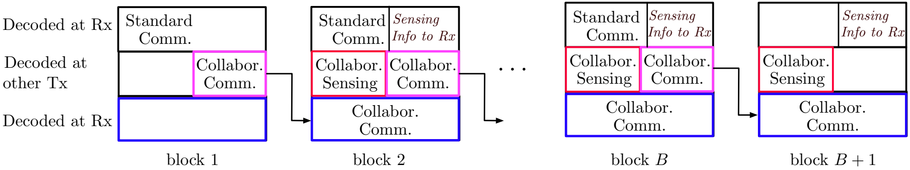
- Choose auxiliaries U_0, U_1, U_2 binary and

$$V_k = \begin{cases} \mathbb{1}\{Z_k = 1\} + 2 \cdot \mathbb{1}\{Z_k = 2\} & \text{if } E_k = 0 \\ \text{“?”} & \text{if } E_k = 1 \end{cases} \qquad \forall k = \{1,2\}$$



Further Improved ISAC MAC Scheme

- Reuse the Ahmadipour & W. multilayer block Markov scheme



- *In each block send sensing information from both Txs to the Rx*

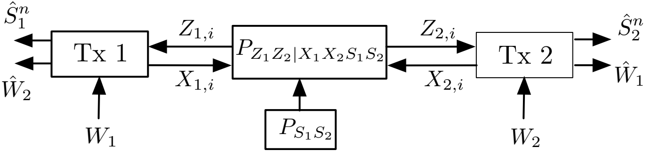
- Receiver can also decode the sensing information exchanged between the Txs

- This further improves the previous schemes

- Further improvements possible by improving Tx-to-Tx comm....

[9] Y. Liu, M. Li, A. Liu, L. Ong, and A. Yener, "Fundamental limits of multiple-access integrated sensing and communication systems,"

Device-to-Device ISAC



- Both Txs sense the state and send data
- Interactive communication between Txs!!
- We propose a coding scheme based on Block-Markov encoding:

  → Improved Han's interactive communication scheme with collaborative sensing!

Take-Aways for the MAC and D2D

- Symbol-by-symbol estimator based on inputs/outputs suboptimal
- Base estimator also on decoded codewords
- Sensing performance improved through collaborative sensing $\rightarrow$ Use the Tx-to-Tx path already used for feedback communication!
- Improved schemes are possible using interactive two-way schemes (Han) and joint source channel coding
- Tradeoff between sensing and communication

Summary

- Presented information-theoretic framework for integrated sensing and communication [Kobayashi,Caire,Kramer'18] and [Joudeh&Willems'22]
- Information-theoretic limits have been derived for various sensing criteria and discrete-memoryless channels/state sequences
- Single Tx: optimal sensing performance depends only on x^n statistics.
- Tradeoff between rates and distortion(s)/exponents .
- Multiple Txs: *Fully integrate coding for collaborative sensing and comm.*

Interesting future research directions

- Simplified capacity-expressions/coding schemes for Channels with memory
- Continuous-time channels
- Other sensing criteria
- Further investigations on secrecy constraints

IT References on ISAC with Distortion

- M. Kobayashi, G. Caire, and G. Kramer, "Joint state sensing and communication: Optimal tradeoff for a memoryless case," *ISIT* 2018.
- M. Ahmadipour, M. Kobayashi, M. W. and G. Caire, "An Information-Theoretic Approach to Joint Sensing and Communication," *Trans. IT*, 2022.
- Y. Xiong, F. Liu, Y. Cui, W. Yuan, T. X. Han, and G. Caire, "On the fundamental tradeoff of integrated sensing and communications under gaussian channels," *Trans. IT*, 2023. "
- H. Nikbakht, M. Wigger, S. Shamai (Shitz), and H. V. Poor, "Integrated Sensing and Communication in the Finite Blocklength Regime," *ISIT 2024*.
- Y. Chen, T. Oechtering, M. Skoglund, and Y. Luo. "On general capacity-distortion formulas of integrated sensing and communication," *Arxiv* , 2023.

