Interleaved and Lifted Reed-Solomon Codes: New Perspectives and Constructions

Henry D. Pfister
Department of Electrical and Computer Engineering
Texas A&M University
College Station Campus

Swiss Federal Institute of Technology, Zurich
July 26th, 2007

Three Questions

- How many equally spaced samples are required to find the amplitude and frequency of the sum of t sinusoids?
- How many errors can be corrected by a $[17, 8, 10]$ shortened Reed-Solomon code over the Galois field $\mathbb{F}_{256}$?
- How are these questions related?

Outline

Outline

Error-Correcting Codes: Background

- Basics
  - Let F^n be the n-dim vector space over the field F
    - The Galois field with q elements will be denoted $\mathbb{F}_q$
  - The Hamming weight $w_H(\mathbf{x})$ of a vector $\mathbf{x} \in F^n$ is the number of non-zero elements or $w_H(\mathbf{x}) = |\{i|x_i \neq 0\}|$
    - Hamming distance between vectors is $d_H(\mathbf{x}, \mathbf{y}) = w_H(\mathbf{x} - \mathbf{y})$
- An $[n, k, d]_F$ linear error-correcting code $\mathcal{C}$
  - is a k-dim subspace of F^n (with q^k elements if $F = \mathbb{F}_q$)
  - with minimum distance $d = \min_{\mathbf{x},\mathbf{y}\in\mathcal{C},\mathbf{x}\neq\mathbf{y}} d_H(\mathbf{x}, \mathbf{y})$
  - Can correct all error patterns with at most $(d-1)/2$ errors

Designing Good Codes

- First order approximation
  - Want codes with large distance (to correct more errors)
    - The Singleton bound asserts that $d \leq n - k + 1$
  - That we can decode efficiently up to $(d-1)/2$ errors
- One answer: Reed-Solomon (RS) Codes
  - $[n, k, d]_F$ codes with $n \leq |F|$ and optimal $d = n - k + 1$
  - Decoder with complexity $O\big(n(n-k) + (n-k)^2\big)$
- Opportunity for improvement
  - Many error patterns with more than $(d-1)/2$ errors can be corrected in theory, but how?

Practical Issues (1)

- Reed-Solomon codes, block length n, and alphabet size $|F|$
  - Main problem is that n is limited by $|F|$
  - For a fixed code rate $R = k/n$, large n helps performance
    - Assume error are i.i.d. with probability p (# errors is binomial)
    - Law of large numbers helps if $p < \frac{1-R}{2}$ $\quad (d \approx (1-R)n)$

- For communication systems with fixed characteristics
  - Increasing the $|F|$ tends to increase the number of errors
  - So a trade-off occurs

Practical Issues (2)

- Some communication systems are exceptions
  - Internet packet channel
    - Alphabet size $|F| \geq 2^{400}$ (≥ 50 bytes) is reasonable
  - Burst error channels (i.e., errors occur in tight groups)
    - For large $|F|$, one burst may affect only 1-2 code symbols
    - Studied by many authors especially for hard disk drives
- For these channels
  - The practical limit on the block size is $\ll |F|$
  - Increasing $|F|$ to improve performance is reasonable

Outline

Definition of a Reed-Solomon (RS) Code

- $[n,k,d]_F$ RS Code via Polynomial Evaluation (RS 1960)
  - Message $(f_0,f_1,\ldots,f_{k-1}) \in F^k \implies f(x) = \sum_{i=0}^{k-1} f_i x^i$
  - Codeword $(c_0,c_1,\ldots,c_{n-1})$ defined by $c_i = f(\alpha^i)$
    - Where $\alpha \in F$ has order n and α^i gives n unique elements
  - Encoder computes the DFT of the message
- Properties
  - Code and encoder are both linear
    - Weighted sum of messages $\rightarrow$ Weighted sum of codewords
  - Minimum distance is $d = n - k + 1$
    - Any non-zero message poly has at most $k-1$ zeros

