Extended Typicality with Feedback and an Application to ISAC

Thomas Sturma, Michèle Wigger
LTCI, *Télécom Paris*, Palaiseau, France
Email: {thomas.sturma, michele.wigger}@telecom-paris.fr

***Abstract*—In this paper we prove that for feedback coding over a channel with memory level f $P_{Z|X_1 \dots X_f}$, with high probability and for large blocklengths, the conditional type of the output given the f preceding inputs remains close to the channel law $P_{Z|X_1 \dots X_f}$. We use this result to establish a novel strong converse result for a memoryless integrated sensing and communication (ISAC) model were the transmitter communicates a message to a receiver over a state-dependent channel and based on the observed generalized feedback signals estimates the channel states. Previous results only managed to obtain the desired results for output feedback or without feedback coding, which in the context of ISAC is a highly limiting assumption.**

***Index Terms*—strong converse, feedback, typicality, memory, change of measure, integrated sensing and communication**

I. Introduction

We study the closed-loop system in Figure 1 where an encoder inputs to a channel with memory $P_{Z|X_1 \cdots X_f}$ for which each output symbol is randomly generated according to the channel's conditional distribution in function of the last f channel inputs, for a given positive integer f. We show that, irrespective of the choice of the encoding that generates the inputs based on past feedback symbols, with probability and for large blocklengths the conditional extended type of the outputs given the last f inputs is close to the $P_{Z|X_1 \cdots X_f}$. This result extends the classical results without feedback, where the desired statement holds simply by the weak law of large numbers, and generalizes our recent result in [1] which only considered memoryless channels, i.e., $f = 1$ in our setup.

To prove our desired result, we show that the difference between the conditional extended type and the channel law $P_{Z|X_1 \cdots X_f}$ is always a martingale. The desired result then follows by applying an inequality for (sub-)martingales, similarly to Kolmogorov's result for random walks used in [1].

In the second part of the paper, we use the established typicality result to prove a converse result for an Integrated Sensing and Communication (ISAC) system, of Figure 2, where the transmitter sends a message over a state-dependent discrete memoryless channel (SDMC) and from the observed generalized feedback signals it reconstructs the state sequence up to a desired distortion. This setup was introduced in [9], which also characterized information-theoretic performance limits of the system, i.e, the set of all simultaneously achievable rate and distortion pairs (see also [10]). These works, showed so called "weak" converse results restricting attention to the regime of vanishing error probabilities. So called "strong" converses for ISAC setups allow that the message decoding error probabilities lie below a certain threshold ϵ and that the distortion constraint be violated with a probability at most δ. The works in [1], [3] proved such "strong" converses and showed that the results in [9] remain valid also under these more relaxed constraints whenever $\epsilon + \delta < 1$. These previous "strong" converses however only hold under the assumption that the transmitter cannot use the feedback for channel coding (but only for the sensing task) [11] or that the feedback observed at the transmitter is the same as the signal observed at the receiver. Both assumptions are highly limiting in ISAC scenarios, and in this paper we prove a "strong" converse result for arbitrary generalized feedback that can be used for the encoding.

Our converse proof is based on change-of-measure arguments [2]–[5], [7], [11] and establishes that the set of achievable rate-distortion pairs in our ISAC system does not depend on the permissible decoding error probability ϵ and the excess distortion probability δ, as long as their sum is below 1, i.e., $\epsilon + \delta < 1$. (It can be shown that for larger values of ϵ and δ the set of achievable rate-distortion pairs increases.) In particular, we show that for $\epsilon + \delta < 1$ the set of achievable rate-distortion pairs coincides with the one presented in [9] and can thus be attained without feedback coding.

Notation. Sets are denoted using calligraphic fronts such as $\mathcal{X}$ and $\mathcal{Y}$. All random variables are assumed finite. For $i \leq j$ two positive integers, sequences of random variables $(X_i, \dots, X_j)$ and realizations $(x_i, \dots, x_j)$ are abbreviated by X_i^j and x_i^j respectively. If $i = 1$, we also use X^j and x^j instead of X_i^j and x_i^j. The probability of an event A and the expectation of a random variable X are denoted by $\mathbf{P}(A)$ and $\mathbf{E}[X]$ respectively. Moreover, throughout this paper, 2^{nR} will denote the integer $\lfloor 2^{nR} \rfloor$, and $[\![a, b]\!]$, for integer numbers $a < b$, denotes the set $\{a, \dots, b\}$. Finally, we will use the Landau notation $o(1)$ to indicate any function that tends to 0 when $n \to +\infty$.

