united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

SMR.1347 - 7

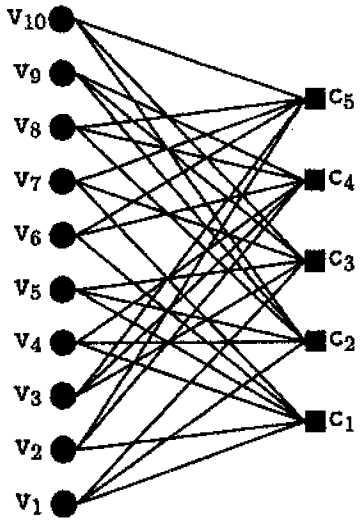
WORKSHOP ON STATISTICAL PHYSICS AND CAPACITY-APPROACHING CODES

AN OVERVIEW OF ITERATIVE CODING

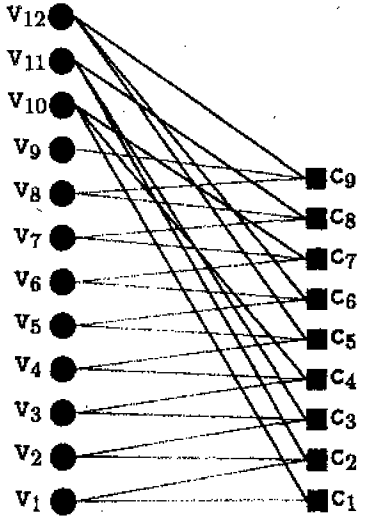
R. URBANKE
Ecole Polytechnique Federal de Lausanne
(E.P.F.L.) DSC/LTHC
CH-1015 Lausanne, Switzerland

Please note: These are preliminary notes intended for internal distribution only.

Code Ensembles



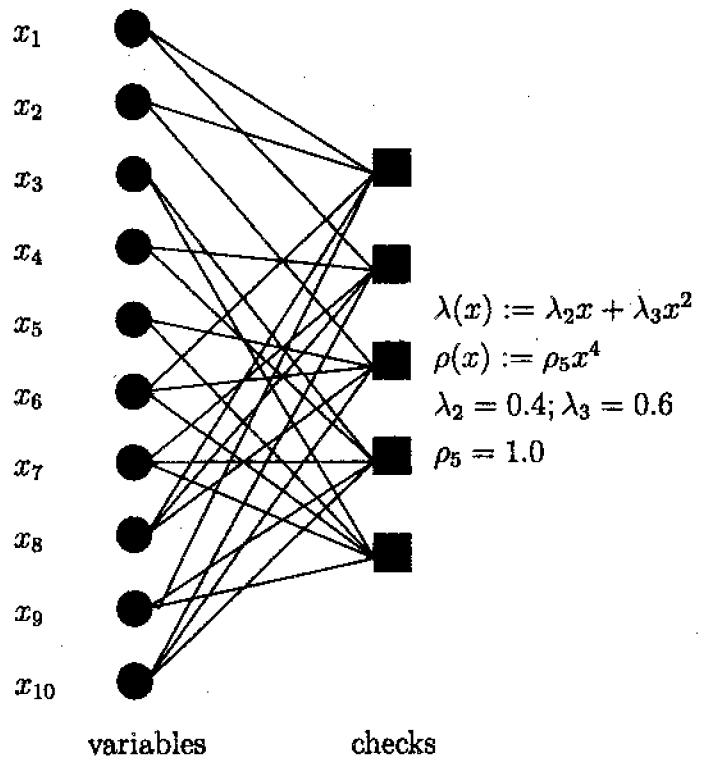
LDPC [Gallager]



RA [Divsalar et. al]

Other Ensembles: Turbo Codes [Berrou et. al.], Woven Codes [Johannesson et. al.], Concatenated Tree Codes [Ping et. al.], MN Codes [Kanter and Saad], ...

The Ensemble $\mathcal{C}(n, \lambda, \rho)$



[LDPC Codes-Gallager][Tanner-Codes on Graphs][McKay, Neal, Wiberg-Rediscovered][Ensembles-Luby et. al.]

The Big Picture

Let $P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G}, \sigma)$ denote the expected *bit error* probability if G is used to transmit over a given binary-input memoryless output-symmetric channel and if the received word is decoded iteratively by a message passing decoder. (fixed # of rounds)

[Concentration Around Ensemble Average] $\exists \alpha(\delta) > 0$ such that

$$\mathsf{Pr}\{|P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G}, \sigma) - \mathbb{E}_{\mathcal{C}(n,\lambda,\rho)}\left[P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G}, \sigma)\right]| > \delta\} \leq e^{-\alpha n}.$$

[Convergence to Cycle-Free Case] $\exists \beta > 0$ such that

$$|\mathbb{E}_{\mathcal{C}(n,\lambda,\rho)}\left[P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G}, \sigma)\right] - \mathbb{E}_{\mathcal{C}(\infty,\lambda,\rho)}\left[P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G}, \sigma)\right]| \leq \frac{\beta}{n}.$$

[Luby et. al-discrete setting; Belief Propagation][Richardson, Urbanke-general case]

Concentration

Idea: Use vertex exposure technique on the bipartite graph to show that important quantities are tightly concentrated.

