A Rate-Distortion Perspective on Multiple Decoding Attempts for Reed-Solomon codes

Phong S. Nguyen, H. D. Pfister, K. R. Narayanan

ECE Department, Texas A&M University

Wireless Communications Lab

Introduction & Motivation

Reed-Solomon (RS) codes are one of the most widely used error-correcting codes. People have put a considerable effort into improving the decoding performance at the expense of complexity.

- Conventional hard-decision (HDD) algorithms such as **Berlekamp-Massey** (BM) algorithm can correct errors up to half the minimum distance ($d_{\min}$) of the code.
- A break-through result of **Guruswami-Sudan** (GS) using bivariate polynomial interpolation and list decoding can correct beyond half the minimum distance.
- **Kötter-Vardy** (KV) later extended the GS decoder to an algebraic soft-decision (**ASD**) decoding algorithm that has a considerable gain over HDD algorithms.

Both GS and KV algorithms however have significant complexity. Multiple runs of error-and-erasure decoding with some low-complexity algorithms, such as the BM, has renewed the interest of researchers.

- The Generalized Minimum Distance (**GMD**) repeats error –and-erasure decoding while successively erasing an even number of the least reliable positions (LRPs).
- More recent work from Lee-Kumar proposes a soft information successive (multiple) error-and-erasure decoding (**SED**) that runs multiple error-and-erasure decoding trials with every combination of an even number $\leq f$ of erasures within the l LRPs.

Question: **How can one construct the "best" set of erasure patterns?**

Multiple Error-and-Erasure Decoding

Classical decoding threshold: Consider an (n,k) RS code. If e symbols are erased, a conventional HDD error-and-erasure decoder such as the BM algorithm can correct v errors in unerased positions if $2v+e<n-k+1$ (1)

Conventional error & erasure patterns: We define $x^n \in \mathbb{Z}_2^n \triangleq \{0,1\}^n$ and $\hat{x}^n \in \mathbb{Z}_2^n$ as an error pattern and an erasure pattern respectively, where $s_i=0$ means that an error occurs and $\hat{s}_i=0$ means that an erasure occurs at index i.

Definition: Given a *letter-by-letter* distortion measure δ, the distortion between an error pattern x^n and an erasure pattern $\hat{x}^n$ is defined by $d(x^n,\hat{x}^n)=\sum_{i=1}^n \delta(x_i,\hat{x}_i)$

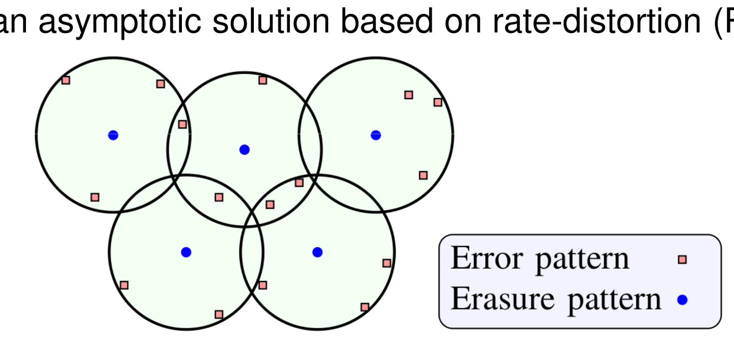
Proposition: If we choose the *letter-by-letter* distortion measure as follows

$$\delta(0,0)=1 \qquad \delta(0,1)=2$$
$$\delta(1,0)=1 \qquad \delta(1,1)=0$$

then the condition (1) reduces to the form $d(x^n,\hat{x}^n)<n-k+1$

Question: **How many decoding attempts needed to achieve a fixed distortion threshold?**

This turns out to be a covering problem where one wants to cover the space of error patterns using a minimum number of balls centered at erasure patterns. This view leads to an asymptotic solution based on rate-distortion (R-D) theory.



Here, error patterns are viewed as source sequences and erasure patterns are viewed as reproduction sequences.

