Strong Converses for Memoryless Bi-Static ISAC

Mehrasa Ahmadipour*, Michèle Wigger†, Shlomo Shamai‡
*UMPA, ENS de Lyon, Email: mehrasa.ahmadipour@ens-lyon.fr
†LTCI Telecom Paris, IP Paris, 91120 Palaiseau, France, Email: michele.wigger@telecom-paris.fr
‡Technion, Haifa, Israel, Email: sshlomo@ee.technion.ac.il

2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10207009

***Abstract*—The paper characterizes the fundamental limits of integrated sensing and communication (ISAC) systems with a bi-static radar, where the radar receiver is located close to the transmitter and estimates or detects the state based on the transmitter's channel inputs and the backscattered signals. Two models are considered. In the first model, the memoryless state sequence is distributed according to a fixed distribution and the goal of the radar receiver is to reconstruct this state-sequence with smallest possible distortion. In the second model, the memoryless state is distributed either according to P_S or to Q_S and the radar's goal is to detect this underlying distribution so that the missed-detection error probability has maximum exponential decay-rate (maximum Stein exponent). Similarly to previous results, our fundamental limits show that the tradeoff between sensing and communication solely stems from the empirical statistics of the transmitted codewords which influences both performances. The main technical contribution are two strong converse proofs that hold for all probabilities of communication error ϵ and excess-distortion probability or false-alarm probability δ summing to less than 1, $\epsilon+\delta < 1$. These proofs are based on two parallel change-of-measure arguments on the sets of typical sequences, one change-of-measure to obtain the desired bound on the communication rate, and the second to bound the sensing performance.**

***Index Terms*—Integrated sensing and communication, strong converse, change of measure arguments.**

I. Introduction

Consider the communication problem in Figure 1, where a transmitter (Tx) sends a message M to a receiver (Rx) over a state-dependent discrete memoryless channel. Moreover, based on a generalized feedback signal, it either attempts to reconstruct the channel state sequence S^n or guesses the hypothesis underlying the distribution of S^n. Such a system has recently been termed integrated sensing and communication (ISAC) and plays a prominent role in the future 6G standard [1], [2], especially in the context of autonomous vehicles or automated manufacturing sites. In fact, in autonomous driving applications or industrial robot applications, backscattered signals from communication can be used for radar sensing to detect hazardous events, infer properties of other terminals (e.g., velocities or directions of other cars), or sense the environment for obstacles. This integration requires significant technological effort but has great potential for improving the performance of wireless systems.

While ISAC has inspired a plethora of works in the signal processing and communications community, see for example [3]–[8] and references therein, only few works were reported from the information-theoretic community [9]–[16]. The first information-theoretic work [9] on ISAC determined the fundamental limits of the rate-distortion version of the ISAC problem in Figure 1, where the radar receiver is colocated with the Tx and uses the feedback signal to reconstruct the state-sequence S^n. Extensions to network scenarios and to

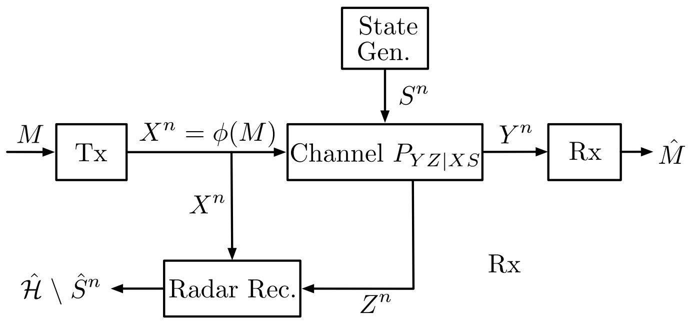

Fig. 1: Bistatic Radar ISAC Model

scenarios with secrecy constraints were subsequently presented in [10]–[13]. The works in [14]–[16] determined the fundamental performance limits of a detection-version of the model in Figure 1, where the state sequence S^n is assumed constant over time and taking on one of multiple possible values depending on a underlying hypothesis. The task of the radar receiver is to guess this hypothesis, and sensing performance was measured in terms of exponential decay-rate of the probability of error, either the minimum exponential decay-rate over all hypotheses [14], [15] or the set of decay-rates that are simultaneously achievable under the different hypotheses [16]. The work [15] also studied a mono-static version of this problem, where the Tx coincides with the radar receiver and thus can use the generalized-feedback signals also for communication purposes. For this mono-static radar scenario however only a coding scheme but no converse was presented. The problem is known to be hard as it relates to the challenging close-loop controlled sensing problem [17].

In this paper, we consider both the rate-distortion and the hypothesis testing versions of the model in Figure 1. In our first model, the state sequence $\{S_t\}_{t\geq 1}$ is independent and identically distributed (i.i.d.) according to a given distribution P_S and the radar wishes to reconstruct this state sequence with smallest possible distortion. In our second model, the state-sequence $\{S_t\}$ depends on a binary hypothesis $\mathcal{H} \in \{0,1\}$. If $\mathcal{H}=0$, then $\{S_t\}$ is i.i.d. according to a distribution P_S or if $\mathcal{H} = 1$, it is i.i.d. according to a distribution Q_S. We measure sensing performance in terms of Stein's exponent, i.e., in terms of the maximum exponential decay-rate of the missed-detection error probability (detecting P_S instead of Q_S) under a permissible threshold on the false-alarm probability (detecting Q_S instead of P_S).

