COVERTRACE-SAT as Disjoint-Subcube Knowledge Compilation: Worst-Case Fragmentation, Conditional PH Collapse, and Connections to Geometric Complexity Theory

Oscar Riveros

https://github.com/maxtuno/

https://independent.academia.edu/oarr

January 2026

Abstract

We study *COVERTRACE-SAT*, an exact SAT/#SAT algorithm that interprets a CNF formula as a union of *forbidden axis-aligned subcubes* in the Boolean hypercube $\Omega_n = \{0,1\}^n$ and maintains this union as a *disjoint* family of subcubes. The disjointness invariant yields immediate exact model counting by volume additivity and enables constructive witness extraction. We formalize the core algorithm, prove correctness, and analyze complexity in terms of *fragmentation* (the number of disjoint subcubes maintained). We establish a tight exponential worst-case lower bound: the odd-parity set in Ω_n requires 2^{n-1} disjoint axis-aligned subcubes, implying that COVERTRACE-style compilation can require exponential space/time on explicit inputs. We then place COVERTRACE within the *knowledge compilation* landscape: the maintained family is precisely a deterministic DNF / disjoint sum of products (DSOP) for $\neg F$. This yields a clean conditional consequence: if every CNF admitted a *uniform* polynomial-time compilation into a polynomial-size disjoint-subcube cover of its forbidden region, then #SAT $\in$ P and the Polynomial-Time Hierarchy collapses to P (via Toda's theorem). Finally, motivated by parity-like obstructions, we propose an *affine* extension (subspaces over $\mathbb{F}_2$) that compresses parity exponentially, and we articulate a GCT-inspired obstruction program by viewing disjoint subcubes as restricted rank-one tensors. We emphasize that lower bounds for a specific representation language do *not* constitute a proof of P $\neq$ NP, but they do suggest concrete, mathematically precise barriers and open problems.

Contents

1 Introduction

COVERTRACE-SAT is an exact algorithmic viewpoint on SAT and #SAT based on an elementary geometric duality: every clause of a CNF forbids exactly the assignments that falsify it, and those assignments form an axis-aligned subcube of the Boolean hypercube $\Omega_n = \{0,1\}^n$. If F is a CNF formula, let $U(F) \subseteq \Omega_n$ denote the set of assignments that falsify at least one clause. Then

$$\#\mathrm{SAT}(F) = 2^n - |U(F)|.$$

COVERTRACE-SAT incrementally constructs a *disjoint* family of forbidden subcubes whose union is exactly $U(F)$. This disjointness invariant transforms inclusion–exclusion into exact additivity.

The price of exactness is *fragmentation*: adding a new forbidden cube to a disjoint union may force splitting existing cubes into many pieces. This paper provides a rigorous treatment of this phenomenon and connects it to three broader frameworks:

1. **Knowledge compilation:** the family maintained by COVERTRACE is a deterministic DNF / disjoint sum of products (DSOP) for $\neg F$.
2. **Complexity theory:** a hypothetical *uniform* polynomial-time compilation of arbitrary CNFs to polynomial-size DSOP would imply #SAT $\in$ P and collapse the Polynomial-Time Hierarchy (PH).
3. **Geometric Complexity Theory (GCT):** disjoint subcubes correspond to a restricted class of rank-one tensors; fragmentation lower bounds become restricted tensor-rank lower bounds, suggesting a discrete "obstruction" viewpoint.

What this paper is *not*. Despite the geometric language, we do *not* claim any proof of P $\neq$ NP. We provide (i) exact correctness results, (ii) unconditional worst-case lower bounds for a specific representation language, and (iii) conditional collapses under strong uniform compilation hypotheses. Bridging such lower bounds to P $\neq$ NP would require universal lower bounds against *all* polynomial-time algorithms, far beyond a single representation.

2 Geometric semantics of CNF

2.1 Patterns and subcubes

Definition 2.1 (Patterns and subcubes)**.** A *pattern* is a vector $p \in \{0, 1, \bullet\}^n$, where $\bullet$ denotes "free". The *subcube* induced by p is

$$Q(p) \;=\; \{x \in \Omega_n : \forall i \in [n],\; p_i \neq \bullet \Rightarrow x_i = p_i\}.$$

The *support* is $\mathrm{supp}(p) = \{i : p_i \in \{0, 1\}\}$ and the *width* is $k(p) = |\mathrm{supp}(p)|$. The volume is $\mathrm{vol}(p) = |Q(p)| = 2^{n-k(p)}$.

