Coding and Cryptography
Department of Electrical and Computer Engineering
Technical University of Munich

Lifted Reed-Solomon Codes with Application to Batch Codes

Lukas Holzbaur[1], Rina Polyanskaya[2], Nikita Polyanskii[13], and Ilya Vorobyev[3]

[1]Technical University of Munich, Germany [2]Institute for Information Transmission Problems, Russia [3]Skolkovo Institute of Science and Technology, Russia

Abstract

Guo, Kopparty and Sudan [1] have initiated the study of error-correcting codes derived by lifting of affine-invariant codes. Lifted Reed-Solomon (RS) codes are defined as the evaluation of polynomials in a vector space over a field by requiring their restriction to every line in the space to be a codeword of the RS code. In this paper, we investigate lifted RS codes and discuss their application to batch codes, a notion introduced in the context of private information retrieval and load-balancing in distributed storage systems. First, we improve the estimate of the code rate of lifted RS codes for lifting parameter $m \geq 4$ and large field size. Second, a new explicit construction of batch codes utilizing lifted RS codes is proposed. For some parameter regimes, our codes have a better trade-off between parameters than previously known batch codes.

Lifted Reed-Solomon Codes

- First defined in [1].
- Generalization of Reed-Muller codes.
- Difficult to determine dimension.

Definition: Lifted Reed-Solomon Codes [1]

The m-*dimensional lift of a Reed-Solomon code* (or $[m, d, q]$-*lifted-RS code*) is a code

- over the field $\mathbb{F}_q$,
- of length $N = q^m$, where each codeword position corresponds to an element of $\mathbb{F}_q^m$,
- where the restriction of each codeword to a line L is a codeword of a length q and dimension d RS code.

In other words, it includes the evaluation in all points of $\mathbb{F}_q^m$ of every m-variate polynomial that is equivalent to an 1-variate polynomial of degree smaller d when restricted to a line in the evaluation space.

Example: Reed-Muller Codes vs. Lifted RS Codes

Let $f(X_1, X_2) = X_1^2 X_2^2$. Then the $[2, 3, 4]$-lifted-RS code includes the codeword $\mathbf{c} = (f(a_1, a_2))|_{(a_1,a_2)\in\mathbb{F}_4^2}$ as for every line L, the degree of $f|_L$ is at most $2 < 3 = d$. Indeed, given a line L parameterized as $(w_1 + v_1 T, w_2 + v_2 T)|_{T\in\mathbb{F}_4}$ in $\mathbb{F}_4^2$, we have

$$\begin{aligned} f|_L &= f(v_1 T + w_1, v_2 T + w_2) = (v_1 T + w_1)^2 (v_2 T + w_2)^2 \\ &\overset{(i)}{=} (v_1^2 T^2 + w_1^2)(v_2^2 T^2 + w_2^2) \\ &\overset{(ii)}{=} (v_1^2 w_2^2 + v_2^2 w_1^2) T^2 + v_1^2 v_2^2 T + w_1^2 w_2^2, \end{aligned}$$

where in (i) we used the property $2v = 0$ for any $v \in \mathbb{F}_4$, and (ii) is implied by the fact that $T^4 = T$ in $\mathbb{F}_4[T]$. On the other hand, the 2-variate RM code of order 3 does not contain $\mathbf{c}$ as the degree of f is 4, which is larger than 3.

Main Result

Theorem: Properties of Lifted RS Codes

Code rate and distance: The rate R and relative distance δ of the $[m, q-r, q]$-lifted-RS code are

$$R = 1 - \Theta\left((q/r)^{\log \lambda_m - m}\right), \ \delta \geq \frac{r}{q} \quad \text{as } q \to \infty,$$

where λ_m is the largest eigenvalue of the $m \times m$ matrix

$$A_m := \begin{pmatrix} \binom{m}{\geq 1} & \binom{m}{0} & 0 & 0 & \dots & 0 \\ \binom{m}{\geq 3} & \binom{m}{2} & \binom{m}{1} & \binom{m}{0} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{m}{\geq 2j+1} & \binom{m}{2j} & \binom{m}{2j-1} & \binom{m}{2j-2} & \dots & \binom{m}{2j-m+2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{m}{\geq 2m-1} & \binom{m}{2m-2} & \binom{m}{2m-3} & \binom{m}{2m-4} & \dots & \binom{m}{m} \end{pmatrix}.$$

Availability: Each symbol of a codeword of the $[m, q-r, q]$-lifted-RS code can be reconstructed in q^{m-1} different ways, each of which involves a disjoint set of coordinates of the codeword with cardinality $q - 1$.

