Achievable Rates for State-Dependent Discrete Memoryless Channels With Coded Sensing

Omkar M. Mujumdar
Dept. of Electrical Engineering
University of Notre Dame
IN, USA
omujumda@nd.edu

Wenyi Zhang
Dept. of Electronics Eng. and Information Science
University of Science and Technology of China
Hefei, China
wenyizha@ustc.edu.cn

J. Nicholas Laneman
Dept. of Electrical Engineering
Universiy of Notre Dame
IN, USA
jnl@nd.edu

***Abstract*—We describe a basic single-user integrated sensing and communication (ISAC) system, considering a state-dependent discrete memoryless channel (DMC) with independent and identically distributed (iid) state sequences that are independent of the input sequence. Furthermore, we assume that no feedback is present and no state information is available to either the transmitter or the receiver. We derive the highest transmission rates achievable for our channel under a given distortion constraint, for three different receiver schemes: *simultaneous decoding and estimation, decoding before estimation, and estimation before decoding*. The state estimator is based on the idea of *coded sensing*, i.e., it uses a lossy source coding codebook to determine the estimate of the state sequence. The results hint towards a trade-off among the performance limits of such ISAC systems.**

***Index Terms*—ISAC, state-dependent DMC, coded sensing, achievable transmission rate, minimum estimation rate.**

I. Introduction

Integrated sensing and communication (ISAC) has been identified as a key technology for many usage scenarios targeted by the upcoming 6G standards [1]. Traditionally, communication (e.g., cellular networks) and sensing networks (e.g., radar) have been developed and deployed independently. However, the burden on the allocated spectrum is increasing with the growing demand for intelligent systems and low-latency applications. Through joint design and optimization, ISAC systems use the same radio hardware and spectrum resources to enable communication and sensing applications to operate jointly within the same network.

Communication systems typically aim to transfer information at a high rate and with high reliability. Radars, on the other hand, detect targets and estimate relevant parameters by comparing the received signal to a known signal. The contrast in the key objectives of the two systems gives rise to performance trade-offs for ISAC. A central challenge of ISAC design lies in understanding the performance trade-offs that arise from fundamental limits dictated by information theory and detection-estimation theory.

State-dependent channels can be used to model the links in an ISAC network mathematically. In the case of a single-user state-dependent channel, conveying information and estimating the state may be required simultaneously.

Continuing the work in [2, Sec. 7.1.1], we consider a single-user state-dependent discrete memoryless channel (DMC) model where neither the transmitter nor the receiver is aware of the state, and no feedback is present. We determine the highest transmission rates achievable on the channel, for three different receiver schemes: *simultaneous decoding and estimation (SDE), decoding before estimation (DBE), and estimation before decoding (EBD)*. The state estimators we consider use a lossy source coding codebook to provide an estimate of the state sequence. We refer to this idea as *coded sensing*.

Note that the channel and setup considered in [3] are the same as ours, but only the DBE scheme is explored. Also, the estimator considered there is a symbol-by-symbol estimator.

This paper presents a problem formulation, states the results, and provides some discussion. Further details, including proofs of the results, are available in [4]. We introduce the channel model along with the definitions of a transmission-estimation code and an achievable rate-distortion pair, in Sec. II. In Sec. III, we derive the achievability results for the three receiver schemes. We borrow the *robust typicality*-based results in [5, Ch. 2] and the idea of indirect rate distortion (see, [6, Sec. 3.5] and [7]), for our derivations. We compare and discuss the achievability results in Sec. IV.

II. Problem Setup

A. Channel Model

We consider a state-dependent DMC with independent and identically distributed (iid) state sequences that are independent of the input. The channel is governed by conditional probability mass functions (pmfs), $\{p_{Y^n|X^n,S^n} = \prod_i p_{Y|X,S}\}_{n=1}^{\infty}$ [1], known to both the transmitter and the receiver. Here, X^n denotes an input sequence, S^n denotes a state sequence, and Y^n denotes a received sequence. The transmitter and the receiver also know the pmfs $\{p_{S^n} = \prod_i p_S\}_{n=1}^{\infty}$ that govern the state process. The input, state, and output alphabets $\mathcal{X}$, $\mathcal{S}$, and $\mathcal{Y}$, respectively, are all considered finite.

To convey a message $M \in \{1, \cdots, M_n\}$ to the receiver, the transmitter sends a length n sequence $X^n(M) \in \mathcal{X}^n$ over the

[1]Due to space constraints, we will use some shortcuts regarding notations related to index variables. The index variable i ranges between $\{1, \cdots, n\}$, l and $\tilde{l}$ between $\{1, \cdots, L_n\}$, m and $\hat{m}$ between $\{1, \cdots, M_n\}$, and $\tilde{m}$ between $\{2, \cdots, M_n\}$. Also, $\sum_i$ refers to $\sum_{i=1}^{n}$, $\prod_i$ refers to $\prod_{i=1}^{n}$, and $\bigcup_i$ or $\bigcap_i$ refer to $\bigcup_{i=1}^{n}$ or $\bigcap_{i=1}^{n}$, unless stated otherwise. The same holds for other index variables.

channel. Upon receiving an output sequence $Y^n \in \mathcal{Y}^n$, the receiver decodes the sent message as $\hat{M} \in \{1, \cdots, M_n\}$ and estimates the state sequence as $\hat{S}^n \in \hat{\mathcal{S}}^n$, where $\hat{\mathcal{S}}$ denotes a finite reconstruction alphabet.

B. *Definitions*

Definition 1 (Transmission Estimation Code)**.** An (n, M_n, D_n) *transmission-estimation code* consists of *message set* $\{1, \cdots, M_n\}$, *encoder* $f_n : \{1, \cdots, M_n\} \to \mathcal{X}^n$, *decoder* $g_n : \mathcal{Y}^n \to \{1, \cdots, M_n\}$, and *estimator* $h_n : \mathcal{Y}^n \to \hat{\mathcal{S}}^n$.