Multiple Error-and-Erasure Decoding (cont.)

We consider a generalization of the conventional error and erasure patterns under the same framework to make better use of the soft information.

Generalized error & erasure patterns: Consider a positive integer l. Let us define $x^n \in \mathbb{Z}_{l+1}^n$ as the generalized error pattern where $x_i=j$ implies that the j-th most likely symbol is correct and $x_i=0$ implies none of the first l most likely symbols is correct. Let $\hat{x}^n \in \mathbb{Z}_{l+1}^n$ be the generalized erasure pattern where $\hat{x}_i=j$ implies that the j-th most likely symbol is used as the hard-decision symbol and $\hat{x}_i=0$ implies that an erasure is used

Theorem: We choose the *letter-by-letter* distortion measure $\delta:\mathbb{Z}_{l+1}\times\mathbb{Z}_{l+1}\to\mathbb{R}_{\geq 0}$ defined by $\delta(x,\hat{x})=[\Delta]_{x,\hat{x}}$ in terms of the $(l+1)\times(l+1)$ matrix

$$\Delta=\begin{pmatrix} 1 & 2 & \ldots & 2 & 2 \\ 1 & 0 & \ldots & 2 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & \ldots & 0 & 2 \\ 1 & 2 & \ldots & 2 & 0 \end{pmatrix}$$

Using this, the condition (1) becomes $d(x^n,\hat{x}^n)<n-k+1$

Proposed Multiple Decoding Algorithm

The first step is designing a distortion measure that converts the condition for a single decoding to succeed to the form where distortion is less than a fixed threshold.

Phase I: Compute rate-distortion function

- **Step 1**: Transmit T (say T=1000) arbitrary test codewords over the channel and compute a set of T matrices $\mathrm{P}_1^{(t)}$ where $[\mathrm{P}_1^{(t)}]_{j,i}$ is the probability of the j-th most likely symbol at codeword position i during time t
- **Step 2**: For each time t, obtain the matrix $\mathrm{P}_2^{(t)}$ from $\mathrm{P}_1^{(t)}$ through a permutation that sort the column in increasing reliability order of codeword positions. Take the entry-wise average of all matrices $\mathrm{P}_2^{(t)}$ to get an average matrix $\bar{\mathrm{P}}$
- **Step 3**: Compute the R-D function of source sequence (error pattern) with probability of source letters derived from $\bar{\mathrm{P}}$. Determine the point on the R-D curve that corresponds to a designed rate R along with the test-channel input-probability distribution vector q that achieves that point.

Phase II: Run actual decoder

- **Step 4**: Based on the actual received signal sequence, sort the codewords positions in increasing reliability order through a permutation π
- **Step 5**: Randomly generate a set of 2^R erasure patterns using the distribution vector q and permute the positions by the permutation π^{-1}
- **Step 6**: Run multiple attempts of the corresponding decoding scheme (e.g. error-and-erasure decoding) using the set of of erasure patterns in Step 5 to produce a list of candidate codewords.
- **Step 7**: Use Maximum-Likelihood (ML) decoding to pick the best codeword on the list.

Rate-Distortion Function

In Step 3 of the proposed algorithm, a numerical computation of the R-D function for a discrete source sequence of n independent but non-identical source components is given by the following theorem.

Theorem: (Factored Blahut-Arimoto algorithm) Consider a single source sequence x^n of n independent but non-identical souce components x_i. Given a parameter s<0, the rate and the distortion for this souce sequence are given by

$$R_s=\sum_{i=1}^n R_{i,s} \quad \text{and} \quad D_s=\sum_{i=1}^n D_{i,s}$$

where the components $R_{i,s}$ and $D_{i,s}$ are computed by the Blahut-Arimoto algorithm with parameter s. This pair of rate-distortion can be achieved by the test-channel input-probability distribution $q_K \triangleq \Pr(\hat{x}^n=K)=\prod_{i=1}^n q_{k_i}$ where the component probability distritubtion $q_{k_i} \triangleq \Pr(\hat{x}_i=k_i)$