IT References on ISAC with Exponents

- H. Joudeh and F. M. J. Willems, "Joint communication and binary state detection," *JSAIT*, 2022.
- H. Wu and H. Joudeh, "On joint communication and channel discrimination," *ISIT 2022*.
- M.-C. Chang, Erdogan, S.-Y. Wang, and M. R. Bloch, "Rate and detection-error exponent tradeoff for joint communication and sensing of fixed channel states," *JSAIT*, 2023.
- M. Ahmadipour, M. W., and S. Shamai (Shitz). "Strong converses for memoryless bi-static ISAC," *ISIT* 2023.
- M. Ahmadipour, M. Wigger, and S. Shamai, "Strong converses for memoryless bi-static ISAC," *IZS 2024*.

IT References on Network ISAC

- M. Kobayashi, H. Hamad, G. Kramer, and G. Caire, “Joint state sensing and communication over memoryless multiple access channels,” *Trans. IT*, July 2019.
- M. Ahmadipour and M. W. “An Information-Theoretic Approach to Collaborative Integrated Sensing and Communication for Two-Transmitter Systems,” *JSAIT*, 2023.
- M. Ahmadipour, M. Kobayashi, M. W. and G. Caire, “An Information-Theoretic Approach to Joint Sensing and Communication,” *Trans. IT*, 2022.
- Y. Liu, M. Li, A. Liu, L. Ong, and A. Yener, “Fundamental limits of multiple-access integrated sensing and communication systems,” *Arxiv*, 2023.
- Y. Liu, M. Li, Y. Han, and L. Ong, “Information-theoretic limits of integrated sensing and communication over interference channels,” *ICC 2024*.
- T. Jiao, Y. Geng, Z. Wei, K. Wan, Z. Yang, and G. Caire, “Information-theoretic limits of bistatic integrated sensing and communication,” *Arxiv*, 2023.

IT References on ISAC with Security Constraints

- O. Günlü, M. Bloch, R. F. Schaefer, and A. Yener, "Secure integrated sensing and communication," *JSAIT* 2023. 2023.
- M. Ahmadipour, M. W., and S. Shamai (Shitz). "Integrated communication and receiver sensing with security constraints on message and state," *ISIT 2023*.
- M. Mittelbach, R. F. Schaefer, M/ Bloch, A. Yener, and O. Günlü, "Secure Integrated Sensing and Communication Under Correlated Rayleigh Fading," *Arxiv* 2024.
- T. Welling, O. Günlü, and A. Yener, "Transmitter actions for secure integrated sensing and communication," *ISIT 2024*.

Related IT Works on State-Communication

- W. Zhang, S. Vedantam, and U. Mitra, "Joint transmission and state estimation: A constrained channel coding approach," *Trans. IT*, 2011.
- A. Salimi, W. Zhang, S. Vedantam, and U. Mitra, "The capacity- distortion function for multihop channels with state," *ISIT 2017*.
- C. Choudhuri, Y. Kim, and U. Mitra, "Causal state communication," *Trans. IT*, 2013.
- S. I. Bross and A. Lapidoth, "The Gaussian source-and-datastreams problem," *Trans. IT*, 2019.
- V. Ramachandran, S. R. B. Pillai, and V. M. Prabhakaran, "Joint state estimation and communication over a state-dependent Gaussian multiple access channel," *Trans. IT*, 2019.
- A. Lapidoth and Y. Steinberg, "The multiple-access channel with causal side information: Double state," *Trans. IT*, 2012.
- A. Lapidoth and Y. Steinberg, "The multiple-access channel with causal side information: Common state," *Trans. IT*, 2012.
- M. Li, O. Simeone, and A. Yener, "Multiple access channels with states causally known at transmitters," *Trans. IT*, 2012.

Related IT Works on Feedback

- F. M. Willems, "Information theoretical results for the discrete memoryless multiple access channel," Ph.D. dissertation 1983.
- O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *Trans. IT*, 2013.
- R. Venkataramanan and S. S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *Trans. IT*, 2013.