On Concentration of Measures for LDPC Code Ensembles

Igal Sason Ronen Eshel

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

2011 IEEE International Symposium on Information Theory
Saint Petersburg, Russia

August 2011.

Concentration of Measures for LDPC Code Ensembles

- Concentration of measures and large deviations analysis are central issues in probability theory, and are strongly related to information theory and coding.
- A fundamental theorem by Richardson and Urbanke (2001) proves a concentration of measure phenomenon for the performance of LDPC codes under iterative message-passing decoding.

Theorem I - [Concentration of performance under iterative message-passing decoding (Richardson and Urbanke, 2001)]

Let $\mathcal{C}$, a code chosen uniformly at random from the ensemble LDPC(n, λ, ρ), be used for transmission over a memoryless binary-input output-symmetric (MBIOS) channel. Assume that the decoder performs l iterations of message-passing decoding, and let $P_{\mathsf{b}}(\mathcal{C}, l)$ denote the resulting bit error probability. Then, for every $\delta > 0$, there exists an $\alpha > 0$ where $\alpha = \alpha(\lambda, \rho, \delta, l)$ (*independent of the block length* n) such that

$$\mathbb{P}\left(|P_{\mathsf{b}}(\mathcal{C}, l) - \mathbb{E}_{\mathsf{LDPC}(n,\lambda,\rho)}[P_{\mathsf{b}}(\mathcal{C}, l)]| \geq \delta\right) \leq e^{-\alpha n}$$

Proof

The proof applies Azuma's inequality to a martingale sequence with bounded differences (IEEE Trans. on IT, Feb. 2001).

Azuma's inequality

Theorem - **[Azuma's inequality]**

Let $X_0, \ldots, X_n$ be a martingale. If the sequence of differences are bounded, i.e.,

$$|X_i - X_{i-1}| \leq d_i \quad \forall\, i = 1, 2, \ldots, n \quad \text{a.s.}$$

then

$$\mathbb{P}(|X_n - X_0| \geq r) \leq 2 \exp\left(-\frac{r^2}{2\sum_{i=1}^{n} d_i^2}\right), \quad \forall\, r > 0.$$

Selected Papers Applying Azuma's Inequality for LDPC Ensembles

- M. Sipser and D. A. Spielman, "Expander codes," *IEEE Trans. on Information Theory*, vol. 42, no. 6, pp. 1710-1722, November 1996.
- M.G. Luby, M. Mitzenmacher, M.A. Shokrollahi and D.A. Spielman, "Efficient erasure correcting codes", *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 569-584, February 2001.
- T. Richardson and R. Urbanke, "The capacity of low-density parity check codes under message-passing decoding." *IEEE Trans. on Information Theory*, vol. 47, pp. 599–618, February 2001.
- A. Kavcic, X. Ma and M. Mitzenmacher, "Binary intersymbol interference channels: Gallager bounds, density evolution, and code performance bounds," *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1636-1652, July 2003.

(Cont.)

- A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding," *IEEE Trans. on Information Theory*, vol. 51, no. 9, pp. 3247–3261, September 2005.
- C. Méasson, A. Montanari and R. Urbanke, "Maxwell construction: The hidden bridge between iterative and maximum a-posteriori decoding," *IEEE Trans. on Information Theory*, vol. 54, pp. 5277–5307, December 2008.
- L.R. Varshney, "Performance of LDPC codes under faulty iterative decoding," *IEEE Trans. on Information Theory*, vol. 57, no. 7, pp. 4427–4444, July 2011.