Multiple ASD decoding

KV offers a condition for successful ASD decoding in terms of two quantities specified as the score S_M and the cost C_M as follows

ASD decoding threshold: The transmitted codeword will be on the list if $\tau(S_M)>C_M$ where $\tau(S_M)=(a+1)[S_M-a(k-1)/2]$ for any $a\in\mathbb{N}$ such that $a(k-1)<S_M<(a+1)(k-1)$

Definition: (Multiplicity type): For some codeword position, let us assign multiplicity m_j to the j-th most likely symbol for $j=1,2,\ldots,l$. The remaining entries are zeros by default. We call the sequence, $(m_1,m_2,\ldots,m_l)$, the column multiplicity type for "top-l" decoding.

Error and erasure patterns for ASD decoding: Consider a multiplicity assignment scheme (MAS) with z multiplicity types. Let be an erasure pattern where $\hat{x}_i=j$ implies that multiplicity type j is used at column i of the multiplicity matrix M. The definition of a generalized error pattern applies unchanged here.

Definition: The set of allowable multiplicity types for "top-l" decoding with maxmimum multiplicity m is defined to be

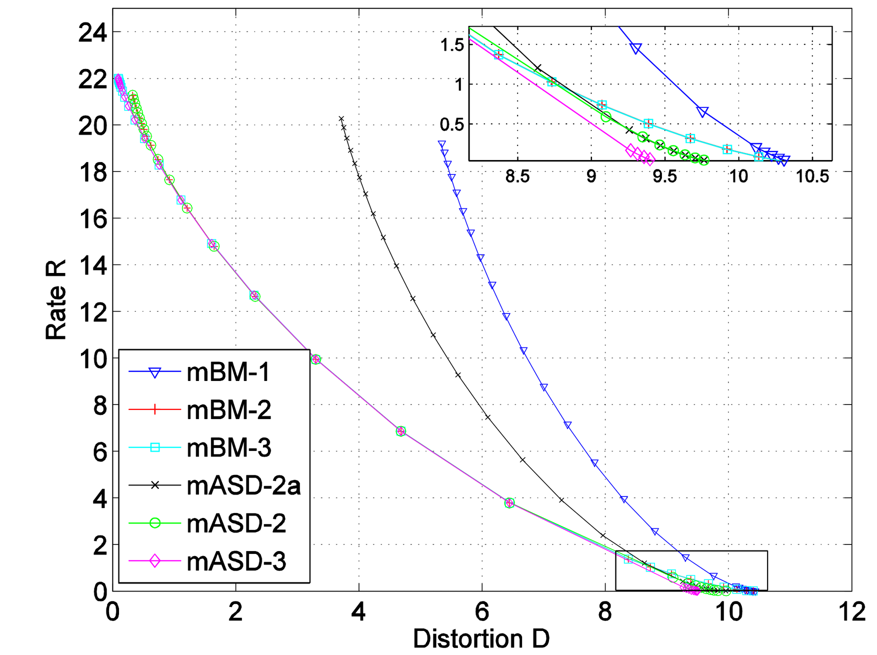
$$A(m,l)=\{(m_1,m_2,\ldots,m_n):\sum_{r=1}^{l} m_r \leq m \text{ and } \sum_{r=1}^{l} m_r(m-m_r)\leq (m+1)(|\{r:m_r\neq 0\}|-1)\min_{r:m_r\neq 0} m_r\}$$

