Capacity-Achieving Ensembles for the Binary Erasure Channel with Bounded Complexity

Henry Pfister
Qualcomm, Inc.
CA 92121, USA
hpfister@qualcomm.com

Igal Sason
Technion
Haifa 32000, Israel
Sason@ee.technion.ac.il

Rüdiger Urbanke
EPFL
Lausanne 1015, Switzerland
Rudiger.Urbanke@epfl.ch

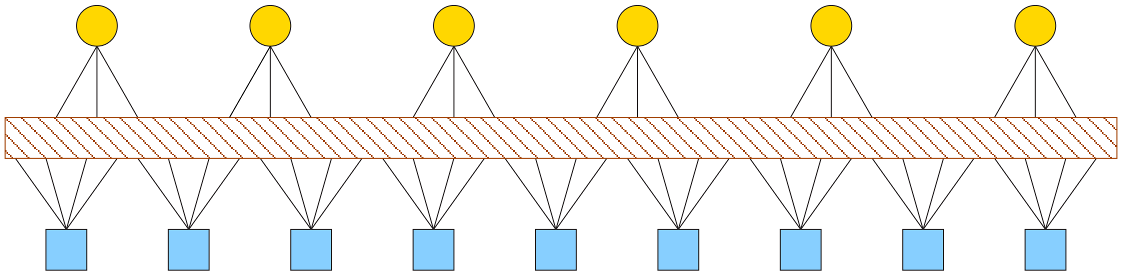
The 23^{rd} IEEE Convention of Electrical and Electronics Engineers in Israel

September 6–7, 2004, Herzlia, Israel

Outline

- Background
  - Codes On Graphs
  - Capacity-Achieving Code Ensembles for the BEC
  - Irregular Repeat Accumulate (IRA) Codes
- Achieving Capacity for the BEC with Bounded Complexity
  - Check-Regular Construction
  - Bit-Regular Construction
- Puncturing Rate Versus Complexity: Information-theoretic bound for punctured codes on graphs. The bound is valid for general memoryless binary-input output-symmetric channels with a refinement for the BEC.
- Simulation Results

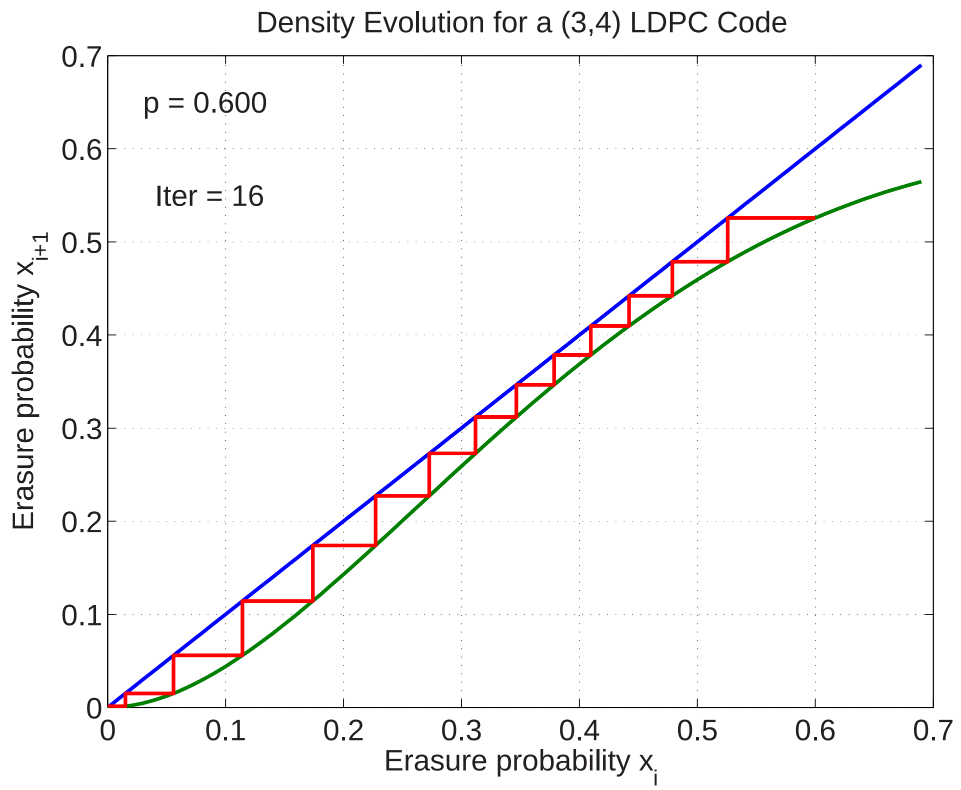
Codes On Graphs (1)