Theorem II - [Concentration of Conditional Entropy of LDPC code ensembles (Méasson et al. 2008)]

Let $\mathcal{C}$ be chosen uniformly at random from the ensemble LDPC(n, λ, ρ). Assume that the transmission of the code $\mathcal{C}$ takes place over an MBIOS channel. Let $H(\mathbf{X}|\mathbf{Y})$ designate the conditional entropy of the transmitted codeword $\mathbf{X}$ given the received sequence $\mathbf{Y}$ from the channel. Then, for any $\xi > 0$,

$$\mathbb{P}\left(\left|H(\mathbf{X}|\mathbf{Y}) - \mathbb{E}_{\mathsf{LDPC}(n,\lambda,\rho)}[H(\mathbf{X}|\mathbf{Y})]\right| \geq \sqrt{n}\,\xi\right) \leq 2\exp(-B\xi^2)$$

where $B \triangleq \frac{1}{2(d_{\mathsf{c}}^{\max}+1)^2(1-R_{\mathsf{d}})}$, $d_{\mathsf{c}}^{\max}$ is the maximal check-node degree, and R_{d} is the design rate of the ensemble.

Proof - **[outline]**

1. Introduction of a martingale sequence with bounded differences:
   - Define the RV $Z = H_{\mathcal{G}}(\mathbf{X}|\mathbf{Y})$, where $\mathcal{G}$ is a graph of a code chosen uniformly at random from the ensemble LDPC(n, λ, ρ)
   - Define the martingale sequence $Z_t = \mathbb{E}[Z|\mathcal{F}_t] \quad t \in \{0, 1, \ldots, m\}$, where the filtration is the sequence of subsets of σ-algebras, generated by revealing each time another parity-check equation of the code.
2. Upper bounds on the differences $|Z_{t+1} - Z_t|$:
   - It was proved that $|Z_{t+1} - Z_t| \leq (r+1)\, H_{\mathcal{G}}(\tilde{X}|\mathbf{Y})$, where r is the degree of parity-check equation revealed at time t, and $\tilde{X} = X_{i_1} \oplus \ldots \oplus X_{i_r}$ (i.e., $\tilde{X}$ is the modulo-2 sum of some r bits in the codeword $\mathbf{X}$).
   - Then $r \leq d_{\mathsf{c}}^{\max}$, and $H_{\mathcal{G}}(\tilde{X}|\mathbf{Y}) \leq 1$.
3. Azuma's inequality was applied to a get a concentration inequality, using $|Z_{t+1} - Z_t| \leq d_{\mathsf{c}}^{\max} + 1$ for every $t = 0, \ldots, m-1$ where $m = n(1 - R_{\mathsf{d}})$ is the number of parity-check nodes.

Improvement 1 - A tightened upper bound on the conditional entropy

- Instead of upper bounding $H_{\mathcal{G}}(\tilde{X}|\mathbf{Y})$ by 1, which is independent of the channel capacity (C), it is proved that

$$H_{\mathcal{G}}(\tilde{X}|\mathbf{Y}) \leq h\left(\frac{1-C^{\frac{r}{2}}}{2}\right)$$

  where h is the binary entropy function to the base 2.

- For a BSC or BEC, this bound can be improved to

$$h\left(\frac{1-\left[1-2h^{-1}(1-C)\right]^{r}}{2}\right)$$

  and

$$1-C^{r}$$

  respectively.

Improvement 2 (trivial)

Instead of taking the trivial bound $r \leq d_{\mathsf{c}}^{\max}$ for all m terms in the Azuma's inequality, one can rely on the degree distribution of the parity-check nodes. The number of parity-check nodes of degree r is $n(1-R_{\mathsf{d}})\Gamma_r$.

Theorem III - [Tightened Expressions for B]

Considering the terms of Theorem II, applying these two improvements yields tightened expressions for B.

