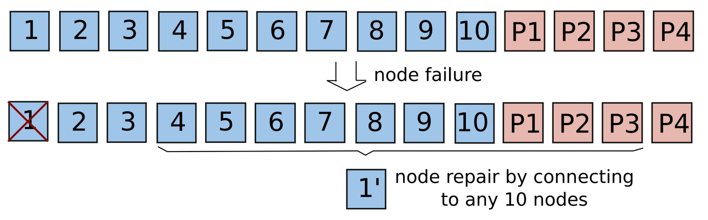
Codes for Efficient Repair: Two Recent Constructions

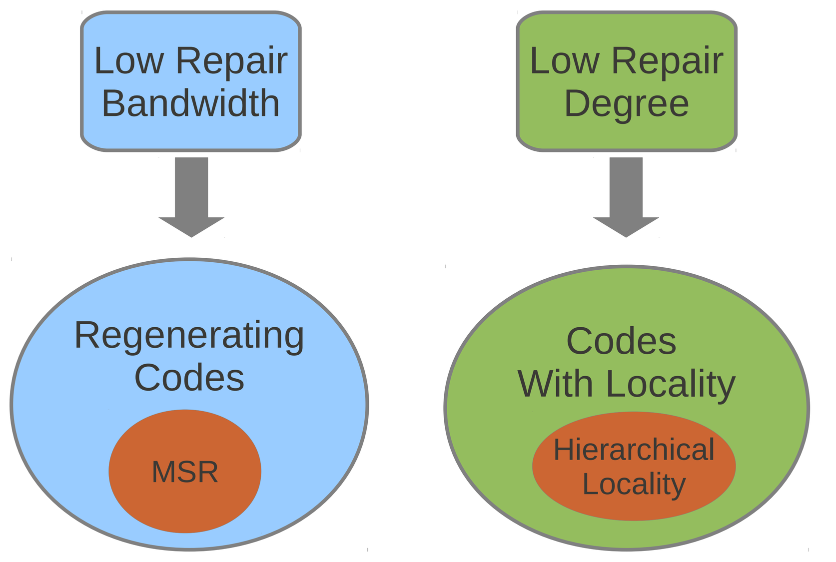
Birenjith Sasidharan, Gaurav Kumar Agarwal and P. Vijay Kumar
Indian Institute of Science

The Need For Efficient Repair

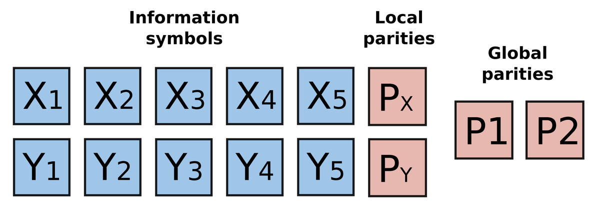
Repair in Reed-Solomon Codes requires large bandwidth and large access!



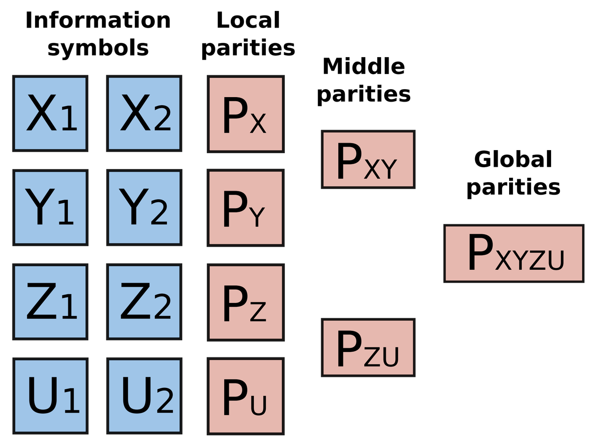
Two approaches to make repair efficient:



Hierarchical Locality



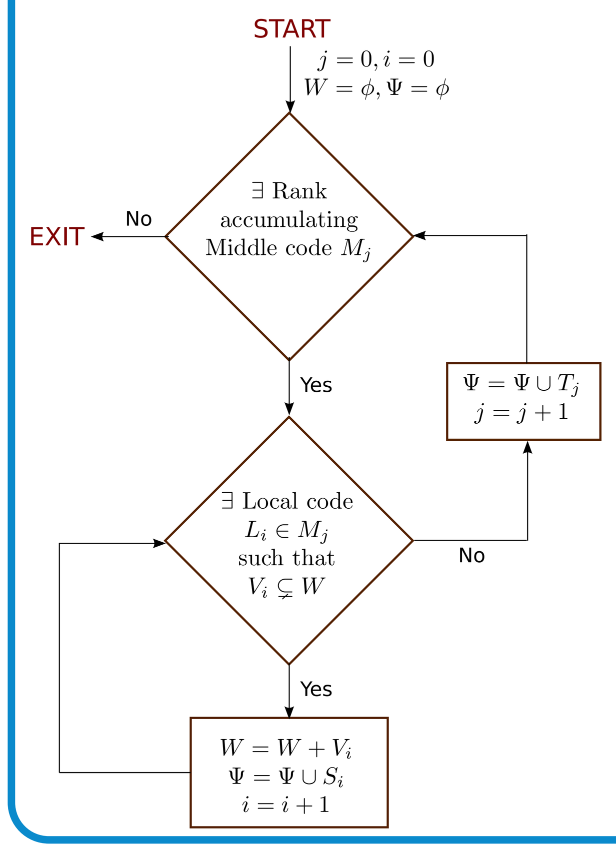
(a) Access 5 nodes to repair a single node-failure.
(b) May access 10 nodes to repair two node-failures.



(c) Hierarchical locality permits efficient repair of two node-failures.
(d) A method to scale to large block-lengths ensuring efficient repair.

Bound On Minimum Distance

$$d \;\leq\; n-k+1-\left(\left\lceil \frac{k}{r_2}\right\rceil -1\right)(\delta_2-1)-\left(\left\lceil \frac{k}{r_1}\right\rceil -1\right)(\delta_1-\delta_2)$$

