

Interpolation and approximation in decoding

Iwan Duursma

March 1998

Abstract

Recently Sudan formulated a decoding procedure for decoding RS-codes beyond the packing radius. The potential of the method for AG-codes was recognized by Shokrollahi and Wasserman. We discuss similarities and differences with some previous algebraic decoding procedures.

1 Introduction

When Goppa introduced his famous construction of linear codes with algebraic curves over a finite field, he suggested that the decoding problem should be regarded as a problem of approximation of differentials. The various contributions to the solution of the decoding problem show that the actual algorithms are more easily formulated in terms of the dual code that is defined in terms of algebraic functions.

Other approaches emphasize interpolation as a way to solve the decoding problem. Thus the code itself may be defined in terms of algebraic functions and the decoding problem can be seen as looking for the best fit through a set of points in which some points are unreliable. The work by Sudan and by Shokrollahi and Wasserman shows the success of this approach in decoding beyond the packing radius of a code.

Characteristic for the approximation algorithms is the use of syndromes and the computation of unknown syndromes. These algorithms typically are fast for high rate codes with few errors. The interpolation method is fast for low rate codes. The decoding beyond the packing radius has so far only been achieved for low rate codes.

2 A class of RS-codes

The code C of length n and dimension k over the field F is said to be of type $[n, k, e, b]$ if any Hamming sphere of radius e contains at most b codewords [1].

For an analysis of the interpolation method we consider RS-codes with carefully chosen parameters. Let $X = X_0 \cup \dots \cup X_b \subset F$ be a partition of coordinates such that X_0 is of size r and the remaining b subsets are of size m . Let f_1, \dots, f_b be b distinct polynomials of degree r that are zero on X_0 . The evaluation on X

of polynomials f of degree at most r defines a (punctured) RS-code of length $n = r + bm$ and dimension $k = r + 1$. The word that is zero on X_0 and that has values $f_j(x)$ on X_j is at distance $(b - 1)m$ of each of the b codewords $f_j(X)$.

Theorem: For given integers $b, r > 0$, let $m = \binom{b}{2}(r + 1)$. let C be a (punctured) RS-code of length $n = r + bm$ and dimension $k = r + 1$. Any Hamming sphere of radius $e = (b - 1)m$ contains *at most* b codewords.

Proof: The codewords can be computed effectively by the interpolation method in [1].

Note that the codes have minimum distance $d = bm$ and that $e/d = (b - 1)/b$.

3 Factorization and syndromes

The equations that are satisfied by the unknown syndromes are in general polynomial and are in general believed to be hard to solve (but see [2]). The interpolation method suggests that these equations can be solved efficiently (that is with the complexity of factorization) in a certain range of parameters. We therefore ask the following.

Problem: Does the vanishing ideal of the unknown syndromes have a special structure in the range of parameters where the interpolation method is successful?

We illustrate the problem for a code of type $[8, 3, 3, 2]$ ($b = r = 2$ in the theorem). For a received word (y_i) , the interpolation method first computes a polynomial

$$(f_0)y^2 + (g_0 + g_1x + g_2x^2)y + (h_0 + h_1x + h_2x^2 + h_3x^3 + h_4)$$

that vanishes in the eight points (x_i, y_i) . Codewords at distance at most three from the received word (y_i) are then obtained from the linear factors of the polynomial.

We assume that the coordinate set $X \subset F$ is the zero set of $x^8 - x$ so that the code is extended cyclic. For known syndromes S_0, S_1, S_2, S_3, S_4 , the unknown syndromes S_5, S_6, S_7 are such that the matrix

$$\begin{pmatrix} S_0 & S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 \\ S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_1 \\ S_2 & S_3 & S_4 & S_5 & S_6 & S_7 & S_1 & S_2 \\ S_3 & S_4 & S_5 & S_6 & S_7 & S_1 & S_2 & S_3 \end{pmatrix}$$

has rank at most three. The full minors of the matrix define the vanishing ideal of the unknown syndromes. From the interpolation method we know that there are *at most two* syndrome extensions. Moreover they can be obtained efficiently.

We claim that the ideal contains the determinant of the matrix

$$\begin{pmatrix} S_0 & S_1 & S_2 & S_0^2 - 2S_3S_4 \\ S_1 & S_2 & S_3 & -S_4^2 \\ S_2 & S_3 & S_4 & -S_1^2 \\ S_3 & S_4 & S_5 & S_5^2 - 2S_1S_2 \end{pmatrix}$$

Indeed, there are *at most two* possible solutions for the first unknown syndrome S_5 .

4 A previous result

In [3], we give a fast erasure decoding scheme for low rate codes. It requires $3kn$ field multiplications and is actually an implementation of the interpolation method for $b = 1$.

References

- [1] M. Sudan, “Decoding of Reed Solomon codes beyond the error correction bound,” *J. Complexity*, vol. 13, no. 1, pp. 180–193, 1997.
- [2] V.M. Sidelnikov, “Decoding the Reed-Solomon code when the number of errors is greater than $(d - 1)/2$, and zeros of polynomials in several variables.” (Russian) *Problemy Peredachi Informatsii*, vol. 30, no. 1, pp. 51–69, 1994; (Translation) *Problems Inform. Transmission*, vol. 30, no. 1, pp. 44–59, 1994.
- [3] I.M. Duursma, “On erasure decoding of AG-codes,” *Proceedings of ITW’94 (Moscow)*. Available at:
<http://www.research.att.com/~duursma/pub>.