- General MBIOS - $B \triangleq \dfrac{1}{2(1-R_{\mathsf{d}})\sum_{i=1}^{d_{\mathsf{c}}^{\max}}(i+1)^2\Gamma_i\left[h\left(\frac{1-C^{\frac{i}{2}}}{2}\right)\right]^2}$
- BSC - $B \triangleq \dfrac{1}{2(1-R_{\mathsf{d}})\sum_{i=1}^{d_{\mathsf{c}}^{\max}}(i+1)^2\Gamma_i\left[h\left(\frac{1-[1-2h^{-1}(1-C)]^i}{2}\right)\right]^2}$
- BEC - $B \triangleq \dfrac{1}{2(1-R_{\mathsf{d}})\sum_{i=1}^{d_{\mathsf{c}}^{\max}}(i+1)^2\Gamma_i\,(1-C^i)^2}$

Comparison of Theorem II Vs. Theorem III

Comparison for the limit where $C \to 1$ bit per channel use

- B should be infinity for a perfect channel.
- Indeed, Theorem III yields $B \to \infty$, in contrast to Theorem II where the parameter B does not depend on C and is finite.

Numerical comparison for BEC and BIAWGN

Let us consider the case where

- $(2, 20)$ regular LDPC code ensemble.
- Communication over a BEC or BIAWGNC with capacity of 0.98 per channel use.

Compared to Theorems II, applying Theorem III results in tighter expressions for B:

- BIAWGN - Improvement by factor $\left[h\left(\frac{1-C^{\frac{d_c}{2}}}{2}\right)\right]^{-2} = 5.134$
- BEC - Improvement by factor $\frac{1}{\left(1-C^{d_c}\right)^2} = 9.051$

Comparison for Heavy-Tail Poisson Distribution (Tornado Codes)

Consider the capacity-achieving Tornado LDPC code ensemble for a BEC with erasure probability p. We wish to design a code ensemble that achieves a fraction $1-\varepsilon$ of the capacity.

- Theorem II - The parity-check degree is Poisson distributed, therefore $d_{\mathsf{c}}^{\max} = \infty$. Hence, $B = 0$ and this result is useless.
- Theorem III - B scales (at least) like $O\left(\frac{1}{\log^2\left(\frac{1}{\varepsilon}\right)}\right)$.

  The parameter B tends to zero slowly as we let the fractional gap ε tend to zero; this demonstrates a rather fast concentration.

Summary

- The use of Azuma's inequality was addressed in the context of proving concentration phenomena for code ensembles defined on graphs and iterative decoding algorithms.
- We tightened a concentration inequality for the conditional entropy of LDPC ensembles by using Azuma's inequality, and deriving an improved upper bound on the jumps of the martingale sequence.
- The improved inequality enables to prove concentration of the conditional entropy for ensembles of Tornado codes (in contrast to the original concentration inequality).
- The conference paper contains another concentration inequality, based on Azuma's inequality, for the cardinality of the fundamental system of cycles in LDPC code ensembles.

Introduction to a recent work (was not submitted to ISIT 2011)

But, Azuma's inequality is not tight !.

For example, if $r > \sum_{i=1}^{n} d_i \quad \Rightarrow \quad \mathbb{P}(|X_n - X_0| \geq r) = 0.$

