

Physics of the Shannon Limits

Neri Merhav

Department of Electrical Engineering
Technion – Israel Institute of Technology
Technion City, Haifa 32000, Israel
Email: merhav@ee.technion.ac.il

Abstract— We provide a simple physical interpretation, in the context of the second law of thermodynamics, to the information inequality (a.k.a. the Gibbs' inequality, which is also equivalent to the log-sum inequality), asserting that the relative entropy between two probability distributions cannot be negative. Since this inequality stands at the basis of the data processing theorem (DPT), and the DPT in turn is at the heart of most, if not all, proofs of converse theorems in Shannon theory, it is observed that conceptually, the roots of fundamental limits of Information Theory can actually be attributed to the laws of physics, in particular, the second law of thermodynamics, and indirectly, also the law of energy conservation. By the same token, in the other direction: one can view the second law as stemming from information-theoretic principles.

Index Terms— Gibbs' inequality, data processing theorem, entropy, second law of thermodynamics, divergence, relative entropy, mutual information.

I. INTRODUCTION

While the laws of physics draw the boundaries between the possible and the impossible in Nature, the coding theorems of Information Theory, or more precisely, their converse parts, draw the boundaries between the possible and the impossible in the design and performance of coded communication systems and in data processing. A natural question that may arise, in view of these two facts, is whether there is any relationship between them. It is the purpose of this work to touch upon this question and to make an attempt to provide at least a partial answer.

Perhaps the most fundamental inequality in Information Theory is the so called *information inequality* (cf. e.g., [1, Theorem 2.6.3, p. 28]), which asserts that the relative entropy (a.k.a. the Kullback–Leibler divergence) between two probability distributions over the same alphabet $P = \{P(x), x \in \mathcal{X}\}$ and $Q = \{Q(x), x \in \mathcal{X}\}$,

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

can never be negative, and a similar fact applies to probability density functions with the summation across \mathcal{X} being replaced by integration.

The *log-sum inequality* (LSI) [1, Theorem 2.7.1, p. 31], which asserts that for two sets of non-negative numbers, (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) :

$$\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq \left(\sum_{i=1}^n a_i \right) \log \left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \right),$$

is completely equivalent¹ to the information inequality, although proved in [1] in a rather different manner.

Yet another name for the same inequality, which is more frequently encountered in the jargon of physicists, is the *Gibbs' inequality*: When the information inequality is applied to two probability distributions of the Boltzmann form (cf. Section IV below), it yields an interesting inequality concerning their corresponding free energies (cf. e.g., [2, Section 5.6, pp. 143–146]), which serves as a useful tool for obtaining good bounds on the free energy of a complex system, when its exact value is difficult to calculate.

In this work, we provide a simple physical interpretation to this inequality of the free energies, and thereby also to the information inequality, or the log-sum inequality. This physical interpretation is directly related to the second law of thermodynamics, which asserts that the entropy of an isolated physical system cannot decrease: According to this interpretation, the divergence between two probability distributions is proportional to the energy dissipated in the system when it undergoes an irreversible process, and hence converts this energy loss into entropy production, or heat. Thus, the non-negativity of the relative entropy is related to the non-negativity of this entropy change, which is, as said, the second law of thermodynamics.

Since the mutual information can be thought of as an instance of the relative entropy, and so can the difference between two mutual informations defined along a Markov chain, then the data processing theorem (DPT) can, of course, also be given the very same physical interpretation. Considering the fact that the DPT is pivotal to most, if not all, converse theorems in Information Theory, this means that, in fact, the fundamental limits of Information Theory can, at least conceptually, be attributed to the laws of physics, in particular, to the second law of thermodynamics.² The rate loss in any suboptimal coded communication system, is given the meaning of irreversibility and entropy production in a corresponding physical system. Optimum (or nearly optimum) communication systems are corresponding to reversible processes (or lack of any process at all) with no entropy

¹The information inequality is obtained from the LSI when (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) both sum to unity, and conversely, the LSI is obtained from the information inequality, by applying the latter to the probability distributions $P_i = a_i / \sum_j a_j$ and $Q_i = b_i / \sum_j b_j$.

²Another law of physics that plays a role here, at least indirectly, is the law of energy conservation, because our derivations are all based on the Boltzmann–Gibbs distribution of equilibrium statistical mechanics, and this distribution, in turn, is derived on the basis of the energy conservation law.