

Subset–sum phase transitions and data compression

Neri Merhav

Department of Electrical Engineering, Technion, Haifa 32000, Israel.

E-mail: merhav@ee.technion.ac.il

Abstract. We propose a rigorous analysis approach for the subset sum problem in the context of lossless data compression, where the phase transition of the subset sum problem is directly related to the passage between ambiguous and non-ambiguous decompression, for a compression scheme that is based on specifying the sequence composition. The proposed analysis lends itself to straightforward extensions in several directions of interest, including non-binary alphabets, incorporation of side information at the decoder (Slepian–Wolf coding), and coding schemes based on multiple subset sums. It is also demonstrated that the proposed technique can be used to analyze the critical behavior in a more involved situation where the sequence composition is not specified by the encoder.

Keywords: Number partitioning, integer partitioning, subset sum, data compression, source coding, entropy, phase transitions.