II. Typicality with Feedback for Certain Channels with Memory

Fix a memory level $f \geq 1$ and a blocklength $n \geq 1$ and consider the setup in Figure 1. Because the output Z_i depends on the f previous inputs $X_i, \dots, X_{i-f+1}$, we ignore Z_1^{f-1} here. We fix the encoding functions $(h_k)_{k \geq f}$ that choose the next input, dependent of all previous outputs, i.e.

$$X_k = h_k(Z_f, \dots, Z_{k-1}), \ \forall k \geq 1. \tag{1}$$

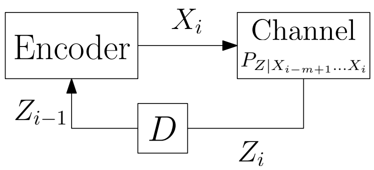

Fig. 1. Closed-loop system with feedback coding.

Given inputs X_{k-f+1}^k, output Z_k follows the conditional distribution $P_{Z|X_1\dots X_f}(\cdot|X_{k-f+1}^k)$.

Define the "extended" joint $(f,1)$-type as:

$$\pi_{X^n,Z^n}^{(f,1)}(a_1^f,b) \triangleq \frac{1}{n}\sum_{k=f}^{n} \mathbf{1}\{X_{k-f+1}^k = a_1^f,\, Z_k = b\} \tag{2}$$

and its marginal as:

$$\pi_{X^n}^{(f)}(a_1^f) \triangleq \frac{1}{n}\sum_{k=f}^{n} \mathbf{1}\{X_{k-f+1}^k = a_1^f\}. \tag{3}$$

The following is the main result of this section which is an extension from our previous result [1], where the result only applied for DMC.

Theorem 1: Irrespective of the choice of the encoding functions $(h_k)_{k\geq 1}$ and for any $\mu > 0$, it holds that:

$$\mathbf{P}\Big(\Big|\pi_{X^n,Z^n}^{(f,1)}(a_1^f,b) \tag{4}$$

$$- \pi_{X^n}^{(f)}(a_1^f)P_{Z|X_1\dots X_f}(b\,|\,a_1^f)\Big| > \mu\Big) \leq \frac{1}{4n\mu^2} \tag{5}$$

Proof: Fix $a_1,\dots,a_f,b$ and we define for $k=1,\dots,n$:

$$\Xi_k \triangleq \mathbf{1}\{X_{k-f+1}^k = a_1^f\}\left(\mathbf{1}\{Z_k=b\} - P_{Z|X_1\dots X_f}(b|a_1^f)\right). \tag{6}$$

Using the definition of the extended types, we can rewrite:

$$|\pi_{x^n,z^n}^{(f,1)}(a_1^f,b) - \pi_{x^n}^{(f)}(a_1^f)P_{Z|X_1\dots X_f}(b\,|\,a_1^f)| = \left|\frac{1}{n}\sum_{k=f}^{n}\Xi_k\right| \tag{7}$$

Define further for $l \geq f$:

$$S_l := \sum_{k=f}^{l} \Xi_k, \tag{8}$$

and notice that the sequence $(S_l)_{l\geq f}$ is a martingale:

$$\begin{aligned}
&\mathbf{E}[S_{l+1}|S_l] - S_l \\
&= \mathbf{E}[S_l + \Xi_{l+1}|S_l] - S_l && (9)\\
&= \mathbf{E}[\Xi_{l+1}|S_l] && (10)\\
&= \mathbf{E}[\Xi_{l+1}|S_l, X_{l-f+2}^{l+1} = a_1^f]\mathbf{P}(X_{l-f+2}^{l+1} = a_1^f|S_l) && (11)\\
&\quad + \mathbf{E}[\Xi_{l+1}|S_l, X_{l-f+2}^{l+1} \neq a_1^f]\mathbf{P}(X_{l-f+2}^{l+1} \neq a_1^f|S_l) && (12)\\
&= 0. && (13)
\end{aligned}$$

The last equality holds because the first expectation is zero: when $X_{l-f+2}^{l+1} = a_1^f$, we have $\mathbb{E}\big[\mathbf{1}\{Z_{l+1} = b\}\big] = P_{Z|X_1,\dots,X_f}\big(b \mid a_1^f\big)$. The second expectation is zero because $\Xi_{l+1} = 0$ when $X_{l-f+2}^{l+1} \neq a_1^f$.