Errors and Syndromes

- Let $c(x) = \sum_{i=0}^{n-1} c_i x^i$ be the codeword polynomial
  - Codeword poly satisfies $c(\alpha^i) = 0$ for $i = 1, 2, \ldots, d-1$
  - Proof follows from examining DFT(DFT(message))
- Let $e(x) = \sum_{i=1}^{t} e_{\sigma(i)} x^{\sigma(i)}$ be the error polynomial
  - for t errors at locations $\sigma(1), \sigma(2), \ldots, \sigma(t)$
- Let $v(x) = c(x) + e(x)$ be the received polynomial
- The syndrome sequence $s_1, s_2, \ldots, s_{n-k}$ is defined by

$$s_j \triangleq v(\alpha^j) = \underbrace{c(\alpha^j)}_{0} + e(\alpha^j) = \sum_{i=1}^{t} e_{\sigma(i)} \alpha^{j\,\sigma(i)}$$

Outline

Connection with Spectral Estimation

- The syndrome sequence is $n-k$ equally spaced samples of the sum of t periodic exponentials

$$s_j = \sum_{i=1}^{t} e_{\sigma(i)} \alpha^{j\,\sigma(i)}$$

  - whose "frequencies" $\sigma(\cdot)$ equal the error locations
  - Finding $\sigma(\cdot)$ and $e_{\sigma(\cdot)}$ (i.e., spectral estimation) gives $e(x)$

- The method of Baron Gaspard De Prony (circa 1795)
  - Computes the amplitude and frequency of the sum of t complex exponentials using $2t$ equally spaced samples
  - Is identical to standard RS decoding (Wolf 1967)

Prony's Method (1)

- One way to understand Prony's method is via filter design
- It simply designs a $t+1$ tap FIR filter: $(b_0, b_1, \ldots, b_t)$
  - Which nulls the syndrome sequence
  - This gives the linear system (assuming $b_0 = 1$ w.o.l.o.g.)

$$\begin{bmatrix} s_1 & s_2 & \cdots & s_t \\ s_2 & s_3 & \cdots & s_{t+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_t & s_{t+1} & \cdots & s_{2t-1} \end{bmatrix} \begin{bmatrix} b_t \\ b_{t-1} \\ \vdots \\ b_1 \end{bmatrix} = \begin{bmatrix} -s_{t+1} \\ -s_{t+2} \\ \vdots \\ -s_{2t} \end{bmatrix}$$

- The transfer function $B(z) = \sum_{i=0}^{t} b_i z^{-i}$ (aka error-locator)
  - Must have zeros at the t frequencies $\sigma(1), \sigma(2), \ldots, \sigma(t)$
  - So, the roots of $B(z)$ give the set $\sigma(\cdot)$ of error locations

Prony's Method (2)

- Once the frequencies (i.e., error locations) are known
  - Errors become erasures and $d-1$ can be corrected
  - The amplitudes are computed with a simple linear system
- Benefits of this perspective
  - RS decoding is limited by finding the error locations
  - Finding more constraints on the filter may allow one to find $B(z)$ with **more roots** and allow decoding beyond $t = \frac{n-k}{2}$
    - Ex: More equations and BCH codes (Hartmann-Tzeng 1974)
  - Correct solution must satisfy all eqns but may not be unique

Outline

Interleaved Reed-Solomon Codes (1)

- Motivation
  - In some systems, there is a chance of long burst errors
  - Interleaving multiple RS codewords together has the effect of spreading the burst error across multiple codes
  - One long RS code would work, but require a larger field
  - Interleaving gives similar results with a smaller field
- Consider m RS codes $[n, k, d]_F$ interleaved together
  - Let $f_i(x) = \sum_{j=0}^{k-1} f_{ij} x^j$ be i-th message polynomial
  - The codeword can be written in matrix form as

$$\begin{bmatrix} f_1(\alpha^0) & f_1(\alpha^1) & \cdots & f_1(\alpha^{n-1}) \\ f_2(\alpha^0) & f_2(\alpha^1) & \cdots & f_2(\alpha^{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ f_m(\alpha^0) & f_m(\alpha^1) & \cdots & f_m(\alpha^{n-1}) \end{bmatrix}$$

