

Magnetization Enumerator for LDPC Codes - a Statistical Physics Approach

Jort van Mourik, David Saad
Neural Computing Research Group,
Aston University,
Birmingham B4 7ET, UK

Yoshiyuki Kabashima
Dept. of Computational Intelligence &
Systems Science,
Tokyo Institute of Technology,
Yokohama 2268502, Japan

Abstract — We propose a method based on the magnetization enumerator to determine the critical noise level for Gallager type low density parity check error correcting codes (LDPC). Our method provides an appealingly simple interpretation to the relation between different decoding schemes, and provides more optimistic critical noise levels than those reported in the information theory literature.

I. INTRODUCTION

Triggered by active investigations on LDPC in both information theory (IT) [1, 2, 3], and statistical physics (SP) [4, 5] communities, there is growing interest in the relationship between the IT and SP analysis. As the two communities investigate similar problems, one may expect techniques known in one framework to bring about new developments in the other. Here we present a direct SP method to determine the critical noise level of Gallager type LDPC codes which allows us to focus on the differences between various decoding criteria and to define the critical noise level for which decoding is theoretically feasible.

II. MAGNETIZATION ENUMERATOR

A Gallager code [1] encodes an K dimensional Boolean message $\vec{s} \in \{0, 1\}^K$ to a transmitted $N (> K)$ dimensional Boolean codeword \vec{t} . Upon transmission via a noisy channel (binary symmetric in this case), a noise \vec{n}° (flip rate p per bit), is added, to yield the received vector \vec{r} , with which a set of $(N - K)$ parity checks for the possible noise vectors is constructed. A (regular) code $\mathbf{A}(k, j)$ is characterized as having k noise components in each parity check, while each noise component is involved in j parity checks, such that the code rate is given by $R = 1 - j/k = K/N$. $\mathcal{I}_{pc}(\mathbf{A}, \vec{n}^\circ)$ is the set of noise vectors satisfying all parity checks. Any general decoding scheme is then based on selecting a vector \vec{n}^* from $\mathcal{I}_{pc}(\mathbf{A}, \vec{n}^\circ)$ using some noise statistics criterion. Upon successful decoding \vec{n}^* will be selected, while a decoding error is declared when a vector $\vec{n}^* \neq \vec{n}^\circ$ is selected. Upon mapping the field $(\{0, 1\}, +\text{mod}(2))$ onto $(\{1, -1\}, \times)$, and adapting the parity checks correspondingly, the statistics of a noise vector \vec{n} can be described by its magnetization $m(\vec{n}) \equiv \frac{1}{N} \sum_{l=1}^N n_l$. We calculate the logarithm (entropy) of the number of noise vectors in $\mathcal{I}_{pc}(\mathbf{A}, \vec{n}^\circ)$ with given magnetization m . Averaging this quantity over all possible codes \mathbf{A} (fixed k, j), and noise vectors \vec{n}° (denoted $\langle \cdot \rangle_{\mathbf{A}, \vec{n}^\circ}$), yields the magnetization enumerator, $\mathcal{M}(m) = \langle \frac{1}{N} \ln [\text{Tr}_{\vec{n} \in \mathcal{I}_{pc}(\mathbf{A}, \vec{n}^\circ)} \delta(m(\vec{n}) - m)] \rangle_{\mathbf{A}, \vec{n}^\circ}$.

$\mathcal{M}(m)$ is calculated using the replica method [5], allowing us to determine the critical noise level p_c [7]: $\mathcal{M}(m) > 0$ indicates typically an exponential number of solutions, while $\mathcal{M}(m) < 0$ indicates vanishing probability for a solution to

exist (for $N \rightarrow \infty$). Hence, if $\mathcal{M}(m(\vec{n}^\circ)) < 0$, decoding is in principle possible. The true noise \vec{n}° typically has the magnetization $m(\vec{n}^\circ) = 1 - 2p$, and $\mathcal{M}(m)$ only takes positive values in the interval $[m_-(p), m_+(p)]$.

The various decoding schemes can be summarized as follows: i) *Maximum likelihood (MAP) decoding* - selects that \vec{n} from \mathcal{I}_{pc} , which has the highest magnetization.

ii) *Typical pairs decoding* [3] - randomly selects an \vec{n} from \mathcal{I}_{pc} , that has the typical magnetization $m(\vec{n}) = 1 - 2p$.

iii) *Finite temperature (MPM) decoding* [4] - attributes an energy $-\frac{1}{2} \ln(\frac{1-p}{p})m(\vec{n})$ to each \vec{n} , and an \vec{n} from \mathcal{I}_{pc} is chosen with the magnetization that minimizes the free energy at temperature $T = 1$ (Nishimori's condition [5], corresponding to a proper choice of prior in the Bayesian framework).

All three decoding schemes can be shown to yield the same critical noise level p_c , and, using Nishimori's gauge theory [5], this can be shown to be exact. The following table compares our results with those reported in the IT literature, and with Shannon's critical noise level [6]:

(k, j)	(6, 3)	(5, 3)	(6, 4)	(4, 3)
Code rate	1/2	2/5	1/3	1/4
IT	0.0915	0.129	0.170	0.205
p_c (\mathcal{M})	0.0998	0.1365	0.1725	0.2095
Shannon	0.109	0.145	0.174	0.214

ACKNOWLEDGMENTS

Support by Grants-in-aid, MEXT (13680400 and 13780208) and JSPS (YK), The Royal Society and EPSRC-GR/N00562 (DS/JvM) is acknowledged.

REFERENCES

- [1] R. G. Gallager, "Low-density Parity-check Codes," *MIT Press, Cambridge, MA*, 1963.
- [2] D. J. C. MacKay, "Good Error-correcting Codes Based on Very Sparse Matrices," *IEEE Trans. Infor. Theory*, vol. 45, pp. 399-431, 1999.
- [3] S. Aji, H. Jin, A. Khandekar, D. J. C. MacKay, R. J. McEliece, "BSC Thresholds for Code Ensembles Based on "Typical Pairs" Decoding," In: Marcus, B., Rosenthal, J. (eds): "Codes, Systems and Graphical Models," *Springer Verlag, New York*, pp. 195-210, 2001.
- [4] Y. Kabashima, T. Murayama, D. Saad, "Typical Performance of Gallager-type Error-Correcting Codes," *Phys. Rev. Lett.*, vol. 84, pp. 1355-1358, 2000.
- [5] H. Nishimori, "Statistical Physics of Spin Glasses and Information Processing," *Oxford University Press, Oxford UK*, 2001.
- [6] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379-423; 623-656, 1948.
- [7] full papers: <http://www.ncrg.aston.ac.uk/People/vanmourj>