Understanding and Reducing Error-Floors of Graphical Codes

Misha Chertkov

Center for Nonlinear Studies & Theory Division, LANL

Sep 30, 2008, LIDS/MIT

The talk is based on joint papers with V. Chernyak (Wayne State, Detroit)
& S.K. Chilappagari, M. Stepanov, B. Vasic (UofA, Tucson)

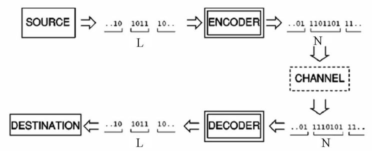
Outline

1. Introduction
   - LDPC. Graphs. Channels. Decoding.
   - Belief Propagation, LP and Loop Series
   - Error-Floor
2. Analysis of BP/LP Error Floor
   - Error-Floor. Pseudo-Codewords and Instantons.
   - Pseudo-codeword/Instanton Search Algorithm
3. Reducing Error Floor (better decoder)
   - Correcting BP along the critical loop
   - Breaking the critical loop
   - Towards Maximum Likelihood
4. Reducing Error Floor (better code)
   - Similarity between instantons for different channels/decoders
   - Instanton Exclusion

Error Correction

Cell phone Hard disk Optical disk Fiber

Scheme:



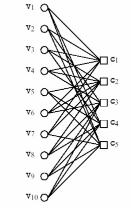
Example of Additive White Gaussian Channel:

$$P(\mathbf{x}_{out}|\mathbf{x}_{in}) = \prod_{i=bits} p(x_{out;i}|x_{in;i})$$

$$p(x|y) \sim \exp(-s^2(x-y)^2/2)$$

- Channel
  is noisy "black box" with only statistical information available
- Encoding:
  use redundancy to redistribute damaging effect of the noise
- Decoding [Algorithm]:
  reconstruct most probable codeword by noisy (polluted) channel

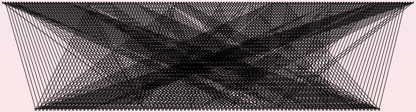
Low Density Parity Check Codes

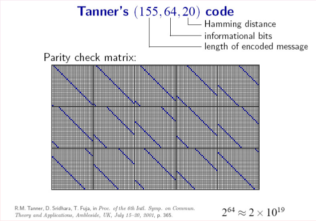


- N bits, M checks, $L = N - M$ information bits
  example: $N = 10$, $M = 5$, $L = 5$
- 2^L codewords of 2^N possible patterns
- Parity check: $\hat{H}\mathbf{v} = \mathbf{c} = \mathbf{0}$
  example:

$$\hat{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

- LDPC = graph (parity check matrix) is sparse





Channels to be discussed

White (not correlated): $P(\mathbf{x}_{out}|\mathbf{x}_{in}) = \prod_{i=bits} p(x_{out;i}|x_{in;i})$

Symmetric, Binary-Input: $p(x|y=0,1) = p(1-x|1-y)$

- Additive White Gaussian Noise (AWGN):

$$p(x|y=0,1) \sim \exp(-s^2(x-y)^2/2)$$

- Binary Erasure Channel (BEC):

$$p(x|y=0,1) = \begin{cases} 1-\epsilon, & x=y \\ \epsilon, & x=*=\text{erasure} \end{cases}$$

- Binary Symmetric Channel (BSC):

$$p(x|y=0,1) = \begin{cases} 1-\epsilon, & x=y \\ \epsilon, & x=1-y \end{cases}$$

Statistical Inference

$\boldsymbol{\sigma}_{\text{orig}}$	$\Rightarrow$	$\mathbf{x}$	$\Rightarrow$	$\boldsymbol{\sigma}$
original *data* $\boldsymbol{\sigma}_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\mathbf{x}\|\boldsymbol{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\boldsymbol{\sigma} \in \mathcal{C}$

Maximum Likelihood

Marginal Probability

$$\arg\max_{\boldsymbol{\sigma}} \mathcal{P}(\boldsymbol{\sigma}|\mathbf{x}) \qquad \arg\max_{\sigma_i} \sum_{\boldsymbol{\sigma}\setminus\sigma_i} \mathcal{P}(\mathbf{x}|\boldsymbol{\sigma})$$

Exhaustive search is generally expensive:
complexity of the algorithm $\sim 2^N$

Statistical Inference

$\boldsymbol{\sigma}_{\text{orig}}$	$\Rightarrow$	$\mathbf{x}$	$\Rightarrow$	$\boldsymbol{\sigma}$
original *data* $\boldsymbol{\sigma}_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\mathbf{x}\|\boldsymbol{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\boldsymbol{\sigma} \in \mathcal{C}$

Maximum Likelihood

Marginal Probability

$$\arg\max_{\boldsymbol{\sigma}} \mathcal{P}(\boldsymbol{\sigma}|\mathbf{x}) \qquad \arg\max_{\sigma_i} \sum_{\boldsymbol{\sigma}\setminus\sigma_i} \mathcal{P}(\mathbf{x}|\boldsymbol{\sigma})$$

Exhaustive search is generally expensive:
complexity of the algorithm $\sim 2^N$

Statistical Inference

σ_{orig}	$\Rightarrow$	$\mathbf{x}$	$\Rightarrow$	$\boldsymbol{\sigma}$
original *data* $\sigma_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\mathbf{x}\|\boldsymbol{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\boldsymbol{\sigma} \in \mathcal{C}$

Maximum Likelihood

Marginal Probability

$$\arg\max_{\boldsymbol{\sigma}} \mathcal{P}(\boldsymbol{\sigma}|\mathbf{x}) \qquad \arg\max_{\sigma_i} \sum_{\boldsymbol{\sigma}\setminus\sigma_i} \mathcal{P}(\mathbf{x}|\boldsymbol{\sigma})$$

Exhaustive search is generally expensive:
complexity of the algorithm $\sim 2^N$

Statistical Inference

$\boldsymbol{\sigma}_{\text{orig}}$	$\Rightarrow$	$\mathbf{x}$	$\Rightarrow$	$\boldsymbol{\sigma}$
original *data* $\boldsymbol{\sigma}_{\text{orig}} \in \mathcal{C}$ codeword	noisy channel $\mathcal{P}(\mathbf{x}\|\boldsymbol{\sigma})$	corrupted data: log-likelihood magnetic field	statistical inference	possible preimage $\boldsymbol{\sigma} \in \mathcal{C}$

$$\boldsymbol{\sigma} = (\sigma_1, \cdots, \sigma_N), \quad N \text{ finite}, \quad \sigma_i = \pm 1 \text{ (example)}$$