Interleaved Reed-Solomon Codes (2)

- Decoding with m syndrome sequences (LK98)
  - Let s_{ij} be the j-th element of the i-th syndrome seq.
  - Syndrome freq. content depends only on column of error
  - All syndromes are used to find one column error-locator
    - A root at α^i implies an error in some row of column i
- Can we correct a set of errors occurring only in t columns?
  - Each syndrome seq. gives $n-k-t$ equations which require the filter to null the last $n-k-t$ elements of the seq.
  - Unique solution possible only if $m(n-k-t)$ eqns $\geq t$ vars
    - Can be rewritten as $t \leq \frac{m}{m+1}(n-k)$
  - Random error model $\rightarrow$ failure prob. $\approx |F|^{-1}$ (CS03,BKY03)

Example: Correcting 2 errors with $n - k = 3$ and $m = 2$

- Assume that columns $\sigma(1)$ and $\sigma(2)$ have errors
  - whose amplitudes in the i-th row are $e_{i,\sigma(1)}$ and $e_{i,\sigma(2)}$
- The column error-locator $B(z)$ can be found by solving

$$\begin{bmatrix} s_{1,1} & s_{1,2} \\ s_{2,1} & s_{2,2} \end{bmatrix} \begin{bmatrix} b_2 \\ b_1 \end{bmatrix} = \begin{bmatrix} -s_{1,3} \\ -s_{2,3} \end{bmatrix}$$

  - where the matrix determinant $D = s_{1,1}s_{2,2} - s_{1,2}s_{2,1}$ is

$$\alpha^{\sigma(1)+\sigma(2)} \left(\alpha^{\sigma(1)} - \alpha^{\sigma(2)}\right) \left(e_{1,\sigma(2)}\, e_{2,\sigma(1)} - e_{1,\sigma(1)}\, e_{2,\sigma(2)}\right)$$

  - Since $\sigma(1) \neq \sigma(2)$, D will be zero iff the last term is zero
  - Easy to see that uniform $e_{i,j}$ implies $Pr(D = 0) \approx |F|^{-1}$.

Outline

Lifted Codes

Definition

An m-lift of an $[n, k, d]_F$ code is an $\left[\frac{n}{m}, \frac{k}{m}, d'\right]_{F^m}$ code where each code symbol in the new code consists of m code symbols from original code.

- The distance satisfies $\frac{d}{m} \leq d' \leq \frac{n-k}{m} + 1$ and the upper bound holds with equality if $d = n - k + 1$
- This definition is motivated by the fact that IRS with column errors really behave like codes over F^m

Upper Bounds for Lifted IRS Codes

- Can we correct $t > \frac{m}{m+1}(n-k)$ errors with an IRS code?
  - Recall that the Hamming bound gives a lower bound on the $n-k$ needed to correct all patterns of t errors

Corollary

The Hamming bound for m interleaved RS codes over $\mathbb{F}_q$ (with $n = q$) lifted into a single code over $\mathbb{F}_{q^m}$ implies that

$$n - k \geq t\left(\frac{m+1}{m} - \frac{\log t - \log(1 - n^{-m})}{m \log n}\right).$$

- This implies $n - k \approx \frac{m+1}{m}t$ is nearly optimal when $\frac{\log t}{m \log n} \ll 1$
  - Similar for high probability decoding and list decoding