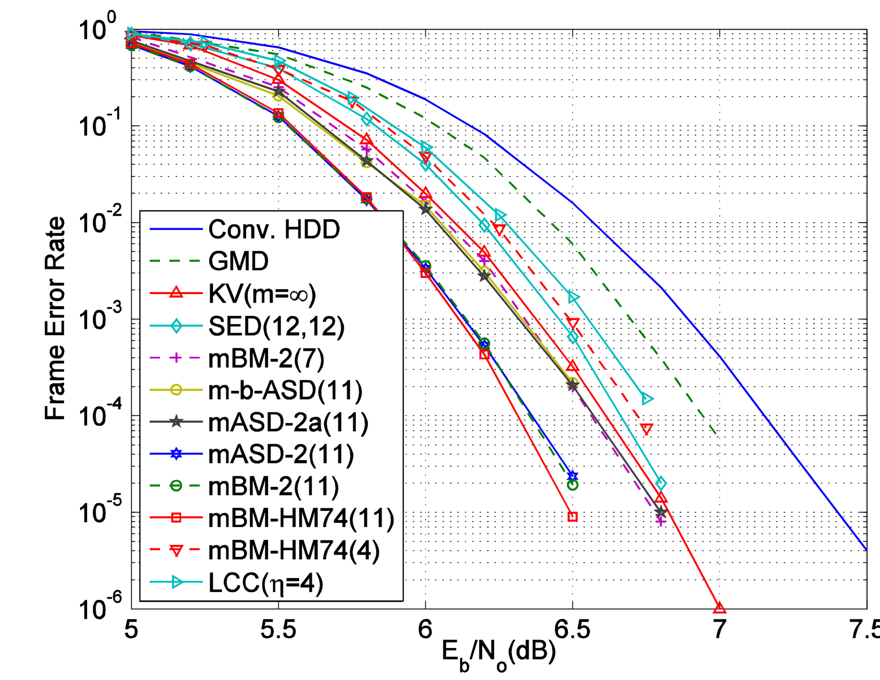
Theorem: Let $z=|A(m,l)|$ be the number of multiplicity types in a MAS for "top-l" decoding with maximum multiplicity m. Let $\delta:\mathbb{Z}_{l+1}\times\mathbb{Z}_{z+1}\backslash\{0\}\to\mathbb{R}_{\geq 0}$ be a letter-by-letter distortion measure defined by $\delta(x,\hat{x})=[\Delta]_{x,\hat{x}}$ where Δ is the $(l+1)\times z$ maxtrix

$$\Delta=\begin{pmatrix} \mu_1 & \mu_2 & \ldots & \mu_z \\ \mu_1-2m_{1,1}/m & \mu_2-2m_{2,1}/m & \ldots & \mu_z-2m_{z,1}/m \\ \mu_1-2m_{1,2}/m & \mu_2-2m_{2,2}/m & \ldots & \mu_z-2m_{z,2}/m \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1-2m_{1,l}/m & \mu_2-2m_{2,l}/m & \ldots & \mu_z-2m_{z,l}/m \end{pmatrix}$$

with $\mu_t=1+\sum_{r=1}^{l}\frac{m_{t,r}(m_{t,r}+1)}{m(m+1)}$. Then the condition for successful ASD decoding of a RS code with rate $\frac{k}{n}\geq\frac{1}{n}+\frac{m(m+3)}{(m+1)(m+2)}$ is equivalent to $d(x^n,\hat{x}^n)<n-k+1$

Simulation Results





Discussion & Future Work

- A R-D approach is proposed as a unified framework to analyzed multiple decoding trials, with various algorithms, of RS in terms of performance and complexity. A connection is made between the complexity and performance (in some asymptotic sense) of these multiple-decoding algorithms and the rate and distortion of an associated R-D problem.
- Covering codes, e.g. perfect codes, can also be combined with R-D approach to mitigate the sub-optimality of random codes when the effective block-length is not large.
- As part of this analysis, we also present numerical computation of the R-D function for sequences of independent but non-identical sources.
- Simulation results shows that our proposed algorithms based on this framework achieve a better performance-versus-complexity trade-off than previously proposed algorithms. One key result is that, for high-rate RS codes, multiple-decoding using the standard BM algorithm is as good as multiple-decoding using more complex ASD algorithms.
- In this work, we only discuss the rate-distortion approach. However, the performance can be further improved by focusing on the rate-distortion error-exponent. The complexity per decoding trial can also be lowered.