- Low Density Parity Check (LDPC) Codes
  - Message passing iterative (MPI) decoding introduced by Gallager
  - Irregular and capacity-achieving codes for the BEC introduced by Luby et al.
  - An ensemble of irregular codes is defined by the degree distribution (d.d.)
  - Let $\lambda(x) = \sum_{n\geq 2} \lambda_n x^{n-1}$ and $\rho(x) = \sum_{n\geq 2} \rho_n x^{n-1}$, where λ_n and ρ_n are the fraction of edges attached to bit and check nodes of degree n

Codes On Graphs (2)

- Density Evolution (DE)
  - Erasure prob. vs. iteration
  - $x_{i+1} = p\,\lambda\,(1 - \rho(1 - x_i))$
- Successful Decoding Rule
  - $p\,\lambda\,(1 - \rho(1 - x)) < x.$
  - Can rewrite for $\lambda(\cdot)$ given $\rho(\cdot)$ as $\lambda(x) < \frac{1}{p}\left(1 - \rho^{-1}(1 - x)\right)$
- Concentration Theorem (R&U)
  - Performance of MPI decoding converges to DE analysis



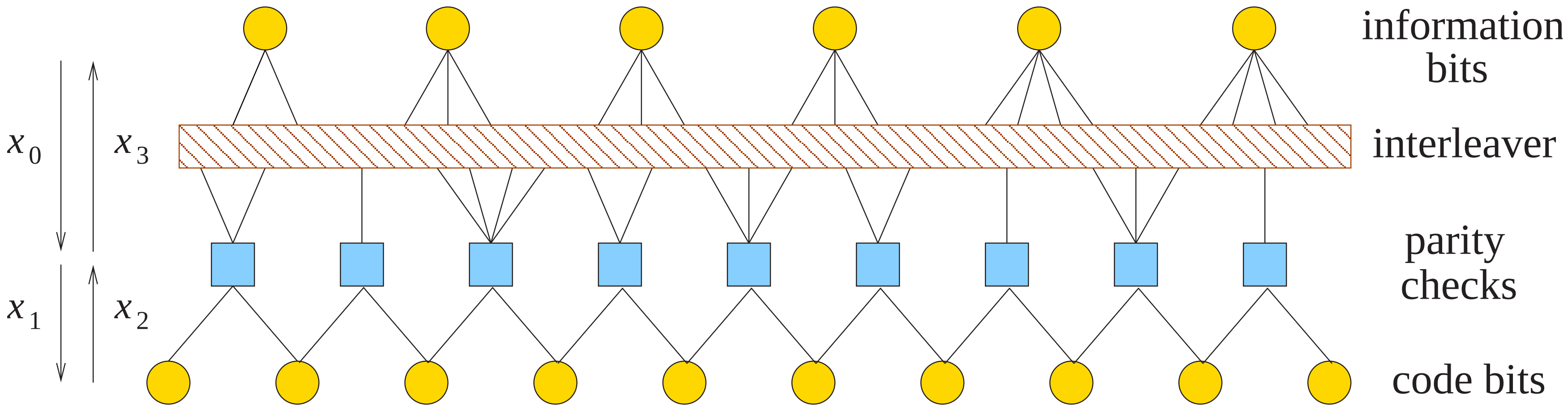
Capacity-Achieving Ensembles (1)

- Sequence of Check-Regular LDPC Codes (Shokrollahi)
  - Check d.d. is regular with degree $k+1$ and given by $\rho^{(k)}(x) = x^k$
  - Bit d.d. given by truncating $\lambda^{(k)}(x) = \frac{1}{p}\left(1 - (1-x)^{1/k}\right)$ so that $\widetilde{\lambda}_k(1) = 1$
- Outline of Proof
  1. DE satisfied with equality before truncation: $p\,\lambda^{(k)}\left(1 - \rho^{(k)}(1-x)\right) = x$
  2. Power series expansion of $\lambda^{(k)}(x)$ is non-negative
  3. Truncated bit d.d. $\widetilde{\lambda}^{(k)}(x)$ satisfies $\widetilde{\lambda}^{(k)}(1) = 1$ and $\widetilde{\lambda}^{(k)}(x) < \lambda^{(k)}(x)$
  4. Decoding condition satisfied: $p\,\widetilde{\lambda}^{(k)}\left(1 - \rho^{(k)}(1-x)\right) < x$ for all $x \in (0,1]$
- Drawback: Achieving $(1-\varepsilon)$ of capacity requires $k \sim \ln \frac{1}{\varepsilon}$