Maximum Likelihood

Marginal Probability

$$\arg\max_{\boldsymbol{\sigma}} \mathcal{P}(\boldsymbol{\sigma}|\mathbf{x}) \qquad \arg\max_{\sigma_i} \sum_{\boldsymbol{\sigma}\setminus\sigma_i} \mathcal{P}(\mathbf{x}|\boldsymbol{\sigma})$$

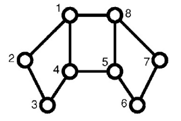
Exhaustive search is generally expensive:
complexity of the algorithm $\sim 2^N$

Graphical models

Factorization (Forney '01, Loeliger '01)

$$\mathcal{P}(\boldsymbol{\sigma}|\mathbf{x}) = Z^{-1}\prod_a f_a(\mathbf{x}_a|\boldsymbol{\sigma}_a)$$

$$\underbrace{Z(\mathbf{x}) = \sum_{\boldsymbol{\sigma}}\prod_a f_a(\mathbf{x}_a|\boldsymbol{\sigma}_a))}_{\text{partition function}}$$

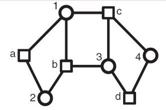


$$f_a \geq 0$$
$$\sigma_{ab} = \sigma_{ba} = \pm 1$$
$$\boldsymbol{\sigma}_1 = (\sigma_{12}, \sigma_{14}, \sigma_{18})$$
$$\boldsymbol{\sigma}_2 = (\sigma_{12}, \sigma_{13})$$

Error-Correction (linear code, bipartite Tanner graph)

$$f_i(h_i|\boldsymbol{\sigma}_i) = \exp(\sigma_i h_i)\cdot\begin{cases} 1, & \forall \alpha,\beta \ni i, \quad \sigma_{i\alpha} = \sigma_{i\beta} \\ 0, & \text{otherwise} \end{cases}$$

$$f_\alpha(\boldsymbol{\sigma}_\alpha) = \delta\left(\prod_{i\in\alpha}\sigma_i, +1\right)$$



h_i - log-likelihoods

Suboptimal but Efficient Decoding

MAP≈BP=Belief-Propagation (Bethe-Pieirls): iterative ⇒ Gallager '61; MacKay '98

- Exact on a tree ▸ Derivation Sketch
- Trading optimality for reduction in complexity: $\sim 2^L \to \sim L$
- BP = solving equations on the graph:

$$\eta_{\alpha j} = h_j + \sum_{\beta \neq \alpha}^{j \in \beta} \tanh^{-1}\left(\prod_{i \neq j}^{i \in \beta} \tanh \eta_{\beta i} \right) \quad \Leftarrow \text{LDPC representation}$$

- Message Passing = iterative BP
- Convergence of MP to minimum of Bethe Free energy can be enforced (Stepanov, Chertkov '06)

Bethe free energy: variational approach (Yedidia,Freeman,Weiss '01 - inspired by Bethe '35, Peierls '36)

$$F = \underbrace{-\sum_a \sum_{\boldsymbol{\sigma}_a} b_a(\boldsymbol{\sigma}_a) \ln f_a(\boldsymbol{\sigma}_a)}_{\text{self-energy}} + \underbrace{\sum_a \sum_{\boldsymbol{\sigma}_a} b_a(\boldsymbol{\sigma}_a) \ln b_a(\boldsymbol{\sigma}_a) - \sum_{(a,c)} b_{ac}(\sigma_{ac}) \ln b_{ac}(\sigma_{ac})}_{\text{configurational entropy}}$$

$$\forall\ a;\ c \in a: \quad \sum_{\boldsymbol{\sigma}_a} b_a(\boldsymbol{\sigma}_a) = 1, \quad b_{ac}(\sigma_{ac}) = \sum_{\boldsymbol{\sigma}_a \setminus \sigma_{ac}} b_a(\boldsymbol{\sigma}_a)$$

Linear Programming version of Belief Propagation

In the limit of large SNR, $\ln f_a \to \pm\infty$: BP→LP

$$\text{Minimize } F \approx E = -\sum_a \sum_{\boldsymbol{\sigma}_a} b_a(\boldsymbol{\sigma}_a) \ln f_a(\boldsymbol{\sigma}_a) = \text{self energy}$$

under set of linear constraints

LP decoding of LDPC codes — Feldman, Wainwright, Karger '03

- ML can be restated as an LP over a codeword polytope
- LP decoding is a "local codewords" relaxation of LP-ML
- Codeword convergence certificate
- Discrete and Nice for Analysis
- Large polytope $\{b_\alpha, b_i\} \Rightarrow$ Small polytope $\{b_i\}$

▸ Linear Programming Decoding

Loop Series:

Chertkov,Chernyak '06

Exact (!!) expression in terms of BP

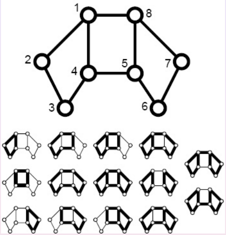
$$Z=\sum_{\boldsymbol{\sigma}}\prod_a f_a(\boldsymbol{\sigma}_a)=Z_0\left(1+\sum_C r(C)\right)$$

$$r(C)=\frac{\prod\limits_{a\in C}\mu_a}{\prod\limits_{(ab)\in C}(1-m_{ab}^2)}=\prod_{a\in C}\tilde{\mu}_a$$

$C\in$ Generalized Loops $=$ Loops without loose ends

$$m_{ab}=\int d\boldsymbol{\sigma}_a b_a^{(bp)}(\boldsymbol{\sigma}_a)\sigma_{ab}$$

$$\mu_a=\int d\boldsymbol{\sigma}_a b_a^{(bp)}(\boldsymbol{\sigma}_a)\prod_{b\in a,C}(\sigma_{ab}-m_{ab})$$



- The Loop Series is finite
- All terms in the series are calculated within BP
- BP is exact on a tree

For more details on BP & Beyond – Attend the Mini-Course:

Special Lecture Series on

Statistical Physics of Algorithms (or Belief Propagation and Beyond)

Michael (Misha) Chertkov

Theory Division, Los Alamos National Laboratory

Day	Monday	Wednesday	Thursday
Date	**Sep 29**	**Oct 1**	**Oct 2**
	Oct 27	**Oct 29**	**Oct 30**
Room	32-141	32-141	36-156
Time	**4-5pm**		

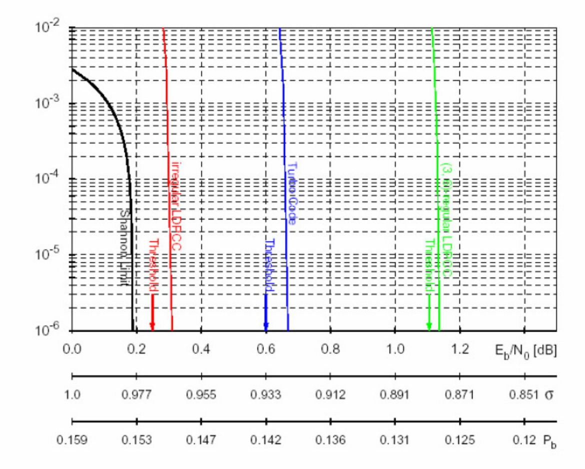
There has been an explosion of interest in the past decade to statistical problems related to computer science and information processing, such as new decoding paradigms for high-volume communication and storage, problems in search and counting through a huge set of combinatorial constraints, etc. Novel ideas in analysis of complexity and development of approximate but systematically improvable algorithms are required for the hard statistical problems. Since the central task of statistical physics is to describe how complex behavior emerges from interaction of large number of basic elements, its tools and concepts are proven valuable in the emerging disciplines.

This mini-course offers an overview of the work by the author and collaborators on analysis and design of efficient algorithms for hard computational problems.

List of subjects includes

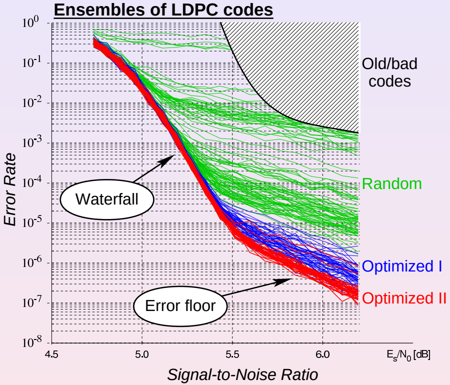
- Statistical Inference (Data Restoration), Combinatorial Optimization and Counting. What is easy (shortest path, dynamical programming, calculations on a tree) and what is difficult (K-SAT, spin-glasses, decoding of graphical codes).
- Graphical Models. Ground State and Partition Function. Belief Propagation, Mean Field and Bethe Free Energy Approximation. Linear Programming and Belief Propagation.
- Belief Propagation and Loop Calculus. Binary case. Graphic and Variational Formulations. Self-avoiding tree (Weitz method). Loop Tower = Extension of Loop Calculus to q-ary alphabet.
- Inference on Planar Graphs: Fisher-Kastelyan-Barahona approach. Loop Calculus for Planar Models. Pfaffians as Grassman Integrals.
- Loop Caculus for Gaussian Graphical Model (continious alphabet). Graphical Gauge Theories. Monomer-Dimer Model and Determinants.
- Example of Synthesis: BP and Beyond for particle tracking. Loop Series as a Cauchi Integral.

Shannon Transition



- Phase Transition
- Ensemble of Codes [analysis & design]
- Thermodynamic limit but ...

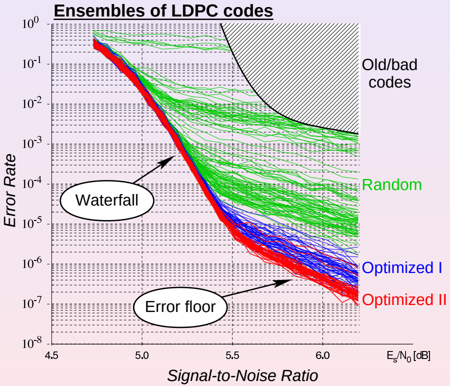
Error-Floor



- T. Richardson '03 (EF)
- Density evolution does not apply (to EF)

- BER vs SNR = measure of performance
- Finite size effects
- Waterfall $\leftrightarrow$ Error-floor
- Error-floor typically emerges due to sub-optimality of decoding, i.e. due to unaccounted loops
- Monte-Carlo is useless at FER $\lesssim 10^{-8}$

Error-floor Challenges



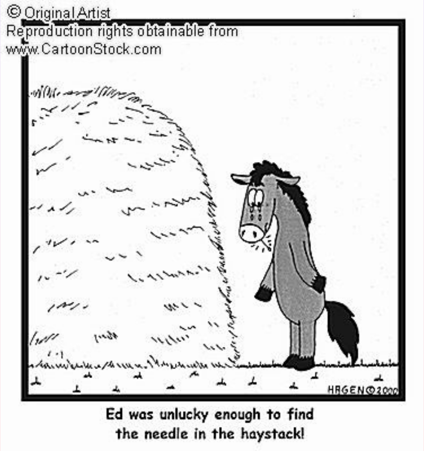
- Understanding the Error Floor (Inflection point, Asymptotics), Need an efficient method to analyze error-floor
- Improving Decoding
- Constructing New Codes

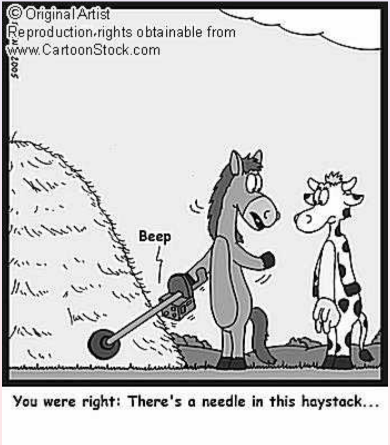
Optimal Fluctuation (Instanton) Approach for Extracting Rare but Dominant Events

Optimal Fluctuation (Instanton) Approach for Extracting Rare but Dominant Events

Ed was unlucky enough to find the needle in the haystack!