Assuming the messages to be uniformly distributed, the average (decoding) *error probability* is defined as $\bar{\varepsilon}_n = \frac{1}{M_n} \sum_m \mathbb{P}\{g_n(Y^n) \neq m | X^n = f_n(m)\}$. *Expected distortion* is defined as $D_n = \mathbb{E}[d_n(S^n, h_n(Y^n))]$, with $d_n(S^n, \hat{S}^n) = \frac{1}{n} \sum_i d(S_i, \hat{S}_i)$, where $d : \mathcal{S} \times \hat{\mathcal{S}} \to [0, \infty)$ is a bounded non-negative *distortion metric*.

Definition 2 (Achievable Rate-Distortion Pair)**.** A rate-distortion pair (R_t, D) is *achievable* on a given channel, if there exists a sequence $\{(n, M_n, D_n)\}_{n=1}^{\infty}$ of transmission estimation codes, that simultaneously satisfies $\lim_{n\to\infty} \bar{\varepsilon}_n = 0$, $\limsup_{n\to\infty} D_n \leq D$, and $\liminf_{n\to\infty} \frac{1}{n} \log_2 M_n \geq R_t$, where D represents a constraint on the expected distortion.

III. Achievability Results

A. *Simultaneous Decoding and Estimation (SDE) Scheme*

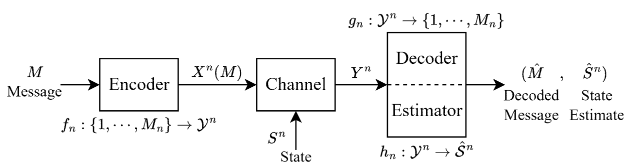

Fig. 1. Simultaneous Decoding and Estimation (SDE) Scheme

The joint pmf on $(X^n, S^n, Y^n, \hat{S}^n)$ is expressed as

$$\begin{aligned} p_{X^n,S^n,Y^n,\hat{S}^n} &= p_{X^n} p_{S^n|X^n} p_{Y^n|X^n,S^n} p_{\hat{S}^n|X^n,S^n,Y^n} \\ &= p_{X^n} p_{S^n} p_{Y^n|X^n,S^n} p_{\hat{S}^n|X^n,Y^n}. \end{aligned} \quad (1)$$

We consider the state reconstruction sequence $\hat{S}^n$ to be conditionally independent of the state S^n, given the input X^n and the output Y^n.

We are given the pmf of the state and the channel, $p_{S^n} = \prod_i p_S$ and $p_{Y^n|X^n,S^n} = \prod_i p_{Y|X,S}$. Let's fix $p_{X^n} = \prod_i p_X$, the pmf of the input, and the pmf of the state reconstruction sequence, given the input and output, $p_{\hat{S}^n|X^n,Y^n} = \prod_i p_{\hat{S}|X,Y}$.

- Generate a transmission codebook $\mathcal{C}_n^t$ consisting of M_n codewords, $x_m^n \in \mathcal{X}^n$, each drawn independently according to the pmf p_{X^n}.
- Also, generate estimation codebook $\mathcal{C}_n^e$ consisting of L_n[2] codewords $\hat{s}_l^n \in \hat{\mathcal{S}}^n$, each drawn independently according to the marginal pmf $p_{\hat{S}^n} \sim \prod_i p_{\hat{S}}$.
- For a message $M = m$, the encoder f_n transmits the m^{th} codeword in the codebook $\mathcal{C}_n^t$, on the channel.
- Upon receiving Y^n, the receiver looks for a pair $(x_{\hat{m}}^n, \hat{s}_l^n)$ in $(\mathcal{C}_n^t, \mathcal{C}_n^e)$ (respectively), that is jointly typical[3] with Y^n. If such a pair exists and $\hat{m}$ is unique, then $\hat{M} = \hat{m}$ is declared as the decoded message, and $\mathcal{C}_n^e(L)$ is declared as the state estimate, where L is chosen as the minimum among the various values obtained for the index l. Otherwise, $\hat{M} = M_n$ is declared as the decoded message, and $\hat{s}_1^n$ is declared as the state estimate.

In other words, for some $y^n \in \mathcal{Y}^n$, the *decoder* output $g_n(y^n) = \hat{m}$ only if following two conditions are simultaneously satisfied for some $\hat{m}$, else $g_n(y^n) = M_n$. First, for at least one l, $(\mathcal{C}_n^t(\hat{m}), y^n, \mathcal{C}_n^e(l)) \in \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$. Second, for each $\tilde{m}$, $(\mathcal{C}_n^t(\tilde{m}), y^n, \mathcal{C}_n^e(\tilde{l})) \notin \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ for each $\tilde{l}$.

For some $y^n \in \mathcal{Y}^n$, the *estimator* output $h_n(y^n) = \hat{s}_L^n$, where $L = \min \mathcal{L}$ when the set $\mathcal{L}$ is not empty, and $L = 1$ otherwise. Here, $\mathcal{L}$ denotes the set of all indices l for which $(\mathcal{C}_n^t(\hat{m}), y^n, \mathcal{C}_n^e(l)) \in \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ for some $\hat{m}$, and for each $\tilde{m}$, $(\mathcal{C}_n^t(\tilde{m}), y^n, \mathcal{C}_n^e(\tilde{l})) \notin \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ for each $\tilde{l}$.

1) Error Probability Analysis: We determine the error probability, averaged over all possible codebook pairs $(\mathcal{C}_n^t, \mathcal{C}_n^e)$. We consider the message $M = 1$ to be sent over the channel.