Theorem 1 [Azuma's Inequality] Let $Z_0, Z_1, \ldots$ be a martingale sequence such that for each $k \geq 1$,

$$|Z_k - Z_{k-1}| \leq \alpha_k,$$

where the constant α_k may depend on k. Then, for all $i \geq 1$ and any $\lambda > 0$

$$\Pr\{|Z_i - Z_0| \geq \lambda\} \leq 2e^{-\frac{\lambda^2}{2\sum_{k=1}^{i} \alpha_k^2}}.$$

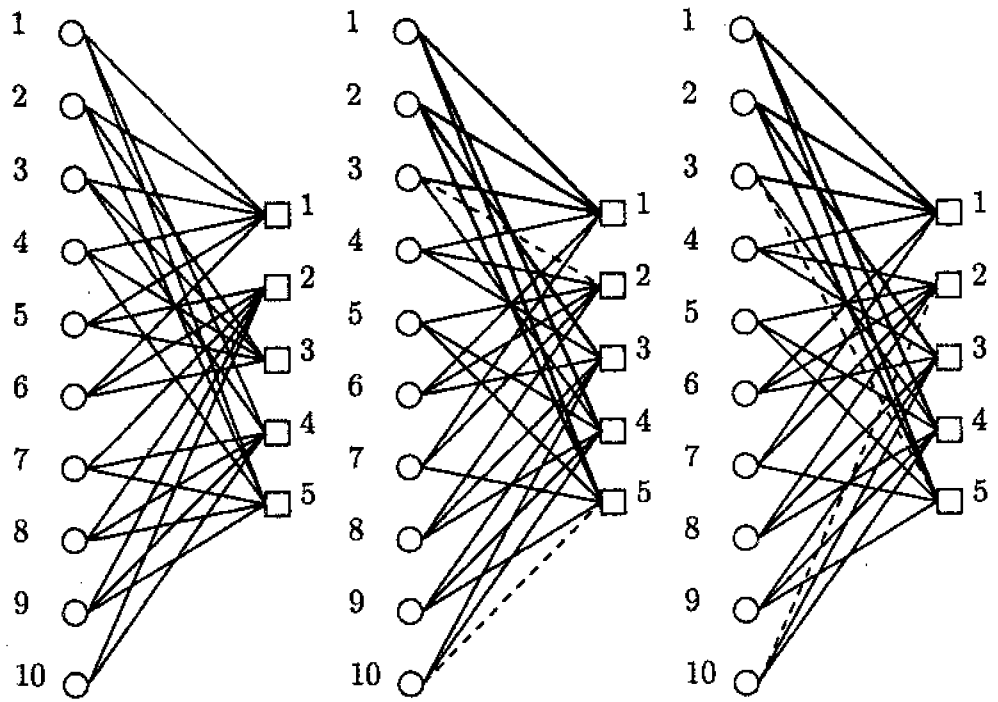


Figure 1: Graph G from the ensemble $\mathcal{C}(10, x^2, x^5)$. The two dashed lines correspond to the two edges whose end points are switched.

Concentration Around Ensemble Average

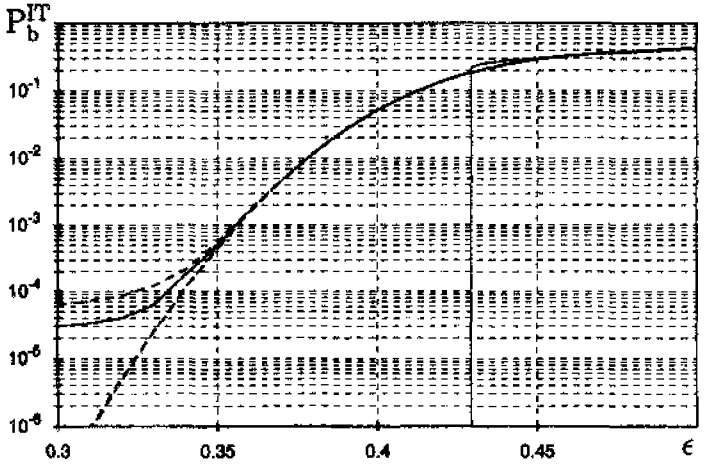


Figure 2: Concentration of the bit erasure probability $P_b^{IT}(G, \epsilon)$ for specific instances $G \in \mathcal{C}(512, x^2, x^5)$ around the ensemble average $\mathbb{E}_{\mathcal{C}(512,x^2,x^5)}\left[P_b^{IT}(G, \epsilon)\right]$ (blue curve). Also shown is the performance of the cycle-free case, $\mathbb{E}_{\mathcal{C}([redacted],x^2,x^5)}\left[P_b^{IT}(G, \epsilon)\right]$ (red curve).

Finite Length Analysis

$$\mathbb{E}_{\mathcal{C}(n,\lambda,\rho)}\left[P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G},\sigma)\right]$$

Finite Length Analysis for the BEC

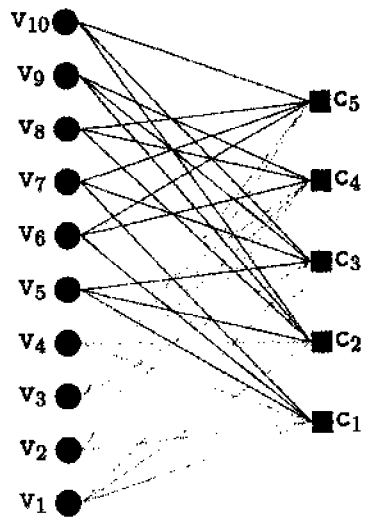


Figure 3: The set $\{v_1, v_2, v_3, v_4\}$ is a stopping set.

Definition 1 [Stopping Sets] A *stopping set* $\mathcal{S}$ is a subset of $\mathcal{V}$, the set of variable nodes, such that all neighbors of $\mathcal{S}$ are connected to $\mathcal{S}$ at least twice.

