Finite-precision Source Resolvability

Yossef Steinberg and Sergio Verdú

C3I Center, George Mason Univ., Fairfax, VA 22030 and Dept. of EE, Princeton Univ., Princeton, NJ 08544, USA

This paper studies the minimum randomness necessary for finite precision simulation of a random source. In random process simulation, the objective of the simulator is to approximate a set of desired statistics. To this end, the simulator has access to a source of pure random bits - a random number generator - and the approximation is achieved by properly mapping the output of the random number generator to the alphabet of the approximated process. An important question that arises is what is the number of pure random bits per source output that the most efficient simulation scheme needs in order to produce every sample path of the approximating process. The answer to this question depends on the statistics of the approximated source and on the sense of approximation. If the objective was to produce - with the aid of only pure random bits - exactly the same statistics (distributions) as that of the desired process, then we could only simulate finite alphabet random processes whose statistics admit finite binary representations. For example, an exact simulation of a binary process with irrational probabilities is not feasible, since the number of fair bits per source output required for accurate simulation is infinite.

The problem appears to have a completely different nature once the requirement to produce exactly the same statistics is relaxed. In [1], Han and Verdú introduced the notion of *resolvability* of a random source. The resolvability of a source is defined as the minimum complexity of the source simulator so that the resulting statistics are *arbitrarily close* to the desired one. The complexity of a simulator is measured in [1] by the number of random bits per source output required to generate every realization of the simulating process, and by arbitrarily close is meant that the variational distance between the desired statistics and the statistics of the simulator output goes to zero with blocklength (n). The fact that the variational distance can be strictly positive for finite n and vanishes only in the limit allows for finite complexity simulation of any finite-alphabet source. Han and Verdú show that the resolvability of any finite-alphabet source equals to its minimum achievable fixed-length source coding rate, and that this quantity is equal to the *sup-entropy rate* of the source as defined in [1], without assuming any restrictions such as stationarity or ergodicity. Of course, the variational distance is not the only criterion by which the quality of approximation can be measured. In [2] it is shown that the same results hold for the weaker $\bar{d}$ distance.

The resolvability - and the sup-entropy rate - of a continuous alphabet random source is infinite, thus its "real world" simulation can never be arbitrarily accurate. Moreover, one may wish to reduce the simulation complexity below the sup-entropy rate at the expense of lower accuracy even in cases where arbitrarily good approximations are feasible. The question addressed in this paper is: what is the minimal complexity of the source simulator if it is required to approximate the desired statistics at a prespecified accuracy level D? We give a precise answer to this question for four different accuracy notions: the variational distance, the Prohorov distance, the general $\bar{\rho}$ distortion measure, and the general ρ_s distortion measure introduced here for the first time. The ρ_s distortion measure is defined by replacing the expectation operation in the $\bar{\rho}$ distance by limsup in probability; it is introduced and used here since we find it useful as a fidelity criterion for approximating nonstationary/nonergodic sources. The minimal complexity of the source simulator so that the resulting statistics approximates the desired statistics at a prespecified accuracy level is called the *finite precision resolvability* of the source.

Strong connections between problems of simulation and noiseless source coding are demonstrated in [1] and [2]. Here we extend this analogy to source coding subject to a fidelity criterion. Given a sequence of distortion measures on sequence spaces, we show that the minimum complexity of source simulation with accuracy level D with respect to the corresponding $\bar{\rho}$ distortion measure is equal to both the minimum achievable source coding rate with average distortion D and to the value of the $\bar{\rho}$ *sup rate-distortion function* at D, $\overline{R}_{\bar{\rho}}(D)$. Similarly, we show that the minimum complexity of source simulation with accuracy level D with respect to the ρ_s distortion measure equals to both the minimum achievable source coding rate with maximal distortion D, and to the value of the *sup rate distortion function* at D, $\overline{R}(D)$. These results hold for arbitrary source (not necessarily stationary/ergodic) and for arbitrary sequence-distortion measures, including non subadditive, context dependent sequence-distortion measures. This generality should be viewed more as a bonus than as the main contribution of this paper, which is to establish a new operational meaning for the rate-distortion function: the complexity of the random number generation necessary to approximate the *distribution* of a source with given tolerance.

At first glance, it may appear that the strong connection found in this paper between rate-distortion theory and source approximation is not surprising. However, recall that the purpose of source coding with a fidelity criterion is to approximate *sample paths* while getting rid of as much randomness as possible, whereas the purpose of approximation theory is to approximate *distributions* while generating as little randomness as possible.

We also provide few examples of the evaluation of the sup rate-distortion functions defined throughout the work. In particular, we show that the finite-precision resolvability with respect to the variational distance of any information stable source is equal to its entropy rate no matter how large the tolerated approximation error is. This phenomenon does not occur with the other (less stringent) approximation measures studied in this work.

REFERENCES

[1] T. S. Han and S. Verdú, "Approximation Theory of Output Statistics," *IEEE Trans. Inform. Theory*, Vol. 39, pp. 752-772, May 1993.

[2] Y. Steinberg and S. Verdú, "Channel Simulation and Coding With Side Information," to appear in *IEEE Trans. Inform. Theory*.