For both our models, we determine the fundamental limits of communication rates and distortion/missed-detection error exponents that are simultaneously achievable. Similarly to previous works [9], [14], [15] our limits exhibit a tradeoff between the sensing and communication performances, which however solely stems from the empirical statistics of the codewords used for communication.

The direct parts of our proofs follow immediately from existing works. Our contributions are the proofs of the converse results. In fact, we present strong converse proofs that hold whenever the maximum allowed probability of communication error ϵ and the maximum allowed distortion-excess probability or false-alarm probability δ satisfy $\epsilon + \delta < 1$. The converse proofs are extensions of the channel coding strong converse proof in [18] to incorporate also the sensing bounds. Interestingly, the same change-of-measure as in [18] can be used to obtain the desired bound on the rate of communication. Different changes-of-measure are used to obtain the desired bounds on the sensing performances.

Strong converse proofs based on change-of-measure arguments go back to Gu and Effros [19], [20] and can be also found in various other works, e.g., [21]. The proof method was formalized and first applied to channel coding by Tyagi and Watanabe [22]. Recent works [18], [23], [24] slightly modified and simplified the technique in [22] by restricting the new measures to sequences on typical or conditionally-typical sets. This feature allows to circumvent resorting to variational characterizations for the multi-letter and single-letter problems as proposed in [22]. Notice that the works [23], [24] also showed the utility of the proposed converse proof method for scenarios with expectation constraints, in which case the fundamental limits depend on the permissible error probabilities.

Notation: Upper-case letters are used for random quantities and lower-case letters for deterministic realizations. Calligraphic font is used for sets. All random variables are assumed finite and discrete. We abbreviate the n-tuples $(X_1, \ldots, X_n)$ and $(x_1, \ldots, x_n)$ as X^n and x^n and the $n-t$ tuples $(X_{t+1}, \ldots, X_n)$ and $(x_{t+1}, \ldots, x_n)$ as X_{t+1}^n and x_{t+1}^n. We further abbreviate *independent and identically distributed* as *i.i.d.* and *probability mass function* as *pmf.*

Entropy, conditional entropy, and mutual information functionals are written as $H(\cdot)$, $H(\cdot|\cdot)$, and $I(\cdot;\cdot)$, where the arguments of these functionals are random variables and whenever their probability mass function (pmf) is not clear from the context, we add it as a subscript to these functionals. The Kullback-Leibler divergence between two pmfs is denoted by $D(\cdot\|\cdot)$. We shall use $\mathcal{T}_\mu^{(n)}(P_{XY})$ to indicate the jointly strongly-typical set with respect to the pmf P_{XY} on the product alphabet $\mathcal{X} \times \mathcal{Y}$ and parameter μ as defined in [25, Definition 2.8]. Specifically, let $n_{x^n,y^n}(a,b)$ denote the number of occurrences of the pair (a,b) in (x^n, y^n):

$$n_{x^n,y^n}(a,b) = |\{t\colon (x_t, y_t) = (a,b)\}|, \tag{1}$$

a pair (x^n, y^n) lies in $\mathcal{T}_\mu^{(n)}(P_{XY})$ if

$$\left|\frac{n_{x^n,y^n}(a,b)}{n} - P_{XY}(a,b)\right| \leq \mu, \qquad \forall (a,b) \in \mathcal{X} \times \mathcal{Y}, \tag{2}$$

and $n_{x^n,y^n}(a,b) = 0$ whenever $P_{XY}(a,b) = 0$. The conditionally strongly-typical set with respect to a conditional pmf $P_{Y|X}$ from $\mathcal{X}$ to $\mathcal{Y}$, parameter $\mu > 0$, and sequence $x^n \in \mathcal{X}^n$ is denoted $\mathcal{T}_\mu^{(n)}(P_{Y|X}, x^n)$ [25, Definition 2.9]. It contains all sequences $y^n \in \mathcal{Y}^n$ satisfying

$$\left|\frac{n_{x^n,y^n}(a,b)}{n} - \frac{n_{x^n}(a)}{n} P_{Y|X}(b|a)\right| \leq \mu, \qquad \forall (a,b) \in \mathcal{X} \times \mathcal{Y}, \tag{3}$$

and $n_{x^n,y^n}(a,b) = 0$ whenever $P_{Y|X}(b|a) = 0$. Here $n_{x^n}(a)$ denotes the number of occurrences of symbol a in x^n. In this paper, we denote the joint type of (x^n, y^n) by $\pi_{x^n y^n}$, i.e.,

$$\pi_{x^n y^n}(a,b) \triangleq \frac{n_{x^n,y^n}(a,b)}{n}. \tag{4}$$

Accordingly, the marginal type of x^n is written as π_{x^n}.