The probability of error, averaged over all possible codebook pairs $(\mathcal{C}_n^t, \mathcal{C}_n^e)$, is expressed as

$$\mathbb{E}[\bar{\varepsilon}_n] = \sum_{\mathcal{C}_n^t, \mathcal{C}_n^e} \mathbb{P}\{\mathcal{C}_n^t\} \mathbb{P}\{\mathcal{C}_n^e\} \mathbb{P}\{g_n(Y^n) \neq 1 | X^n = f_n(1)\}.$$

For each m and l, let us denote the event $(X_m^n, Y^n, \hat{S}_l^n) \notin \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ by $\mathcal{E}_{m,l}$, where X_m^n represents the m^{th} codeword in a transmission codebook, and $\hat{S}_l^n$ irepresents the l^{th} estimate in an estimation codebook.

Given the codeword X_1^n is transmitted, a decoding error occurs when either of the following two events occurs. First, $(X_1^n, Y^n, \hat{S}_l^n) \notin \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ for each l. Second, There exists at least one $\tilde{m}$ for which, $(X_{\tilde{m}}^n, Y^n, \hat{S}_{\tilde{l}}^n) \in \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ for at least one $\tilde{l}$. Let's denote the first error event by $\mathcal{E}_A$ and the second by $\mathcal{E}_B$.

For n large enough, consider $M_n = 2^{nR_t}$ and $L_n = 2^{nR_e}$. In [4], we show that $\mathbb{P}\{\mathcal{E}_A\} \leq e^{-2^{n(R_e - I(X,Y;\hat{S}) - \delta_n(\varepsilon))}}$, and $\mathbb{P}\{\mathcal{E}_B\} \leq 2^{-n[I(X;Y) - R_t - \delta(\varepsilon)]}$, where $\delta_n(\varepsilon)$ and $\delta(\varepsilon)$ tend to 0 as $\varepsilon \to 0$.

Thus, applying the union bound,

$$\begin{aligned} \mathbb{E}[\bar{\varepsilon}_n] &\leq \mathbb{P}\{\mathcal{E}_A\} + \mathbb{P}\{\mathcal{E}_B\} \\ &\leq e^{-2^{n(R_e - I(X,Y;\hat{S}) - \delta_n(\varepsilon))}} + 2^{-n[I(X;Y) - R_t - \delta(\varepsilon)]}. \end{aligned} \quad (2)$$

The above term tends to 0 as $n \to \infty$ and $\varepsilon \to 0$, provided $R_e > I(X, Y; \hat{S})$ and $R_t < I(X; Y)$.

2) Expected Distortion Calculation: A joint pmf on $(X^n, S^n, Y^n, \hat{S}^n)$ is expressed as $p_{X^n,S^n,Y^n,\hat{S}^n} = p_{S^n} p_{X^n,Y^n|S^n} p_{\hat{S}^n|X^n,S^n,Y^n}$. But, for the SDE scheme, since $\hat{S}^n$ is conditionally independent of S^n given X^n and Y^n, $p_{\hat{S}^n|X^n,S^n,Y^n} = p_{\hat{S}^n|X^n,Y^n}$. Thus, $S^n \to (X^n, Y^n) \to \hat{S}^n$ forms a Markov chain.

[2]The quantity $R_e = \frac{1}{n} \log_2 L_n$ could be interpreted as *estimation rate*, where L_n denotes the number of state estimates we consider.

[3]The notion of typicality, known as *robust typicality*, is adopted from [5, Sec. 2.4, 2.5].

The expected distortion of a transmission-estimation code, $D_n = \mathbb{E}[d_n(S^n, h_n(Y^n))]$, can be expressed as [4],

$$D_n = \frac{1}{M_n}\sum_m \sum_{s^n,y^n} p_{S^n,Y^n|X^n}(s^n, y^n|x_m^n) d_n(s^n, h_n(y^n))$$

$$= \frac{1}{M_n}\sum_m \sum_{y^n} p_{Y^n|X^n}(y^n|x_m^n)\tilde{d}_n((x_m^n, y^n), h_n(y^n)). \quad (3)$$

Where, $\tilde{d}_n((x^n, y^n), \hat{s}^n) = \frac{1}{n}\sum_i \tilde{d}((x_i, y_i), \hat{s}_i)$ with

$$\tilde{d}((x, y), \hat{s}) = \sum_s p_{S|X,Y}(s|x, y)d(s, \hat{s})$$
$$= \mathbb{E}[d(S, \hat{s})|X = x, Y = y].$$

Using the modified additive distortion metric above, we can reduce an indirect rate distortion problem (see e.g., [5, Problem 3.19], [6, Sec. 3.5], and [7]) to a standard rate distortion (lossy source coding) problem (see e.g., [5, Sec. 3.6]).

The expected distortion, averaged over all possible codebook pairs, can be expressed as

$$\mathbb{E}[D_n] = \sum_{\mathcal{C}_n^t, \mathcal{C}_n^e} \mathbb{P}\{\mathcal{C}_n^t\}\mathbb{P}\{\mathcal{C}_n^e\}\mathbb{E}[d_n(S^n, h_n(Y^n))|X^n = x_1^n]$$

$$= \sum_{\mathcal{C}_n^t, \mathcal{C}_n^e} \mathbb{P}\{\mathcal{C}_n^t\}\mathbb{P}\{\mathcal{C}_n^e\}\mathbb{E}[\tilde{d}_n((x_1^n, Y^n), h_n(Y^n))|X^n = x_1^n]$$

$$= \Big[\sum_{(\mathcal{C}_n^t, y^n, \mathcal{C}_n^e)\in\mathcal{E}_A\bigcup\mathcal{E}_B} + \sum_{(\mathcal{C}_n^t, y^n, \mathcal{C}_n^e)\notin\mathcal{E}_A\bigcup\mathcal{E}_B}\Big]\mathbb{P}\{\mathcal{C}_n^t\}\mathbb{P}\{\mathcal{C}_n^e\}\cdot$$
$$\cdot\, p_{Y^n|X^n}(y^n|x_1^n)\tilde{d}_n((x_1^n, y^n), h_n(y^n)). \quad (4)$$

Let us define $\tilde{d}_{\max} = \max_{(\mathcal{X}\times\mathcal{Y})\times\hat{\mathcal{S}}} \tilde{d}((x, y), \hat{s})$. Since we consider the metric $d(s, \hat{s})$ to be bounded, the distortion metric $\tilde{d}((x, y), \hat{s})$ will also be bounded.

