UCLP($\mathcal{P}$) for an Explicit Constant-Width SAT Family: Local Conditioning Bounds and a Linear Transcript-Capacity Lower Bound

1 Objects and conventions

All logarithms are base 2. Entropies and mutual informations are measured in bits. A *CNF* formula F is a conjunction of clauses over Boolean variables. A *width-w CNF* has every clause of size at most w.

Let X be a random variable uniformly distributed over the satisfying assignments of F. For any random variable Y, define Shannon entropy $\mathrm{H}(Y)$ and mutual information $\mathrm{I}(X;Y)$ in the standard way. For a filtration (transcript) $(S_t)_{t=0}^m$, define the *transcript capacity*

$$\mathrm{TC}(S_{0:m}) \ := \ \sum_{t=1}^{m} \mathrm{I}(X; S_t \mid S_{t-1}).$$

1.1 Canonical r-local transcripts

Fix an integer radius $r \geq 1$. A *canonical r-local transcript* for inputs of size n is a sequence $T = (S_0, S_1, \ldots, S_m)$ where each transition $S_{t-1} \to S_t$ is computed from the current state by inspecting only an r-neighborhood in a bounded-degree representation graph of the state. The only property used in this document is the following locality-to-conditioning principle.

Definition 1 (Local conditioning proxy)**.** *Fix a bounded-degree bipartite factor graph representation $\mathcal{G}(F)$ of a CNF F (variables–clauses incidence with bounded degrees). Let $\mathrm{Ball}_{\mathcal{G}}(u,r)$ denote the radius-r ball around a node u in $\mathcal{G}(F)$. A transcript step is* r-local *if, conditioned on the prior state, its new information is measurable with respect to the sigma-algebra generated by the labels/constraints within some radius-r ball $\mathrm{Ball}_{\mathcal{G}}(u,r)$.*

Accordingly, it suffices to bound the mutual information about X revealed by conditioning on any fixed radius-r ball in $\mathcal{G}(F)$.

2 An explicit constant-width SAT family from LDPC/Tseitin constraints

2.1 LDPC parity-check instances

Fix constants $\Delta_v, \Delta_c \geq 3$. For each n, let H_n be a bipartite graph with:

- variable nodes V_n with $|V_n| = n$,
- check nodes C_n with $|C_n| = \Theta(n)$,

- every $v \in V_n$ has degree Δ_v, every $c \in C_n$ has degree Δ_c,
- girth $\mathrm{girth}(H_n) \geq 4r + 4$.

Such families exist (explicit constructions are known) and can be taken to be expanders; expansion is not used for the linear lower bound, only bounded degree and girth.

Let $A_n \in \mathbb{F}_2^{|C_n| \times n}$ be the incidence matrix of H_n (row for each check, column for each variable). Fix a right-hand side $b_n \in \mathbb{F}_2^{|C_n|}$ such that the linear system

$$A_n x = b_n \quad \text{over } \mathbb{F}_2$$

is consistent. Define $\mathsf{Sol}_n := \{x \in \mathbb{F}_2^n : A_n x = b_n\}$.

Proposition 2 (Solution space size)**.** *If $A_n x = b_n$ is consistent then Sol_n is an affine subspace of $\mathbb{F}_2^n$ of size $|\mathsf{Sol}_n| = 2^{n - \mathrm{rank}(A_n)}$.*

2.2 Constant-width Tseitin-to-CNF encoding

We encode each check equation (a parity constraint on Δ_c bits) using a standard constant-size CNF gadget with auxiliary variables.

Fix Δ_c. For a check node $c \in C_n$ with neighbors $N(c) = \{v_1, \dots, v_{\Delta_c}\}$, introduce auxiliary variables $y_{c,1}, \dots, y_{c,\Delta_c - 1}$ and enforce:

$$y_{c,1} = x_{v_1} \oplus x_{v_2}, \quad y_{c,2} = y_{c,1} \oplus x_{v_3}, \quad \dots, \quad y_{c,\Delta_c-1} = y_{c,\Delta_c-2} \oplus x_{v_{\Delta_c}},$$

and finally $y_{c,\Delta_c-1} = b_n(c)$. Each XOR relation $z = u \oplus v$ can be expressed by a width-3 CNF using four clauses:

$$(z \vee u \vee v) \wedge (z \vee \neg u \vee \neg v) \wedge (\neg z \vee u \vee \neg v) \wedge (\neg z \vee \neg u \vee v).$$

Also $z = b$ is enforced by a unit clause.

Let F_n be the conjunction of all such gadget clauses over all checks $c \in C_n$.