Optimal Fluctuation (Instanton) Approach for Extracting Rare but Dominant Events





Pseudo-codewords and Instantons

Error-floor is caused by Pseudo-codewords:

Wiberg '96; Forney et.al'99; Frey et.al '01;
Richardson '03; Vontobel, Koetter '04-'06

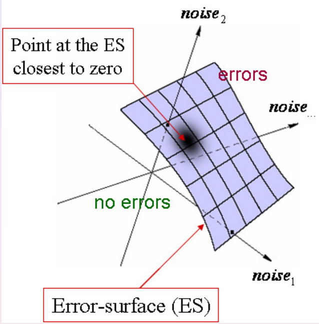
Instanton = optimal conf of the noise

$$BER = \int d(noise)\ WEIGHT(noise)$$

$$BER \sim WEIGHT \left(\begin{array}{c} optimal\ \ conf \\ of\ \ the\ \ noise \end{array} \right)$$

$$\begin{array}{c} optimal\ \ conf \\ of\ \ the\ \ noise \end{array} = \begin{array}{c} Point\ \ at\ \ the\ \ ES \\ closest\ \ to\ \ "0" \end{array}$$

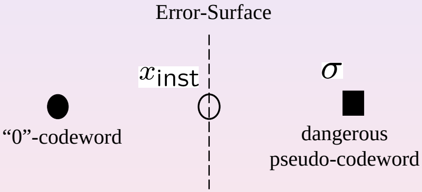
Instantons are decoded to Pseudo-Codewords



Instanton-amoeba

= optimization algorithm
Stepanov, et.al '04,'05
Stepanov, Chertkov '06

Idea of a Smarter Search Strategy



Weighted Median: [for AWGN]

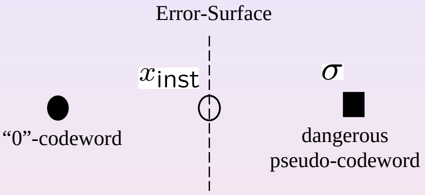
$$\mathbf{x}_{\text{inst}} = \frac{\boldsymbol{\sigma}}{2}\frac{\sum_i \sigma_i}{\sum_i \sigma_i^2}, \quad d = \frac{\left(\sum_i \sigma_i\right)^2}{\sum_i \sigma_i^2}$$

$$\text{FER} \sim \exp(-d \cdot s^2/2)$$

Wiberg '96; Forney et.al '01

Vontobel, Koetter '03,'05

Idea of a Smarter Search Strategy



Weighted Median: [for AWGN]

$$\mathbf{x}_{\text{inst}} = \frac{\boldsymbol{\sigma}}{2}\frac{\sum_i \sigma_i}{\sum_i \sigma_i^2}, \quad d = \frac{\left(\sum_i \sigma_i\right)^2}{\sum_i \sigma_i^2}$$

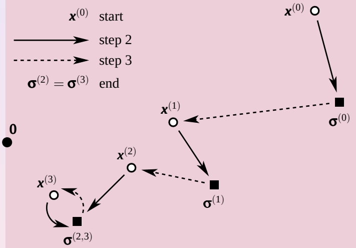
$$\text{FER} \sim \exp(-d \cdot s^2/2)$$

Wiberg '96; Forney et.al '01
Vontobel, Koetter '03,'05

Getting to the instanton in few shots:

Pseudo-Codeword Search Algorithm [Continuous Channel]

Chertkov, Stepanov '06

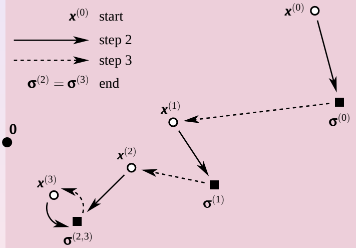


- **Start:** Initiate $\mathbf{x}^{(0)}$.
- **Step 1:** $\mathbf{x}^{(k)}$ is decoded to $\boldsymbol{\sigma}^{(k)}$.
- **Step 2:** Find $\mathbf{y}^{(k)}$ - weighted median between $\boldsymbol{\sigma}^{(k)}$, and "0"
- **Step 3:**
  If $\mathbf{y}^{(k)} = \mathbf{y}^{(k-1)}$, $k_* = k$ End.
  Otherwise go to **Step 2** with $\mathbf{x}^{(k+1)} = \mathbf{y}^{(k)} + 0$.

- Monotonicity of Iterations (e.g. observed empirically) is not proved for the AWGN version of the algorithm

Pseudo-Codeword Search Algorithm [Continuous Channel]

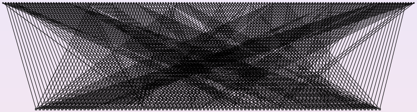
Chertkov, Stepanov '06

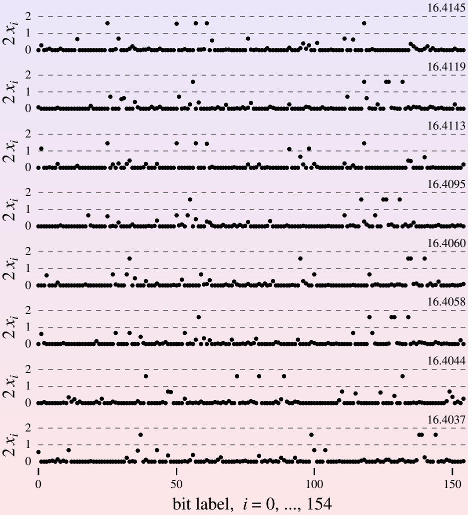


- **Start:** Initiate $\mathbf{x}^{(0)}$.
- **Step 1:** $\mathbf{x}^{(k)}$ is decoded to $\boldsymbol{\sigma}^{(k)}$.
- **Step 2:** Find $\mathbf{y}^{(k)}$ - weighted median between $\boldsymbol{\sigma}^{(k)}$, and "0"
- **Step 3:**
  If $\mathbf{y}^{(k)} = \mathbf{y}^{(k-1)}$, $k_* = k$ End.
  Otherwise go to **Step 2** with $\mathbf{x}^{(k+1)} = \mathbf{y}^{(k)} + 0$.