II. Memoryless State and Average Distortion as a Sensing Measure

Consider the bistatic radar receiver model over a memoryless channel in Fig. 1. A transmitter (Tx) that wishes to communicate a random message M to a receiver (Rx) over a state-dependent channel. The message M is uniformly distributed over the set $\{1, \ldots, 2^{nR}\}$ with $R > 0$ and $n > 0$ denoting the rate and blocklength of communication, respectively. The channel from the Tx to the Rx depends on a state-sequence $S^n = (S_1, \ldots, S_n)$ which is i.i.d. according to a given pmf P_S.

For a given blocklength n, the Tx thus produces the n-length sequence of channel inputs

$$X^n = \phi^{(n)}(M) \tag{5}$$

for some choice of the encoding function $\phi^{(n)}\colon \{1, \ldots, 2^{nR}\} \to \mathcal{X}^n$.

Based on X^n and S^n the channel produces the sequences Y^n observed at the Rx and the backscattered signal Z^n. The channel is assumed memoryless and described by the stationary transition law $P_{YZ|XS}$ implying that the pair (Y_t, Z_t) is produced according to the channel law $P_{YZ|XS}$ based on the time-t symbols (X_t, S_t).

The Rx attempts to guess message M based on the sequence of channel outputs Y^n:

$$\hat{M} = g^{(n)}(Y^n) \tag{6}$$

using a decoding function $g^{(n)}\colon \mathcal{Y}^n \to \{1, \ldots, 2^{nR}\}$.

Performance of communication is measured in terms of average error probability

$$p^{(n)}(\text{error}) := \Pr[\hat{M} \neq M] \tag{7}$$

The radar receiver produces as a guess a reconstruction of the state sequence

$$\hat{S}^n = h^{(n)}(X^n, Z^n), \tag{8}$$

based on the inputs and backscattered signals. Radar sensing performance is measured as time-averaged distortion

$$\text{dist}^{(n)}\Big(\hat{S}^n, S^n\Big) \triangleq \frac{1}{n}\sum_{i=1}^{n} d\Big(\hat{S}_i, S_i\Big), \tag{9}$$

for a given and bounded distortion function $d(\cdot,\cdot)$.

In this context we have the following definition and result.

Definition 1: A rate-distortion pair (R, D) is (ϵ, δ)-achievable over the state-dependent channel $(\mathcal{X}, \mathcal{Y}, P_{Y|XS})$ with state-distribution P_S, if there exists a sequence of encoding, decoding, and estimation functions $\{(\phi^{(n)}, g^{(n)}, h^{(n)})\}$ such that the average probability of error satisfies

$$\varlimsup_{n\to\infty} p^{(n)}(\text{error}) \leq \epsilon \tag{10}$$

and the excess distortion probability

$$\varlimsup_{n\to\infty} \Pr\Big[\text{dist}^{(n)}\Big(\hat{S}^n, S^n\Big) > D\Big] \leq \delta. \tag{11}$$

Theorem 1: For any $\epsilon+\delta<1$, a rate-distortion pair (R,D) is (ϵ,δ)-achievable, if and only if, there exists a pmf P_X satisfying

$$R = I_{P_X P_S P_{Y|XS}}(X;Y) \tag{12}$$

and

$$D \geq \mathbb{E}_{P_X P_S P_{Z|XS}}[d(\hat{s}(X,Z),\ S)] \tag{13}$$

where

$$\hat{s}(x,z) := \min_{\hat{s}\in\hat{\mathcal{S}}} \sum_s P_{S|XZ}(s|x,z) d(\hat{s},s) \tag{14}$$

and

$$P_{S|XZ}(s|x,z) := \frac{P_S(s)P_{Z|XS}(z|x,s)}{\sum_{s'} P_S(s')P_{Z|XS}(z|x,s')} \tag{15}$$

Proof: The limiting case $\epsilon,\delta\downarrow 0$ of the theorem was already proved in [11]. Achievability of the theorem follows thus directly from this previous result. The converse is proved in the following subsection, also using the next lemma, which is from [11]. ■

Lemma 1 (From [11]): Without loss in optimality, one can restrict to the per-symbol estimator

$$h^{(n)}(x^n,z^n) = (\hat{s}(x_1,z_1),\ldots,\hat{s}(x_n,z_n)). \tag{16}$$

A. Strong Converse Proof

Fix a sequence of encoding and decoding functions $\{(\phi^{(n)},g^{(n)})\}_{n=1}^{\infty}$ and consider the optimal estimator $h^{(n)}$ in Lemma 1. Assume that (10) and (11) are satisfied. For readability, we will also write $x^n(\cdot)$ for the function $\phi^{(n)}(\cdot)$. Choose a sequence of small positive numbers $\{\mu_n\}_{n=1}^{\infty}$ satisfying

$$\lim_{n\to\infty} \mu_n = 0 \tag{17}$$

$$\lim_{n\to\infty} \left(n\cdot\mu_n^2\right)^{-1} = 0. \tag{18}$$

Expurgation: Fix $\eta\in(0,1-\epsilon-\delta]$ and let $\tilde{\mathcal{M}}$ be the set of messages m that satisfy the following two conditions:

$$\Pr\left[\hat{M}\neq M|M=m\right] \leq 1-\eta \tag{19a}$$

$$\Pr\left[\text{dist}^{(n)}\left(\hat{S}^n,S^n\right) > D|M=m\right] \leq 1-\eta. \tag{19b}$$