Notice that $(X_1^n, Y^n, h_n(Y^n)) \in \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$ when $(\mathcal{C}_n^t, y^n, \mathcal{C}_n^e) \notin \mathcal{E}_A\bigcup\mathcal{E}_B$. Thus, according to the *typical average lemma* [5, p. 26] and (2),

$$\mathbb{E}[D_n] \le \tilde{d}_{\max}(\mathbb{P}\{\mathcal{E}_A\} + \mathbb{P}\{\mathcal{E}_B\})$$
$$+ (1 + \varepsilon)\mathbb{E}[\tilde{d}((X, Y), \hat{S})]\mathbb{P}\{\mathcal{E}_A^c \cap \mathcal{E}_B^c\}$$
$$\le \tilde{d}_{\max}(e^{-2^{n(R_e - I(X,Y;\hat{S}) - \delta_n(\varepsilon))}} + 2^{-n[I(X;Y) - R_t - \delta(\varepsilon)]})$$
$$+ (1 + \varepsilon)\mathbb{E}[\tilde{d}((X, Y), \hat{S})]. \quad (5)$$

Hence, if $\mathbb{E}[\tilde{d}((X, Y), \hat{S})] \le \frac{D}{(1+\varepsilon)}$, where D denotes the distortion constraint, then $\mathbb{E}[D_n] \le D$ as $n \to \infty$ and $\varepsilon \to 0$, provided $R_e > I(X, Y; \hat{S})$ and $R_t < I(X; Y)$.

3) Result: It can be easily verified that $\mathbb{E}[\tilde{d}((X, Y), \hat{S})] = \mathbb{E}[d(S, \hat{S})]$. Let $\mathcal{P}_1(D)$ denote the set of all joint pmfs $p_{X,S,Y,\hat{S}} = p_X p_S p_{Y|X,S} p_{\hat{S}|X,Y}$ that satisfy the inequality

$$\mathbb{E}[d(S, \hat{S})] = \sum_{x,s,y,\hat{s}} p_{X,S,Y,\hat{S}}(x, s, y, \hat{s})d(s, \hat{s}) \le D. \quad (6)$$

Given a state pmf p_S, a channel pmf $p_{Y|X,S}$, a non-negative bounded distortion metric $d(S, \hat{S})$, and a valid distortion constraint D[4]; if we choose pmfs p_X and $p_{\hat{S}|X,Y}$ such that the joint pmf $p_{X,S,Y,\hat{S}} = p_X p_S p_{Y|X,S} p_{\hat{S}|X,Y} \in \mathcal{P}_1(D)$; then for n sufficiently large and ε sufficiently small, provided $R_e > I(X, Y; \hat{S})$ and $R_t < I(X; Y)$, we see that

1) The average expected distortion, considering all possible codebook pairs $(\mathcal{C}_n^t, \mathcal{C}_n^e)$, is upper bounded by D.
2) The average error probability, considering all possible codebook pairs, is arbitrarily close to 0.

Following the first observation above, there exists a codebook pair $(\tilde{\mathcal{C}}_n^t, \tilde{\mathcal{C}}_n^e)$ for which the expected distortion D_n is upper bounded by D.

According to the Markov inequality [8, p. 64], the probability of error $\overline{\varepsilon}_n$ for a randomly generated codebook pair, tends to 0 in probability as $n \to \infty$ [5, Remark 3.3]. Hence, it is likely, with probability arbitrarily close to 1, that the probability of error for $(\tilde{\mathcal{C}}_n^t, \tilde{\mathcal{C}}_n^e)$ is arbitrarily close to 0.

We state the achievability result for the SDE scheme below.

Theorem 1. *Given a state pmf p_S, a channel conditional pmf $p_{Y|X,S}$, a non-negative bounded distortion metric $d(S, \hat{S})$, and a valid distortion constraint D, the highest transmission rate achievable using the SDE scheme is*

$$R_t^*(D) = \max_{p_{X,S,Y,\hat{S}}\in\mathcal{P}_1(D)} I(X; Y).$$

Remark 1. The minimum estimation rate necessary for transmission rate R_t to be achievable, using the SDE scheme, is

$$R_e^*(D) = \min_{p_{X,S,Y,\hat{S}}\in\mathcal{P}_1(D)} I(X, Y; \hat{S}).$$

B. Decoding Before Estimation (DBE) Scheme

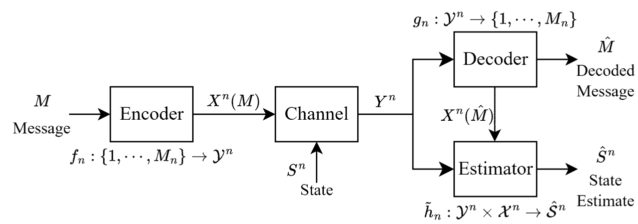

Fig. 2. Decoding Before Estimation (DBE) Scheme

The joint pmf on $(X^n, S^n, Y^n, \hat{S}^n)$ is exactly the same as in (1). Let's fix pmfs $p_{X^n} = \prod_i p_X$ and $p_{\hat{S}^n|X^n,Y^n} = \prod_i p_{\hat{S}|X,Y}$. The processes of codebook pair generation and encoding are the same as before.