Capacity-Achieving Ensembles (2)

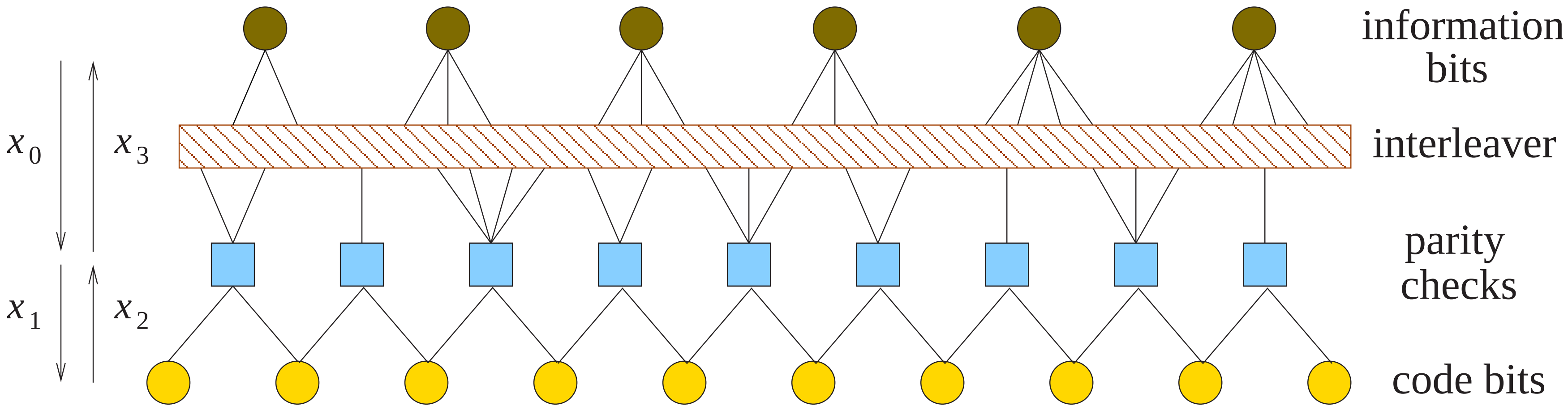
- Complexity to Achieve a Fraction $(1-\varepsilon)$ of BEC Capacity
  - MPI decoding complexity proportional to number of edges in graph
  - Shokrollahi showed number of edges $\sim \ln \frac{1}{\varepsilon}$ for LDPC codes
- Complexity for More General Channels
  - Define minimum complexity of encoding and decoding as $\chi_E(\varepsilon)$ and $\chi_D(\varepsilon)$
  - Based on analysis, Khandekar et al. conjectured: $\chi_D(\varepsilon) = O\left(\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon}\right)$
  - Edges in graph proportional to parity-check matrix density
  - How sparse can the parity-check matrix be in terms of ε ?
  - Sason and Urbanke showed density must grow like $\frac{K_1+K_2 \ln \frac{1}{\varepsilon}}{1-\varepsilon}$ for LDPC codes
  - Question: Can we get better trade-offs with other graphical models?

Systematic IRA Codes



$$
\begin{aligned}
x_1 &= 1 - (1 - x_2)\, R(1 - x_0), \\
x_2 &= p\, x_1, \\
x_3 &= 1 - (1 - x_2)^2\, \rho(1 - x_0), \\
x_0 &= p\, \lambda(x_3)
\end{aligned}
$$

Non-Systematic IRA Codes



$$
\begin{aligned}
x_1 &= 1 - (1 - x_2)\, R(1 - x_0), \\
x_2 &= p\, x_1, \\
x_3 &= 1 - (1 - x_2)^2\, \rho(1 - x_0), \\
x_0 &= \lambda(x_3)
\end{aligned}
$$

