Massage-Passing Algorithms and Improved LP Decoding

Sanjeev Arora Constantinos Daskalakis David Steuer

presented by Tomáš Gavenčiak

Parity check code $\mathbb{C}(G)$ defined by a bipartite graph $G = (V_L \cup V_R, E)$ is the set of 0/1 assignments to V_L such that for all $j \in V_R$, $\sum_{i \in N(j)} \equiv 0 \bmod 2$. V_L are the variable nodes, V_R the check nodes. $|V_L| = n$, $|V_R| = m$.

Low-density parity check code is a parity check code defined by a graph with constant (or bounded) degrees, d_L and d_R.

Nearest codeword problem. Given $y \in \{0,1\}^n$, find $x \in \mathbb{C}(G)$ with minimal $|x - y|_1$.

LP decoding program for the nearest codeword problem. Given $y \in \{0,1\}^n$, minimize $|x - y|_1$ subject to

$$x \in \bigcap_{j \in V_R} \mathrm{Conv}\mathbb{C}_j.$$

Main theorem 1. Let G be a $(3,6)$-regular bipartite graph of girth $\Omega(\log n)$, $p \leq 0.05$ and $x \in \{0,1\}^n$ a codeword of $\mathbb{C}(G)$. Suppose y was obtained from x by flipping each bit independently with probability p. Then with probability at least $1 - \exp(-n^\gamma)$ for some $\gamma > 0$

- x is the optimal solution to the LP decoding algorithm.
- a simple message-passing (dynamic programming) algorithm computes the codeword x from y and certifies that it is the codeword nearest to y in time $O(n \log n)$.

Note. The girth requirement can be lowered to $\Omega(\log \log n)$ making the decoding probability $1 - 1/\mathrm{poly}(n)$ and running time $O(n \log \log n)$.

T-local deviation at $i_0 \in V_L$ is an assignment $\beta \in \{0,1\}^n$ with $\beta_{i_0} = 1$ and satisfying all the checks in $N^{2T}(i_0)$. A T-local deviation is *minimal* if all check nodes in $N^{2T}(i_0)$ have 0 or 2 neighbours set to 1 and all variable nodes outside $N^{2T}(i_0)$ are set to 0.
For a minimal T-local deviation at i_0 β and weights $w = (w_1, \ldots, w_T)$, the *w-weighted deviation* $\beta^{(w)}$ has $\beta_i^{(w)} = \beta_i w_t$ if $\mathrm{dist}(i_0, i) = 2t, 1 \leq t \leq T$ and $\beta_i^{(w)} = 0$ otherwise.

A codeword $x \in \{0,1\}^n$ is (T, w)-locally optimal for $y \in \{0,1\}^n$ if for all T-local deviations β,

$$|x \oplus \beta^{(w)} - y|_1 > |x - y|_1,$$

where $(a \oplus b)_i = |a_i - b_i|_1$. Note that the T-local deviation at i minimizing the left side can be computed by dynamic programming.

Theorems 2-4. Let $T < \frac{1}{4}$ girth(G) and $w = (w_1, \ldots, w_T) \geq 0$. If x is a (T, w)-locally optimal codeword for $y \in \{0,1\}^n$, then

(2) x is the unique nearest codeword for y. *(local optimality certificate)*
(3) the w-weighted *min-sum algorithm* computes x in T iterations.
(4) x is the unique optimal solution to the LP decoding program.

Theorem 5 (local optimality of a codeword). Let G be a (d_L, d_R)-regular bipartite graph and $T < \frac{1}{4}\,\mathrm{girth}(G)$. Let $0 < p < 1$ and $x \in \{0,1\}^n$ a codeword of $\mathbb{C}(G)$. Suppose y was obtained from x by flipping every bit independently with probability p.

1. If d_L, d_R and p satisfy a technical condition (which is met for $d_L = 3$, $d_R = 6$ and $p \leq 0.02$) then x is $(T,1)$-locally optimal with probability at least $1 - nc^{-(d_L-1)^T}$ for some constant $c > 1$.

2. If d_L, d_R and p satisfy another technical condition (which is met for $d_L = 3$, $d_R = 6$ and $p \leq 0.0247$) then there is $w \in [0,1]^T$ such that x is (T,w)-locally optimal with probability at least $1 - nc^{-(d_L-1)^T}$ for some constant $c > 1$.

3. If $d_L = 3$, $d_R = 6$ and $p \leq 0.05$ then x is (T,w)-locally optimal with probability at least $1 - nc^{-2^T}$ for some weight-vector $w \leq 0$ and some constant $c > 1$. x is $(T,\nVdash)$-locally optimal with probability at least $1 - nc^{-(d_L-1)^T}$ for some $c > 1$.

Lemma 1. Let $T < \frac{1}{4}\,\mathrm{girth}(G)$. Then for every codeword $z \neq 0$ there is a distribution over minimal T-local deviations β, such that for every weight-vector $w \in [0,1]^T$,

$$\mathbb{E}\beta^{(w)} = \alpha z$$

for some scaling-constant $\alpha \leq 0$.

Lemma 2. Let $T < \frac{1}{4}\,\mathrm{girth}(G)$ and $w \in [0,1]^T$. Then for every non-zero LP solution $z \in [0,1]^n$, there is a distribution over minimal T-local deviations β such that

$$\mathbb{E}\beta^{(w)} = \alpha z$$

for some scaling-constant $\alpha \leq 0$.

Lemma 3. Let x be a codeword and x' an LP solution. Then $x \oplus x'$ is also an LP solution.

Lemma 4. Let z be a non-zero LP solution. There are functions p_j for every $j \in V_R$, $p_j : N(j) \times N(j) \to [0,1]$ such that for every $i \in N(j)$,

$$z_i = \sum_{i' \in N(j) \setminus \{i\}} p_j(i, i')$$

and for $i, i' \in N(j)$ symmetric, i.e. $p_j(i,i') = p_j(i', i)$.