Outline

Spectral Estimation Revisited

- Can we use these ideas to improve Prony's method?
  - Spectral estimation with fewer samples needs more eqns
  - Consider the sum of t complex sinusoids ($\gamma_i = \mathrm{e}^{\omega_i \sqrt{-1}}$) and its time reversed complex conjugate

$$s_j = \sum_{i=1}^{t} a_i \gamma_i^j \qquad s^*_{r-j} = \sum_{i=1}^{t} \left(a_i^* \gamma_i^{-r}\right) \gamma_i^j$$

- Since s^*_{r-j} has the same frequency content as s_j
  - Gives a new method for t sinusoids using $\lceil 3t/2 \rceil$ samples
  - Example: $t = 2$, $s_j = a_1\gamma_1^j + a_2\gamma_2^j$, and 3 samples gives

$$\begin{bmatrix} s_1 & s_2 \\ s_3^* & s_2^* \end{bmatrix} \begin{bmatrix} b_2 \\ b_1 \end{bmatrix} = \begin{bmatrix} -s_3 \\ -s_1^* \end{bmatrix}$$

  - If the amplitudes a_1, a_2 have a uniform random phase, then this method fails on a set of measure zero

Have We Cheated the First Question?

- Prony's method
  - Works for arbitrary complex exponential rotators
  - For example, $\gamma_1 = 2e^{3\sqrt{-1}}$
- The new method based on conjugacy
  - Requires that $(\gamma_1^*)^{-1} = \gamma_1$ which implies $|\gamma_1| = 1$
- Yes, we have cheated!
  - We can find t "sinusoids" using $\lceil 3t/2 \rceil$ samples
    - but not t "complex exponentials"
  - The gain comes from reducing γ_1 from 2-dim to 1-dim

Generalization to Galois Fields

- Let $F = \mathbb{F}_{q^2}$ and let $\alpha \in F$ be a primitive element
  - Choose $\beta = \alpha^{q-1}$ and consider its conjugate β^q (w.r.t. $\mathbb{F}_q$)
    - A little algebra shows $\beta^q = \alpha^{q^2-q} = \alpha^{1-q} = \beta^{-1}$
- Since $\beta^{q+1} = \alpha^{(q-1)(q+1)} = \alpha^{q^2-1}$, we see β has order $q+1$
  - And we can construct a $[q+1, k, q+2-k]_F$ shortened RS code by evaluating $f(x) = \sum_{i=0}^{k-1} f_i x^i$ at β^i for $i = 0, \ldots, q$
  - Syndrome seq. will have time reversal conjugate symmetry
  - Decoding trick allows attempts up to $\lfloor 2(n-k)/3 \rfloor$ errors

Now for the Second Question

- Applying the previous construction with $k = 8$ and $q = 16$
  - Gives a $[17, 8, 10]$ shortened RS code over $\mathbb{F}_{256}$
  - Which we attempt to decode up to $\lfloor 2(n-k)/3 \rfloor = 6$ errors
    - Classical decoding corrects only $\lfloor (d-1)/2 \rfloor = 4$
  - Evaluating the Hamming bound shows that this is optimal
    - Any $n = 17$ code over $\mathbb{F}_{256}$ which corrects most errors of weight ≤ 6 must have fewer than 256^9 codewords
- This partially answers the second and third questions

Outline

Folded Reed-Solomon Codes

- Given a $[n,k,d]_F$ RS code $c_i = f(\alpha^i)$ with $n = \text{ord}(\alpha)$, $m \mid n$
  - Lift to F^m by grouping columns of $N = \frac{n}{m}$ by m matrix

$$\begin{bmatrix} f(\alpha^0) & f(\alpha^1) & \cdots & f(\alpha^{N-1}) \\ f(\alpha^N) & f(\alpha^{N+1}) & \cdots & f(\alpha^{2N-1}) \\ \vdots & \vdots & \ddots & \vdots \\ f(\alpha^{(m-1)N}) & f(\alpha^{(m-1)N+1}) & \cdots & f(\alpha^{mN-1}) \end{bmatrix}$$