Theorem Let $P_{\mathrm{b}}^{\mathrm{IT}}(\mathtt{G},\epsilon)$ denote the *bit erasure* probability when transmitting over a BEC with erasure probability ϵ using a code $\mathtt{G}$, $\mathtt{G} \in \mathcal{C}(n, x^{\mathtt{l}-1}, x^{\mathtt{r}-1})$, and a belief propagation decoder. In a similar manner, let $P_{\mathrm{B}}^{\mathrm{IT}}(\mathtt{G},\epsilon)$ denote the *block erasure* error probability. Define the functions $T(v,c,d)$, $N(v,c,d)$, $M(v,c,d)$ and $O(v,s,c,d)$ by the recursions

$$\begin{aligned}
T(v,c,d) &:= \binom{d+c\mathtt{r}}{v\mathtt{l}}(v\mathtt{l})!, \\
N(v,c,d) &:= T(v,c,d) - M(v,c,d), \\
M(v,c,d) &:= \sum_s \binom{v}{s} O(v,s,c,d), \\
O(v,s,c,d) &:= \sum_k \binom{c}{k} \mathrm{coef}(((1+x)^{\mathtt{r}} - 1 - \mathtt{r}x)^k (1+x)^d, x^{s\mathtt{l}})(s\mathtt{l})! \\
&\qquad N(v-s, c-k, d+k\mathtt{r}-s\mathtt{l}),
\end{aligned}$$

and the boundary condition

$$O(v,s,c,d) = 0 \text{ if } s \le 0 \text{ or } v\mathtt{l} > c\mathtt{r} + d.$$

Then

$$\begin{aligned}
\mathbb{E}_{\mathtt{G}}\left[P_{\mathrm{b}}^{\mathrm{IT}}(\mathtt{G},\epsilon)\right] &= \sum_e \binom{n}{e} \epsilon^e (\bar{\epsilon})^{n-e} \sum_v \frac{v}{n} \frac{\binom{e}{v} O(e,v,n\frac{\mathtt{l}}{\mathtt{r}},0)}{(e\mathtt{l})!\binom{n\mathtt{l}}{e\mathtt{l}}}, \\
\mathbb{E}_{\mathtt{G}}\left[P_{\mathrm{B}}^{\mathrm{IT}}(\mathtt{G},\epsilon)\right] &= \sum_e \binom{n}{e} \epsilon^e (\bar{\epsilon})^{n-e} \sum_v \frac{\binom{e}{v} O(e,v,n\frac{\mathtt{l}}{\mathtt{r}},0)}{(e\mathtt{l})!\binom{n\mathtt{l}}{e\mathtt{l}}}.
\end{aligned}$$

[Di,Proietti,Telatar,Richardson,Urbanke]

Finite Length Analysis for the BEC

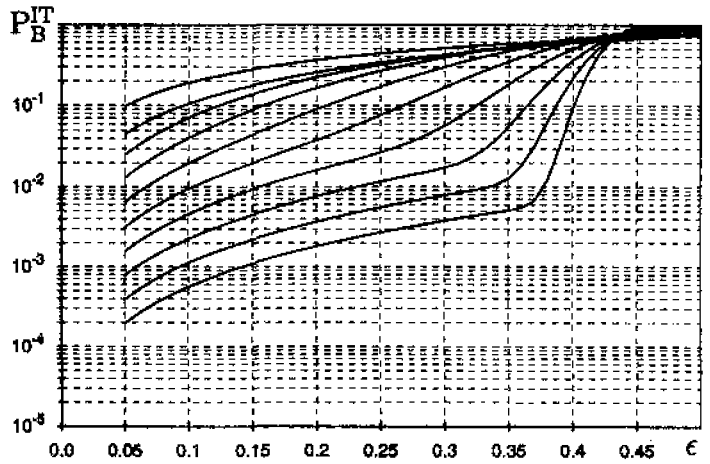


Figure 4: $\mathbb{E}_{\mathcal{C}(n,\mathcal{A}^3,\mathcal{A}^5)}\left[\mathrm{P}_{\mathrm{B}}^{\mathrm{IT}}(\mathtt{G},\epsilon)\right]$ as a function of ϵ for $n = 2^i$, $i \in [10]$.

Finite Length Analysis Open Questions

1. Expurgated Ensembles?
2. Find simpler expressions for $\mathbb{E}_{\mathcal{C}(n,x^2,x^5)}\left[\mathsf{P}_{\mathsf{b}}^{\mathsf{IT}}(\mathtt{G},\epsilon)\right]$ and $\mathbb{E}_{\mathcal{C}(n,x^2,x^6)}\left[\mathsf{P}_{\mathsf{B}}^{\mathsf{IT}}(\mathtt{G},\epsilon)\right]$.
3. Irregular Case.
4. Optimization.
5. Other Ensembles.
6. Other Channels.
7. Distribution of number of iterations [McKay, Kanter]

Asymptotic Analysis

$$\mathbb{E}_{\mathcal{C}(\infty,\lambda,\rho)}\left[P_{\mathsf{b}}^{\mathsf{IT}}(\mathsf{G},\sigma)\right]$$

Gallager's Bound

Lemma 1 Arbitrarily reliable transmission over the BSC with parameter ϵ using a linear code with "maximal degree" d is not possibly at rates above $1 - h(\epsilon) - \frac{h(\epsilon)}{2\ln 2}(1-2\epsilon)^{2d}$.

