Miscorrection Probability Beyond The Minimum Distance[1]

Yuval Cassuto
California Institute of Technology
Electrical Engineering Department
MC 136-93 Pasadena, CA 91125, U.S.A.
E-mail: ycassuto@paradise.caltech.edu

Jehoshua Bruck
California Institute of Technology
Electrical Engineering Department
MC 136-93 Pasadena, CA 91125, U.S.A.
E-mail: bruck@paradise.caltech.edu

Abstract — **The micorrection probability of a list decoder is the probability that the decoder will have at least one non-causal codeword in its decoding sphere. Evaluating this probability is important when using a list-decoder as a conventional decoder since in that case we require the list to contain at most one codeword for most of the errors. A lower bound on the miscorrection is the main result. The key ingredient in the proof is a new combinatorial upper bound on the list-size for a general $q-$ary block code. This bound is tighter than the best known on large alphabets, and it is shown to be very close to the algebraic bound for Reed-Solomon codes. Finally we discuss two known upper bounds on the miscorrection probability and unify them for linear MDS codes.**

I. Combinatorial List-Size Bound

A simple closed form bound on the list size is proposed for a general (n, d) block code over $q-$ary alphabet, decoded to radius t. This bound is independent of q, and for $q > q_0$, where q_0 depends on n, d, t, it is tighter than the best known combinatorial bound that can be found e.g in [3, 4]. If we fix the asymptotic distance by $\gamma = 1 - \frac{d}{n}$ and decoding radius by $\delta = 1 - \frac{t}{n}$, the proposed bound suggests that

$$L \leq \frac{\delta - \gamma}{\delta^2 - \gamma} \tag{1}$$

whereas for large q, [3] tends to a looser bound of $\frac{1-\gamma}{\delta^2-\gamma}$.When using (1) with the parameters of Reed-Solomon codes, it is shown to coincide with the algebraic bound of [2] when δ tends to $\sqrt{\gamma}$, for all rates γ. This δ corresponds to the maximum achievable decoding radius of the Guruswami-Sudan algorithm. For general values of δ, the difference between (1) and the algebraic bound is bounded above by $\frac{1}{4}\left[1 + \frac{2}{1-\sqrt{\gamma}}\right]$. Using that bound, the difference is less than 2 for $\gamma \leq \frac{1}{2}$ and less than 10 for $\gamma \leq 0.9$.

II. Lower Bound on Miscorrection

The following theorem states a lower bound on the miscorrection probability.

Theorem 1 *Let $\mathcal{C}^*$ be an (n, k, d) linear code with alphabet size q and weight distribution $A(w)$. t is the decoding radius, $t_0 \equiv \left\lfloor \frac{d-1}{2} \right\rfloor < t < n(1-\sqrt{1-d/n})$. C_0 will be the transmitted codeword, E the error word and $R = C_0 + E$, the received word. The miscorrection probability given a weight u error is defined as $P_e(u) = Pr(\exists C \in \mathcal{C}^* \backslash C_0 : D(R, C_0) = u, D(R, C) \leq t)$. Then*

$$P_e(u) \geq \frac{t^2 - 2nt + dn}{n(d-t)} \cdot \frac{\sum_{w=d}^{n} A(w) \sum_{s=0}^{t} N(w, u; s)}{\binom{n}{u}(q-1)^u} \tag{2}$$

and $N(w, u; s)$ is an efficiently computable function, independent of $\mathcal{C}^*$ for all $(w, u; s)$.

The first term in the right side of (2) is the inverse of the combinatorial bound on the list size discussed in I and the second term is a known upper bound on the number of weight u decodable words discussed in III. Sample results for a linear MDS code with parameters $n = 31, k = 15, q = 32$ are shown in figure 1 for decoding radius of $t = t_0 + 1 = 9$. The curves from top to bottom are: i) upper bound using the method from [5]. ii) improved lower bound given in (2). iii) lower bound that counts only in the t_0 sphere. The true value of the miscorrection is proved to be between the two upper curves.

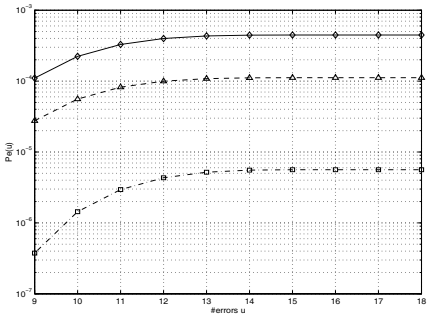

Figure 1: Bounds on the miscorrection probability for a (31,15) MDS code, decoded to radius 9

III. Counting Decodable Words

For symmetric channels, the miscorrection probability is calculated by counting decodable words. When decoding beyond half the minimum distance, known methods for counting decodable words no longer provide exact solutions. We discuss two methods to count decodable words, the direct method [5] and the inclusion-exclusion method [6]. By simplifying the expression of [6], we show that the two methods give the same upper bound on the number of decodable words for all t.

References

[1] Y. Cassuto and J. Bruck, "Miscorrection Probability Beyond the Minimum Distance", electronic technical report, http://www.paradise.caltech.edu/papers/etr058.pdf, 2004.

[2] R.J McEliece, "The Guruswami-Sudan Algorithm for Decoding Reed-Solomon Codes", IPN progress report 42-153, JPL , 2003.

[3] O. Goldreich, R. Rubinfeld, M. Sudan, "Learning Polynomials with Queries: The Highly Noisy Case", proc. of FOCS, 1995.

[4] V. Guruswami and M. Sudan "Extensions to the Johnson Bound", Manuscript, 2001.

[5] Z.M Huntoon and A.M Michelson, "On the Computation of the Probability of Post-Decoding Error Events for Block Codes", *IEEE Trans. on Inform. Theory* IT-23, May 1977.

[6] K.M Cheung, "More on the Decoder Error Probability for Reed-Solomon Code", *IEEE Trans. on Inform. Theory* IT-35, July 1989.

[1]This work was supported in part by the Lee Center for Advanced Networking at the California Institute of Technology.