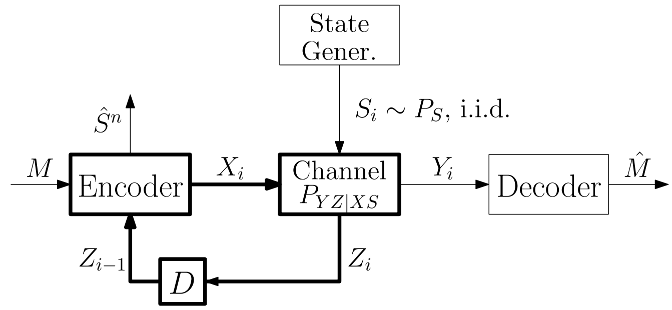

Fig. 2. ISAC Model

So, applying Doob's inequality for L^2 [13] to our martingale (S_l), we obtain:

$$\mathbf{P}\left(\sup_{f\leq l\leq n}|S_l| \geq n\mu\right) \leq \frac{\mathbf{E}[S_n^2]}{n^2\mu^2}. \tag{14}$$

We continue to bound $\mathbf{E}[S_n^2]$. To this end, notice that

$$\mathbf{E}[S_n^2] = \sum_{k=f}^{n}\mathbf{E}[\Xi_k^2] + 2\sum_{f\leq k<l\leq n}\mathbf{E}[\Xi_k\Xi_l] \tag{15}$$

We again notice that

$$\mathbf{E}[\Xi_k\Xi_l|X_{l-f+1}^l \neq a_1^f] = 0 \tag{16}$$

because $\Xi_l = 0$. Therefore:

$$\begin{aligned}
\mathbf{E}[\Xi_k\Xi_l] &= \mathbf{E}[\Xi_k\Xi_l|X_{l-f+1}^l = a_1^f]\mathbf{P}(X_{l-f+1}^l = a_1^f) && (17)\\
&\overset{(a)}{=} \big(\mathbf{E}[\Xi_k|X_{l-f+1}^l = a_1^f] \\
&\qquad \cdot \mathbf{E}[\Xi_l|X_{l-f+1}^l = a_1^f]\big)\mathbf{P}(X_{l-f+1}^l = a_1^f) && (18)\\
&\overset{(b)}{=} 0, && (19)
\end{aligned}$$

where (a) holds because $l > k$ and given $X_{l-f+1}^l = a_1^f$ the law of Ξ_l is a function of a_1^f irrespective of the realization of Ξ_k, and (b) holds again because conditioned on the event $X_{l-f+1}^l = a_1^f$ the random variable Ξ_l has zero mean.

Using that $\mathbf{E}[\Xi_k^2] \leq 1/4$, we then obtain:

$$\mathbf{P}\left(|S_n| \geq n\mu_n\right) \leq \mathbf{P}\left(\sup_{f\leq l\leq n}|S_l| \geq n\mu_n\right) \leq \frac{1}{4n\mu^2}. \tag{20}$$

■

III. Converse Theorem for an ISAC Model

We use our result on typicality in the special case $f=1$ to prove a converse for the integrated sensing and communication (ISAC) model in Figure 2.

A transmitter seeks to communicate a random message M to a receiver over a state-dependent discrete memoryless channel (SDMC) $P_{YZ|XS}$. The message M is assumed to be uniformly distributed over the set $[\![1, 2^{nR}]\!]$, where $R > 0$ denotes the communication rate and $n > 0$ the blocklength.

The SDMC is influenced by a state sequence S^n, which is independent and identically distributed (i.i.d.) according to a given probability mass function (pmf) P_S. At each time i, the channel outputs (Y_i, Z_i) are generated based on the current input X_i and the state S_i, according to the given SDMC transition probability $P_{YZ|XS}$. Here, output Y_i is observed at the receiver, while Z_i denotes the signal that is backscattered.

The transmitter creates its channel inputs based on the message and the past generalized feedback signals $Z_1, \ldots, Z_{i-1}$:

$$X_i = f_i^{(n)}(M, Z_1, \ldots, Z_{i-1}), \qquad i = 1, \ldots, n, \tag{21}$$

for some sequence of encoding functions $\{f_i^{(n)}\}_{i=1}^n$. It also uses the feedback outputs Z^n and the produced input sequence X^n to reconstruct the channel's state sequence S:

$$\hat{S}^n = h^{(n)}(X^n, Z^n). \tag{22}$$

The receiver attempts to guess the message M based on the channel outputs Y^n using a decoding function $g^{(n)}$:

$$\hat{M} = g^{(n)}(Y^n). \tag{23}$$

Reliability of communication is evaluated via the average probability of decoding error at the receiver

$$P_e^{(n)} = \mathbf{P}\left(\hat{M} \neq M\right). \tag{24}$$

Sensing performance at the transmitter is measured by the average expected distortion between the true state sequence S^n and its reconstruction $\hat{S}^n$:

$$\text{dist}^n(\hat{S}^n, S^n) = \frac{1}{n}\sum_{i=1}^n d(\hat{S}_i, S_i), \tag{25}$$

where d denotes a positive and bounded distortion function.