- Monotonicity of Iterations (e.g. observed empirically) is not proved for the AWGN version of the algorithm

PCS Test $(155, 64, 20)$, LP, AWGN

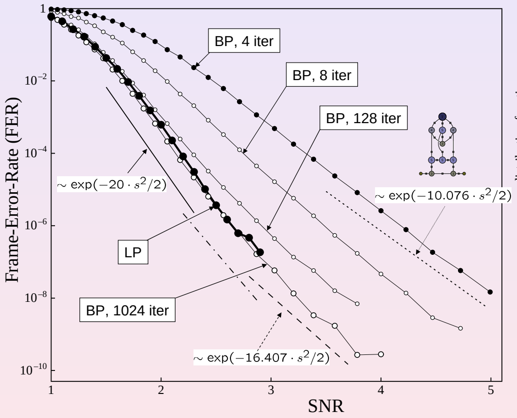


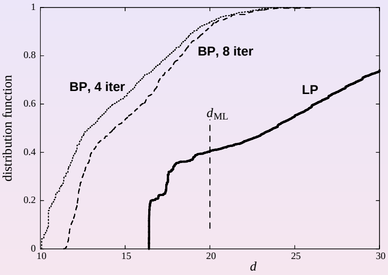


- Fast Convergence ($5 - 10$ iterations)
- ~ 200 pseudo-codewords within $16.4037 < d < 20$

Reducing Complexity of LP

FER vs SNR AWGN





Instanton-amoeba:
Stepanov, et.al '04,'05,'06
LP-based PCS-search:
Chertkov, Stepanov '06,'07

Instanton-Search Algorithm for BSC/LP

Instantons for BSC/LP

Explaining the Error-Floor: Results and Challenges

Results:

- Instantons are responsible for the Error-Floor
- Finding an instanton is an optimization problem
- Instanton-Search Toolbox is Efficient for getting correct Error-Floor asymptotic

Challenges:

- Need an efficient algorithm constructing the entire FER vs SNR curve at once, at least finding the inflection point
- Improve decoder
- Improve code

Explaining the Error-Floor: Results and Challenges

Results:

- Instantons are responsible for the Error-Floor
- Finding an instanton is an optimization problem
- Instanton-Search Toolbox is Efficient for getting correct Error-Floor asymptotic

Challenges:

- Need an efficient algorithm constructing the entire FER vs SNR curve at once, at least finding the inflection point
- Improve decoder
- Improve code

Loop Calculus & Pseudo-Codeword Analysis

Chertkov, Chernyak '06

Single loop truncation

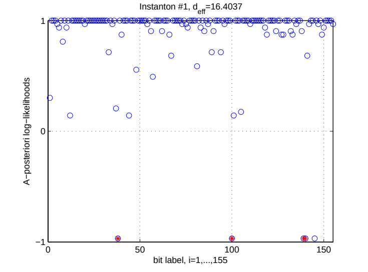
$$Z = Z_0(1 + \sum_C r_C) \approx Z_0(1 + r(\Gamma))$$

Synthesis of Pseudo-Codeword Search Algorithm (Chertkov, Stepanov '06) & Loop Calculus

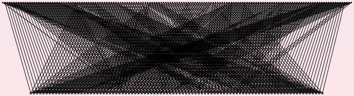
- Consider pseudo-codewords one after another
- For an individual pseudo-codeword/instanton identify a critical loop, Γ, giving major contribution to the loop series.
- Hint: look for single connected loops and use local "triad" contributions as a tester: $r(\Gamma) = \prod_{\alpha \in \Gamma} \tilde{\mu}_\alpha^{(bp)}$

Proof-of-Concept test [(155, 64, 20) code over AWGN]

- $\forall$ pseudo-codewords with $16.4037 < d < 20$ ($\sim$ 200 found) there always exists a simple single-connected critical loop(s) with $r(\Gamma) \sim 1$.
- Pseudo-codewords with the lowest d show $r(\Gamma) = 1$
- Invariant with respect to other choices of the original codeword



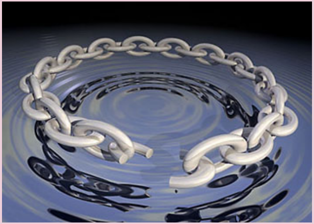
Bigger Set



Correcting along the critical loop

- Decode with BP/LP
- Find the critical loop
- Correct/Improve Decoding

... by breaking the loop



Breaking the critical loop locally

Chertkov '07

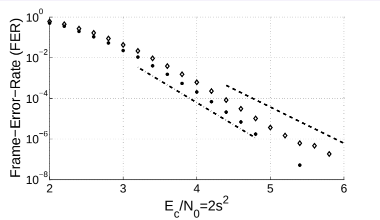
Loop Guided Guessing (LGG)

- **1.** Run the LP algorithm. Terminate if LP succeeds.
- **2.** If LP fails, find the critical loop, Γ.
- **3.** Pick any bit along the critical loop and "fix the bit" running two corrected LP schemes. Terminate if any of LPs succeeds.
- **4.** If not return to **Step 3** selecting another bit along the critical loop or to **Step 2** for an improved selection principle for Γ.

Loop-Corrected BP

LP-erasure

[155, 64, 20] test of LGG



- Complexity of LGG is the same as of LP
- LGG corrects 9 out of 10 errors at $E_b/N_0 = 4.8$!!
- Error Floor is Reduced !!

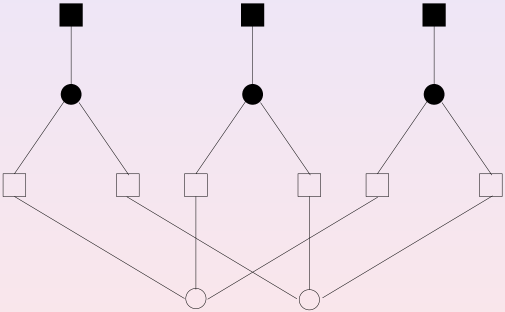
What to do with the remaining 1/10 ?

- Draper, Yedidia, Wang ISIT'07: Fixing $1, 2, ..., k$ bits $= 2^k$ LPs till decode to a codeword (ML certificate enforced).
- Weiss, Yanover, Meltzer '07: Sufficient condition for bits decoded by the bare LP to integers to show the right values.