  - Introduced by Krachkovsky in 2003

- Can be decoded like an interleaved RS code
  - Decimating the syndrome to m subsequences aliases all errors in the same column to a single error frequency
  - The m subseq. are treated as syndromes of an IRS code

Conjugacy Class Reed-Solomon Codes

- Given a $[n, k, d]_F$ RS code with $E = \mathbb{F}_q$ and $F = E^m = \mathbb{F}_{q^m}$
  - Any element in F has at most m conjugates w.r.t E
  - Let $\widetilde{F} \subset F$ contain all elements with exactly m conjugates
  - So, $\widetilde{F}$ breaks into $N = \frac{|\widetilde{F}|}{m}$ disjoint conjugacy classes
  - $\beta_{i,j}$ will denote the j-th member of the i-th conjugacy class
- Conjugacy class RS (CCRS) code from the $N \times m$ matrix

$$\begin{bmatrix} f(\beta_{0,0}) & f(\beta_{1,0}) & \cdots & f(\beta_{N-1,0}) \\ f(\beta_{0,1}) & f(\beta_{1,1}) & \cdots & f(\beta_{N-1,1}) \\ \vdots & \vdots & \ddots & \vdots \\ f(\beta_{0,m-1}) & f(\beta_{1,m-1}) & \cdots & f(\beta_{N-1,m-1}) \end{bmatrix}$$

Comparison of Folded and CC RS Codes

- Consider the base field $F = \mathbb{F}_{64}$
  - Supports a $[21, k, 21 - k + 1]_{F^3}$ folded RS code (3-lift)
    - Works with $m = 3$ because $3 \cdot 21 = 63$
  - Supports a $[20, 20 - (d - 1), d]_{F^3}$ CCRS code (3-lift)
    - Works with $m = 3$ because $4^3 = 64$
    - Length is $\frac{|\widetilde{F}|}{3}$ and $\left|\widetilde{F}\right| = |\mathbb{F}_{64} - \mathbb{F}_4| = 60$
- Folded RS codes always slightly longer than CCRS codes
- Benefit lies in flexibility and algebraic connections
  - Folded RS codes with $m = 2$ over binary fields don't exist

Outline

Algebraic List Decoding

- First Sudan (1996) and then with Guruswami (1999)
  - RS List decoding up to $\tau = n(1 - \sqrt{R})$ errors in poly time
    - Returns all codewords within radius τ of received
  - Largest gains at low rates
- What is the hidden structure that allows this to work?
  - Perhaps it is hidden syndromes (Schmidt et al. 2006)
  - Consider the m-th power of symbols received correctly

$$(v_i)^m = (c_i + e_i)^m = \Big(f(\alpha^i) + 0\Big)^m = (f^m)(\alpha^i)$$

  - New received seq. for "virtual" RS code with msg $f^m(x)$
    - Where the error locations unchanged so the associated new syndromes have the same frequency content

List Decoding of Interleaved RS Codes

- Clever generalization of GS by Parvaresh-Vardy (PV04)
  - IRS List decoding up to $\tau = n(1 - R^{m/m+1})$ errors
    - Possibility of failure, but simulations show this unlikely
  - Gains over previous decoder at low rates
- What is the hidden structure that allows this to work?
  - This is an interesting open question
- Big news: PV decoding for folded RS Codes (GR05)
  - List decoder provably returns all codewords within radius τ
  - By Guruswami and Rudra in 2005

Outline

Summary

- Interleaved Reed-Solomon Codes
  - Allow algebraic decoding beyond $(d-1)/2$ errors
  - Led to new improved Prony's method for spectral estimation
  - Led to new shortened RS with improved decoding
- Lifted Codes
  - Applied Hamming bound to IRS code performance
  - New Conjugacy Class Reed-Solomon Codes
- Algebraic List Decoding
  - Prony/filter perspective even provides insights for GS
  - Future work: Achieve GS/PV gains with simple decoder