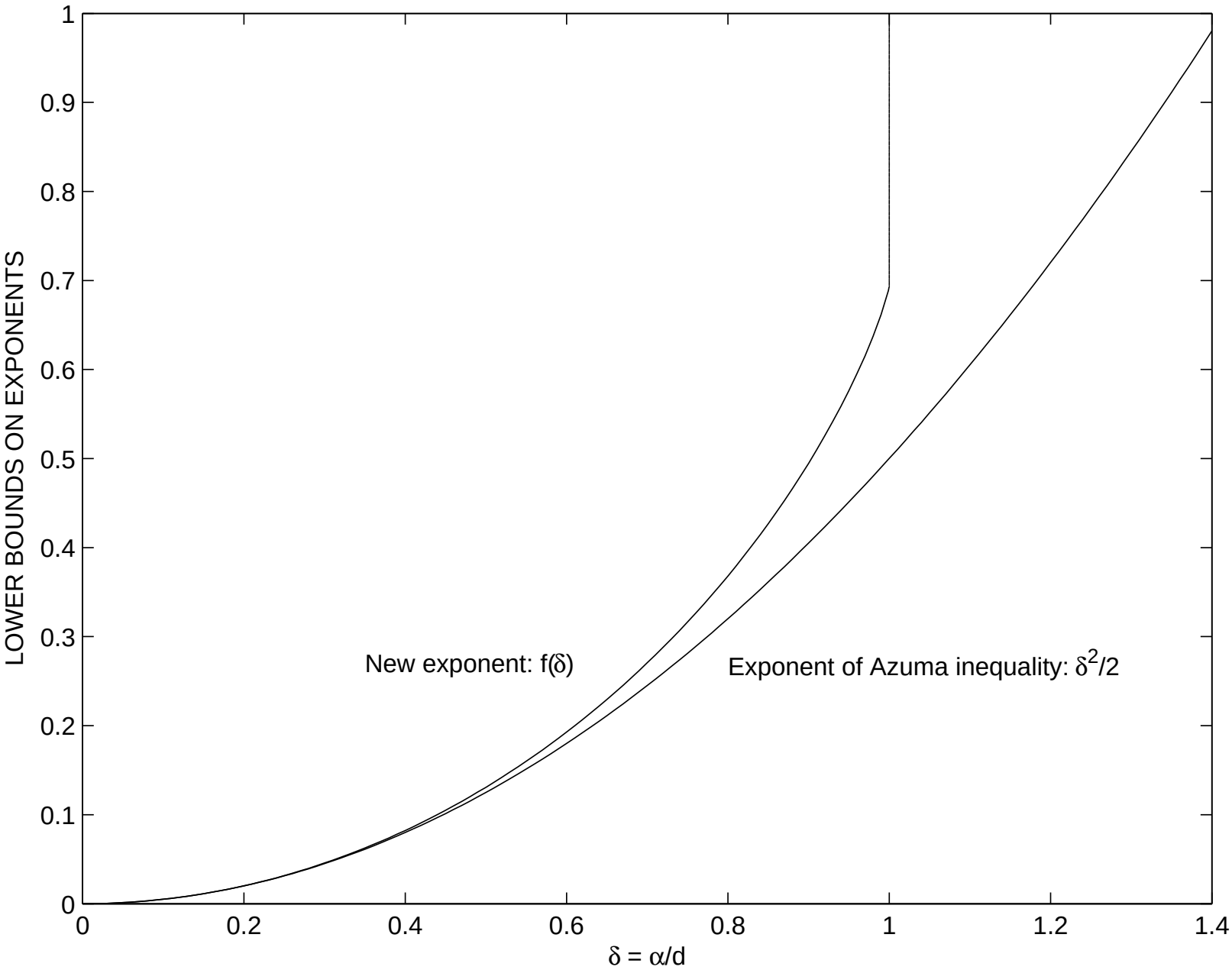
One possible tightening of Azuma's inequality

If $d_i = d$ for $i \in \{1, \ldots, n\}$, then by revisiting the proof of the Azuma-Hoeffding inequality, it is possible to improve it and get

$$\mathbb{P}(X_n - X_0 \geq n\alpha) \leq \exp\big(-n f(\delta)\big)$$

where $\delta \triangleq \frac{\alpha}{d}$, and

$$f(\delta) = \begin{cases} \ln(2)\Big[1 - h\left(\frac{1-\delta}{2}\right)\Big], & 0 \leq \delta \leq 1 \\ +\infty, & \delta > 1 \end{cases}$$


LOWER BOUNDS ON EXPONENTS
δ = α/d
New exponent: f(δ)
Exponent of Azuma inequality: δ²/2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4

This was just a **first step** towards refining Azuma's inequality.

More on refined versions of the Azuma-Hoeffding inequality, and on some of their applications in information theory and related topics are addressed in the paper:

I. Sason, "On Refined Versions of the Azuma-Hoeffding Inequality with Applications in Information Theory," a survey with some original results (an un-published work). See `http://arxiv.org/abs/1111.1977`.

The motivation

The survey paper is meant to stimulate the use of refined versions of the Azuma-Hoeffding inequality in information-theoretic aspects.