Proof Outline:

$\underline{X}$	transmitted word
$\underline{Y}$	received word
$\underline{Y} = (\underline{U}, \underline{V})$	$\underline{U}$ is an *information set*
$\underline{S}^T = \underline{HY}^T$	syndrome

$$
\begin{aligned}
H(\underline{X}) &= nR \\
H(\underline{Y}|\underline{X}) &= nh(\epsilon) \\
H(\underline{Y}) &= H(\underline{U},\underline{S}) = H(\underline{U}) + H(\underline{S}|\underline{U}) \leq H(\underline{U}) + H(\underline{S}) \\
&= \sum_{i=1}^{k} H(\underline{U}_i|\underline{U}_1,\cdots,\underline{U}_{i-1}) + \sum_{i=1}^{n-k} H(\underline{S}_i|\underline{S}_1,\cdots,\underline{S}_{i-1}) \\
&\leq \sum_{i=1}^{k} H(\underline{U}_i) + \sum_{i=1}^{n-k} H(\underline{S}_i) \\
&\leq k + (n-k)h\left(\frac{1+(1-2\epsilon)^d}{2}\right) \\
\frac{1}{n}H(\underline{U}|\underline{Y}) &= \frac{1}{n}\left[H(\underline{X}) - H(\underline{Y}) + H(\underline{Y}|\underline{X})\right] \geq -h(\epsilon) + (1-R)h\left(\frac{1+(1-2\epsilon)^d}{2}\right)
\end{aligned}
$$

Asymptotic Analysis for the Binary Erasure Channel

$$\epsilon_i := \mathbb{E}[\text{fraction of erasure messages passed in } i\text{-th round}]$$

$$\epsilon_i = \epsilon_0 \lambda(1 - \rho(1 - \epsilon_{i-1}))$$

Asymptotic Analysis - Capacity Achieving Codes for the BEC

Given: BEC with parameter ϵ.
Task: For any $\delta > 0$ find degree distribution pair $(\lambda(x), \rho(x))$ such that

$$r(\lambda, \rho) := 1 - \frac{\int \rho}{\int \lambda} \geq 1 - \epsilon - \delta.$$

and such that

$$\epsilon\lambda(1 - \rho(1 - x)) > x, \forall 0 < x < 1.$$

Solution: [Heavy Tail Poisson Sequence]

$$\begin{aligned} \lambda_\alpha(x) &:= -\frac{1}{\alpha}\ln(1-x) = \frac{1}{\alpha}\sum_{i=1}^{\infty}\frac{x^i}{i}, \text{ and} \\ \rho_\alpha(x) &:= e^{\alpha(x-1)} = e^{-\alpha}\sum_{i=0}^{\infty}\frac{\alpha^i x^i}{i!}. \end{aligned}$$

[Luby, Mitzenmacher, Shokrollahi, Spielman, Stehmann - LDPC codes][McEliece et. al. for RA codes]

Asymptotic Analysis – General Case

Density Evolution: Determine the distribution of the messages passed in the i-th iteration.

[concept, discrete setting-Gallager][general case, efficient algorithm for BP-Richardson, Shokrollahi, Urbanke]

Different Representations

Variable	Density	Representation	Domain
$\mathtt{v}$	ν	$\log \frac{p_0}{p_1}$	$\mathtt{v} \in [-\infty, \infty]$
$\mathtt{d}$	$\eth$	$p_0 - p_1$	$\mathtt{d} \in [-1, 1]$
$\mathtt{a}$	α	$\lvert p_0 - p_1 \rvert$	$\mathtt{a} \in [0, 1]$
$\mathtt{c}$	γ	$-\log \lvert p_0 - p_1 \rvert$	$\mathtt{c} \in [0, \infty]$
$\mathtt{s}$	σ	$\mathsf{sign}\,(p_0 - p_1)$	$\mathtt{s} \in \{0, 1\}$

Density Evolution for Belief Propagation Algorithm

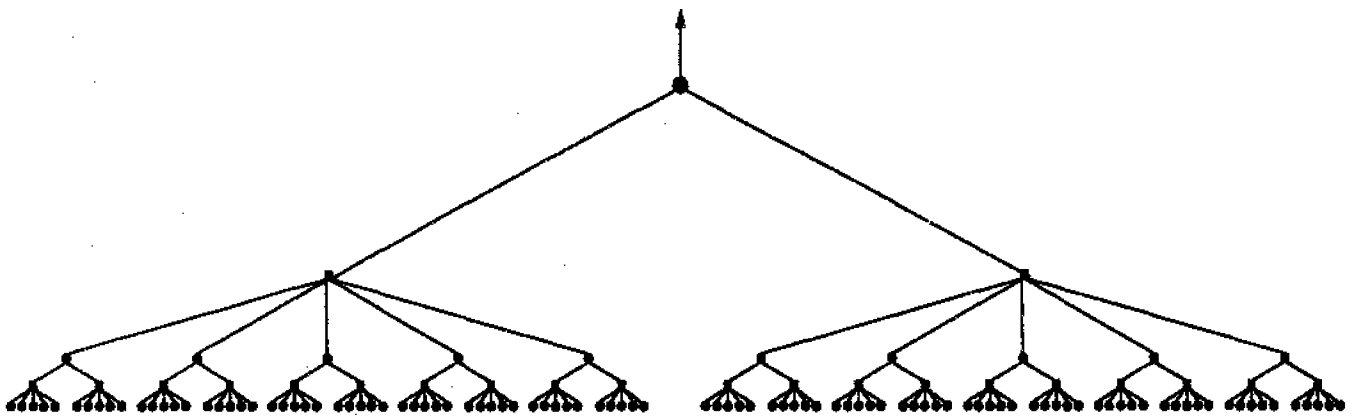
	Variable node side	Check node side
Representation	$\mathbf{v} = \log \frac{p_0}{p_1}$	$(\mathbf{c}, \mathbf{s}) = (-\log \lvert p_0 - p_1 \rvert, \mathrm{sign}(p_0 - p_1))$
Message evolution	$m_{out} = m_{rec} + \sum m_{in}$	$(\mathbf{c}, \mathbf{s})_{out} = \sum (\mathbf{c}, \mathbf{s})_{in}$
Density evolution	$\nu_{out} = (\otimes \nu_{in}) \otimes \nu_{rec}$	$(\gamma, \sigma)_{out} = \otimes (\gamma, \sigma)_{in}$

But density evolution is NOT limited to BP algorithm! Any *message passing* algorithm can be analysed in terms of the evolution of its message density.