Our further strategy:

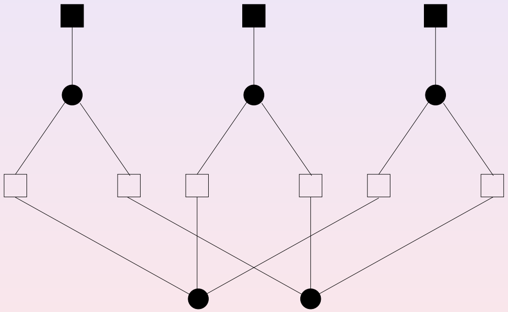
- Use Loop Calculus in sequential selection of the fixed bits
- Longer codes
- Back to iterative BP

Relation Between Instantons for Different Channels and Decoders $[155, 64, 20]$



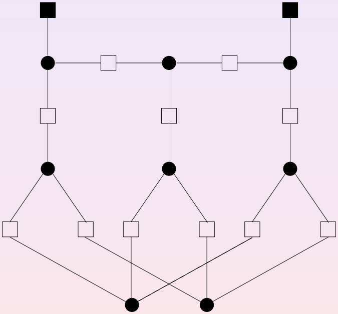
$(5, 3)$ trapping set of Gallager A.

Relation Between Instantons for Different Channels and Decoders $[155, 64, 20]$



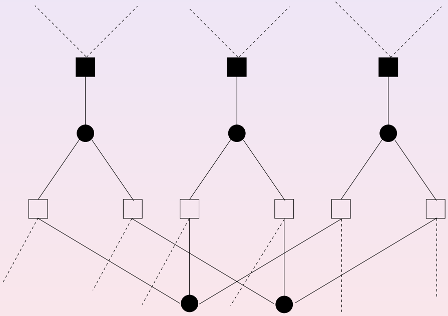
Support structure of the BSC/LP instanton.

Relation Between Instantons for Different Channels and Decoders $[155, 64, 20]$



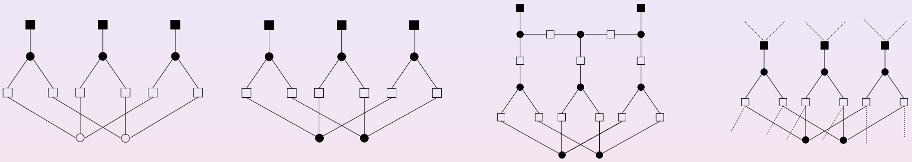
Trapping set, $(8, 2)$, over BEC.

Relation Between Instantons for Different Channels and Decoders $[155, 64, 20]$



Essential part of the AWGN-instanton.

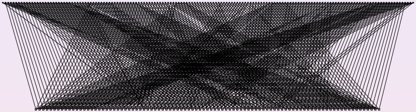
Relation Between Instantons for Different Channels and Decoders $[155, 64, 20]$



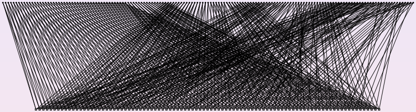
All have a common backbone!!

Chilappagari, Chertkov, Stepanov, Vasic '08

Excluding the “bad” structure



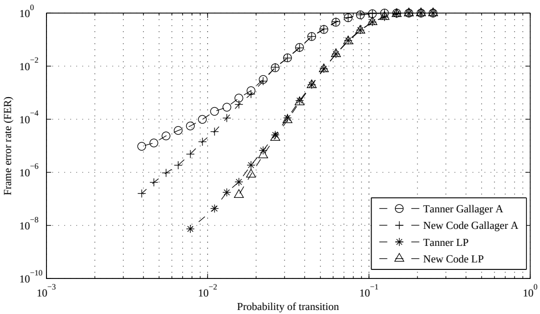
Tanner $[155, 64, 20]$ code

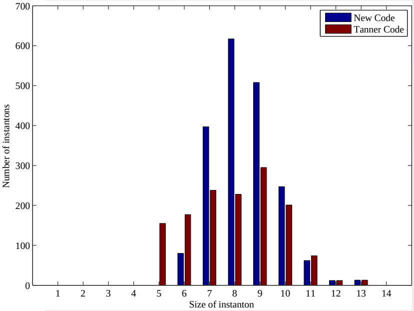


Similar code with the $(5, 3)$ backbone excluded

Random construction in the spirit of [Chilappagari, Krishnan, Vasic ’08]

Old and New Codes – Comparison BSC





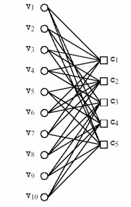
Chilappagari, Chertkov, Stepanov, Vasic '08

Conclusions

- Error floor is due to low-weight (dangerous) pseudo-codewords
- Instanton-amoeba & Pseudo-codeword/instanton search algorithms are efficient techniques allowing to find dangerous instantons/pseudo-codewords and reconstruct FER vs SNR curve efficiently.
- Critical loops in the Loop Series signify wrong decoding. The Loop Series based analysis offers efficient guiding principle for decoding improvement.
- Instantons found for the same code but different decoders and channels are similar. This observation allows efficient code construction excluding backbones on the most dangerous (probable) instantons.
- Reducing the error floor may be not that difficult ... after all

Thank You !!

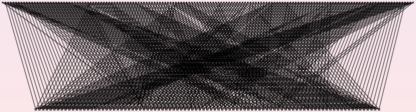
Low Density Parity Check Codes

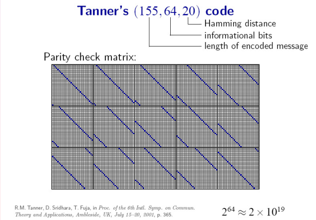


- N bits, M checks, $L = N - M$ information bits
  example: $N = 10$, $M = 5$, $L = 5$
- 2^L codewords of 2^N possible patterns
- Parity check: $\hat{H}\mathbf{v} = \mathbf{c} = \mathbf{0}$
  example:

$$\hat{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

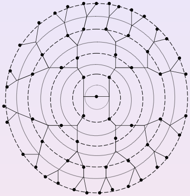
- LDPC = graph (parity check matrix) is sparse





Coding-Decoding

BP is Exact on a Tree (LDPC)



$$Z(\mathbf{h}) = \sum_{\boldsymbol{\sigma}} \prod_{\alpha=1}^{M} \delta\left(\prod_{i\in\alpha} \sigma_i, 1\right) \exp\left(\sum_{i=1}^{N} h_i\sigma_i\right)$$

