

# Minimal Randomness and Information Theory

Sergio Verdú

Dept. Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA

*Abstract* — This is a tutorial survey of recent information theoretic results dealing with the minimal randomness necessary for the generation of random processes with prescribed distributions.

## I. INTRODUCTION

Shannon Theory explores the fundamental limits on the size of codes that enable the reliable reproduction or transmission of information. Reliability is typically quantified by the probability that the decoded message is equal to the original one, or by some measure of the distance between the original and decoded messages.

In this paper we are not interested in the reproduction or transmission of information but rather in the generation of random processes with prescribed distributions, and associated problems. For example, we may want to simulate a “real-world” random process, or the response of a system to such an input. Random process generation is accomplished by adequately mapping a source of pure random bits. A key question that quantifies the “complexity” of the random process is the *minimal randomness* of the source of pure bits necessary to accomplish the task. As in conventional Shannon theory, a rich theory arises when some distance (often arbitrarily small) is allowed between the desired and the resulting probability distributions. To this end, several distance measures have been considered in the literature, such as variational distance, divergence,  $\rho$ -distance, etc.

## II. SOURCE RESOLVABILITY

The resolvability of a source is defined [1] as the minimal number of random bits per sample it takes to reproduce the  $n$ -dimensional distributions with arbitrary accuracy as  $n$  tends to infinity. A general formula is shown in [1] for the resolvability of a source. For the special case of stationary ergodic sources and variational distance it is equal to the entropy rate. Reference [7] considers the problem of finite-precision resolvability where the approximation distance need not be arbitrarily small, and shows that for any information stable source  $D$ -resolvability is independent of  $D$  in the special case of variational distance. However, with less stringent approximation measures such as the Prohorov and  $\rho$ -distance, the  $D$ -resolvability is shown in [7] to be given by the rate distortion function evaluated with a sample-path distortion metric derived from the distribution distance measure.

## III. CHANNEL RESOLVABILITY

In system simulation, the objective is to induce the same output distributions as those that would obtain with a “real-world” input. The channel (or system) resolvability defined as the minimal randomness required to generate any desired input so that the output distributions are approximated with arbitrary accuracy. Naturally, the more “random” a system is, the lower its resolvability, as it does not pay to reproduce fine details in the input distributions. It is shown in [1] that the channel resolvability is equal to its capacity for most discrete

channels (those that satisfy the strong converse). The complementary problem where the input is given but the channel is to be simulated is studied in [5], where it is shown that the minimal randomness required to simulate the system for a specific input is equal to the conditional entropy rate of the output given the input.

## IV. INTRINSIC RANDOMNESS

A problem which is dual to source resolvability is the *maximal* randomness rate that can be extracted from an arbitrary source. The *intrinsic randomness rate* of a source is defined in [9] as the largest rate of *almost-fair* coin flips that can be extracted by a deterministic mapping of the source. For stationary ergodic sources and variational distance the intrinsic randomness rate is equal to the entropy rate [9]. However there are nonstationary sources for which the intrinsic randomness rate is not equal to the minimal noiseless source coding rate. The more general problem of finite precision intrinsic randomness is studied in [10]. Using variational distance, [10] shows that the finite precision intrinsic randomness rate is given, essentially, by the inverse asymptotic distribution of the entropy density.

## REFERENCES

- [1] T. S. Han, S. Verdú, “Approximation Theory of Output Statistics,” *IEEE Trans. on Information Theory*, vol. IT-39, pp. 752-772, May 1993.
- [2] T. S. Han, S. Verdú, “Spectrum Invariance under Output Approximation for Discrete Memoryless Channels with Full Rank,” *Problemy Peredachi Informatsii*, (in Russian), vol. 29, no. 2, p. 9-27, 1993, translated in *Problems of Information Transmission*, Apr.-June 1993, p. 101-118.
- [3] T. S. Han and S. Verdú, “The Resolvability and the Capacity of the AWGN Channel are equal,” *Proc. 1994 IEEE Int. Symp. Information Theory*, Trondheim, Norway, June 1994, p. 463.
- [4] M. Burnashev and S. Verdú, “Measures separated in  $L_1$  Metrics and ID-codes,” *Problemy Peredachi Informatsii*, (in Russian), to appear.
- [5] Y. Steinberg, S. Verdú, “Channel Simulation and Coding with Side Information,” *IEEE Trans. on Information Theory*, vol. 40, no. 3, pp. 634-646, May 1994.
- [6] Y. Steinberg and S. Verdú, “The Random Bit Rate required for Channel Simulation,” *Proc. Sixth Joint Swedish-Russian International Workshop on Information Theory*, pp. 447-451, Moelle, Sweden, Aug. 22-27, 1993.
- [7] Y. Steinberg and S. Verdú, “Coarse Approximations of Source Statistics and Rate-Distortion Theory,” *IEEE Trans. on Information Theory*, submitted.
- [8] Y. Steinberg and S. Verdú, “Finite Precision Source Resolvability,” *Proc. 1994 IEEE Int. Symp. Information Theory*, Trondheim, Norway, June 1994, p. 296.
- [9] S. Vembu and S. Verdú, “Generating Random Bits from an Arbitrary Source: Fundamental Limits,” *IEEE Trans. on Information Theory*, submitted.
- [10] Y. Steinberg and S. Verdú, “Finite Precision Intrinsic Randomness and Source Resolvability,” *Proc. IT/STAT Workshop '94*, Alexandria, VA, Oct. 1994.