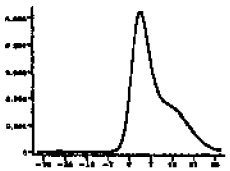
Density Evolution

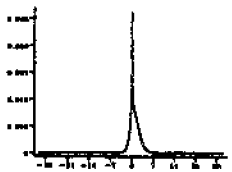
outgoing message

How Does the Expected Value Evolve?



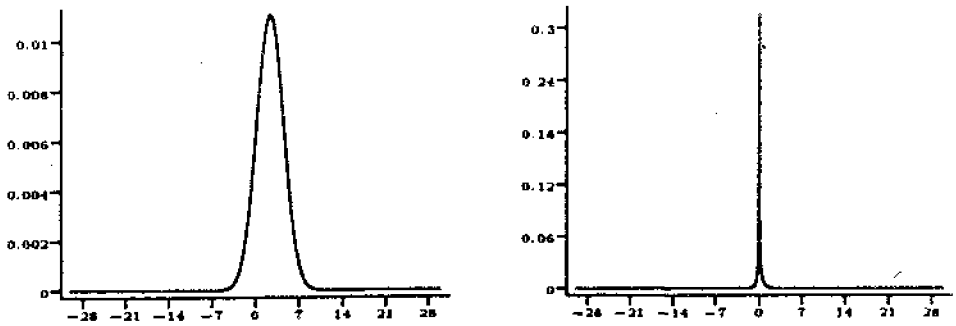
messages received from channel



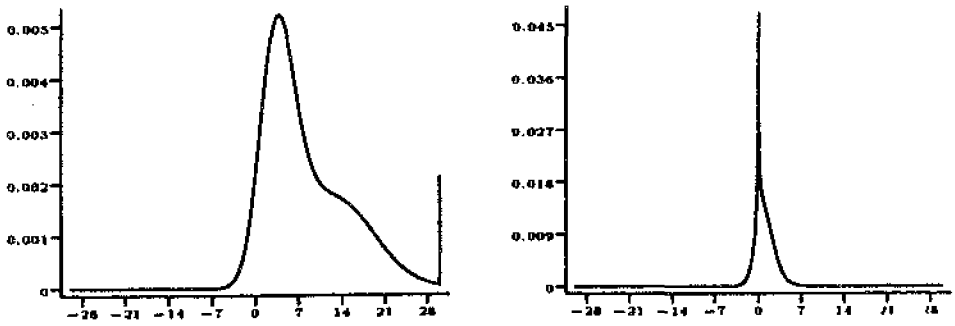


Density Evolution Example

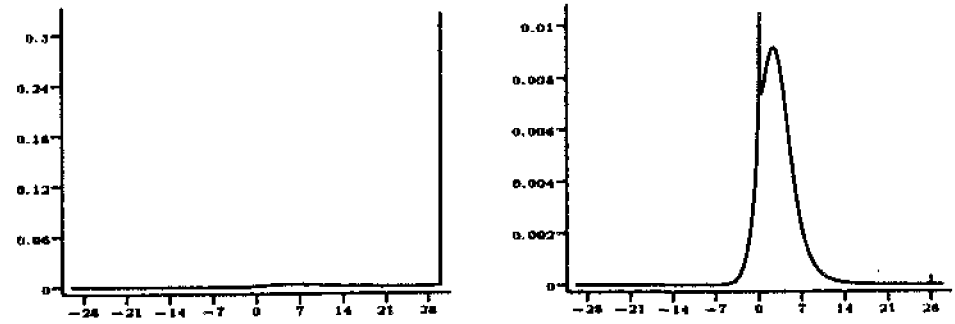
Iter. 1



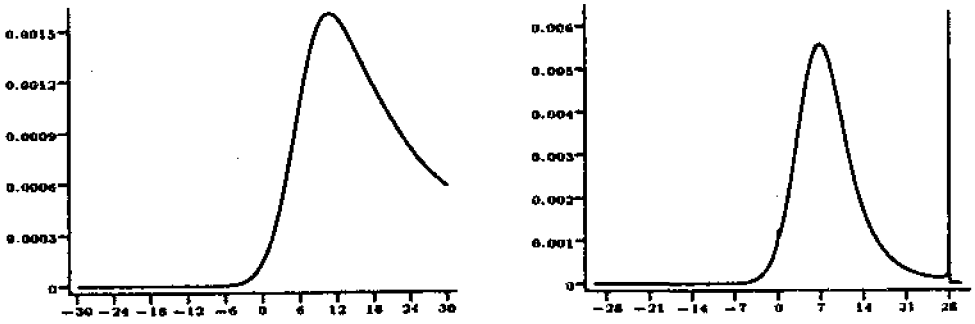
Iter. 15



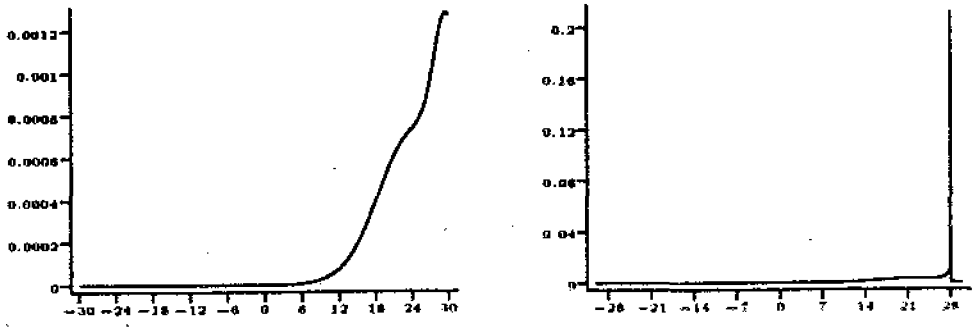
Iter. 30



Iter. 50



Iter. 100



Gallager Algorithm A

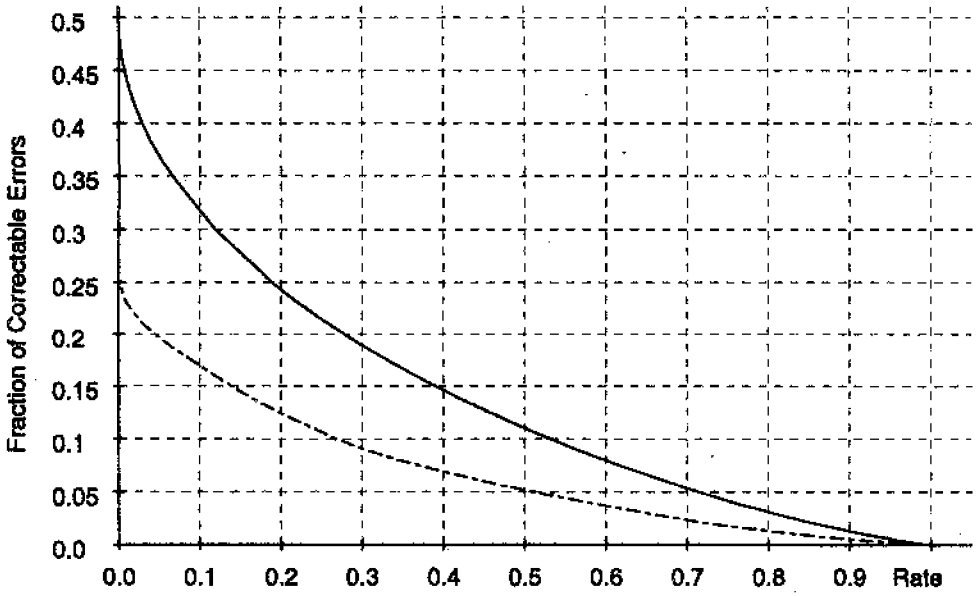


