Good Codes Singleton Bound

David Ponarovsky

October 5, 2022

Abstract

We propose a new asymptotic upper bound on the trade-off between the rate and the distance of a good error correction code.

1 Preambles

Coding theory has emerged by the need to transfer information in noisy communication channels. By embedding a message in higher dimension space, one can guarantee robustness against possible faults. The ratio of the original content length to the passed message length is the rate of the code, and it measures how consuming our communication protocol is. Furthermore, the distance of the code quantifies how many faults the scheme can absorb such that the receiver could recover the original message.

Even though it is obvious that any construction resilient to a large number of faults should have a complexity price, The exact relation between the rate and the code distance is still unknown. However, we do know un-tight upper bounds. The first one was proved by Singelton and set the linear constraint: $\rho+\delta \leq 1-\frac{1}{\Delta}$ for any $[\Delta, \rho\Delta, \delta\Delta]$ linear code [Sin64].

Non-formally, we say that code is good if its distance and rate are scaled linearly in the encoded message length. Besides the fact that good codes are considered efficient in terms of robustness and overhead, they are also vital components in establishing secure multiparty computation [BGW19] and have a deep connection to probabilistic proofs.

In this work, we show a new upper bound $\rho+\frac{5}{16}\delta \leq 1+\frac{1}{16}$ which tighter than the Singelton bound and holds for any good code. First, we state the notations, definitions, and formal theorem in section 2. Then in sections 3 and 4, we review past results and provide their proofs in order to make this paper self-contained. Readers familiar with the basic concepts of LDPC codes and the Tanner code construction should consider skipping directly to section 5, in which we provide our proof.

2 Notations, Definitions, And Our Contribution

Here we focus only on linear binary codes, which one could think about as linear subspaces of $\mathbb{F}_2^n$. A common way to measure resilience is to ask how many bits an evil entity needs to flip such that the corrupted vector will be closer to another vector in that space than the original one. Those ideas were formulated by Hamming [Ham50], who presented the following definitions.

Definition. Let $n \in \mathbb{N}$ and $\rho, \delta \in (0,1)$. We say that C is a *binary linear code* with parameters $[n, \rho n, \delta n]$. If C is a subspace of $\mathbb{F}_2^n$, and the dimension of C is at least ρn. In addition, we call the vectors belong to C *codewords* and define the distance of C to be the minimal number of different bits between any codewords pair of C.

From now on, we will use the term code to refer to linear binary codes, as we don't deal with any other types of codes. Also, even though it is customary to use the above parameters to analyze codes, we will use their percent forms called the relative distance and the rate of code, matching δ and ρ correspondingly.

3 Singleton Bound

4 Tanner Code

5 Construction

6 Comparing To Known Bounds In A Variety Of Regimes

References

[Ham50] R. W. Hamming. "Error detecting and error correcting codes". In: *The Bell System Technical Journal* 29.2 (1950), pp. 147–160. DOI: 10.1002/j.1538-7305.1950.tb00463.x.

[Sin64] R. Singleton. "Maximum distanceq-nary codes". In: *IEEE Transactions on Information Theory* 10.2 (Apr. 1964), pp. 116–118. ISSN: 1557-9654. DOI: 10.1109/TIT.1964.1053661.

[BGW19] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. "Completeness Theorems for Non-Cryptographic Fault-Tolerant Distributed Computation". In: *Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali*. New

York, NY, USA: Association for Computing Machinery, 2019, pp. 351–371. ISBN: 9781450372664. URL: https://doi.org/10.1145/3335741.3335756.