Our goal is to have small decoding error probability and small excess distortion probability.

Definition 1: A rate-distortion pair (R, D) is (ϵ, δ)-achievable if there exist sequences of encoding, decoding and estimation functions $\{\{f_k^{(n)}\}_{k=1}^n, g^{(n)}, h^{(n)}\}_{n=1}^{+\infty}$ satisfying

$$\overline{\lim_{n\to+\infty}}\, P_e^{(n)} \leq \epsilon \tag{26}$$

$$\overline{\lim_{n\to\infty}}\, P_D^{(n)} \leq \delta. \tag{27}$$

where $P_D^{(n)} = \mathbf{P}\left(\text{dist}^n\left(\hat{S}^n, S^n\right) > D\right)$.

Theorem 2: For any $\varepsilon + \delta < 1$, if a rate-distortion pair (R, D) is (ϵ, δ)-achievable then there exists P_X satisfying

$$R = I_{P_X P_S P_{YZ|XS}}(X; Y) \tag{28}$$

$$D \geq \mathbf{E}_{P_X P_S P_{YZ|XS}}\left[d(\hat{s}(X, Z), S)\right] \tag{29}$$

where

$$\hat{s}(x, z) = \min_{\hat{s}\in\mathcal{S}} \sum_s P_{S|XZ}(s|x, z) d(\hat{s}, s). \tag{30}$$

Achievability for $\epsilon = \delta = 0$ follows directly from the results in [9], [10], where it was also shown that the transmitter does not need to rely on the feedback outputs $Z_1, \ldots, Z_{i-1}$ to produce the next input X_i. The contribution of our article lies in proving the converse to this statement.

Remark 1: Notice that better performances can be achieved when $\epsilon + \delta \geq 1$.

A. Proof of Theorem 2

Fix $\epsilon, \delta > 0$ with sum $\epsilon + \delta < 1$. Fix also a sequence of encoding and decoding functions $\{\{f_i^{(n)}\}_{i=1}^n, g^{(n)}\}_{n=1}^\infty$ satisfying (26) and (27) and consider the optimal state estimator (see [10, Lemma 1]) defined by $h^n(x^n, z^n) = (\hat{s}(x_1, z_1), \ldots, \hat{s}(x_n, z_n))$, where $\hat{s}(x, z)$ is defined in (30).

Define for any blocklength n the parameter $\mu_n = n^{-1/4}$. Consider now the three following conditions:

$$g^{(n)}(y^n) = m \tag{31}$$

$$\text{dist}^n(h^{(n)}(x^n, z^n), s^n) \leq D \tag{32}$$

$$\begin{aligned}\big|\pi_{x^n,s^n,y^n,z^n}&(a, b, c, d)\\ &-\pi_{x^n}(a)P_S(b)P_{YZ|XS}(c, d|a, b)\big| \leq \mu_n,\end{aligned} \tag{33}$$

where in the above, $x_i = f_i(m, z^{i-1})$ for $i = 1, \ldots, n$.

Fix a constant η such that $0 < \eta < 1 - \varepsilon - \delta$, and define the subset $\tilde{\mathcal{M}}_n \subseteq [\![1, 2^{nR}]\!]$ to consist of all messages m satisfying

$$\mathbf{P}\Big(\hat{M} \neq M \text{ or } \text{dist}^n\big(\hat{S}^n, S^n\big) > D \mid M = m\Big) \leq 1 - \eta, \tag{34}$$

and let $\tilde{M}$ be uniformly distributed over $\tilde{\mathcal{M}}_n$. Notice that we have the following inequality:

$$\frac{|\tilde{\mathcal{M}}_n|}{2^{nR}} \geq \left(1 - \frac{P_e^{(n)} + P_D^{(n)}}{1 - \eta}\right) =: \gamma_n, \tag{35}$$

This inequality holds because by the union bound, we have:

$$\begin{aligned}&\mathbf{E}_M\left[\mathbf{P}\left[\hat{M} \neq M \text{ or } \text{dist}^n\left(\hat{S}^n, S^n\right) > D \mid M\right]\right]\\ &\leq P_e^{(n)} + P_D^{(n)}.\end{aligned} \tag{36}$$

Considering the conditional probability in above equation as a nonnegative random variable, by Markov's inequality we have:

$$\begin{aligned}&\mathbf{P}_M\left[\mathbf{P}\left[\hat{M} \neq M \text{ or } \text{dist}^n\left(\hat{S}^n, S^n\right) > D \,\middle|\, M\right] > 1 - \eta\right]\\ &\leq \frac{P_e^{(n)} + P_D^{(n)}}{1 - \eta}\end{aligned} \tag{37}$$

The desired inequality follows from the fact that the left-hand side of the above inequality can be written as $1 - \frac{|\tilde{\mathcal{M}}_n|}{2^{nR}}$, since all messages are equiprobable with probability 2^{-nR}.

