CCIT REPORT #742 August 2009

Another Look at the Physics of Large Deviations With Application to Rate–Distortion Theory

Neri Merhav

***Abstract*—We revisit and extend the physical interpretation recently given to a certain identity between large–deviations rate–functions (as well as applications of this identity to Information Theory), as an instance of thermal equilibrium between several physical systems that are brought into contact. Our new interpretation, of mechanical equilibrium between these systems, is shown to have several advantages relative to that of thermal equilibrium. This physical point of view also provides a trigger to the development of certain alternative representations of the rate–distortion function and channel capacity, which are new to the best knowledge of the author.**

***Index Terms*—Large deviations theory, Chernoff bound, statistical physics, free energy mechanical equilibrium, rate–distortion theory.**

I. Introduction

RELATIONSHIPS between information theory and statistical physics have been widely recognized in the last few decades, from a wide spectrum of aspects. These include conceptual aspects, of parallelisms and analogies between theoretical principles in the two disciplines, as well as technical aspects, of mapping between mathematical formalisms in both fields and borrowing analysis techniques from one field to the other. One example of such a mapping, is between the paradigm of random codes for channel coding and certain models of magnetic materials, most notably, Ising models and spin glass models (cf. e.g., [10] and many references therein). Today, it is quite widely believed that research in the intersection between information theory and statistical physics may have the potential of fertilizing both disciplines.

This paper is more related to the former aspect mentioned above, namely, the relationships between the two areas in the conceptual level. In particular, we revisit results of a recent work [9], and propose a somewhat different perspective, which as we believe, has certain advantages, that will be explained and shown in the sequel.

More specifically, in [9], an identity between two forms of the rate function of a certain large deviations event was established, with several applications in information theory. Inspired by a few earlier works (cf. e.g., [8], [12], [14]), this identity was interpreted as *thermal equilibrium* between several many–particle physical systems that are brought in contact. In particular, the parameter that undergoes optimization of the Chernoff bound, henceforth referred to as the *Chernoff parameter*, plays a role that is intimately related to the equilibrium temperature: in fact, it is the reciprocal of the temperature, called the *inverse temperature*. The corresponding large deviations rate function is then identified with the entropy of the system.

While this physical interpretation is fairly reasonable, it turns out, as we show in this paper, that it leaves quite some room for improvement, and we will mention here just two points. The first, is that this interpretation does not generalize to rate functions of combinations of two or more rare events, where the number of Chernoff parameters is as the number of events. This is because there is only one temperature parameter in physics. The other point, which is on a more technical level, is the following (more details and clarifications will follow in Subsection 2B below): while the log–moment generating function, pertaining to the large deviations rate function, naturally includes weighting by probabilities, its physical analogue, which is the *partition function*, does not. If these probabilities are subjected to optimization (e.g., optimization of random coding distributions), they may depend on the Chernoff parameter, i.e., on the temperature, in a rather complicated manner, and then the resulting expression can no longer really be viewed as a partition function.

In this paper, we propose to interpret the above–mentioned identity of rate functions as an instance of *mechanical equilibrium* (i.e., balance between mechanical forces), rather than thermal equilibrium, and then the Chernoff parameter plays the physical role of an external *force*, or *field*, applied to the physical system in consideration. In this paradigm, the large deviations rate function has a natural interpretation as the (Helmholtz) *free energy* of the system, rather than as entropy. Accordingly, since the rate–distortion function (and similarly, also channel capacity) can be thought of as a large deviations rate function, it can also be interpreted as the free energy of a certain system.

This interpretation has several advantages. First, it is consistent with the analogy between the free energy in physics and the Kullback–Leibler divergence in information theory (see, e.g., [1],[11]), which is well known to play a role as a rate function when the large deviations analysis is approached by the method of types [4]. Second, it is free of the limitations mentioned in the previous paragraph, as we will see in the sequel. Third, it serves as a trigger to develop certain representations of the rate–distortion function (and analogously, the channel capacity), which are new to the best knowledge of the author.

Since the rate–distortion function can be thought of as free energy, as mentioned above, one of the representations of the rate–distortion function expresses it as (the minimum achievable) mechanical work carried out by the aforementioned external force, along a 'distance' that is measured in terms of the distortion. Another representation, which follows

N. Merhav is with the Department of Electrical Engineering, Technion – Israel Institute of Technology, Haifa, 32000, Israel. E-mail: merhav@ee.technion.ac.il.