Since the set of messages not satisfying (19a) is at most of size

$$\frac{\epsilon}{(1-\eta)}2^{nR}, \tag{20}$$

and similarly also the set of messages not satisfying (19b) is of size at most $\frac{\delta}{(1-\eta)}2^{nR}$, we can deduce that the set $\tilde{\mathcal{M}}$ (which is the complement of the union of these two sets) is of size at least

$$\left(1-\frac{\epsilon+\delta}{1-\eta}\right)2^{nR} = \frac{(1-\eta-\epsilon-\delta)}{(1-\eta)}2^{nR}. \tag{21}$$

Define the random variable $\tilde{M}$ to be uniform over the set $\tilde{\mathcal{M}}$ and let

$$\tilde{X}^n = x^n(\tilde{M}), \tag{22}$$

and thus

$$\frac{|\tilde{\mathcal{M}}|}{2^{nR}} \geq \left(1-\frac{\epsilon+\delta}{1-\eta}\right) =: \gamma. \tag{23}$$

Let T be a uniform random variable over $\{1,\ldots,n\}$, independent of all other random variables and notice that

$$P_{\tilde{X}_T}(x) = \frac{1}{n}\sum_{t=1}^{n} P_{\tilde{X}_t(x)} \tag{24}$$

$$= \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}[\mathbb{1}\{\tilde{X}_t(\tilde{M})=x\}] \tag{25}$$

$$= \mathbb{E}[\pi_{x^n(\tilde{M})}(x)]. \tag{26}$$

Let now $\{n_i\}$ be an increasing subsequence of blocklengths so that the probability vector $P_{\tilde{X}_T}$ converges and denote the convergence point by P_X:

$$\lim_{n_i\to\infty} \frac{1}{|\tilde{\mathcal{M}}|}\sum_{m\in\tilde{M}} \pi_{x^{n_i}(m)}(x) =: P_X(x), \qquad \forall x\in\mathcal{X}. \tag{27}$$

In the remainder of this proof, we restrict attention to this subsequence of blocklengths $\{n_i\}$.

Proof of Channel Coding Bound: We first prove the converse bound for channel coding. Consider the conditions

$$g^{(n)}(y^n) = m \tag{28a}$$

$$\left|\pi_{s^n,x^n(m),y^n}(a,b,c) - P_S(a)\pi_{x^n(m)}(b)P_{Y|XS}(c|a,b)\right| \leq \mu_n, \tag{28b}$$

and define for each message $m\in\tilde{\mathcal{M}}$ the set

$$\mathcal{D}_{\mathcal{C},m} := \{(s^n,y^n)\colon \quad \text{(28a) and (28b)}\}. \tag{29}$$

Introduce the new random variables $(S^n_{\mathcal{C}},Y^n_{\mathcal{C}})$ of joint conditional pmf

$$P_{S^n_{\mathcal{C}}Y^n_{\mathcal{C}}|\tilde{M}}(s^n,y^n|m) = \frac{P_S^{\otimes n}(s^n)\cdot P_{Y|XS}^{\otimes n}(y^n|x^n(m),s^n)}{\Delta_{\mathcal{C},m}}\cdot\mathbb{1}\{(s^n,y^n)\in\mathcal{D}_{\mathcal{C},m}\}, \tag{30}$$

for

$$\Delta_{\mathcal{C},m} := \sum_{s^n,y^n} P_S^{\otimes n}(s^n)\cdot P_{Y|XS}^{\otimes n}(y^n|x^n(m),s^n) \cdot\mathbb{1}\{(s^n,y^n)\in\mathcal{D}_{\mathcal{C},m}\}. \tag{31}$$

By using Chebyshev's inequality, see [25, Remark to Lemma 2.12] and Conditions (19a) and (28), we have:

$$\Delta_{\mathcal{C},m} \geq \eta - \frac{|\mathcal{S}||\mathcal{X}||\mathcal{Y}|}{4\mu_n^2 n}, \quad \forall m\in\tilde{\mathcal{M}}. \tag{32}$$

Moreover, for $\tilde{M}=m$:

$$P_{Y^n_{\mathcal{C}}|\tilde{M}=m}(y^n) = \sum_{s^n}\frac{P_S^{\otimes n}(s^n)\cdot P_{Y|XS}^{\otimes n}(y^n|x^n(m),s^n)}{\Delta_{\mathcal{C},m}}\cdot\mathbb{1}\{(s^n,y^n)\in\mathcal{D}_{\mathcal{C},m}\} \tag{33}$$

$$\leq \sum_{s^n}\frac{P_S^{\otimes n}(s^n)\cdot P_{Y|XS}^{\otimes n}(y^n|x^n(m),s^n)}{\Delta_{\mathcal{C},m}} \tag{34}$$

$$= \frac{P_{Y|X}^{\otimes n}(y^n|x^n(m))}{\Delta_{\mathcal{C},m}}. \tag{35}$$

Continue to notice that:

$$R = \frac{1}{n}H(\tilde{M}) - \frac{1}{n}\log\gamma \tag{36}$$

$$\overset{(a)}{=} \frac{1}{n} I(\tilde{M}; Y_{\mathcal{C}}^n) - \frac{1}{n}\log\gamma \tag{37}$$
$$= \frac{1}{n} H(Y_{\mathcal{C}}^n) - \frac{1}{n} H\Big(Y_{\mathcal{C}}^n \Big| \tilde{M}\Big) - \frac{1}{n}\log\gamma \tag{38}$$
$$\leq \frac{1}{n}\sum_{i=1}^{n} H(Y_{\mathcal{C},i}) - \frac{1}{n} H\Big(Y_{\mathcal{C}}^n \Big| \tilde{M}\Big) - \frac{1}{n}\log\gamma \tag{39}$$
$$= H\Big(Y_{\mathcal{C},T}\Big|T\Big) - \frac{1}{n} H\Big(Y_{\mathcal{C}}^n \Big| \tilde{M}\Big) - \frac{1}{n}\log\gamma \tag{40}$$
$$\leq H(Y_{\mathcal{C},T}) - \frac{1}{n} H\Big(Y_{\mathcal{C}}^n \Big| \tilde{M}\Big) - \frac{1}{n}\log\gamma, \tag{41}$$

where we defined the random variable T to be uniform over $\{1, \ldots, n\}$ independent of the other random variables. Here, (a) holds because $\tilde{M} = g(Y_{\mathcal{C}}^n)$ by Condition (28a).

Notice next that

$$P_{\tilde{X}_T S_{\mathcal{C},T} Y_{\mathcal{C},T}(x,s,y)} \tag{42}$$
$$= \frac{1}{n}\sum_{t=1}^{n} P_{\tilde{X}_t S_{\mathcal{C},t} Y_{\mathcal{C},t}}(x,s,y) \tag{43}$$
$$= \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\Big[\mathbb{1}\Big\{\tilde{X}_t, S_{\mathcal{C},t}, Y_{\mathcal{C},t}) = (x,s,y)\Big\}\Big] \tag{44}$$
$$= \mathbb{E}\Big[\pi_{x^n(\tilde{M}) S_{\mathcal{C}}^n Y_{\mathcal{C}}^n}(x,s,y)\Big] \tag{45}$$

However, by Condition (28b) for any triple (x, s, y) with positive $P_S(s)P_{Y|XS}(y|x, s)$ the following inequality is satisfied with probability 1:

$$\Big|\pi_{x^n(m) S_{\mathcal{C}}^n Y_{\mathcal{C}}^n}(x,s,y) - \pi_{x^n(m)}(x) P_{Y|XS}(y|x,s) P_S(s)\Big| \leq \mu_n. \tag{46}$$

By (27) and (46) and since $\mu_{n_i} \to 0$ as $n_i \to \infty$:

$$\lim_{i\to\infty} P_{\tilde{X}_T S_{\mathcal{C},T} Y_{\mathcal{C},T}}(x,s,y) = P_X(x) P_S(s) P_{Y|XS}(y|x,s), \tag{47}$$

which by continuity of the entropy functional implies

$$\lim_{n_i\to\infty} H(Y_{\mathcal{C},T}) = H_{P_X P_S P_{Y|XS}}(Y). \tag{48}$$

Next, by definition and by (35):

$$\frac{1}{n_i} H(Y_{\mathcal{C}}^{n_i} | \tilde{M} = m)$$
$$= -\frac{1}{n_i} \sum_{y^{n_i}\in\mathcal{D}_{\mathcal{C},m}} P_{Y_{\mathcal{C}}^{n_i}|\tilde{M}=m}(y^{n_i}) \log P_{Y_{\mathcal{C}}^{n_i}|\tilde{M}=m}(y^{n_i}) \tag{49}$$
$$\geq -\frac{1}{n_i} \sum_{y^{n_i}\in\mathcal{D}_{\mathcal{C},m}} P_{Y_{\mathcal{C}}^{n_i}|\tilde{M}=m}(y^{n_i}) \log \frac{P_{Y|X}^{\otimes n}(y^{n_i}|x^{n_i}(m))}{\Delta_{\mathcal{C},m}} \tag{50}$$
$$= -\frac{1}{n_i}\sum_{t=1}^{n_i} \sum_{y^{n_i}\in\mathcal{D}_{\mathcal{C},m}} P_{Y_{\mathcal{C}}^{n_i}|\tilde{M}=m}(y^{n_i}) \log P_{Y|X}(y_t|x_t(m)) + \frac{1}{n_i}\log\Delta_{\mathcal{C},m}, \tag{51}$$
$$= -\frac{1}{n_i}\sum_{t=1}^{n_i}\sum_{y_t\in\mathcal{Y}} P_{Y_{\mathcal{C},t}|\tilde{M}=m}(y_t) \log P_{Y|X}(y_t|x_t(m)) + \frac{1}{n_i}\log\Delta_{\mathcal{C},m}, \tag{52}$$
$$= -\frac{1}{n_i}\sum_{t=1}^{n_i}\sum_{y\in\mathcal{Y}} \mathbb{E}\Big[\mathbb{1}\{Y_{\mathcal{C},t} = y\}\Big|\tilde{M} = m\Big] \log P_{Y|X}(y|x_t(m)) + \frac{1}{n_i}\log\Delta_{\mathcal{C},m}, \tag{53}$$
$$= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathbb{E}\left[\frac{1}{n_i}\sum_{t=1}^{n_i}\mathbb{1}\{x_t(m) = x, Y_{\mathcal{C},t} = y\}\Big|\tilde{M} = m\right] \cdot \log P_{Y|X}(y|x) + \frac{1}{n_i}\log\Delta_{\mathcal{C},m}, \tag{54}$$
$$= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\sum_{s\in\mathcal{S}} \mathbb{E}\Big[\pi_{x^{n_i}(m) S_{\mathcal{C}}^{n_i} Y_{\mathcal{C}}^{n_i}}(x,s,y)\Big|\tilde{M} = m\Big] \cdot \log P_{Y|X}(y|x) + \frac{1}{n_i}\log\Delta_{\mathcal{C},m}, \tag{55}$$