- Upon receiving Y^n, the decoder $g_n : \mathcal{Y}^n \to \{1, \cdots, M_n\}$ looks for a codeword in $\mathcal{C}_n^t$, that is jointly typical with Y^n. If such a codeword exists and is unique, $g_n(Y^n) = \hat{m}$ is declared as the decoded message, else $g_n(Y^n) = M_n$.
- The corresponding decoded codeword $\mathcal{C}_n^t(\hat{M})$ and Y^n are the inputs to the estimator[5] $\tilde{h}_n : \mathcal{Y}^n \times \mathcal{X}^n \to \hat{\mathcal{S}}^n$. The estimator looks for an estimate in $\mathcal{C}_n^e$, that is jointly typical with $(\mathcal{C}_n^t(\hat{M}), Y^n)$. If such an estimate exists, then $\tilde{h}_n(Y^n, \mathcal{C}_n^t(\hat{M})) = \mathcal{C}_n^e(L)$ is declared as the state estimate, where L is chosen as the minimum value among all the indices l for which $(\mathcal{C}_n^t(\hat{M}), Y^n, \mathcal{C}_n^e(l))$ are jointly typical. Otherwise, $\tilde{h}_n(Y^n, \mathcal{C}_n^t(\hat{M})) = \mathcal{C}_n^e(1)$.

1) Error Probability Analysis: Again, we determine the error probability, averaged over all possible codebook pairs, and consider message $M = 1$ to be sent over the channel.

[4]For D to be a valid constraint on the expected distortion, we require $0 \le D \le \max_{\mathcal{S}\times\hat{\mathcal{S}}} d(s, \hat{s})$, and given p_S and $p_{Y|X,S}$, $\mathcal{P}_1(D)$ not to be empty.

[5]Considering Def. 1, the function $h_n(Y^n) \triangleq \tilde{h}_n(Y^n, f_n(g_n(Y^n)))$ defines the estimator.

For each m, let us denote the event $(X_m^n, Y^n) \notin \mathcal{T}_\varepsilon^n(X,Y)$ by $\mathcal{E}_m$. Given codeword X_1^n is transmitted, a decoding error occurs either when $(X_1^n, Y^n) \notin \mathcal{T}_\varepsilon^n(X,Y)$, or when there exists at least one $\tilde{m}$ for which $(X_{\tilde{m}}^n, Y^n) \in \mathcal{T}_\varepsilon^n(X,Y)$.

According to the definition of joint typicality [5, Sec. 2.5], $\mathbb{P}\{\mathcal{E}_1\} \to 0$ as $n \to \infty$. Also, $\mathbb{P}\{\bigcup_{\tilde{m}} \mathcal{E}_{\tilde{m}}^c\} \leq 2^{-n[I(X;Y)-R_t-\delta(\varepsilon)]}$ for $M_n = 2^{nR_t}$ [4].

Thus, the probability of decoding error, averaged over all possible codebook pairs $(\mathcal{C}_n^t, \mathcal{C}_n^e)$,

$$\mathbb{E}[\bar{\varepsilon}_n] \leq \mathbb{P}\{\mathcal{E}_1\} + \mathbb{P}\{\bigcup\nolimits_{\tilde{m}} \mathcal{E}_{\tilde{m}}^c\}, \tag{7}$$

which tends to 0 as $n \to \infty$ and $\varepsilon \to 0$, given $R_t < I(X;Y)$.

2) Expected Distortion Calculation: As in the SDE scheme, $S^n \to (X^n, Y^n) \to \hat{S}^n$ forms a Markov chain for this scheme as well. Thus, similar to (3),

$$\begin{aligned} D_n &= \mathbb{E}[d_n(S^n, \tilde{h}_n(Y^n, f_n(g_n(Y^n))))] \\ &= \frac{1}{M_n}\sum\nolimits_m \mathbb{E}[d_n(S^n, \tilde{h}_n(Y^n, f_n(g_n(Y^n))))|X^n = f_n(m)] \\ &= \frac{1}{M_n}\sum\nolimits_m \mathbb{E}[\tilde{d}_n(x_m^n, Y^n, \tilde{h}_n(Y^n, f_n(g_n Y^n)))|X^n = x_m^n]. \end{aligned}$$

Let us denote the event $\mathcal{E}_1 \bigcup (\bigcup_{\tilde{m}} \mathcal{E}_{\tilde{m}}^c)$ by $\mathcal{A}$. The expected distortion, averaged over all possible codebook pairs,

$$\begin{aligned} \mathbb{E}[D_n] = \Big[\sum_{(\mathcal{C}_n^t, y^n)\in\mathcal{A}} + \sum_{(\mathcal{C}_n^t, y^n)\notin\mathcal{A}} \Big] \sum_{\mathcal{C}_n^e} \mathbb{P}\{\mathcal{C}_n^t\} \cdot \\ \cdot\, \mathbb{P}\{\mathcal{C}_n^e\} p_{Y^n|X^n}(y^n|x_1^n) \tilde{d}_n(x_1^n, y^n, \tilde{h}_n(y^n, f_n(g_n(y^n)))). \end{aligned} \tag{8}$$

In [4], we show that (8) is bounded as follows

$$\begin{aligned} \mathbb{E}[D_n] \leq \tilde{d}_{\max}(\mathbb{P}\{\mathcal{E}_1\} + \mathbb{P}\{\bigcup\nolimits_{\tilde{m}} \mathcal{E}_{\tilde{m}}^c\} \\ + e^{-2^{n[R_e - I(X,Y;\hat{S}) - \delta_n(\varepsilon)]}}) + (1+\varepsilon)\mathbb{E}[\tilde{d}((X,Y),\hat{S})]. \end{aligned} \tag{9}$$

Following (7) and because $\tilde{d}_{\max} < \infty$, if $\mathbb{E}[\tilde{d}((X,Y),\hat{S})] \leq \frac{D}{(1+\varepsilon)}$, then $\mathbb{E}[D_n] \leq D$ as $n \to \infty$ and $\varepsilon \to 0$, provided $R_t < I(X;Y)$ and $R_e > I(X,Y;\hat{S})$.