Figure 5: The solid curve corresponds to the capacity formula $r = 1 - h(x_0^*)$ whereas the dashed curve corresponds to the optimal right-concentrated degree distribution. For rates above 0.4, right concentrated degree distribution pairs are optimal.

[Bazzi, Richardson, Urbanke]

Suboptimal Decoders

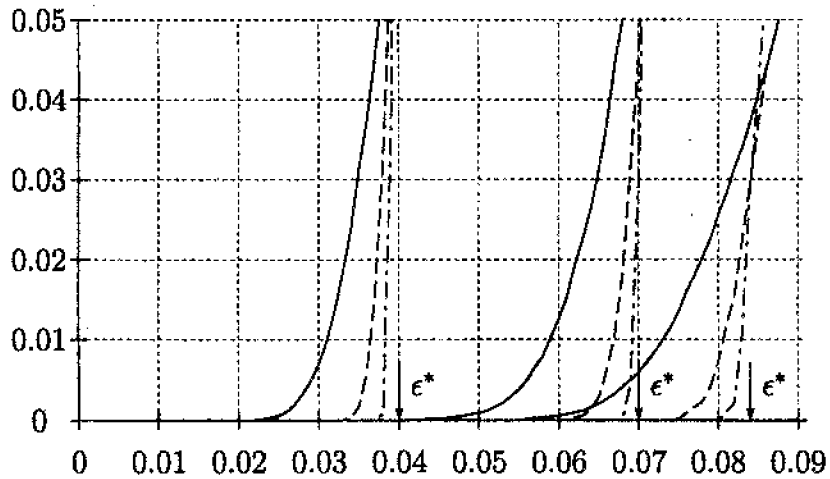


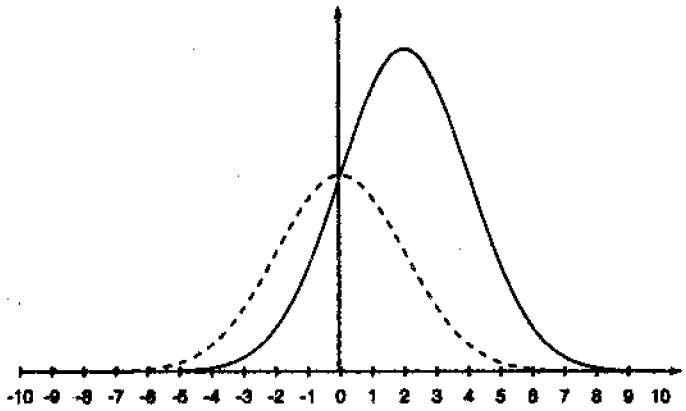
Figure 6: Bit error probability versus parameter ϵ for the $(3, 6)$-regular ensemble transmitting over the BSC channel for three decoding algorithms. The solid curves correspond to a codeword length of 1000, whereas the dashed and the dotted-dashed curves correspond to codeword lengths of 10000 and 100000, respectively.