Convergence rate $m - \log(\lambda_m)$ for different values of m:

m	2	3	4	5	6
$m - \log(\lambda_m)$	4.1504×10^{-1}	1.4479×10^{-1}	4.1747×10^{-2}	9.6043×10^{-3}	1.7653×10^{-3}

Batch Codes

- First defined in [5].
- Many disjoint recovery sets for each message symbol.
- We construct high rate batch codes.

Definition: Batch Code [5]

A primitive multiset k-batch code of

- length N,
- dimension n,
- each set of k message symbols (possibly with repetition) can be recovered from k disjoint sets of coordinates $R_1, \dots, R_k$.

Example: Simplex Code as Batch Code

Consider the simplex code given by the generator matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

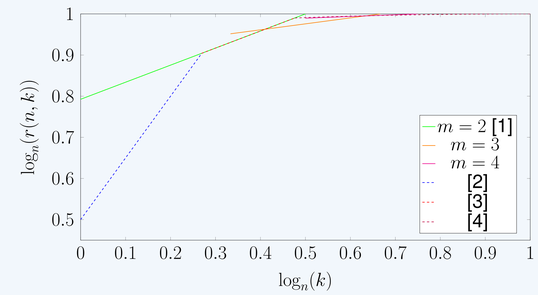
This is a $k = 3$ batch code. For example, the triple (x_1, x_1, x_2) can be recovered from $R_1 = \{1\}$, $R_2 = \{5, 6\}$ and $R_3 = \{2\}$ or the triple (x_3, x_3, x_3) from $R_1 = \{3\}$, $R_2 = \{1, 4\}$, and $R_3 = \{2, 6\}$.

Batch Codes from Lifted RS Codes

Theorem: Parameters of Batch Codes from Lifted RS Codes

Fix integers q, m and $r < q$. The $[m, q-r, q]$-lifted-RS code has the following properties:

- The length of the code is q^m.
- The rate of the code is $1 - \Theta\left((q/r)^{\log \lambda_m - m}\right)$ as $q \to \infty$.
- The code is a k-batch code for $k = q^{m-2} r$.



Lifted Multiplicity Codes

- Multiplicity codes are a generalization of RM codes, which include the evaluation of all derivatives up to a given order [6].
- Many disjoint recovery sets.
- Lifted multiplicity codes combine lifted RS codes and multiplicity codes [7].
- The rate of lifted multiplicity codes can be derived based from our bound on the rate of lifted RS codes [8].

References

[1] A. Guo, S. Kopparty, and M. Sudan, "New affine-invariant codes from lifting," in Proc. 4th Conf. Innov. Theor. Computer Sci., 2013, pp. 529–540.

[2] N. Polyanskii and I. Vorobyev, "Constructions of batch codes via finite geometry," in Proc. IEEE Int. Symp. Inf. Theory, July 2019, pp. 360–364.

[3] R. Polyanskaya and N. Polyanskii, "Batch codes based on lifted multiplicity codes," in Proc. IEEE XVI Int. Symp. Probl. Redund. Inf. Contr. Syst., 2019, pp. 69–74.

[4] H. Asi and E. Yaakobi, "Nearly optimal constructions of PIR and batch codes," IEEE Trans. Inf. Theory, vol. 65, no. 2, pp. 947–964, 2018.

[5] Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai, "Batch codes and their applications," in Proc. 36th Annu. ACM Symp. Theory Comput. (STOC), 2004, pp. 262–271.

[6] S. Kopparty, S. Saraf, and S. Yekhanin, "High-rate codes with sublinear-time decoding," J. Assoc. Comput. Mach., vol. 61, no. 5, p. 28, 2014.

[7] L. Wu, "Revisiting the multiplicity codes: A new class of high-rate locally correctable codes," in Proc. IEEE 53rd Annu. Allerton Conf. Commun. Contr. Comput. (Allerton), 2015, pp. 509–513.

[8] R. Li and M. Wootters, "Lifted multiplicity codes and the disjoint repair group property," in Proc. Approx. Randomiz. Combinat. Optim. Algor. Techn. (APPROX/RANDOM), vol. 145, 2019, pp. 38:1–38:18.