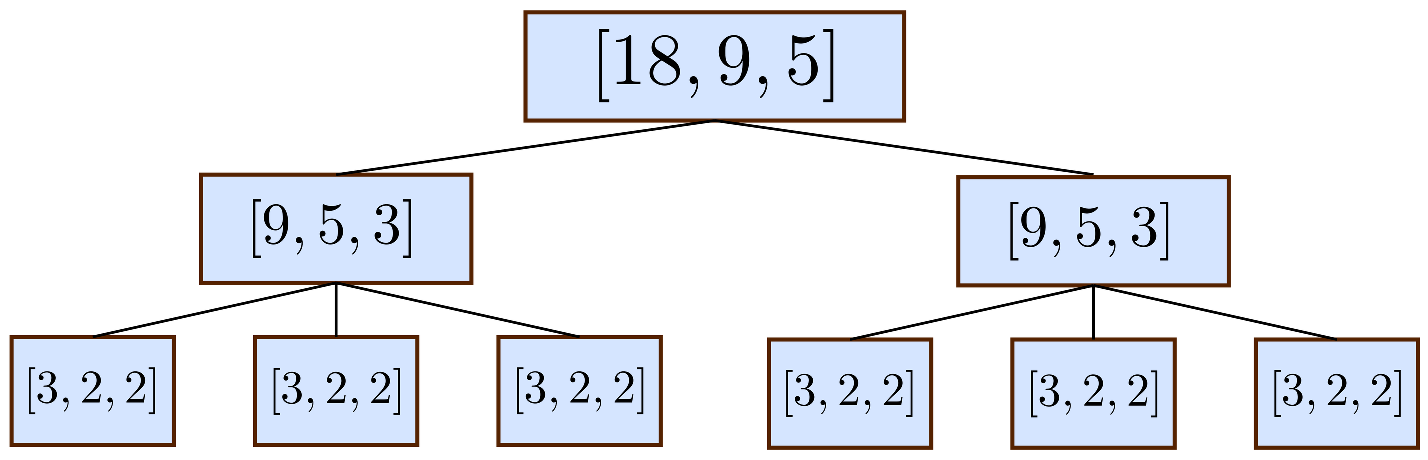


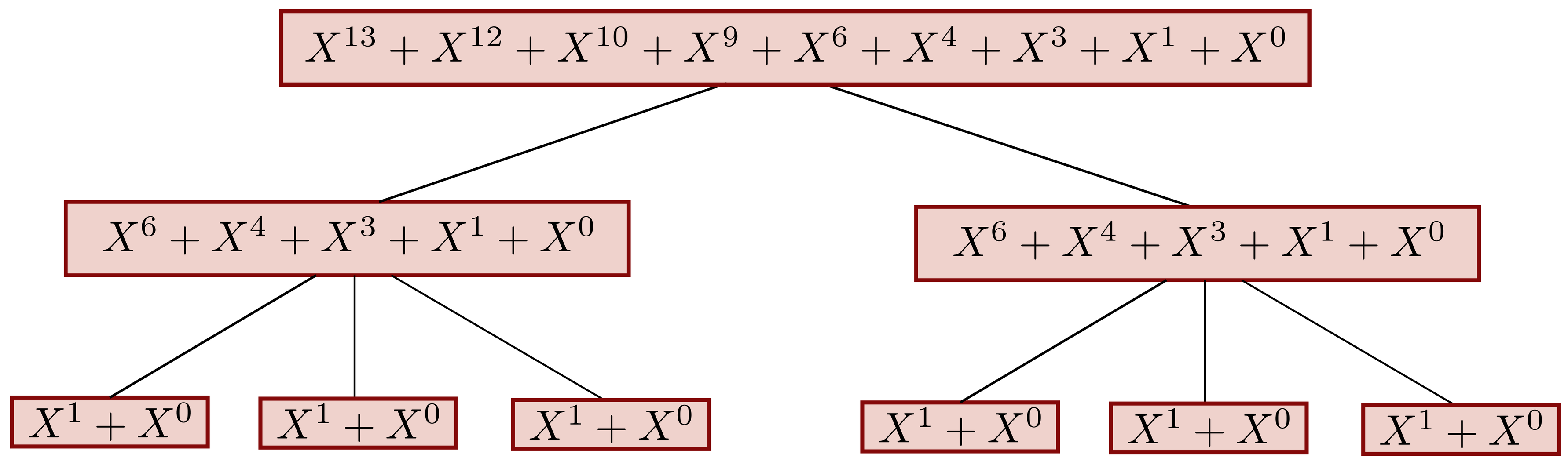
M_j is a middle code of dimension r_1 and minimum distance δ_1.

L_i is a local code of dimension r_2 and minimum distance δ_2.