IRA Code Comparison

- Systematic IRA Codes (Jin, Khandekar, McEliece)
  - Capacity-achieving d.d. sequences with complexity $\sim \ln \frac{1}{\varepsilon}$ (S&U)
  - DE fixed point condition for $x \in (0,1]$

$$p_0 \, \lambda \left(1 - \left[\frac{1-p}{1-pR(1-x)} \right]^2 \rho(1-x) \right) = x \quad \text{where} \quad R(x) = \frac{\int_0^x \rho(t)\, \mathrm{d}t}{\int_0^1 \rho(t)\, \mathrm{d}t}$$

  - If we assume $\rho(0) = 0$, then this implies that $\lambda(1) = 1/p_0$

- Non-Systematic IRA Codes
  - Analysis above implies that a properly normalized $\lambda(\cdot)$ must have $p_0 = 1$
  - Non-sys IRA codes satisfy the DE equation with $\rho(1) = 1$ and $\lambda(1) = 1$

Non-Systematic IRA Code Issues

- Getting Decoding Started
  - DE update has a fixed point at $x = 1$

$$\lambda\left(1-\left[\frac{1-p}{1-pR(1-x)}\right]^2\rho(1-x)\right) < x$$

- Solutions
  - Systematic bits, degree 1 checks, and/or pilot bits
  - LT codes and Bi-Regular IRA codes (ten Brink) use degree 1 checks
  - Pilots bits are really the same as doping

Bit-Regular Construction

- Ensemble of bit-regular non-sys IRA codes with $\lambda(x) = x^{q-1}$
  - The parity-check d.d. which satisfies the DE equality for this $\lambda(x)$ is

$$\rho(x) = \frac{1 - (1-x)^{\frac{1}{q-1}}}{\left[1 - p\left(1 - qx + (q-1)\left[1 - (1-x)^{\frac{q}{q-1}}\right]\right)\right]^2}$$

  - For $q = 3$, the power series expansion of $\rho(x)$ is non-negative iff $p \in [0, 1/13]$

- Truncating the check d.d. to degree $M(\varepsilon)$ (via degree 1 checks)
  - Let $\rho_\varepsilon(x) = \left(1 - \sum_{n=1}^{M(\varepsilon)} \rho_n\right) + \sum_{n=1}^{M(\varepsilon)} \rho_n x^{n-1}$ where

$$\sum_{n=M(\varepsilon)+1}^{\infty} \rho_n < \frac{\varepsilon}{q(1-p)}$$

Bit-Regular Construction (Cont.)

- In this case, bit erasure probability converges to zero and

$$R^{\mathrm{IRA}} \geq (1-\varepsilon)(1-p)\ .$$

- Complexity (edges per info bit) upper bounded by $q + \frac{2}{(1-p)(1-\varepsilon)}$.

Check-Regular Construction

- Ensemble of check-regular non-sys IRA codes with $\rho(x) = x^2$.
  - The information-bit d.d. which satisfies the DE equality for this $\rho(x)$ is

$$\lambda(x) = 1 + \frac{2p(1-x)^2 \sin\left(\frac{1}{3}\arcsin\left(\sqrt{-\frac{27p(1-x)^{\frac{3}{2}}}{4(1-p)^3}}\right)\right)}{\sqrt{3}\,(1-p)^4\left(-\frac{p(1-x)^{\frac{3}{2}}}{(1-p)^3}\right)^{\frac{3}{2}}}.$$

  - Can show the power series expansion of $\lambda(x)$ is non-negative for $p \in [0, 0.95]$.
- Truncating the bit d.d. to degree $M(\varepsilon)$ (via pilot bits).
  - Treat all information bits with degree $> M(\varepsilon)$ as pilot bits.

Check-Regular Construction (Cont.)

- Effective bit d.d. $\lambda_\varepsilon(x) = \sum_{n=2}^{M(\varepsilon)} \lambda_n x^{n-1}$ where

$$\sum_{n=M(\varepsilon)+1}^{\infty} \frac{\lambda_n}{n} < \frac{(1-p)\varepsilon}{3}$$

- Again, bit erasure probability converges to zero and $R^{\text{IRA}} \geq (1-\varepsilon)(1-p)$
- Complexity (edges per info bit) upper bounded by $\frac{5}{1-p}$ (this bound is tight when the gap to capacity vanishes, i.e., $\varepsilon \to 0$).

$\Rightarrow$ Achieving capacity of the BEC with bounded complexity per information bit.