Define now for any $m \in \tilde{\mathcal{M}}_n$ the set $\mathcal{D}_{n,m}$ consisting of all tuples (s^n, y^n, z^n) such that (31)–(33) hold, and introduce random variables $(\tilde{S}^n, \tilde{Y}^n, \tilde{Z}^n)$ jointly distributed with $\tilde{M}$ through the conditional pmfs:

$$\begin{aligned}&P_{\tilde{S}^n,\tilde{Y}^n,\tilde{Z}^n|\tilde{M}=m}(s^n, y^n, z^n)\\ &\quad\triangleq \frac{P_S^{\otimes n}(s^n)\prod_{i=1}^n P_{YZ|XS}(y_i z_i|x_i(m, z^{i-1}), s_i)}{\Delta_{n,m}}\\ &\quad\quad\cdot\mathbf{1}\{(s^n, y^n, z^n) \in \mathcal{D}_{n,m}\},\end{aligned} \tag{38}$$

where $\Delta_{n,m} \triangleq \mathbf{P}\left((S^n, Y^n, Z^n) \in \mathcal{D}_{n,m} \mid M = m\right)$.

Notice that:

$$\begin{aligned}
&\Delta_{n,m}\\
&= 1-\mathbf{P}\left((S^n,Y^n,Z^n)\notin\mathcal{D}_{n,m}\,|M=m\right) && (39)\\
&= 1-\mathbf{P}\big(M\neq\hat{M}\text{ or }\mathrm{dist}^n(\hat{S}^n,S^n)>D\text{ or }\exists(a,b,c,d):\\
&\qquad |\pi_{X^n,S^n,Y^n,Z^n}(a,b,c,d) && (40)\\
&\qquad -\pi_{X^n}(a)P_S(b)P_{YZ|XS}(c,d|a,b)|>\mu_n\mid M=m\big)\\
&\geq \eta-\sum_{a,b,c,d}\mathbf{P}\left(|\pi_{X^n,S^n,Y^n,Z^n}(a,b,c,d)\right.\\
&\qquad \left.-\pi_{X^n}(a)P_{YZ|XS}(c,d|a,b)P_S(b)|>\mu_n\,|\,M=m\right) && (41)
\end{aligned}$$

By Theorem 1, we obtain for any m and (a,b,c,d):

$$\begin{aligned}
\mathbf{P}\Big(&|\pi_{X^nS^nY^nZ^n}(a,b,c,d)\\
&-\pi_{X^n}(a)\,P_{YZ|XS}(c,d|a,b)P_S(b)\Big|>\mu_n\\
&\Big|M=m\Big)\leq\frac{1}{4n\mu_n^2}, \qquad (42)
\end{aligned}$$

and thus, by (41):

$$\Delta_{n,m}\geq\eta-\frac{|\mathcal{X}||\mathcal{Y}||\mathcal{S}||\mathcal{Z}|}{4n\mu_n^2},\quad\forall m\in\tilde{\mathcal{M}}_n. \qquad (43)$$

We now prove the desired bounds on the rate and the distortion. Starting with the rate, by definition of $\tilde{M}$ and (35), we have $P_{\tilde{M}}(m)\leq 1/\gamma_n 2^{nR}$, thus:

$$\begin{aligned}
H(\tilde{M})&\geq\sum_{m\in\tilde{\mathcal{M}}_n}P_{\tilde{M}}(m)\log\left(\gamma_n 2^{nR}\right) && (44)\\
&=\log(\gamma_n)+nR && (45)
\end{aligned}$$

We then obtain the upper bound for R:

$$\begin{aligned}
R&\leq\frac{1}{n}H(\tilde{M})-\frac{1}{n}\log(\gamma_n) && (46)\\
&\overset{(a)}{=}\frac{1}{n}I(\tilde{Y}^n;\tilde{M})+o(1) && (47)\\
&=\frac{1}{n}H(\tilde{Y}^n)-\frac{1}{n}H(\tilde{Y}^n\,|\,\tilde{M})+o(1) && (48)
\end{aligned}$$

where (a) follows from the fact that $\tilde{M}$ is a deterministic function of $\tilde{Y}^n$, as condition (31) ensures error-free decoding, i.e., $\tilde{M}=g^{(n)}(\tilde{Y}^n)$. Moreover, $\frac{1}{n}\log(\gamma_n)$ vanishes as the blocklength grows, due to assumptions (26) and (27), and because $\epsilon+\delta<1-\eta$.