Construction Of Codes With Hierarchical Locality Having Optimal Minimum Distance

$[18, 9, 5]$**-code with hierarchical locality:**





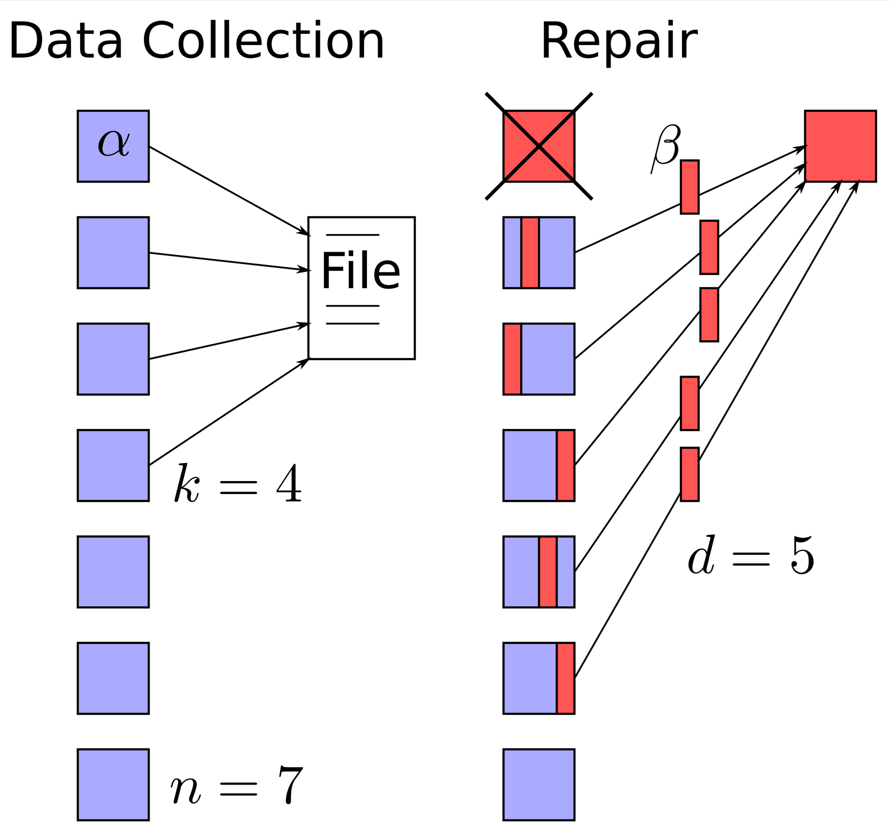
1. The global code, the middle codes and the local codes together form a tree $\mathcal{T}$.
2. For every vertex $v \in \mathcal{T}$, an encoding polynomial $p_v(X)$ is identified such that the encoding polynomials corresponding to any set of siblings are pair-wise coprime.
3. Message polynomial of degree $(r_2-1)=1$ are attached to leaves of $\mathcal{T}$.
4. The encoding polynomials are used to "lift" the message polynomials up to the root of the tree. We use principles of Chinese Remainder Theorem in the process of lifting. The set of all possible values of the polynomial $c(X)$ generated at the root of the tree, forms the message book.
5. At first and zeroth levels of the tree, the message book is shrinked by eliminating some polynomials to adjust the dimension to the desired dimension of the middle code and the global code respectively.
6. The polynomial $c(X)$ at the root of the tree is evaluated at n points to obtain codeword symbols.
7. The construction is extendable to multiple levels.

List of codes in comparison with appropriate codes with locality:

Code W. 2-level Hier. Locality (C_2)	Best Code W. Locality (C_1)	Block Len.	Sto. O/h	Access (1-fail.)	Access (C_2) (2-fail.)	Access (C_1) (2-fail.)
$[18,9,5,(9,5),(3,2)]$	$[18,9,6,(3,2)]$	18	2x	2	5	9
$[102,65,5,(51,33),(3,2)]$	$[102,65,6,(3,2)]$	102	1.57x	2	33	65
$[546,361,5,(273,181),(3,2)]$	$[546,361,6,(3,2)]$	546	1.51x	2	181	361

[1] B. Sasidharan, G. K. Agarwal, P. V. Kumar, "Codes With Hierarchical Locality," arXiv 1501.06683

Minimum Storage Regenerating (MSR) Codes



- A file of size $B=k\alpha$ symbols from $\mathbb{F}_q$ is encoded into $n\alpha$ code symbols. These symbols are stored among n nodes, each storing α symbols.
- The file can be reconstructed downloading symbols from any set of k nodes.
- Thus MSR codes are Maximum-Distance-Separable codes over the vector alphabet $\mathbb{F}_q^{\alpha}$. The rate of the code is $\frac{k}{n}$.
- In addition, a failed node can be repaired by downloading symbols from any set of d helper nodes, each contributing β symbols each.
- The repair bandwidth $\gamma = d\beta = \frac{d\alpha}{d-k+1}$ can be achieved. This is the best possible. When $d=n-1$, $\gamma = \left(\frac{n-1}{n-k}\right)\alpha$.
- The quantity α is called the sub-packetization level.