Proposition 3 (Width, bounded degree, and satisfiable assignments)**.** *Each F_n is a width-3 CNF. Moreover, there is a bijection between Sol_n and satisfying assignments of F_n restricted to the original variables x_v: every $x \in \mathsf{Sol}_n$ extends uniquely to a satisfying assignment of F_n over all variables (including the auxiliaries), and every satisfying assignment restricts to an $x \in \mathsf{Sol}_n$.*

Proof. Width-3 is immediate from the XOR CNF encoding above. Given $x \in \mathsf{Sol}_n$, the sequential definitions force each $y_{c,j}$ uniquely, and the gadget clauses are satisfiable. Conversely, the XOR gadgets enforce the parity equalities, hence the restriction x must satisfy $A_n x = b_n$. □

2.3 Factor graph and locality

Let $\mathcal{G}(F_n)$ be the variable–clause incidence graph of the CNF F_n. Because H_n has bounded degrees and each check contributes a constant-size gadget, $\mathcal{G}(F_n)$ has bounded degree (independent of n), and the girth condition on H_n implies that radius-r neighborhoods in $\mathcal{G}(F_n)$ are acyclic (trees) after possibly increasing r by a constant depending only on Δ_c.

3 Local conditioning reveals only constant information

Let X be uniform over satisfying assignments of F_n (equivalently, uniform over Sol_n pushed forward through the unique extension map).

3.1 Linear-algebraic view

Because satisfying assignments correspond to an affine $\mathbb{F}_2$-subspace on the original variables, X induces an affine distribution on those variables, and the auxiliary variables are deterministic functions of the originals. Hence mutual information about X revealed by any local ball is bounded by the number of independent linear constraints that the ball imposes.

Lemma 4 (Tree-local constraints have constant rank)**.** *Fix $r \geq 1$. There exists a constant $R = R(r, \Delta_v, \Delta_c)$ such that for every n, and every radius-r ball B in the factor graph $\mathcal{G}(F_n)$, the set of parity constraints on the original variables implied by the clauses inside B has $\mathbb{F}_2$-rank at most R.*

Proof. Because $\mathcal{G}(F_n)$ has bounded degree, the ball B contains at most $O(1)$ variables and clauses. Each XOR-gadget clause set inside B enforces a constant number of parity relations among the variables present in B. All such relations live in a vector space over $\mathbb{F}_2$ of dimension at most the number of original variables in B. Therefore the rank is bounded by a constant R. □

Lemma 5 (Local conditioning leaks $O(1)$ bits)**.** *Fix $r \geq 1$ and let B be any radius-r ball in $\mathcal{G}(F_n)$. Let Z_B be the complete induced labeled sub-instance inside B. Then*

$$\mathrm{I}(X; Z_B) \ \leq \ R(r, \Delta_v, \Delta_c).$$

Proof. Auxiliary variables are deterministic functions of original variables under satisfiable assignments, so Z_B is a deterministic function of a constant-sized restriction of X. Thus $\mathrm{H}(Z_B) \leq R$ and $\mathrm{I}(X; Z_B) \leq \mathrm{H}(Z_B) \leq R$. □

3.2 Per-step bound

Corollary 6 (Per-step capacity bound)**.** *Fix $r \geq 1$. For any canonical r-local transcript $T = (S_t)_{t=0}^m$ operating on F_n and any $t \geq 1$,*

$$\mathrm{I}(X; S_t \mid S_{t-1}) \leq R(r, \Delta_v, \Delta_c).$$

Proof. By r-locality, conditioned on S_{t-1}, the new information in S_t is measurable with respect to some Z_B. Hence $\mathrm{I}(X; S_t \mid S_{t-1}) \leq \mathrm{I}(X; Z_B) \leq R$. □

4 A linear transcript-capacity lower bound

Lemma 7 (Entropy requirement)**.** *Let X be uniform over satisfying assignments of F_n. Then*

$$\mathrm{H}(X) = \log_2 |\mathsf{Sol}_n| = n - \mathrm{rank}(A_n).$$

Theorem 8 (Linear lower bound on transcript capacity)**.** *Fix $r \geq 1$. Assume* $\mathrm{rank}(A_n) \leq (1-\gamma)n$ *for some constant $\gamma > 0$. Let $T = (S_t)_{t=0}^m$ be any canonical r-local transcript that determines X uniquely. Then*

$$\mathrm{TC}(T) \ \geq \ \gamma n.$$

Proof. If X is determined by S_m, then $\mathrm{H}(X \mid S_m) = 0$ and

$$\mathrm{TC}(T) = \mathrm{I}(X; S_m) = \mathrm{H}(X) \geq \gamma n.$$

□

Corollary 9 (UCLP($\mathcal{P}$) is NO)**.** *Let $\mathcal{P} = \{F_n\}$ be the SAT family above. Then any fixed-radius local transcript solving $\mathcal{P}$ must have transcript capacity $\Omega(n)$.*

5 Remarks on $\Omega(n \log n)$ strengthening

The present argument establishes an unconditional linear lower bound. An $\Omega(n \log n)$ bound requires additional information-diffusion or entropy-decay mechanisms not assumed here.