Coding Unstable Scalar Markov Processes Into Two Streams

Anant Sahai
UC Berkeley - EECS
e-mail: sahai@eecs.berkeley.edu

I. Introduction

Source-coding theorems for stable Markov and auto-regressive processes under mean-squared-error distortion are well established. [3, 4] Markov processes:

$$X_{t+1} = AX_t + W_t \tag{1}$$

where W_t white and $X_0 = 0$ contain the essence of the problem, which is to minimize the average distortion subject to a rate constraint on the encoding.

The unstable cases $A \geq 1$ are substantially more difficult to deal with since they are neither ergodic nor stationary and furthermore their variance grows unboundedly with time. As a result, Gray was able to prove only finite horizon results for such nonstationary processes and the general infinite-horizon unstable case had remained essentially open since Gray's 1970 paper.[4] On the computational side, Hashimoto and Arimoto gave a parametric form for computing the $R(D)$ function for unstable auto-regressive Gaussian processes.[5] Toby Berger gave an explicit coding theorem for an important sub-case: the marginally unstable Wiener process ($A = 1$) by introducing an ingenious parallel stream methodology and noticing that although the Wiener process is nonstationary, it is an independent increments process.[2, 1]

In earlier work, I gave a variable rate coding theorem that showed that we can achieve the $R(D)$ bound in the infinite-horizon case if we were allowed the use of variable rate codes.[8, 9] The question of whether or not fixed rate codes could be made to work was left open.

In this paper, I show that it is possible to encode a scalar unstable Markov process into two parallel fixed-rate bit streams. While in this paper I assume scalar Markov processes with bounded support for the driving noise, it is possible to extend the arguments to cover the general Gaussian ARMA case.[10]

II. R(D) Bound

For each N, we can calculate a standard information-theoretic rate-distortion function for the finite horizon problem as follows:

$$R_N^X(d) = \inf_{\{p_{Y_1^N|X_1^N} : \frac{1}{N}\sum_{i=1}^N E\left[(X_i - Y_i)^2\right] \leq d\}} \frac{1}{N} I(X_1^N; Y_1^N) \tag{2}$$

The infinite horizon case is defined as the limit of the finite horizon cases:

$$R_\infty^X(d) = \liminf_{N\to\infty} R_N^X(d) \tag{3}$$

III. Two stream encoding

Our strategy for approaching the $R_\infty^X(D)$ bound will be:

- Look at blocks of size N and encode the values of checkpoints X_{kN} recursively to very high precision using rate $N(\log_2 A + \epsilon_1)$ per value.
- Use the checkpoints at the start of the blocks to transform the process in between (the history) so that it looks close to an i.i.d. sequence of $\vec{X}$ from the finite horizon problems.
- Use the checkpoints at the end of the blocks to conditionally encode the history to fidelity D at a rate close to $N(R_\infty^X(D) + \epsilon_2)$.

The source decoding proceeds in the reverse manner. The checkpoints are decoded first, and then used to decode the history. Finally the two will be recombined to give a reconstruction of the original source to the desired fidelity. The above strategy follows the spirit of Berger's encoding[1] and that used for our variable rate code earlier[8, 9]. The key difference with Berger's scheme for the Wiener process is that the checkpoint encoding is of non-negligible rate relative to the main stream. The difference with our earlier variable rate scheme is that the history is encoded conditioned on the checkpoints rather than the other way around.

Lemma III.1 *It is possible to causally and recursively encode checkpoints spaced by N to arbitrarily high fidelity with average rate arbitrarily close to $\log_2 A$ by choosing N large enough.*

Theorem III.1 *Given an unstable scalar Markov process driven by noise with bounded support and having access to infinite common randomness between encoder and decoder, it is possible to encode the process to average fidelity close to D using two fixed-rate bitstreams. The first of which has rate arbitrarily close to $\log_2 A$ while the second has rate arbitrarily close to $R_\infty^X(D) - \log_2 A$.*

References

[1] T. Berger, "Information Rates of Wiener Processes." IEEE Transactions on Information Theory, 16(2):134–139, 1970.

[2] T. Berger, *Rate Distortion Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[3] R. Gallager, *Information Theory and Reliable Communication.* New York, NY: John Wiley and Sons, 1971.

[4] R. Gray, "Information Rates of Autoregressive Processes." IEEE Transactions on Information Theory, 16(4):412–421, 1970.

[5] T. Hashimoto and S. Arimoto, "On the Rate-Distortion Function for the Nonstationary Autoregressive Process." IEEE Transactions on Information Theory, 26(4):478–480, 1980.

[6] S.K. Mitter and N. J. Newton, "Information Flow and Entropy Production in the Kalman-Bucy Filter", In preparation.

[7] A. Sahai, ""Any-time" Capacity and A Separation Theorem For Tracking Unstable Processes", ISIT 2000.

[8] A. Sahai, "A Variable Rate Source-Coding Theorem for Unstable Scalar Markov Processes.", ISIT 2001.

[9] A. Sahai, ""Any-time" Information Theory, PhD Dissertation, MIT 2001.

[10] A. Sahai, "Source coding for unstable processes", In preparation.