3) Result: Given state pmf p_S, channel pmf $p_{Y|X,S}$, distortion metric $d(S,\hat{S})$, and valid distortion constraint D; if we choose p_X and $p_{\hat{S}|X,Y}$ such that joint pmf $p_{X,S,Y,\hat{S}} \in \mathcal{P}_1(D)$; then for n sufficiently large and ε sufficiently small, provided $R_t < I(X;Y)$ and $R_e > I(X,Y;\hat{S})$,

- There exists a codebook pair $(\tilde{\mathcal{C}}_n^t, \tilde{\mathcal{C}}_n^e)$ for which the expected distortion D_n is upper bounded by D.
- Using the Markov inequality-based argument in III-A3, it is highly likely that the probability of error for the codebook $\tilde{\mathcal{C}}_n^t$ is arbitrarily close to 0.

We state the achievability result for the DBE scheme below.

Theorem 2. *Given a state pmf p_S, a channel conditional pmf $p_{Y|X,S}$, a non-negative bounded distortion metric $d(S,\hat{S})$, and a valid distortion constraint D, the highest transmission rate achievable using the DBE scheme is*

$$R_t^*(D) = \max\nolimits_{p_{X,S,Y,\hat{S}} \in \mathcal{P}_1(D)} I(X;Y).$$

Remark 2. The minimum estimation rate necessary for transmission rate R_t to be achievable, using the DBE scheme, is

$$R_e^*(D) = \min\nolimits_{p_{X,S,Y,\hat{S}} \in \mathcal{P}_1(D)} I(X,Y;\hat{S}).$$

Remark 3. The DBE scheme yields the same rates $R_t^*(D)$ and $R_e^*(D)$ as the SDE scheme (Thm. 1 and Remark 1).

C. Estimation Before Decoding (EBD) Scheme

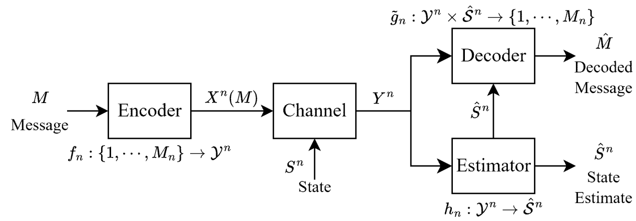

Fig. 3. Estimation Before Decoding (EBD) Scheme

For this scheme, the joint pmf on $(X^n, S^n, Y^n, \hat{S}^n)$ is $p_{X^n} p_{S^n} p_{Y^n|X^n,S^n} p_{\hat{S}^n|Y^n}$. Here, we consider the state reconstruction sequence $\hat{S}^n$ to be conditionally independent of both the state S^n and the input X^n, given the output Y^n.

Let us fix the pmfs $p_{X^n} = \prod_i p_X$ and $p_{\hat{S}^n|Y^n} = \prod_i p_{\hat{S}|Y}$. The processes of codebook pair generation and encoding are the same as before.

- Upon receiving Y^n, the estimator $h_n : \mathcal{Y}^n \to \hat{\mathcal{S}}^n$ looks for an estimate in $\mathcal{C}_n^e$ that is jointly typical with Y^n. If such an estimate exists, then $h_n(Y^n) = \mathcal{C}_n^e(L)$ is declared as the state estimate, where L is chosen as the minimum value among all the indices l for which $(Y^n, \mathcal{C}_n^e(l))$ are jointly typical, else $h_n(Y^n) = \mathcal{C}_n^e(1)$.
- The state estimate $\mathcal{C}_n^e(L)$ and Y^n are the inputs to the decoder[6], $\tilde{g}_n : \mathcal{Y}^n \times \hat{\mathcal{S}}^n \to \{1, \cdots, M_n\}$. The decoder looks for a codeword in $\mathcal{C}_n^t$ which is jointly typical with the pair $(Y^n, \mathcal{C}_n^e(L))$. If such a codeword exists and is unique, then $\tilde{g}_n(Y^n, \mathcal{C}_n^e(L)) = \hat{m}$ is declared as the decoded message. Otherwise, $\tilde{g}_n(Y^n, \mathcal{C}_n^e(L)) = M_n$.

1) Expected Distortion Calculation: In contrast to the earlier two schemes, in the DBE scheme, $S^n \to Y^n \to \hat{S}^n$ forms a Markov chain. In [4], we show that

$$\begin{aligned} D_n &= \sum\nolimits_{y^n} \sum\nolimits_{s^n} p_{S^n,Y^n}(s^n, y^n) \frac{1}{n} \sum\nolimits_i d(s_i, [h_n(y^n)]_i) \\ &= \sum\nolimits_{y^n} p_{Y^n}(y^n) \sum\nolimits_i \frac{1}{n} \sum\nolimits_{s_i} p_{S|Y}(s_i|y_i) d(s_i, [h_n(y^n)]_i) \\ &= \sum\nolimits_{y^n} p_{Y^n}(y^n) \frac{1}{n} \sum\nolimits_i \hat{d}(y_i, [h_n(y^n)]_i) \\ &= \sum\nolimits_{y^n} p_{Y^n}(y^n) \hat{d}_n(y^n, h_n(y^n)) = \mathbb{E}[\hat{d}_n(Y^n, h_n(Y^n))], \end{aligned}$$

where $\hat{d}(y,\hat{s}) = \mathbb{E}[d(S,\hat{s})|Y=y] = \sum_s p_{S|Y}(s|y) d(s,\hat{s})$.