Before studying the first term, we defined $T=T_n$ to be independent of $(\tilde{M},\tilde{X}^n,\tilde{Y}^n,\tilde{Z}^n,\tilde{S}^n)$ and uniform over $[\![1,n]\!]$.

$$\begin{aligned}
\frac{1}{n}H(\tilde{Y}^n)&=\frac{1}{n}\sum_{i=1}^n H(\tilde{Y}_i|\tilde{Y}^{i-1}) && (49)\\
&\leq\frac{1}{n}\sum_{i=1}^n H(\tilde{Y}_i) && (50)\\
&=\sum_{i=1}^n P_{T_n}(i)H(\tilde{Y}_T\,|\,T=i) && (51)\\
&=H(\tilde{Y}_T\,|\,T) && (52)\\
&\leq H(\tilde{Y}_T) && (53)
\end{aligned}$$

We now require the following lemma to complete the bound on the rate:

Lemma 1: There exists a subsequence $\{n_i\}_{i\geq1}$ and a pmf P_X on $\mathcal{X}$ such that:

$$\begin{aligned}
&\lim_{i\to+\infty}P_{\tilde{X}_T,\tilde{Y}_T,\tilde{Z}_T\tilde{S}_T}(x,y,z,s)\\
&\qquad\qquad=P_X(x)P_S(s)P_{YZ|XS}(y,z|x,s) && (54)\\
&\underline{\lim}_{i\to+\infty}\frac{1}{n_i}H(\tilde{Y}^{n_i}|\tilde{M})\geq H(Y|X) && (55)
\end{aligned}$$

where the random variables (X,Y) in the entropy term are distributed according to the joint marginal of pmf $P_XP_SP_{Y|XS}$.

Proof: To prove the first limit (54) we fix a quadruple $(x,y,z,s)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\times\mathcal{S}$ and $n\geq1$. Then,

$$\begin{aligned}
&P_{\tilde{X}_{T_n}\tilde{Y}_{T_n}\tilde{Z}_{T_n}\tilde{S}_{T_n}}(x,y,z,s)\\
&=\frac{1}{n}\sum_{i=1}^n P_{\tilde{X}_i\tilde{Y}_i\tilde{Z}_i\tilde{S}_i}(x,y,z,s) && (56)\\
&=\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^n\mathbf{1}\{\tilde{X}_i=x,\tilde{Y}_i=y,\tilde{Z}_i=z,\tilde{S}_i=s\}\right] && (57)\\
&=\mathbf{E}[\pi_{\tilde{X}^n\tilde{Y}^n\tilde{Z}^n\tilde{S}^n}(x,y,z,s)] && (58)\\
&\leq\mathbf{E}[\pi_{\tilde{X}^n}(x)]P_{YZ|XS}(y,z|x,s)P_S(s)+\mu_n, && (59)
\end{aligned}$$

where the inequality holds by Condition (32). In a similar way,

$$\begin{aligned}
&P_{\tilde{X}_{T_n}\tilde{Y}_{T_n}\tilde{Z}_{T_n}\tilde{S}_{T_n}}(x,y,z,s)\\
&\geq\mathbf{E}[\pi_{\tilde{X}^n}(x)]P_{YZ|XS}(y,z|x,s)P_S(s)-\mu_n. && (60)
\end{aligned}$$

Notice that $\{\mathbf{E}[\pi_{\tilde{X}^n}(x)]\}_{n\geq1}$ is a bounded sequence, so by the Bolzano-Weierstrass theorem, there exists a subsequence $\{n_i\}$ and an accumulation point P_X such that:

$$\lim_{i\to+\infty}\mathbf{E}[\pi_{\tilde{X}^{n_i}}(x)]=P_X(x),\qquad\forall x\in\mathcal{X}. \qquad (61)$$

With (59)–(60) and since μ_n tends to 0 as n approaches infinity, it proves the first limit (54).

In the following we restrict to blocklengths $n\in\{n_i\}_i$. To prove the second limit (55), we start by noting:

$$\frac{1}{n}H(\tilde{Y}^n|\tilde{M})=\frac{1}{n}H(\tilde{Y}^n,\tilde{Z}^n|\tilde{M})-\frac{1}{n}H(\tilde{Z}^n|\tilde{Y}^n,\tilde{M}). \qquad (62)$$

and

$$\begin{aligned}
\frac{1}{n}H(\tilde{Z}^n|\tilde{Y}^n,\tilde{M})&=\frac{1}{n}\sum_{i=1}^n H(\tilde{Z}_i|\tilde{Z}^{i-1},\tilde{Y}^n,\tilde{M}) && (63)\\
&=H(\tilde{Z}_{T_n}|\tilde{Z}^{T_n-1},\tilde{Y}^n,\tilde{M},T_n) && (64)\\
&\leq H(\tilde{Z}_{T_n}|\tilde{X}_{T_n},\tilde{Y}_{T_n}) && (65)\\
&=H(Z|X,Y)+o(1), && (66)
\end{aligned}$$

where the inequality holds because $\tilde{X}_{T_n}$ is determine by $\tilde{M}$ and $\tilde{Z}^{T_n-1}$, and the last equation holds by (54).

