Sparse Parities for Large Scale Discrete Integration and Counting

Abstract—**Model counting and discrete integration over exponentially large sets are common tasks in machine learning. Ermon *et al.* have developed an approximation algorithm with provable guarantees based on (Average) Universal Hashing: the algorithm solves a tractable number of instances of an (NP-hard) combinatorial optimization problems subject to randomly generated linear parity constraints. Empirical evidence suggests that sparser parity constraints, involving a smaller number of variables, or parities with additional structure can substantially improve the performance of the combinatorial solvers in practice. But, excessive sparsity deteriorates the theoretical approximation guarantees.**

In this work, we establish a connection between universal hashing and linear error correcting codes; we show how the desired statistical properties of Average Universal (AU) hash families translate to properties of the weight enumerator of a random linear code ensemble. Finally, we consider a simple random construction of linear parities where each constraint independently includes a variable according to some vanishing probability. We show that when number of constructed parities grows linearly in the number of variables, then parities of expected logarithmic degree suffice to achieve the desired properties.