Let us define $\hat{d}_{\max} = \max_{\mathcal{Y}\times\hat{\mathcal{S}}} \hat{d}(y,\hat{s})$. Since we consider $d(s,\hat{s})$ to be bounded, the metric $\hat{d}(y,\hat{s})$ will also be bounded.

The expected distortion, averaged over all possible codebook pairs, is bounded as follows [4],

$$\begin{aligned} \mathbb{E}[D_n] \leq \hat{d}_{\max}(\mathbb{P}\{Y^n \notin \mathcal{T}_{\tilde{\varepsilon}}^n(Y)\} + e^{-2^{n[R_e - I(Y;\hat{S}) - \delta_n(\tilde{\varepsilon})]}}) \\ + (1+\tilde{\varepsilon})\mathbb{E}[\hat{d}(Y,\hat{S})]. \end{aligned} \tag{10}$$

Since $\tilde{d}_{\max} < \infty$, if $\mathbb{E}[\hat{d}(Y,\hat{S})] \leq \frac{D}{(1+\tilde{\varepsilon})}$, then $\mathbb{E}[D_n] \leq D$ as $n \to \infty$ and $\tilde{\varepsilon} \to 0$, provided $R_e > I(Y;\hat{S})$.

[6]Considering Def. 1, $g_n(Y^n) \triangleq \tilde{g}_n(Y^n, h_n(Y^n))$ defines the decoder.

2) Error Probability Analysis: We are given the pmf $p_{S^n} = \prod_i p_S$ of the state and the pmf $p_{Y^n|X^n,S^n} = \prod_i p_{Y|X,S}$ of the channel. From this, we can derive the conditional pmf $p_{Y^n|X^n} = \prod_i p_{Y|X}$ of the channel between input X^n and the output Y^n. Since we consider $\hat{S}^n$ to be conditionally independent of X^n given Y^n, if we further consider iid pmfs $p_{X^n} = \prod_i p_X$ and $p_{\hat{S}^n|Y^n} = \prod_i p_{\hat{S}|Y}$, the joint pmf on $(X^n, Y^n, \hat{S}^n)$ can be expressed as $p_{X^n,Y^n,\hat{S}^n} = p_{X^n} p_{Y^n|X^n} p_{\hat{S}^n|Y^n} = \prod_i p_X p_{Y|X} p_{\hat{S}|Y}$.

Thus, $X^n \to Y^n \to \hat{S}^n$ forms a Markov chain. Hence, we know that $\hat{S}^n \to Y^n \to X^n$ also forms a Markov chain.

Again, we determine the error probability, averaged over all possible codebook pairs $(\mathcal{C}_n^t, \mathcal{C}_n^e)$, and consider message $M = 1$ to be sent over the channel. Given codeword X_1^n is transmitted, a decoding error occurs when either $(X_1^n, Y^n, h_n(Y^n)) \notin \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$, or there exists at least one $\tilde{m}$ for which $(X_{\tilde{m}}^n, Y^n, h_n(Y^n)) \in \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$.

Lets define a set $\mathcal{B}$ that consists of all possible pairs $(\mathcal{C}_n^e, y^n)$ such that $(y^n, h_n(y^n)) \in \mathcal{T}_{\tilde{\varepsilon}}^n(Y, \hat{S})$. Then, the probability of decoding error, averaged over all possible codebook pairs, is

$$\mathbb{E}[\bar{\varepsilon}_n] = \Big[\sum_{(\mathcal{C}_n^e, y^n) \in \mathcal{B}} + \sum_{(\mathcal{C}_n^e, y^n) \notin \mathcal{B}} \Big] \mathbb{P}\{\mathcal{C}_n^e\} \sum_{\mathcal{C}_n^t} \mathbb{P}\{\mathcal{C}_n^t\} \cdot \\ \cdot p_{Y^n|X^n}(y^n|x_1^n) \mathbf{1}\{\tilde{g}_n(y^n, h_n(y^n)) \neq 1\}, \tag{11}$$

where $\mathbf{1}\{\cdot\}$ denotes an indicator function.

Now, for each $(y^n, \hat{s}^n) \in \mathcal{Y}^n \times \hat{\mathcal{S}}^n$ and $m \in \{1, \cdots, M_n\}$, let $\mathcal{E}_m(y^n, \hat{s}^n)$ denote $(X_m^n, y^n, h_n(y^n)) \notin \mathcal{T}_\varepsilon^n(X, Y, \hat{S})$.

In [4], we show that (11) is bounded as follows,

$$\mathbb{E}[\bar{\varepsilon}_n] \leq \mathbb{P}\{(Y^n, h_n(Y^n)) \notin \mathcal{T}_{\tilde{\varepsilon}}^n(Y, \hat{S})\} \\ + \sum_{(\mathcal{C}_n^e, y^n) \in \mathcal{B}} \mathbb{P}\{\mathcal{C}_n^e\} p_{Y^n}(y^n) \big[\mathbb{P}\{\mathcal{E}_1(y^n, h_n(y^n))\} \\ + \mathbb{P}\{\bigcup_{\tilde{m}} \mathcal{E}_{\tilde{m}}^c(y^n, h_n(y^n))\}\big], \tag{12}$$

tends to 0 as $n \to \infty$, given $R_e > I(Y; \hat{S})$, $R_t < I(X; Y)$.