where $P_{Y|X}(y|x) = \sum_{s\in\mathcal{S}} P_{Y|XS}(y|x, s)P_S(s)$. Averaging over all messages $m \in \tilde{\mathcal{M}}$, we obtain:

$$\frac{1}{n_i} H(Y_{\mathcal{C}}^{n_i}|\tilde{M}) \tag{56}$$
$$\geq -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\sum_{s\in\mathcal{S}} \mathbb{E}\Big[\pi_{x^{n_i}(\tilde{M}) S_{\mathcal{C}}^{n_i} Y_{\mathcal{C}}^{n_i}}(x,s,y)\Big] \cdot \log P_{Y|X}(y|x) + \mathbb{E}\left[\frac{1}{n_i}\log\Delta_{\mathcal{C},\tilde{M}}\right]. \tag{57}$$

By (32) the term $E\Big[\frac{1}{n_i}\log\Delta_{\mathcal{C},\tilde{M}}\Big]$ vanishes for increasing blocklengths, and thus using the definition of P_X in (27), fone can follow the same bounding steps as leading to (46) to obtain:

$$\lim_{i\to\infty}\frac{1}{n_i} H(\tilde{Y}^{n_i}|\tilde{M})$$
$$= -\sum_{x\in\mathcal{X}} P_X(x) \sum_{y\in\mathcal{Y}}\sum_{s\in\mathcal{S}} P_S(s) P_{Y|XS}(y|x,s) \log P_{Y|X}(y|x)$$
$$= H_{P_X P_S P_{Y|XS}}(Y|X). \tag{58}$$

Combining (41) with (48) and (58), and since $\frac{1}{n_i}\log\gamma \to 0$ as $n \to \infty$, we can conclude that

$$R \leq H_{P_X P_S P_{Y|XS}}(Y) - H_{P_X P_S P_{Y|XS}}(Y|X) \tag{59}$$
$$= I_{P_X P_S P_{Y|XS}}(X;Y). \tag{60}$$

Proof of Distortion Bound: Consider the two conditions

$$\mathrm{dist}^{(n)}\Big(h^{(n)}(x^n(m), z^n), s^n\Big) \leq D \tag{61a}$$
$$\Big|\pi_{s^n,x^n(m),z^n}(a,b,c) - P_S(a)\pi_{x^n(m)}(b) P_{Z|XS}(c|a,b)\Big| \leq \mu_n, \tag{61b}$$

and define for each message $m \in \tilde{\mathcal{M}}$ the set

$$\mathcal{D}_{\mathcal{S},m} := \{(s^n, z^n)\colon \quad \text{(61a) and (61b)}\}. \tag{62}$$

Recall the definition $\tilde{X}^n = x^n(\tilde{M})$ and the limit in (27). Define the new random variables $(S_{\mathcal{S}}^n, Z_{\mathcal{S}}^n)$ of joint conditional pmf

$$P_{S_{\mathcal{S}}^n Z_{\mathcal{S}}^n|\tilde{M}}(s^n, z^n|m) = \frac{P_S^{\otimes n}(s^n)\cdot P_{Z|XS}^{\otimes n}(z^n|x^n(m), s^n)}{\Delta_{\mathcal{S},m}} \cdot \mathbb{1}\{(s^n, z^n) \in \mathcal{D}_{\mathcal{S},m}\}, \tag{63}$$

for

$$\Delta_{\mathcal{S},m} := \sum_{s^n,z^n} P_S^{\otimes n}(s^n)\cdot P_{Z|XS}^{\otimes n}(z^n|x^n(m), s^n)$$

$$\cdot \mathbb{1}\{(s^n, z^n) \in \mathcal{D}_{\mathcal{S},m}\}. \tag{64}$$

Notice that by using Chebyshev's inequality, see [25, Remark to Lemma 2.12] and Conditions (19b) and (61), we have:

$$\Delta_{\mathcal{S},m} \geq \eta - \frac{|\mathcal{S}||\mathcal{X}||\mathcal{Z}|}{4\mu_n^2 n}, \qquad \forall m \in \tilde{\mathcal{M}}. \tag{65}$$

Following similar steps to (44)–(47), by (61b) and definition (27), we can conclude that

$$\lim_{n_i \to \infty} P_{\tilde{X}_T S_{\mathcal{S},T} Z_{\mathcal{S},T}}(x,s,z) = P_X(x) P_S(s) P_{Z|XS}(z|x,s). \tag{66}$$