For the second term of (62), fix now a message $m\in\tilde{\mathcal{M}}_n$ and define:

$$\mathcal{D}'_{n,m}=\{(y^n,z^n)\in\mathcal{Y}^n\,|\,\exists\,s^n\colon(s^n,y^n,z^n)\in\mathcal{D}_{n,m}\}, \qquad (67)$$

Then:

$$
\begin{aligned}
&\frac{1}{n}H(\tilde{Y}^n, \tilde{Z}^n \mid \tilde{M} = m) \\
&= -\frac{1}{n} \sum_{(y^n,z^n)\in\mathcal{D}'_{n,m}} P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n) \\
&\qquad \cdot \log\Big(P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n)\Big) && (68)\\
&\geq -\frac{1}{n} \sum_{(y^n,z^n)\in\mathcal{D}'_{n,m}} P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n) \\
&\qquad \cdot \log\left(\frac{P_{Y^nZ^n|M=m}(y^n, z^n)}{\Delta_{n,m}}\right) && (69)\\
&= -\frac{1}{n} \sum_{(y^n,z^n)\in\mathcal{D}'_{n,m}} P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n) \\
&\qquad \cdot \sum_{i=1}^{n} \log\Big(P_{Y_iZ_i|M,Y^{i-1},Z^{i-1}}(y_i, z_i \mid m, y^{i-1}, z^{i-1})\Big) \\
&\qquad + \frac{1}{n}\log(\Delta_{n,m}) && (70)\\
&\overset{(a)}{\geq} -\frac{1}{n} \sum_{(y^n,z^n)\in\mathcal{D}'_{n,m}} P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n) \\
&\qquad \cdot \sum_{i=1}^{n} \log\Big(P_{YZ|X}(y_i, z_i \mid x_i(m, z^{i-1}))\Big) + C_n && (71)\\
&= -\sum_{(y^n,z^n)\in\mathcal{D}'_{n,m}} P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n) \\
&\qquad \cdot \sum_{a,b,c} \pi_{x^n(m,z^n),y^n,z^n}(a, b, c) \log\Big(P_{YZ|X}(b, c \mid a)\Big) \\
&\qquad + C_n && (72)
\end{aligned}
$$

where (a) holds by (43) and:

$$
C_n \triangleq \frac{1}{n}\log\left(\eta - \frac{|\mathcal{X}||\mathcal{Y}||\mathcal{S}||\mathcal{Z}|}{4n\mu_n^2}\right) \xrightarrow[n\to+\infty]{} 0 \qquad (73)
$$

Notice next that by summing Condition (33) over $\mathcal{S}$, we obtain a similar inequality:

$$
\begin{aligned}
&\pi_{x^n(m,z^n),y^n,z^n}(a, b, c) \\
&\qquad \geq \pi_{x^n(m,z^n)}(a)P_{YZ|X}(b, c|a) - |\mathcal{S}|\mu_n. && (74)
\end{aligned}
$$

By (72) and (74), we may deduce that:

$$
\begin{aligned}
&\frac{1}{n}H(\tilde{Y}^n, \tilde{Z}^n|\tilde{M} = m) \\
&\quad \geq \sum_{y^n,z^n\in\mathcal{D}'_{n,m}} P_{\tilde{Y}^n\tilde{Z}^n|\tilde{M}=m}(y^n, z^n) \\
&\qquad\qquad \cdot \left(\sum_{a\in\mathcal{X}} \pi_{x^n(m,z^n)}(a)H(Y, Z|X = a)\right) \\
&\quad + \sum_{\substack{(a,b,c):\\ P_{YZ|X}(b,c|a)>0}} |\mathcal{S}|\mu_n \log P_{YZ|X}(b, c|a) \\
&\quad + C_n && (75)
\end{aligned}
$$

By exchanging the order of summation:

$$
\begin{aligned}
&\frac{1}{n}H(\tilde{Y}^n, \tilde{Z}^n|\tilde{M} = m) \\
&\quad = \sum_{a\in\mathcal{X}} H(Y, Z|X = a)\mathbf{E}[\pi_{\tilde{X}^n}(a)|\tilde{M} = m] \\
&\quad + \sum_{\substack{(a,b,c):\\ P_{YZ|X}(b,c|a)>0}} |\mathcal{S}|\mu_n \log P_{YZ|X}(b, c|a) \\
&\quad + C_n. && (76)
\end{aligned}
$$