Symmetry Condition for BP Densities

Assumptions: **binary signalling (extension possible), output symmetric channel, messages are in log-likelihood ratio form;**

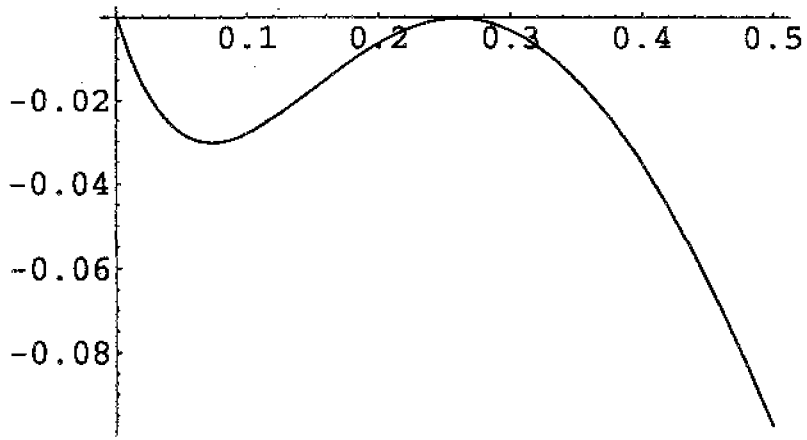
The densities which occur at any iteration fulfill the symmetry condition

$$f(x)e^{-x/2} = f(-x)e^{x/2}$$



Asymptotic Analysis - Stability for BEC

Progress Per Iteration	$\epsilon^*\lambda(1-\rho(1-\epsilon))-\epsilon$
Progress around Zero	$(\epsilon^*\lambda'(0)\rho'(1)-1)\epsilon$
Stability Condition	$\lambda'(0)\rho'(1) < \frac{1}{\epsilon^*}$



[Luby et. al.]

Asymptotic Analysis - General Stability Condition

Theorem 2 [General Stability Condition] Assume we are given a degree distribution pair (λ, ρ) and a symmetric density P_0. For $\ell \geq 1$ define $\mathrm{P}_\ell := \mathrm{P}_0 \otimes l(\Gamma^{-1}((\Gamma(\mathrm{P}_{\ell-1}))))$. Let $r := -\ln\left(\int_{\mathbb{R}} \mathrm{P}_0(x)^{-\frac{x}{2}}\right)$ and assume that $\int_{\mathbb{R}} e^{sx} \mathrm{d}(\int \mathrm{P}_0)(x) < \infty$ for all s in some neighborhood of zero.

[Necessity] If $\lambda'(0)\rho'(1) > e^r$ then there exists a constant $\xi = \xi(\lambda, \rho, \mathrm{P}_0)$, $\xi > 0$, such that for all $\ell \in \mathbb{N}$, $\mathrm{P_e}(\mathrm{P}_\ell) > \xi$.

[Sufficiency] If $\lambda'(0)\rho'(1) < e^r$ then there exists a constant $\xi = \xi(\lambda, \rho, \mathrm{P}_0)$, $\xi > 0$, such that if for some $\ell \in \mathbb{N}$, $\mathrm{P_e}(\mathrm{P}_\ell(\mathrm{P}_0)) \leq \xi$ then $\mathrm{P_e}(\mathrm{P}_\ell)$ converges to zero as ℓ tends to infinity.

[Richardson, Shokrollahi, Urbanke]

The Final Reward...

[Chung, Forney, Richardson, Urbanke]

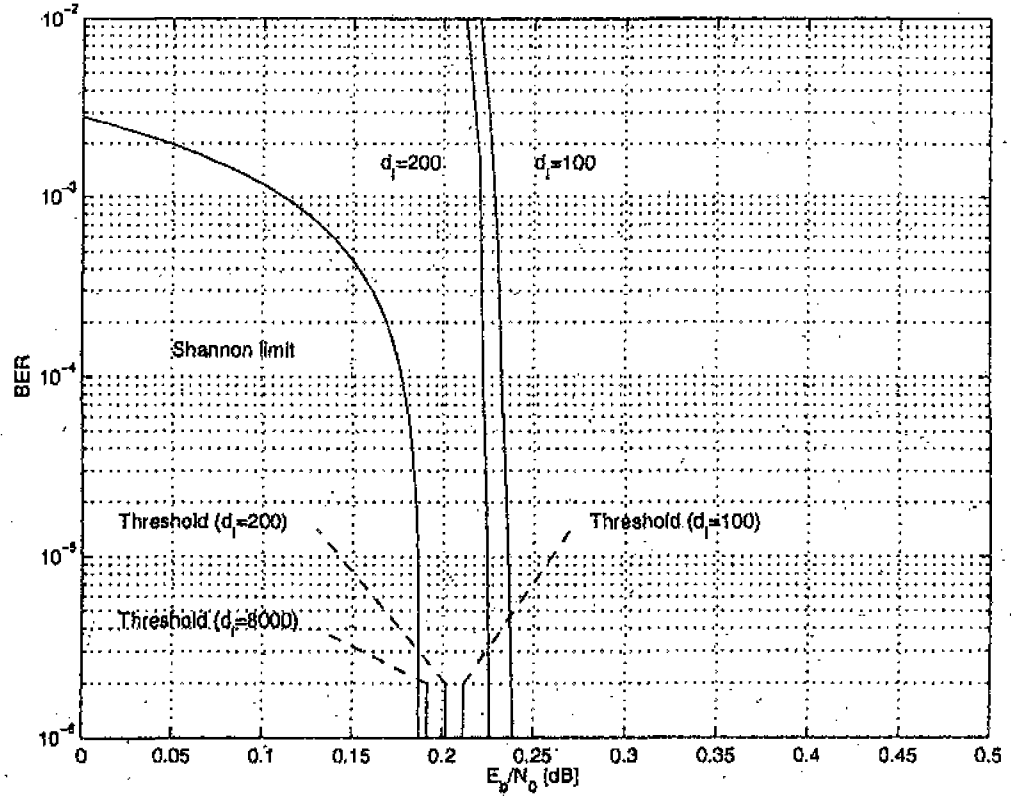


Fig. 2. Simulation results for $d_l = 100, 200$ codes using a block length of 10^7.