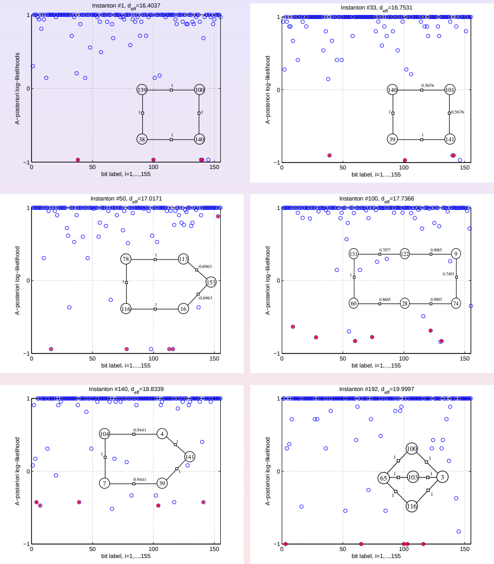
h_i is a log-likelihood at a bit (outcome of the channel)

$$Z^{\pm}_{j\alpha}(\mathbf{h}^{>}) \equiv \sum_{\boldsymbol{\sigma}^{>}}^{\sigma_j=\pm 1} \prod_{\beta>} \delta\left(\prod_{i\in\beta} \sigma_i, 1\right) \exp\left(\sum_{i>} h_i\sigma_i\right)$$

$$Z^{\pm}_{j\alpha} = \exp(\pm h_j) \prod_{\beta\neq\alpha}^{j\in\beta} \frac{1}{2}\left(\prod_{i\neq j}^{i\in\beta}(Z^{+}_{i\beta}+Z^{-}_{i\beta}) \pm \prod_{i\neq j}^{i\in\beta}(Z^{+}_{i\beta}-Z^{-}_{i\beta})\right)$$

$$\eta_{j\alpha} \equiv \frac{1}{2}\ln\left(\frac{Z^{+}_{j\alpha}}{Z^{-}_{j\alpha}}\right), \quad \eta_{j\alpha} = h_j + \sum_{\beta\neq\alpha}^{j\in\beta} \tanh^{-1}\left(\prod_{i\neq j}^{i\in\beta} \tanh \eta_{i\beta}\right)$$

Bethe Free Energy


Instanton #1, d_eff=16.4037
Instanton #33, d_eff=16.7531
Instanton #50, d_eff=17.0171
Instanton #100, d_eff=17.7366
Instanton #140, d_eff=18.8339
Instanton #192, d_eff=19.9997
A−posteriori log−likelihoods
A−posteriori log−likelihood
bit label, i=1,...,155

Back

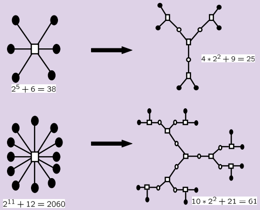
Reducing complexity of LP

Complexity of the bare LP grows exponentially with check degree

Current solutions:

- Adaptive LP (Taghavi, Siegel '06)
- BP-style relaxation of LP (Vontobel, Koetter '06)

Dendro-trick = Graph Modification (our solution) Chertkov,Stepanov'07



- MAP solutions are identical
- Set of Pseudo-codewords are identical
- Instanton spectra are very alike, $\approx$

Pseudo-Codeword Search

Minimize, $E=\sum_\alpha\sum_{\sigma_\alpha} b_\alpha(\sigma_\alpha)\sum_{i\in\alpha}\sigma_i(1-2x_i)/q_i$,
under the conditions:
$\forall i,\alpha \quad 0\le b_i(\sigma_i), b_\alpha(\sigma_\alpha)\le 1$
$\forall\ i: \quad \sum_{\sigma_i} b_i(\sigma_i)=1,$
$\forall\ i\ \forall\ \alpha\ni i: \quad b_i(\sigma_i)=\sum_{\sigma_\alpha\setminus\sigma_i} b_\alpha(\sigma_\alpha)$

Linear Programming

Extended Variational Principe & Loop-Corrected BP

Bare BP Variational Principe: $$\left.\frac{\partial Z_0}{\partial \eta_{ab}}\right|_{\eta^{(bp)}} = 0$$

New choice of Gauges guided by the knowledge of the critical loop Γ

$$\left.\frac{\partial \exp(-\mathcal{F})}{\partial \eta_{ab}}\right|_{\eta_{\rm eff}} = 0, \quad \mathcal{F} \equiv -\ln(Z_0 + Z_\Gamma)$$

BP-equations are modified along the critical loop Γ

$$\left.\frac{\sum_{\boldsymbol{\sigma}_a}(\tanh(\eta_{ab}+\eta_{ba})-\sigma_{ab})P_a(\boldsymbol{\sigma}_a)}{\sum_{\boldsymbol{\sigma}_a} P_a(\boldsymbol{\sigma}_a)}\right|_{\eta_{\rm eff}} = \left.\text{explicitly known contribution}\right|_{\eta_{\rm eff}} \neq 0 \quad [\text{along } \Gamma]$$

Loop-Corrected BP Algorithm

- **1.** Run bare BP algorithm. Terminate if BP succeeds (i.e. a valid code word is found).
- **2.** If BP fails find the most relevant loop Γ that corresponds to the maximal $|r_\Gamma|$. Triad search is helping.
- **3.** Solve the modified-BP equations for the given Γ. Terminate if the improved-BP succeeds.
- **4.** Return to **Step 2** with an improved Γ-loop selection.

Breaking the Loop

LP-erasure = simple heuristics

- **1.** Run LP algorithm. Terminate if LP succeeds (i.e. a valid code word is found).
- **2.** If LP fails, find the most relevant loop Γ that corresponds to the maximal amplitude $r(\Gamma)$.
- **3.** Modify the log-likelihoods along the loop Γ introducing a shift towards zero, i.e. introduce a complete or partial erasure of the log-likelihoods at the bits. Run LP with modified log-likelihoods. Terminate if the modified LP succeeds.
- **4.** Return to **Step 2** with an improved selection principle for the critical loop.

$(155, 64, 20)$ Test

- **IT WORKS!**
  All troublemakers ($\sim$ 200 of them) previously found by LP-based Pseudo-Codeword-Search Algorithm method were successfully corrected by the LP-erasure algorithm.
- Method is invariant with respect the choice of the codeword (used to generate pseudo-codewords).

