Ming Hsieh Department of Electrical Engineering

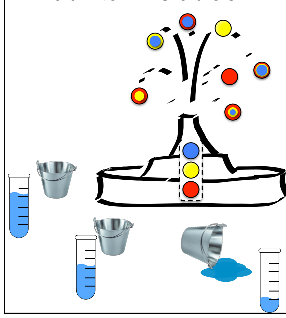
Repairable Fountain Codes

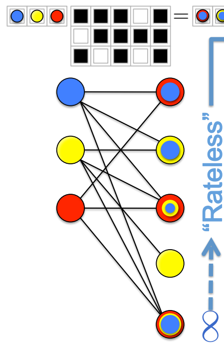
M. Asteris & A. G. Dimakis

USC Viterbi School of Engineering

Introduction

Fountain Codes



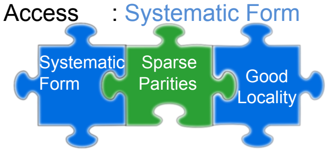


- Distributed encoding.
- Dynamic adjusting of n.

↳ Appealing to Distributed Storage!

Q What else would we like?

- Easy Repair. A lot of metrics.
  Our Focus : Good Locality
- Max Reliability : MDS
- Easy Access : Systematic Form



If parities depend on "few" input symbols, a systematic code, can have good locality.

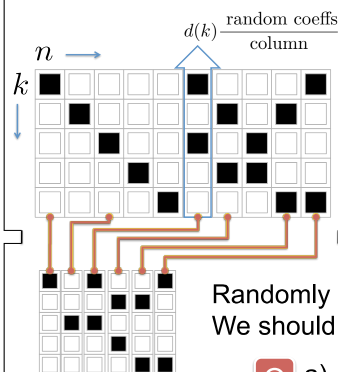
! •But MDS not possible. Why?
•Drop. Instead: $(1+\epsilon)k$ symbols should suffice w.h.p.

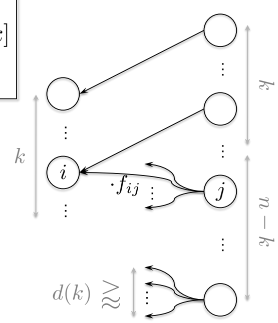
Systematic + Fountain ?

A Simple Randomized Construction

$\forall$ encoded symbol:

- $\mathrm{d(k)}$ input symbols, $i.i.d., \sim \mathcal{U}[k]$
- $\mathrm{d(k)}$ coeffs, $i.i.d., \sim \mathcal{U}[q]$





Randomly select $(1+\epsilon)k$ columns.
We should be able to decode *w.h.p.*

Q a) Can we decode?
b) How small can $d(k)$ be?

Theorem 1

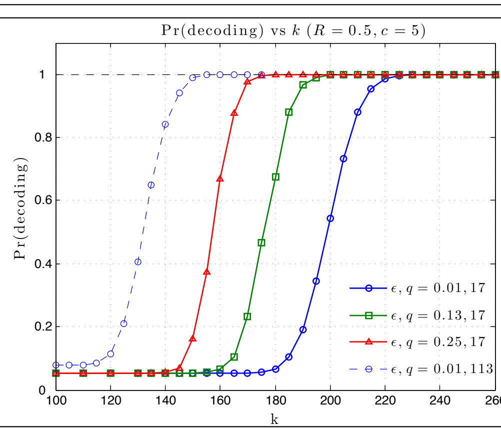
" Decoding $w.h.p.$ $\Rightarrow$ $d(k) = \Omega(\ln(k))$ "

Coupon Collector

Theorem 2

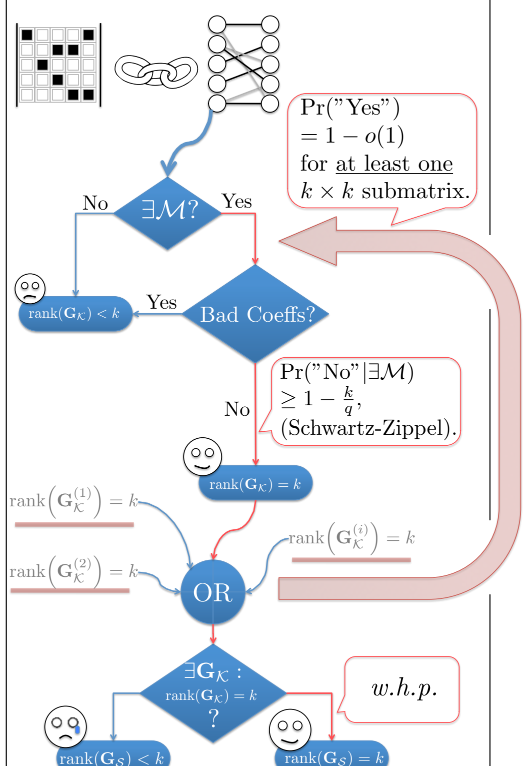
$d(k) = c\ln(k) \quad (c \propto 1/\epsilon)$ $\Rightarrow$ $(1+\epsilon)\,k$ random columns of $\mathbf{G}$ are lin. indep. w.p. k/q close to 1.

Simulation



How to prove?

$\forall$ $k \times k$ submatrix $\mathbf{G}_{\mathcal{K}}$, of our $k \times (1+\epsilon)k$ $\mathbf{G}_{\mathcal{S}}$:



asteris@usc.edu

Ming Hsieh Institute
Ming Hsieh Department of Electrical Engineering