3) Result: It can be easily verified that $\mathbb{E}[\hat{d}(Y, \hat{S})] = \mathbb{E}[d(S, \hat{S})]$. Let $\mathcal{P}_2(D)$ denote the set of all joint pmfs $p_{X,S,Y,\hat{S}} = p_X p_S p_{Y|X,S} p_{\hat{S}|Y}$ that satisfy the inequality

$$\mathbb{E}[d(S, \hat{S})] = \sum\nolimits_{x,s,y,\hat{s}} p_{X,S,Y,\hat{S}}(x, s, y, \hat{s}) d(s, \hat{s}) \leq D. \tag{13}$$

Given state pmf p_S, channel pmf $p_{Y|X,S}$, distortion metric $d(S, \hat{S})$, and valid distortion constraint D; if we choose p_X and $p_{\hat{S}|Y}$ such that joint pmf $p_{X,S,Y,\hat{S}} \in \mathcal{P}_2(D)$; then for n sufficiently large and ε sufficiently small, provided $R_e > I(Y; \hat{S})$ and $R_t < I(X; Y)$,

- There exists a codebook $\tilde{\mathcal{C}}_n^e$ for which the expected distortion D_n is upper bounded by D.
- Using the Markov inequality-based argument in III-A3, it is highly likely that the probability of error for a codebook pair $(\mathcal{C}_n^t, \tilde{\mathcal{C}}_n^e)$ is arbitrarily close to 0.

We state the achievability result for the EBD scheme below.

Theorem 3. *Given a state pmf p_S, a channel conditional pmf $p_{Y|X,S}$, a non-negative bounded distortion metric $d(S, \hat{S})$, and a valid distortion constraint D, the highest transmission rate achievable using the EBD scheme is*

$$R_t^*(D) = \max\nolimits_{p_{X,S,Y,\hat{S}} \in \mathcal{P}_2(D)} I(X; Y).$$

***Remark* 4.** The minimum estimation rate necessary for transmission rate R_t to be achievable, using the EBD scheme, is

$$R_e^*(D) = \min\nolimits_{p_{X,S,Y,\hat{S}} \in \mathcal{P}_2(D)} I(Y; \hat{S}).$$

***Remark* 5.** Separate decoding and estimation at the receiver (Fig. 4) yields the same rates $R_t^*(D)$ and $R_e^*(D)$ as the EBD scheme (Thm. 3 and Remark 4).

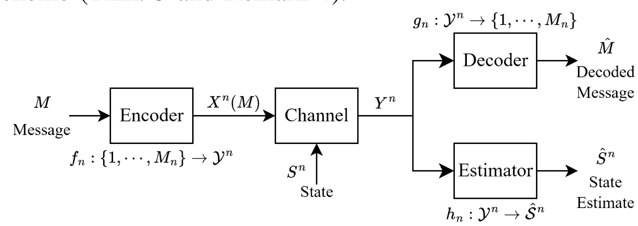

Fig. 4. Separate Decoding and Estimation Scheme

IV. Discussion and Conclusion

Given a distortion constraint D that is valid for all three schemes, we observe that the highest transmission rate achievable for the *EBD* scheme is lower than the one for the other two schemes. This is because, the set of joint pmfs that satisfy the distortion constraint, $\mathcal{P}_2(D)$ (13), when compared to the corresponding set for the other two schemes, $\mathcal{P}_1(D)$ (6), is more restrictive.

Since conditioning reduces entropy [8, Thm. 2.6.5], $I(Y; \hat{S}) = H(\hat{S}) - H(\hat{S}|Y) \leq H(\hat{S}) - H(\hat{S}|X, Y) = I(X, Y; \hat{S})$, however, this does not necessarily mean that the minimum estimation rate required for the *EBD* scheme is lower than that for the other two schemes. Again, the reason is $\mathcal{P}_2(D)$ being more restrictive than $\mathcal{P}_1(D)$.

Regardless of the scheme, the highest transmission rate achievable, $R_t^*(D)$, is a non-decreasing function of the distortion constraint D, while the minimum estimation rate required, $R_e^*(D)$, is a non-increasing function of D.

Both the optimal rates $R_t^*(D)$ and $R_e^*(D)$ can be attained simultaneously only when the joint pmf that attains $R_t^*(D)$ is the joint pmf that attains $R_e^*(D)$. Due to this strict requirement, when choosing a joint pmf for either of the three schemes, we may have to compromise on one of the two rates. This can be interpreted as a performance trade-off for an ISAC system described by our model.

V. Future Directions

We want to explore the relationship between *coded sensing* and the one-shot estimator used in [3] to see whether our approach can achieve the same performance as in [3]. It would also be interesting to study how coded sensing performs for general state-dependent channels, where either the state, the channel, or both, have memory.

Acknowledgment

This work has been supported in part by NSF through award 21-32700.

REFERENCES

[1] "M.2160 : Framework and overall objectives of the future development of IMT for 2030 and beyond." [Online]. Available: https://www.itu.int/rec/R-REC-M.2160/en

[2] S. P. Kotagiri, "State-dependent networks with side information and partial state recovery," Ph.D. Thesis, University of Notre Dame, 2007.

[3] W. Zhang, S. Vedantam, and U. Mitra, "Joint transmission and state estimation: A constrained channel coding approach," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7084–7095, 2011.

[4] O. M. Mujumdar, "Fundamental tradeoffs in integrated sensing and communication with coded sensing," Ph.D. Thesis, University of Notre Dame, 2025, in preparation.

[5] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[6] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, 1971.

[7] H. Witsenhausen, "Indirect rate distortion problems," *IEEE Transactions on Information Theory*, vol. 26, no. 5, pp. 518–521, 1980.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. USA: Wiley-Interscience, 2006.