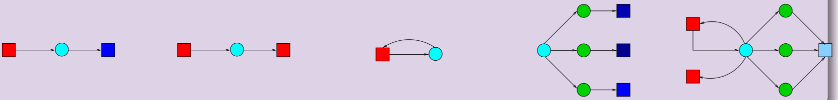
General Conjecture:

- Loop-erasure algorithm is capable of reducing the error-floor
- Local adjustment of the algorithm, anywhere along the critical loop, in the spirit of the Facet Guessing (Dimakis, Wainwright '06), may be sufficient $\Rightarrow$

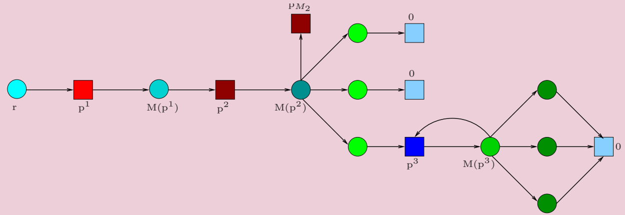
Breaking the Loop

Instanton-Search Algorithm for BSC

Required discrete channel adjustment



Typical Sequence for $[155, 64, 20]$

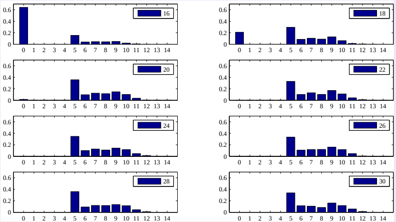


Chilappagari, Chertkov, Vasic '08

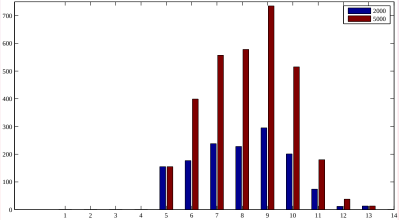
FER under AWGN

Instantons for [155, 64, 20], BSC, LP

Frequency of instanton sizes



Instanton-Bar-Graph



Chilappagari, Chertkov, Vasic '08

Bibliography (A)

Loop Calculus, Series, Tower (general, not coding specific)

- V. Chernyak and M. Chertkov, "Loop Calculus and Belief Propagation for q-ary Alphabet: Loop Tower," *Proceedings of IEEE ISIT 2007*, June 2007, Nice, `arXiv:cs.IT/0701086`.
- M. Chertkov and V. Chernyak, "Loop series for discrete statistical models on graphs," JSTAT/2006/P06009, `arXiv:cond-mat/0603189`.
- M. Chertkov and V. Chernyak, "Loop Calculus in Statistical Physics and Information Science," *Phys. Rev. E*, **73**, 065102(R) (2006), `arXiv:cond-mat/0601487`.

Reducing the Error Floor (better decoder)

- M. Chertkov, "Reducing the Error Floor", invited talk at the *Information Theory Workshop '07 on "Frontiers in Coding"*, September 2-6, 2007.
- M. Chertkov and V. Chernyak, "Loop Calculus Helps to Improve Belief Propagation and Linear Programming Decodings of Low-Density-Parity-Check Codes," invited talk at 44^{th} *Allerton Conference*, September 27-29, 2006, Allerton, IL, `arXiv:cs.IT/0609154`.

All papers are available at http://cnls.lanl.gov/∼chertkov/pub.htm

Bibliography (B)

Instantons & Pseudo-Codewords. Analyzing the Error-Floor.

- S.K. Chilappagari, M. Chertkov. M. Stepanov and B. Vasic, "Efficient and Universal Techniques for Analysis and Reduction of LDPC Error-floors", In preparation for submission to IEEE JSAC on Capacity Approaching Codes, October 2008.
- S.K. Chilappagari, M. Chertkov and B. Vasic, "Provably Efficient Instanton Search Algorithm for LP-decoding over the BSC", submitted to IEEE IT, `arXiv:0808.2515`.
- M. Chertkov and M. Stepanov, "Searching for low weight pseudo-codewords," invited talk at the *2007 Information Theory and Application Workshop, proceedings*, ITA CALIT2, UCSD, `arXiv:cs.IT/0702024`.
- M. Chertkov and M. Stepanov, "Pseudo-codeword Landscape," *Proceedings of IEEE ISIT 2007*, June 2007, Nice, `arXiv:cs.IT/0701084`.
- M. Chertkov and M. Stepanov, "An Efficient Pseudo-Codeword-Search Algorithm for Linear Programming Decoding of LDPC Codes," to appear in *IEEE Transactions on Information Theory*, `arXiv:cs.IT/0601113`.
- M. Stepanov and M. Chertkov, "Instanton analysis of Low-Density-Parity-Check codes in the error-floor regime," *Proceeding of IEEE ISIT 2006*, July 2006 Seattle, `arXiv:cs.IT/0601070`.
- M. Stepanov and M. Chertkov, "The error-floor of LDPC codes in the Laplacian channel," Proceedings of 43rd Allerton Conference (September 28-30, 2005, Allerton, IL), `arXiv:cs.IT/0507031`.
- M. Stepanov, V. Chernyak, M. Chertkov and B. Vasic, "Diagnosis of weakness in error correction: a physics approach to error floor analysis," Phys. Rev. Lett. **95**, 228701 (2005),`cond-mat/0506037`.
- V. Chernyak, M. Chertkov, M. Stepanov and B. Vasic, "Error correction on a tree: An instanton approach" , Phys. Rev. Lett. **93**, 198702-1 (2004).

Bibliography (C)

Other subjects related to LDPC+ decoding

- J. A. Anguita, M. Chertkov, B. Vasic and M. A. Neifeld, "Bethe-Free-Energy Based Decoding of Low-Density Parity-Check Codes on Partial Response Channels," submitted to *IEEE Journal of Selected Areas in Communications*.
- M. Stepanov and M. Chertkov, "Improving convergence of belief propagation decoding," *Proceedings of 44th Allerton Conference*, September 27-29, 2006, Allerton, IL, `arXiv:cs.IT/0607112`.

All papers are available at http://cnls.lanl.gov/∼chertkov/pub.htm