

Average Redundancy of the Shannon Code for Markov Sources

Neri Merhav*

Wojciech Szpankowski†

August 6, 2012

Abstract

It is known that for memoryless sources, the average and maximal redundancy of fixed-to-variable length codes, such as the Shannon and Huffman codes, exhibit two modes of behavior for long blocks. It either converges to a limit or it has an oscillatory pattern, depending on the irrationality or rationality, respectively, of certain parameters that depend on the source. In this paper, we extend these findings, concerning the Shannon code, to the case of a Markov source, which is considerably more involved. While this dichotomy, of convergent vs. oscillatory behavior, is well known in other contexts (including renewal theory, ergodic theory, local limit theorems and large deviations of discrete distributions), in information theory (e.g., in redundancy analysis) it was recognized relatively recently. To the best of our knowledge, no results of this type were reported thus far for Markov sources. We provide a precise characterization of the convergent vs. oscillatory behavior of the Shannon code redundancy for a class of irreducible, periodic and aperiodic, Markov sources. These findings are obtained by analytic methods, such as Fourier/Fejér series analysis and spectral analysis of matrices.

Index Terms: Shannon code, average redundancy, Fourier series, uniform convergence, spectral analysis, analytic information theory.

*N. Merhav is with the Department of Electrical Engineering, Technion – Israel Institute of Technology, Technion City, Haifa 32000, Israel. E-mail: merhav@ee.technion.ac.il. N. Merhav's work was supported by the Israeli Science Foundation (ISF), grant no. 208/08.

†W. Szpankowski is with the Department of Computer Science, Purdue University, W. Lafayette, IN 47907, U.S.A. E-mail: spa@cs.purdue.edu. He is also a Visiting Professor at ETI, Gdańsk University of Technology, Poland. W. Szpankowski's work was supported in part by the NSF Science and Technology Center for Science of Information Grant CCF-0939370, NSF Grant CCF-0830140, AFOSR Grant FA8655-11-1-3076, and NSA Grant H98230-11-1-0141.