Puncturing Rate Versus Complexity

for the BEC

- Let $\{\mathcal{C}'_m\}$ be a sequence of binary linear block codes, and let $\{\mathcal{C}_m\}$ be a sequence of codes which is constructed by randomly puncturing information bits from the codes in $\{\mathcal{C}'_m\}$.
  - The communication of the punctured codes takes place over a BEC. The erasure probability of the BEC is p.
  - Assume the sequence $\{\mathcal{C}_m\}$ achieves a fraction $1-\varepsilon$ of the channel capacity with vanishing bit erasure probability.
  - Let P_{pct} designate the puncturing rate of the information bits, and let

$$P_{\text{eff}} \triangleq 1-(1-P_{\text{pct}})(1-p).$$

  - Let $l_{\min}$ designate the minimum number of edges which connect a parity-check node with the nodes of the parity bits.

Puncturing Rate Versus Complexity for the BEC (Cont.)

By information-theoretic tools, we prove that:

With probability 1 w.r.t. the random puncturing patterns, and for an arbitrary representation of the sequence of codes $\{\mathcal{C}'_m\}$ by Tanner graphs, the asymptotic decoding complexity under MPI decoding satisfies

$$\liminf_{m\to\infty} \chi_D(\mathcal{C}_m) \geq \frac{p}{1-p}\left(\frac{\ln\left(\frac{P_{\text{eff}}}{\varepsilon}\right)}{\ln\left(\frac{1}{1-P_{\text{eff}}}\right)} + l_{\min}\right)$$

To achieve capacity with bounded complexity requires $P_{\text{pct}} = 1 - O(\varepsilon)$, i.e., the puncturing rate of the information bits should go to 1.

Puncturing Rate Versus Complexity

for Memoryless Binary-Input Output-Symmetric Channels

- Let $\{\mathcal{C}'_m\}$ be a sequence of binary linear block codes, and let $\{\mathcal{C}_m\}$ be a sequence of codes which is constructed by randomly puncturing information bits from the codes in $\{\mathcal{C}'_m\}$.
  - The communication of the punctured codes takes place over a memoryless binary-input output-symmetric (MBIOS) channel with capacity C bits per channel use.
  - The sequence $\{\mathcal{C}_m\}$ achieves a fraction $1-\varepsilon$ of the channel capacity with vanishing bit error probability.
  - Let P_{pct} designate the puncturing rate of the information bits.

Puncturing Rate Versus Complexity for MBIOS Channels (Cont.)

With probability 1 w.r.t. the random puncturing patterns, and for an arbitrary representation of the sequence of codes $\{\mathcal{C}'_m\}$ by Tanner graphs, the asymptotic decoding complexity per iteration under MPI decoding satisfies

$$\liminf_{m\to\infty} \chi_D(\mathcal{C}_m) \geq \frac{1-C}{2C} \frac{\ln\left(\frac{1}{\varepsilon}\,\frac{1-(1-P_{\text{pct}})C}{2C\ln 2}\right)}{\ln\left(\frac{1}{(1-P_{\text{pct}})(1-2w)}\right)}$$

where

$$w \triangleq \frac{1}{2}\int_{-\infty}^{+\infty} \min\left(f(y), f(-y)\right)\,\mathrm{d}y$$

and $f(y) \triangleq p(y|x=1)$ designates the conditional *pdf* of the channel, given the input is $x = 1$.

Puncturing Rate Versus Complexity (Cont.)

