

Low-Density Parity-Check Codes with Rates Very Close to the Capacity of the q -ary Symmetric Channel for Large q

Amin Shokrollahi
Laboratoire d'algorithmique
EPFL

1015 Lausanne, Switzerland
e-mail: amin.shokrollahi@epfl.ch

Wei Wang¹
Dept. of EECS
UC Berkeley

Berkeley, CA 94720-1770, U.S.A.
e-mail: wangwei@eecs.berkeley.edu

Abstract — Transmission of packets over computer networks is subject to packet-level errors, which appear as “bursts” of bit-level errors and are not well modeled by memoryless binary channels. Using a standard scrambling technique [1] transmission of packets can be modeled by the q -ary symmetric channel (q -SC) with alphabet size q and error probability p . Furthermore, significant improvements in computational efficiency can be obtained by codes that operate at the packet-level instead of the bit-level. The capacity of the q -SC is $1 + p \log_q(p) + (1-p) \log_q(1-p) - p \log_q(1-q)$, which is close to $1-p$ for large q . [1] first designed an efficient decoding algorithm for LDPC codes on the q -SC, and showed that it can afford rates arbitrarily close to $1-2p$. We improve the analysis of this decoding algorithm to show that LDPC codes with the Tornado edge distribution, together with an erasure pre-code, can achieve rates ϵ -close to capacity. We also extend this decoder into a family of decoding algorithms which become progressively more powerful but also more complex. We show that in the limit, when the decoder is allowed to look infinitely deep into the decoding tree, it can achieve capacity without pre-coding. However computational requirements make this decoder impractical.

Verification decoding on LDPC graphs relies on the idea that if the sum of all neighboring variable nodes of a check node is zero, then with probability $1 - 1/(q-1)$ all the neighboring variable nodes are correct. The message passing decoding algorithm passes a proposed value and a 0/1-state indicating whether the value is verified. Consider the edge (v, c) between a variable node v and a check node c , and the tree associated with the neighborhood of v . Initially v has a received value r_v from the q -SC, and an unverified state $\xi = 0$.

1. v sends to c the message $(\omega, \xi) = (\omega_j, 1)$ if \exists neighbor j s.t. $\xi_j = 1$ or $\omega_j = r_v$, or if $\exists i \neq j$ s.t. $\omega_i = \omega_j$. Otherwise $(\omega, \xi) = (r_v, 0)$.
2. c sends to v the message $(\omega, \xi) = (\sum_i \omega_j, 1)$ if $\xi_j = 1 \forall j$. Otherwise $(\omega, \xi) = (\sum_i \omega_j, 0)$.

Let p_i and q_i be the probabilities that messages passed from variable to check nodes at the i th round of the decoding algorithm are {wrong and unverified}, or {correct and unverified}, respectively. Let $p_0 = p$ and $q_0 = 1 - p$. Then the decoding error recursions are [1] $p_{i+1} = p_0(\lambda(1 - \rho(1 - p_i)) + (\rho(1 - p_i) - \rho(1 - p_i - q_i))\lambda'(1 - \rho(1 - p_i)))$ and $q_{i+1} = (1 - p_0)\lambda(1 - \rho(1 - p_i))$. The Tornado distribution [2] with parameters $D \geq 1$ and $\alpha \geq 1$ is defined by the degree polynomials $\lambda(x) = \frac{1}{H_D} \sum_{k=1}^D \frac{x^k}{k}$

¹Work done while visiting EPFL, supported by NSF Graduate Fellowship and Lucent GRPW.

and $\rho(x) = e^{\alpha(x-1)}$, where $H_D = \sum_{i=1}^D 1/i$. The error recursions become $p_{i+1} \leq \theta p_i + \frac{\theta}{\alpha}(1 - e^{-\alpha q_i})$ and $q_{i+1} \leq \theta \eta p_i$, where $\theta = \alpha p / H_D$ and $\eta = (1 - p) / p$.