Asymptotic Analysis - Linear Time Encodable Codes

Theorem 3 [Codes with Linear Encoding Complexity] Let (λ, ρ) be a degree distribution pair satisfying $\alpha^*(\rho, \lambda) = 1$, with minimum right degree at least three and satisfying the *strict* inequality $\lambda'(0)\rho'(1) > 1$. Let $\mathtt{G}$ be chosen at random from the ensemble $\mathcal{C}(n, \lambda, \rho))$. Then $\mathtt{G}$ is encodable in linear time with probability at least $1 - bc^{\sqrt{n}}$ for some positive constants b and c, where $c < 1$.

Conclusion: All optimized codes are linear time encodable.

[Richardson, Urbanke]

Asymptotic Analysis - Linear Time Encodable Codes

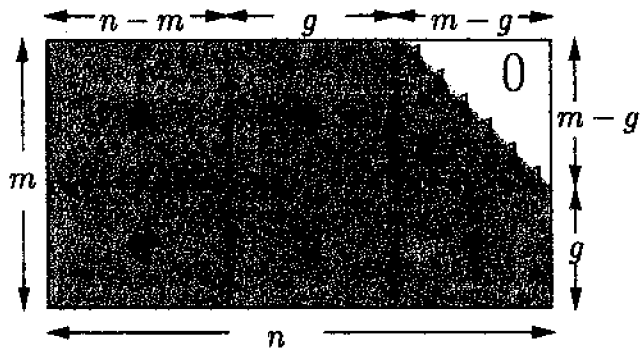
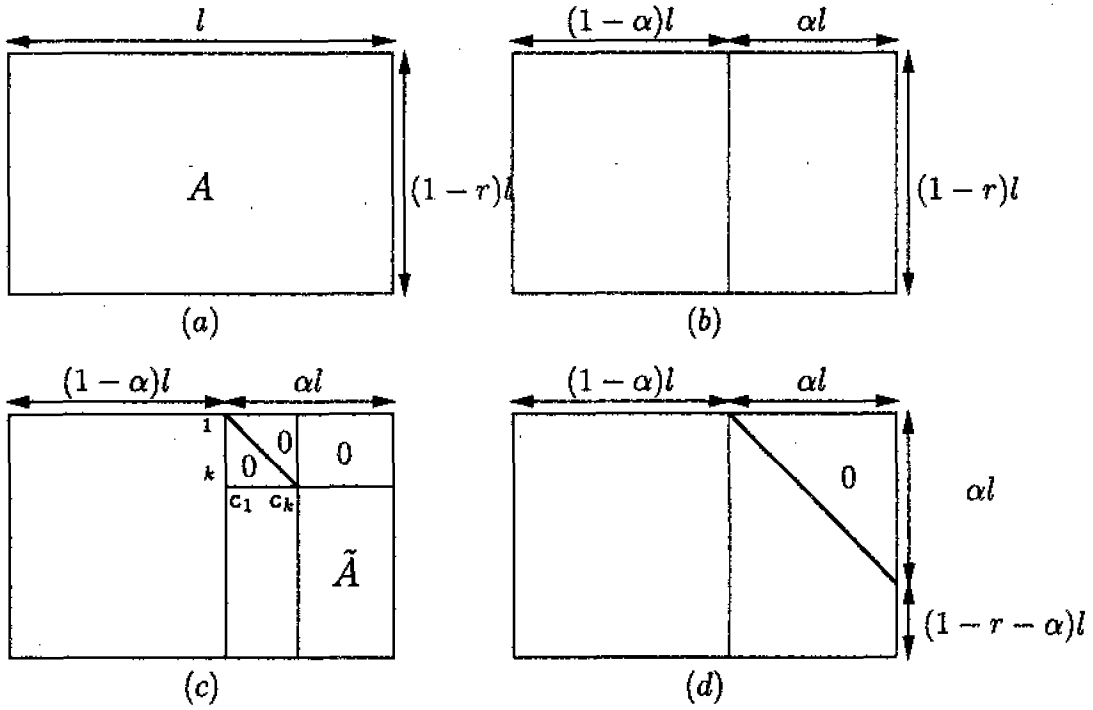


Figure 7: The parity-check matrix in approximate lower triangular form.

Asymptotic Analysis - Linear Time Encodable Codes



Asymptotic Analysis - Linear Time Encodable Codes

Consider the subgraph induced by degree two variable nodes. Pick a (degree two) variable node at random and look at all its neighbors. Look at neighbors of neighbors ...

$$Y_t = Y_{t-1} + X_t - 1$$

Asymptotic Analysis - Good Codes Have Bad Minimum Distance

Flatness Condition [Shokrollahi] Consider the BEC with erasure probability ϵ. Let (λ_i, ρ_i) be a sequence of degree distribution pairs with a threshold of at least ϵ and rate converging to $1-\epsilon$. Then

$$\lambda_i'(0)\rho_i'(1) \to \frac{1}{\epsilon} > 1.$$

Growth of Minimum Distance of LDPC Ensembles Let (λ, ρ) be a degree distribution pair and for each $n \in \mathbb{N}$ let $\mathcal{C}(n, \lambda, \rho)$ denote the ensemble of LDPC codes of length n and with degree distribution pair (λ, ρ).

(i) If $\lambda'(0)\rho'(1) < 1$, then there exist constants α and β, both strictly positive, such that at most a fraction $e^{-n\beta}$ of codes in $\mathcal{C}(n, \lambda, \rho)$ have a minimum distance below αn.

(ii) If on the other hand $\lambda'(0)\rho'(1) > 1$, then there exist constants α and β, both strictly positive, such that at most a fraction $e^{-n\beta}$ of codes in $\mathcal{C}(n, \lambda, \rho)$ have a minimum distance exceeding $\alpha \ln(n)$.

Conclusion Capacity achieving LDCP ensembles can not have large minimum distance.

[Di, Richardson, Urbanke]