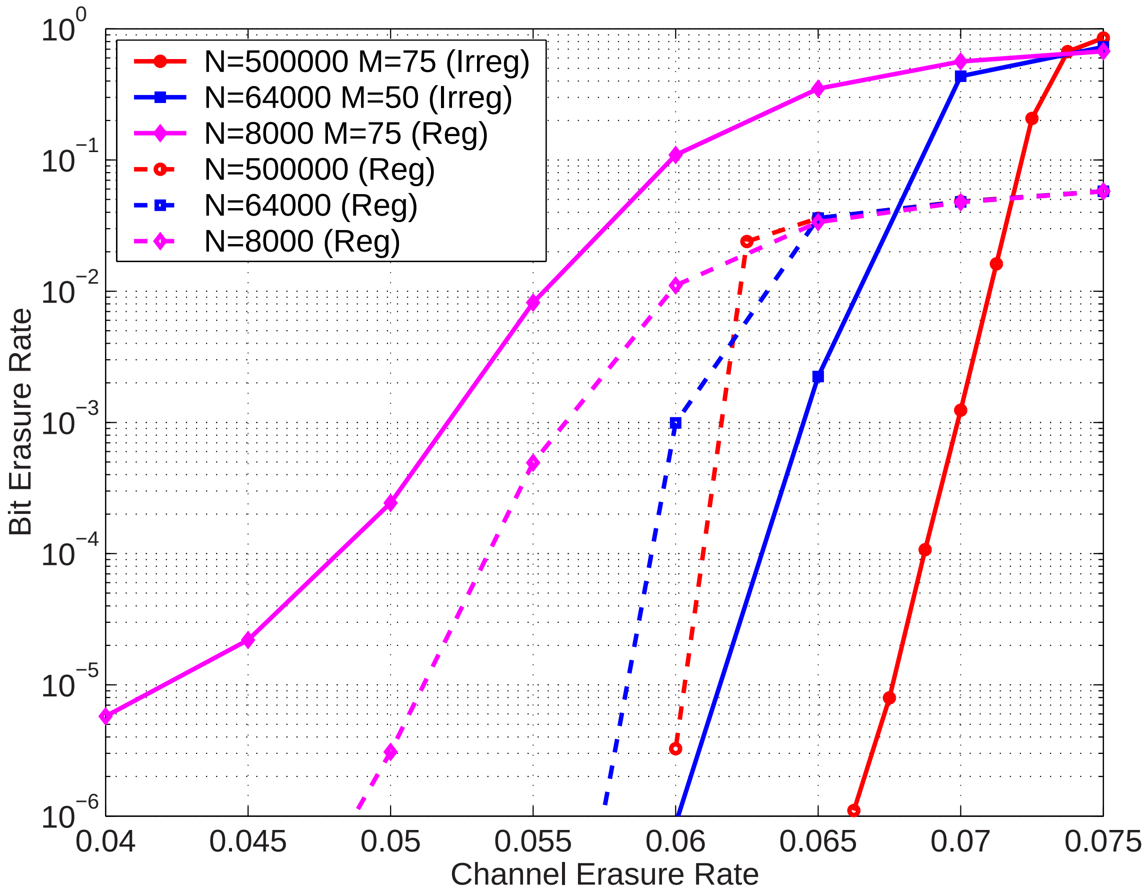
- We assume *random puncturing* of information bits.
  For achieving capacity of an arbitrary MBIOS channel with bounded complexity per iteration, the puncturing rate of the information bits should go to 1.
- The lower bounds on the decoding complexity in the last two theorems clearly also hold if we require vanishing block error/ erasure probability.
- The lower bound on the decoding complexity that we get for the BEC is at least twice larger than the lower bound for the BEC which we get from the theorem which applies to general MBIOS channels.

Simulation Setup

- Code Design
  - Pick one d.d. and compute the other via power series truncation
  - Bit-regular truncation: Set $\rho_n = 0$ for $n > M$ and renormalize
    Then add some systematic bits to get decoding started (e.g., 100-200)
  - Check-regular truncation: Force bits of degree $> M$ to be pilot bits
  - Vary ”design” p to get the desired code rate
- Code Construction
  - Quantize the algebraic d.d. to integers based on block length
  - First, construct by randomly matching bit and check edges
  - Next, swap ”bad” edges randomly to remove mult. edges and 4-cycles

