

THE STRONG CONVERSE TO THE IDENTIFICATION CODING THEOREM

Te Sun Han
Department of Information Systems
Senshu University
Kawasaki, Japan

Sergio Verdú
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

The identification coding theorem of Ahlswede and Dueck [1] for single-user discrete memoryless channels (DMC) without feedback states that the identification capacity of a noisy channel is equal to its Shannon capacity. The identification capacity is the maximal iterated logarithm of the number of messages divided by the blocklength which can be reliably transmitted when the receiver is only interested in deciding whether a specific message was transmitted or not. The main result of [1] was a so-called soft-converse result where the probabilities of missed and false identification are required to vanish exponentially with the blocklength. This paper shows that for any fixed pair of probabilities of missed and false identification (whose sum is less than unity), the identification capacity is upper bounded by the Shannon capacity.

Let A and B be the input and output alphabets of a noisy channel, and let $W^{(n)}(\cdot | x^n)$ be the conditional distribution on B^n , given that the input is equal to the codeword $x^n \in A^n$.

The Ahlswede-Dueck [1] identification codes are defined as

Definition An $(n, N, \lambda_1, \lambda_2)$ ID code is a collection $\{Q_a, D_a, a = 1, \dots, N\}$ such that

- 1) Q_a is a probability distribution (PD) on A^n
- 2) $D_a \subset B^n$
- 3) $Q_a W^{(n)}(D_a) \geq 1 - \lambda_1, \quad a = 1, \dots, N$
- 4) $Q_a W^{(n)}(D_b) \leq \lambda_2, \quad \text{for all } a \neq b$

where we have used the notation for unconditional output distributions:

$$QW(D) = \int W(D | x) dQ(x)$$

The collection $\{Q_a, a = 1, \dots, N\}$ will be referred as the *codebook*, and each of its elements will be referred to as a *codistribution* (as the counterparts to the *codewords* in a channel transmission code are now distributions on the set of codewords). The *rate* of an $(n, N, \lambda_1, \lambda_2)$ ID code is defined as $\frac{1}{n} \log \log N$. The counterpart to the standard definitions of achievable rates and capacity of transmission codes is

Definition R is a (λ_1, λ_2) -achievable ID rate if for every $\gamma > 0$ and for all sufficiently large n , there exist $(n, N, \lambda_1, \lambda_2)$ ID codes whose rate satisfies

$$\frac{1}{n} \log \log N > R - \gamma$$

The supremum of the (λ_1, λ_2) -achievable ID rates is called the (λ_1, λ_2) -ID capacity of the channel.

The direct (positive) part of the identification coding theorem holds for arbitrary channels:

Theorem 1. Fix $0 \leq \lambda_1 \leq 1, 0 \leq \lambda_2 \leq 1$, and denote $\lambda = \min\{\lambda_1, \lambda_2\}$. For any arbitrary channel $\{W^{(n)}: A^n \rightarrow B^n, n=1, \dots, \infty\}$, the λ -capacity (in the maximal error probability sense) is a (λ_1, λ_2) -achievable ID rate.

If either probability of error is allowed to be so large that $\lambda_1 + \lambda_2 \geq 1$, it is possible to sharpen Theorem 1 because then any $R > 0$ is an (λ_1, λ_2) -achievable ID rate. To see this simply choose $Q^{(n)}$ and $D^{(n)}$ such that $Q^{(n)} W^{(n)}(D^{(n)}) = 1 - \lambda_1$, and construct an ID code where every codistribution and decoding set are equal to $Q^{(n)}$ and $D^{(n)}$ respectively. Then, $Q_a W^{(n)}(D_b) = 1 - \lambda_1 \leq \lambda_2$, and therefore we have obtained $(n, N, \lambda_1, \lambda_2)$ ID codes for all n and N .

In the nontrivial setting where $\lambda_1 + \lambda_2 < 1$, the objective is to show the converse identification coding theorem namely, that Theorem 1 cannot be improved. Such a result has been shown by Ahlswede and Dueck [1] for discrete-memoryless channels in the soft-converse version (a weaker form than the weak converse to the Shannon theorem), in which λ_1 and λ_2 are required to decrease to zero exponentially fast. The contribution of this paper is the strong converse in Theorem 2 which, together with Theorem 1 and the direct part of the Shannon Theory for DMCs (i.e., that $\max_P I(P, W)$ is λ -achievable for $0 < \lambda < 1$), implies that the (λ_1, λ_2) -ID capacity of a DMC is equal to its Shannon capacity if $0 < \lambda_1 + \lambda_2 < 1$.

Theorem 2 Consider a discrete memoryless channel with transition probability matrix $W: A \rightarrow B$. If $\lambda_1 + \lambda_2 < 1$, then the (λ_1, λ_2) -ID capacity of the channel is upperbounded by $C = \max_P I(P, W)$.

The proof is an application of the achievability theorem in the theory of approximation of output statistics developed by the authors. The results can be generalized to channels with arbitrary memory structure.

References

1. R. Ahlswede and G. Dueck, "Identification via channels," *IEEE Trans. Information Theory*, vol. IT-35, pp. 15-29, Jan. 1989.

This work was supported by the US Office of Naval Research under Grant N00014-90-J-1734.