By Condition (61a), we have with probability 1:

$$D \geq \frac{1}{n} \sum_{t=1}^{n} d\Big(\hat{s}\Big(\tilde{X}_t, Z_{\mathcal{S},t}\Big)\ ,\ S_{\mathcal{S},t}\Big). \tag{67}$$

Therefore, for any blocklength n_i:

$$D \geq \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbb{E}\Big[d\Big(\hat{s}\Big(\tilde{X}_j, Z_{\mathcal{S},j}\Big)\ ,\ S_{\mathcal{S},j}\Big)\Big] \tag{68}$$

$$= \mathbb{E}\Big[d\Big(\hat{s}\Big(\tilde{X}_T, Z_{\mathcal{S},T}\Big)\ ,\ S_{\mathcal{S},T}\Big)\Big], \tag{69}$$

and by (66) in the limit as $n_i \to \infty$:

$$D \geq \mathbb{E}_{P_X P_S P_{Z|XS}}[d(\hat{s}(X,S)\ ,\ S)]. \tag{70}$$

This concludes the proof of the converse.

III. Stein's Exponent as a Sensing Measure

In this section we assume that the state-sequence S^n depends on a binary hypothesis $\mathcal{H} \in \{0,1\}$. Under the null hypothesis $\mathcal{H} = 0$ it is i.i.d. according to the pmf P_S and under the alternative hypothesis $\mathcal{H} = 1$ it is i.i.d. according to the pmf Q_S. The radar receiver attempts to guess the underlying hypothesis based on the inputs and backscattered signals, so it produces a guess of the form

$$\hat{\mathcal{H}} = h^{(n)}(X^n, Z^n) \in \{0,1\}. \tag{71}$$

Radar sensing performance is measured in terms of Stein's exponent. That means, it is required that the type-I error probability

$$\alpha_n := \Pr\Big[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0\Big] \tag{72}$$

stays below a given threshold, while the type-II error probability

$$\beta_n := \Pr\Big[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\Big] \tag{73}$$

should decay exponentially fast to 0 with largest possible exponent.

Definition 2: A rate-exponent pair (R,E) is (ϵ,δ)-achievable over the state-dependent DMC $(\mathcal{X},\mathcal{Y},P_{Y|XS})$ with state-distribution P_S, if there exists a sequence of encoding, decoding, and estimation functions $\{(\phi^{(n)}, g^{(n)}, h^{(n)})\}$ such that for each blocklength n the average probability of error satisfies

$$\overline{\lim_{n\to\infty}}\, p^{(n)}(\text{error}) \leq \epsilon, \qquad \mathcal{H} \in \{0,1\}, \tag{74}$$

while the detection error probabilities satisfy:

$$\overline{\lim_{n\to\infty}}\, \alpha_n \leq \delta, \tag{75}$$

and

$$-\underline{\lim_{n\to\infty}}\, \frac{1}{n} \log \beta_n \geq E. \tag{76}$$

Theorem 2: For any $\epsilon, \delta \geq 0$ satisfying $\epsilon + \delta < 1$, a rate-exponent pair (R,E) is (ϵ,δ)-achievable, if and only if, there exists a pmf P_X satisfying

$$R \leq \min\{I_{P_X P_S P_{Y|XS}}(X;Y), I_{P_X Q_S P_{Y|XS}}(X;Y)\}, \tag{77}$$

and

$$E \leq \mathbb{E}_{P_X}\big[D(P_{Z|X} \| Q_{Z|X})\big] \tag{78}$$

where $P_{Z|X}$ and $Q_{Z|X}$ denote the conditional marginals of $P_S P_{Z|XS}$ and $Q_S P_{Z|XS}$, respectively.

Proof: Achievability follows by standard random coding for a compound channel and by applying a Neyman-Pearson test at the radar receiver. The converse is proved in the long version of this paper [26]. ■

The works in [15], [16] consider degenerate state-distributions where P_S and Q_S are deterministic distributions. In this case, our Theorem simplifies as follows.[1]

Corollary 3: Assume degenerate state-distributions $P_S(s_0) = 1$ and $Q_S(s_1) = 1$ for two distinct symbols $s_0, s_1 \in \mathcal{S}$. Then, for any $\epsilon, \delta \geq 0$ satisfying $\epsilon + \delta < 1$, a rate-exponent pair (R,E) is (ϵ,δ)-achievable, if and only if, there exists a pmf P_X satisfying

$$R \leq \min\Big\{I_{P_X P_{Y|X}^{(s_0)}}(X;Y),\ I_{P_X P_{Y|X}^{(s_1)}}(X;Y)\Big\}, \tag{79}$$

and

$$E \leq \mathbb{E}_{P_X}\big[D(P_{Z|XS}(\cdot|X,s_0) \| P_{Z|XS}(\cdot|X,s_1)\big], \tag{80}$$

where $P_{Y|X}^{(s)}(y|x) \triangleq P_{Y|XS}(y|x,s)$ for any triple (x,s,y).