Savings On Sub-Packetization

First known construction of high-rate MSR code achieving a sub-packetization that is polynomial in k!

	Prod. Matrix Code [4]	Zigzag Code [3]	Present Code [2]
$R=\frac{2}{5}$	$k-1$	$(n-k)^{k+1}$	–
$R=\frac{1}{2}$	$k-1$	$(n-k)^{k+1}$	$(n-k)^2$
$R=\frac{2}{3}$	–	$(n-k)^{k+1}$	$(n-k)^3$
$R=\frac{3}{4}$	–	$(n-k)^{k+1}$	$(n-k)^4$
$R=\frac{4}{5}$	–	$(n-k)^{k+1}$	$(n-k)^5$

[3] I. Tamo, Z. Wang, J. Bruck, "Zigzag Codes: MDS Array Codes With Optimal Rebuilding," T-IT, Mar 2013
[4] K. V. Rashmi, N. B. Shah, P. V. Kumar, "Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction," T-IT, Aug 2011

MSR Code Construction With Low Sub-Packetization

$(n=6, k=4, d=5, \alpha=8)$**-MSR code:**

1. The construction permits rates of the form $R=\frac{t-1}{t}$. In this example, $R=\frac{2}{3}$, i.e., $t=3$.
2. Each row of the codeword array is indexed by a $(t=3)$-tuple (x,y,z) with each coordinate taking values from $\mathbb{F}_2$. Each column is indexed by (ℓ,θ), $\ell=0,1,2$ and $\theta \in \mathbb{F}_2$.
3. The code is defined by $(n-k)\alpha = 2\alpha$ parity-check (p-c) constraints satisfied by the code-symbols. The parity-check constraints that a code-symbol satisfies is determined by its index $(x,y,z,(\ell,\theta))$.
4. The first set of α p-c constraints are the row parity constraints. The second set of α p-c constraints are the jump parity constraints.
5. The repair is by help-by-transfer, i.e., no computation required at helper nodes during repair.

Illustration of Row Parity-Check Constraints

Col Index / Row Index	$\ell=0$, $\theta=0$, Node 1	$\ell=0$, $\theta=1$, Node 2	$\ell=1$, $\theta=0$, Node 3	$\ell=1$, $\theta=1$, Node 4	$\ell=2$, $\theta=0$, Node 5	$\ell=2$, $\theta=1$, Node 6
(000)	A	A	A	A	A	A
(001)	B	B	B	B	B	B
(010)	C	C	C	C	C	C
(011)	D	D	D	D	D	D
(100)						
(101)						
(110)						
(111)						

Illustration of Jump Parity-Check Constraints

Col Index / Row Index	$\ell=0$, $\theta=0$, Node 1	$\ell=0$, $\theta=1$, Node 2	$\ell=1$, $\theta=0$, Node 3	$\ell=1$, $\theta=1$, Node 4	$\ell=2$, $\theta=0$, Node 5	$\ell=2$, $\theta=1$, Node 6
(000)	P	P	P	PR	P	PQ
(001)	Q	Q	Q	QS	PQ	Q
(010)	R	R	PR	R	R	RS
(011)	S	S	QS	S	RS	S
(100)	P					
(101)	Q					
(110)	R					
(111)	S					

[2] B. Sasidharan, G. K. Agarwal, P. V. Kumar, "A High-Rate MSR Code With Polynomial Sub-Packetization Level," arXiv 1501.06662