We then take the expectation with respect to $P_{\tilde{M}}$ and use the convergence of $(\mathbf{E}[\pi_{\tilde{X}^{n_i}}(x)])_{n_i\geq 1}$ to P_X to conclude:

$$
\varliminf_{i\to+\infty} \frac{1}{n_i}H(\tilde{Y}^{n_i}\tilde{Z}^{n_i}|\tilde{M}) \geq \sum_{a\in\mathcal{X}} H(Y, Z|X = a)P_X(a). \qquad (77)
$$

■

Subsequently, by virtue of this lemma, and as the block-length n_i tends to $+\infty$, it must hold that:

$$
R \leq H(Y) - H(Y|X) = I(X; Y). \qquad (78)
$$

We prove the distortion bound. By (32), with probability 1:

$$
\begin{aligned}
D &\geq \frac{1}{n_i}\sum_{j=1}^{n_i} d\left(\hat{s}\left(\tilde{X}_j, \tilde{Z}_j\right), \tilde{S}_j\right) && (79)\\
&= \mathbf{E}\left[d\left(\hat{s}\left(\tilde{X}_{T_n}, \tilde{Z}_{T_n}\right), \tilde{S}_{T_n}\right)\right] && (80)
\end{aligned}
$$

and by (54), we can conclude by taking $n_i \to +\infty$ that:

$$
D \geq \mathbf{E}\left[d(\hat{s}(X, Z), S)\right]. \qquad (81)
$$

IV. Conclusion

In the first part we considered channels with memory $P_{Z|X_1\cdots X_f}$ where each output depends on the previous f inputs. We proved that for arbitrary encoding functions, even relying on feedback, the extended conditional type of the outputs given past f inputs is arbitrary close to the channel law $P_{Z|X_1\cdots X_f}$ with probability tending to 1 as the number of channel uses tends to infinity. This extends our previous results for memoryless channels in [1].

In the second part of the paper, we use above technical result to establish a strong converse result for a canonical memoryless ISAC system. Previous results only managed to prove the desired converse when the transmitter cannot use the feedback for coding or when the backscattered signal is assumed to be equal to the receiver's output signal. In this article, we consider arbitrary backscattering signals and the transmitter is allowed to use feedback coding.

V. Acknowledgment

The authors acknowledge funding from the ERC under Grant Agreement 101125691.

REFERENCES

[1] T. Sturma and M. Wigger, "Typicality with Feedback," *2025 IEEE Information Theory Workshop*, Sydney, Australia, 2025, pp. 1-6.

[2] 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.

[3] M. Hamad, M. Wigger, and M. Sarkiss, "Strong converses using typical changes of measures and asymptotic Markov chains," *IEEE Trans. Inf. Theory,* vol. 70, no. 3, pp. 1720-–1737, Mar. 2024.

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

[5] W. Gu and M. Effros, "A strong converse in source coding for super-source networks," *2011 IEEE International Symposium on Information Theory Proceedings*, St. Petersburg, Russia, 2011, pp. 395–399.

[6] H. Tyagi and S. Watanabe, "Strong Converse Using Change of Measure Arguments," in *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 689–703, Feb. 2020.

[7] D. Takeuchi and S. Watanabe, "Tight Exponential Strong Converse for Source Coding Problem With Encoded Side Information," in *IEEE Transactions on Information Theory*, vol. 71, no. 3, pp. 1533–1545, March 2025

[8] J. Wolfowitz, "The coding of messages subject to chance errors," *Illinois Journal of Mathematics*, vol. 1, no. 4, pp. 591–606, 1957.

[9] M. Kobayashi, G. Caire, and G. Kramer, "Joint state sensing and communication: Optimal tradeoff for a memoryless case," *2018 IEEE International Symposium on Information Theory Proceedings*, Vail, CO, USA, 2018, pp. 111–115.

[10] M. Ahmadipour, M. Kobayashi, M. Wigger and G. Caire, "An Information-Theoretic Approach to Joint Sensing and Communication," in *IEEE Transactions on Information Theory*, vol. 70, no. 2, pp. 1124–1146, Feb. 2024.

[11] M. Ahmadipour, M. Wigger and S. Shamai, "Strong Converses for Memoryless Bi-Static ISAC," 2023 *IEEE International Symposium on Information Theory* , Taipei, Taiwan, 2023, pp. 1818–1823.

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

[13] R. Durrett, *Probability – theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Vol. 49 (Fifth edition of 1991 original ed.). Cambridge University Press, 2019.