IV. Conclusion and Future Directions

In this paper we established the strong converse for two ISAC problems with bi-static radar whenever $\epsilon + \delta < 1$. Interesting future research directions include extensions to mono-static radar systems where the transmitter can apply closed-loop encodings depending also on past generalized feedback systems or systems with memory. Analyzing other sensing criteria is also of interest, such as the minimum exponential decay-rate over all hypotheses or the estimation error when the distribution of the state-sequence depends on a single continuous-valued parameter. The setup where only part of the state or a noisy version of the state is to be estimated is also of interest, for instance, in scenarios in state-dependent fading channels where one has no interest in estimating the fading. Notice that this setup is included in the rate-distortion model through an appropriate definition of the distortion measure. On a related note, our model also includes as special case the setups where the receiver has perfect or imperfect channel-state information by including this state-information as part of the output.

Acknowledgement

The work of S. Shamai has been supported by the European Union's Horizon 2020 Research And Innovation Programme, grant agreement no. 694630.

[1]Recall that [14]–[16] required exponential decrease both for the type-I and type-II error probabilities α_n and β_n. Their result is thus not a special case of ours.

References

[1] O. Li, J. He, K. Zeng, Z. Yu, X. Du, Z. Zhou, Y. Liang, G. Wang, Y. Chen, P. Zhu, W. Tong, D. Lister, and L. Ibbetson, "Integrated sensing and communication in 6G: a prototype of high resolution multichannel THz sensing on portable device," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 1, p. 106, 2022.

[2] Z. Wei, H. Qu, Y. Wang, X. Yuan, H. Wu, Y. Du, K. Han, N. Zhang, and Z. Feng, "Integrated sensing and communication signals towards 5g-a and 6g: A survey," *IEEE Internet of Things Journal*, pp. 1–1, 2023.

[3] S. H. Dokhanchi, M. B. Shankar, M. Alaee-Kerahroodi, T. Stifter, and B. Ottersten, "Adaptive waveform design for automotive joint radar-communications system," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2019, pp. 4280–4284.

[4] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[5] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1236–1259, 2011.

[6] M. P. Roberton and E. Brown, "Integrated radar and communications based on chirped spread-spectrum techniques," *IEEE MTT-S International Microwave Symposium Digest, 2003*, vol. 1, pp. 611–614, 2003.

[7] G. N. Saddik, R. S. Singh, and E. R. Brown, "Ultra-wideband multifunctional communications/radar system," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 7, pp. 1431–1437, 2007.

[8] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty, "A survey on fundamental limits of integrated sensing and communication," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 994–1034, 2022.

[9] M. Kobayashi, G. Caire, and G. Kramer, "Joint state sensing and communication: Optimal tradeoff for a memoryless case," 2018, pp. 111–115.

[10] M. Kobayashi, H. Hamad, G. Kramer, and G. Caire, "Joint state sensing and communication over memoryless multiple access channels," 2019, pp. 270–274.

[11] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Transactions on Information Theory*, pp. 1–1, 2022.

[12] M. Ahmadipour and M. Wigger, "An information-theoretic approach to collaborative integrated sensing and communication for two-transmitter systems," 2023. [Online]. Available: https://arxiv.org/abs/2304.01972

[13] O. Günlü, R. F. Schaefer, H. Boche, and H. V. Poor, "Secure and private source coding with private key and decoder side information," 2022. [Online]. Available: https://arxiv.org/abs/2205.05068

[14] H. Joudeh and F. M. J. Willems, "Joint communication and binary state detection," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 1, pp. 113–124, 2022.

[15] M.-C. Chang, S.-Y. Wang, T. Erdogan, and M. R. Bloch, "Rate and detection-error exponent tradeoff for joint communication and sensing of fixed channel states," 2022. [Online]. Available: https://arxiv.org/abs/2210.07963

[16] H. Wu and H. Joudeh, "Joint communication and channel discrimination," 2022. [Online]. Available: https://arxiv.org/abs/2208.07450

[17] S. Nitinawarat, G. Atia, and V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Transactions on Automatic Control*, vol. 58, 05 2012.

[18] M. Hamad, M. Wigger, and M. Sarkiss, "Strong converses using typical changes of measures and asymptotic Markov chains," 2023. [Online]. Available: https://arxiv.org/abs/2301.06289

[19] W. Gu and M. Effros, "A strong converse for a collection of network source coding problems," in *2009 IEEE International Symposium on Information Theory*. IEEE, 2009, pp. 2316–2320.

[20] ——, "A strong converse in source coding for super-source networks," in *2011 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2011, pp. 395–399.

[21] T. Han and K. Kobayashi, "Exponential-type error probabilities for multiterminal hypothesis testing," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 2–14, 1989.

[22] H. Tyagi and S. Watanabe, "Strong converse using change of measure arguments," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 689–703, 2019.

[23] M. Hamad, M. Wigger, and M. Sarkiss, "Multi-hop network with multiple decision centers under expected-rate constraints," 2022. [Online]. Available: https://arxiv.org/abs/2208.14243

[24] S. Faour, M. Hamad, M. Sarkiss, and M. Wigger, "Testing against independence with an eavesdropper," 2022. [Online]. Available: https://arxiv.org/abs/2211.03475

[25] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

[26] M. Ahmadipour, M. Wigger, and S. Shamai, "Strong converses for memoryless bi-static isac," 2023. [Online]. Available: https://arxiv.org/abs/2303.06636