(p_i, q_i) converge if $p_{i+1} \leq \theta p_i + \frac{\theta}{\alpha}(1 - e^{-\alpha \theta \eta p_{i-1}})$ converges. By Lemma 1 a sufficient condition for convergence of p_i to a nonzero error probability x is $\theta + \theta^2 \eta e^{-\alpha \theta \eta x} < 1$. The remaining uncorrected errors can be handled by an erasure code, such as the Tornado code [2]. The combined rate of the LDPC and erasure codes is $R = (1 - (1 + \theta \eta)x) \left(1 - \frac{p}{\theta} \frac{(1 - e^{-\alpha})}{(1 - 1/(D+1))}\right) \approx (1 - \epsilon)(1 - p)$ if $\alpha \gg 1$, $D \gg 1$ and $\theta = 1 - \delta$. The condition for an error recursion fixed point at $x > 0$ then implies that $\alpha = O(1/\epsilon)$. Thus the complexity of the verification decoding, which is proportional to the number of edges in the LDPC graph with n variable nodes, is $O(n/\epsilon)$. The probability of decoding error is the probability of false verification.

Lemma 1. Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ be a continuous function with partial derivatives $D_1 f$ and $D_2 f$. Call $x \in \mathbb{R}$ a fixed point of f if $f(x, x) = x$. Let $x_0 \in \mathbb{R}$, $x_1 = f(x_0, 1 - x_0)$, and $x_{n+1} = f(x_n, x_{n-1})$ for $n \geq 1$. If there are constants A and B such that $A + B < 1$, $|D_1 f| \leq A$ and $|D_2 f| \leq B$, then a unique fixed point x of f exists, and $x = \lim_{n \rightarrow \infty} x_n$.

Theorem 1. Over n uses of the q -SC with error probability p , an erasure pre-code and a LDPC code with the Tornado distribution can together achieve a rate of $(1 - \epsilon)(1 - p)$. The computational complexity of the encoder and the decoder is $O(\frac{1}{\epsilon} n \log q)$, and the decoding error probability is $O(n/q)$.

Now consider a message passing decoding algorithm that looks infinitely deep into the decoding tree.

1. v sends to c the message $(\omega, \xi) = (\omega_j^k, 1)$ if $\exists j$ s.t. $\xi_j = 1$, or if $\exists j, k$ s.t. $\omega_j^k = r_v$, or if $\exists i \neq j, k, l$ s.t. $\omega_i^k = \omega_j^l$. Otherwise $(\omega, \xi) = ([r_v, \omega_1, \dots, \omega_{d_v-1}], 0)$.
2. c sends to v the message $(\omega, \xi) = (\sum_i \omega_j^1, 1)$ if $\xi_j = 1 \forall j$. Otherwise $(\omega, \xi) = ([\Sigma], 0)$, where $[\Sigma]$ denotes the vector of all possible sums of one element from each ω_j .

The error recursions for this decoding algorithm are $p_{i+1} = p\lambda(1 - \rho(1 - p_i))$ and $q_{i+1} = p(\rho(1 - p_i) - \rho(1 - p_i - q_i))\lambda'(1 - \rho(1 - p_i)) + (1 - p)\lambda(1 - \rho(1 - p_i))$. Applying the Tornado distribution, the error recursions become $p_{i+1} \leq \theta p_i$ and $q_{i+1} \leq \theta \eta p_i + \frac{\theta}{\alpha}(1 - e^{-\alpha q_i})$. Thus $(p_i, q_i) \rightarrow (0, 0)$ if $\theta < 1$.

Theorem 2. A LDPC code with the Tornado distribution can achieve a rate of $1 - p$ over the q -SC with error probability p .

REFERENCES

- [1] M. Luby and M. Mitzenmacher. Verification codes. *Proc. Allerton Conf. on Communication, Control, and Computing*, 2002.
- [2] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman. Efficient erasure correcting codes. *IEEE Trans. Inform. Theory*, 47:569–584, 2001.