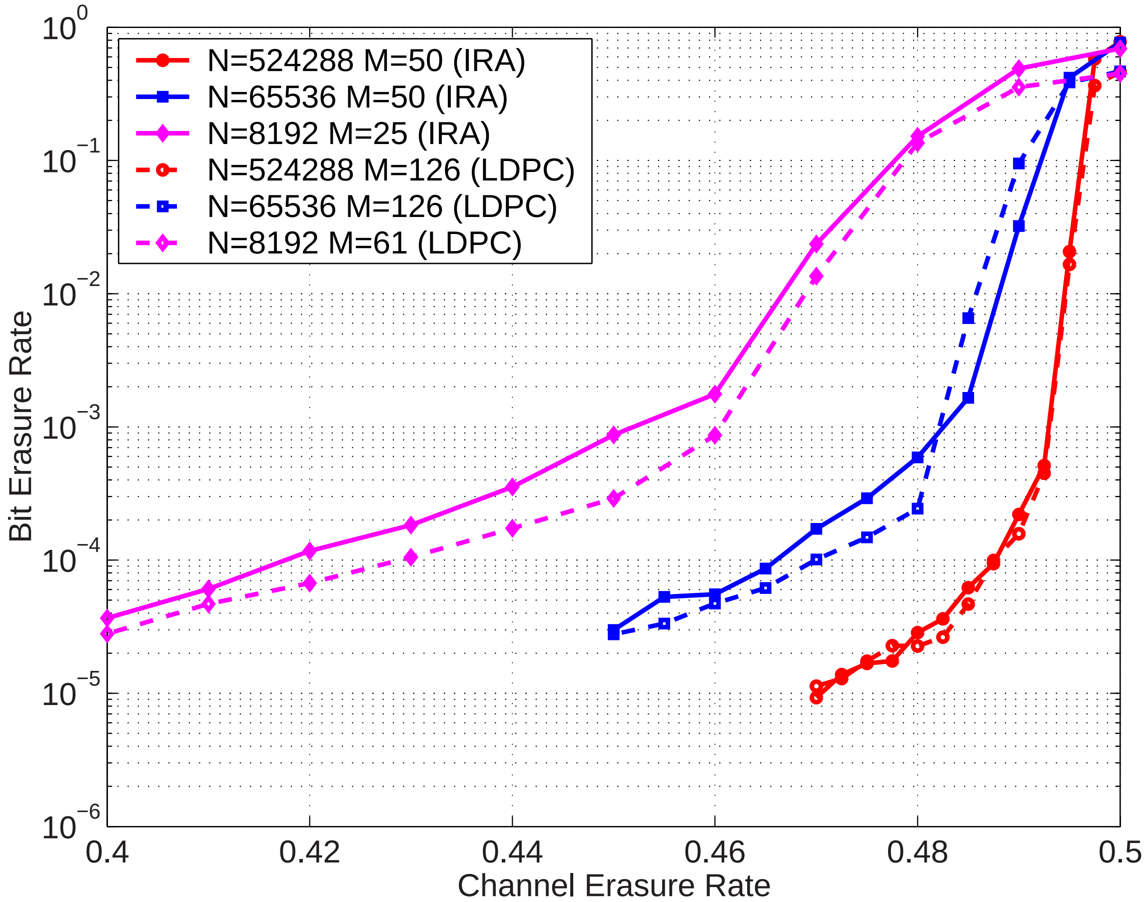
Simulation Results: Bit-Regular

- Design Details (Rate=0.925)
  - "Irreg": Best of $M = 25, 50, 75$
  - "Reg": Sys-IRA d.d. (3,37)
- Observations
  - No apparent error floor
  - Number of sys bits required doesn't grow with length
  - Rate loss is small for large N and large for small N



Simulation Results: Check-Regular

- Design Details (Rate=0.5)
  - "IRA": Best of $M = 25, 50, 75$
  - "LDPC": Check-reg $q = 8, 9$
- Observations
  - Performance very similar
  - Error floor in both ensembles due to marginal stability



Summary

- Previous constructions for the BEC have provably unbounded complexity which grows at least like $O(\ln \frac{1}{\varepsilon})$
- Our main results are:
  - Showing the existence of capacity-achieving codes for the BEC with bounded complexity. We show that under message-passing iterative (MPI) decoding, this new bounded complexity result is only possible because we allow a sufficient number of state nodes in the Tanner graph representing a code ensemble. The state nodes in the Tanner graph of the examined IRA ensembles are introduced by puncturing all the information bits.
  - Derivation of an information-theoretic lower bound on the decoding complexity of randomly punctured codes on graphs. The bound holds for every memoryless binary-input output-symmetric channel with a refinement for the BEC.

Summary (Cont.)

- The central point in this paper is that by allowing state nodes in the Tanner graph, one may obtain a significantly better tradeoff between performance and complexity as the gap to capacity vanishes.

- Under MPI decoding and the random puncturing assumption, it follows from the information-theoretic bound that a necessary condition to achieve the capacity with bounded complexity (or with bounded complexity per iteration for a general MBIOS channel) is that the puncturing rate of the information bits goes to one.

Summary (Cont.)

- For *fixed* complexity, the new codes eventually (for n large enough) outperform any code proposed to date. On the other hand, the *convergence speed* to the ultimate performance limit happens to be quite slow, so for small to moderate block lengths, the new codes are not necessarily record breaking.
- Further research into the construction of codes with bounded complexity is likely to produce codes with better performance for small to moderate block lengths.

Full Paper

- The full paper is submitted to *IEEE Transactions on Information Theory*.
- It is at `http://www.ee.technion.ac.il/people/sason/PSU.pdf`