Definition 2.2 (Disjoint families)**.** A family $\mathcal{U} \subseteq \{0, 1, \bullet\}^n$ is *disjoint* if $Q(u) \cap Q(v) = \emptyset$ for all distinct $u, v \in \mathcal{U}$. Its covered set is $\mathrm{Cov}(\mathcal{U}) = \bigcup_{u\in\mathcal{U}} Q(u)$.

2.2 Clauses induce forbidden subcubes

Let F be a CNF formula on variables $x_1, \ldots, x_n$ with clauses $C_1, \ldots, C_m$.

Definition 2.3 (Clause-to-cube map)**.** For a clause C, define the pattern $p(C) \in \{0, 1, \bullet\}^n$ coordinate-wise:

$$p(C)_i \;=\; \begin{cases} 0 & \text{if } x_i \in C, \\ 1 & \text{if } \neg x_i \in C, \\ \bullet & \text{if } x_i \notin C \text{ and } \neg x_i \notin C. \end{cases}$$

Proposition 2.4 (Falsifying assignments form a subcube)**.** *An assignment $x \in \Omega_n$ falsifies C if and only if $x \in Q(p(C))$.*

Definition 2.5 (Forbidden region)**.** The *forbidden region* of F is

$$U(F) \;=\; \bigcup_{j=1}^{m} Q(p(C_j)).$$

Then $\#\mathrm{SAT}(F) = 2^n - |U(F)|$ and F is satisfiable iff $U(F) \neq \Omega_n$.

3 The COVERTRACE core algorithm

3.1 Elementary cube relations

We use two fundamental relations on patterns $p, r \in \{0, 1, \bullet\}^n$.

- **Intersection:** $Q(p) \cap Q(r) \neq \emptyset$ iff p and r do not disagree on any jointly-fixed coordinate.
- **Containment:** $Q(p) \subseteq Q(r)$ iff every coordinate fixed by r is also fixed by p with the same value.

3.2 Disjoint difference CubeDiff

The engine of COVERTRACE is a disjoint decomposition of a cube difference.

Problem 3.1 (CubeDiff)**.** Given patterns $p, r \in \{0, 1, \bullet\}^n$, compute a disjoint family $\mathcal{D} \subseteq \{0, 1, \bullet\}^n$ such that

$$Q(p) \setminus Q(r) \;=\; \bigsqcup_{d \in \mathcal{D}} Q(d),$$

where $\bigsqcup$ denotes disjoint union.

A standard constructive solution splits p along a coordinate fixed in r but free in p.

Algorithm 1 $\mathrm{CubeDiff}(p, r)$: disjoint decomposition of $Q(p) \setminus Q(r)$

Require: $p, r \in \{0, 1, \bullet\}^n$
Ensure: disjoint family $\mathcal{D}$ with $\bigsqcup_{d \in \mathcal{D}} Q(d) = Q(p) \setminus Q(r)$

if $Q(p) \cap Q(r) = \emptyset$ **then return** $\{p\}$
end if
if $Q(p) \subseteq Q(r)$ **then return** $\emptyset$
end if
choose an index i such that $p_i = \bullet$ and $r_i \in \{0, 1\}$
let $b \leftarrow r_i$
define $p^{\neq}$ by fixing $p_i \leftarrow 1 - b$
define $p^{=}$ by fixing $p_i \leftarrow b$
return $\{p^{\neq}\} \;\cup\; \mathrm{CubeDiff}(p^{=}, r)$

Lemma 3.2 (Size bound for a single difference)**.** *Let $p, r \in \{0, 1, \bullet\}^n$ with $Q(p) \cap Q(r) \neq \emptyset$ and $Q(p) \not\subseteq Q(r)$. Then* $\mathrm{CubeDiff}(p, r)$ *returns at most* $|\mathrm{supp}(r) \setminus \mathrm{supp}(p)|$ *subcubes, and runs in* $O(n \cdot |\mathrm{supp}(r) \setminus \mathrm{supp}(p)|)$ *time in a direct pattern representation.*

Proof. Each recursive step fixes one coordinate i that is fixed in r but free in p. Thus the recursion depth is at most $|\mathrm{supp}(r) \setminus \mathrm{supp}(p)|$. Exactly one cube is emitted per recursion level (the branch $p^{\neq}$), and the recursion terminates when the remaining branch becomes a subset of r. □

3.3 Incremental disjointization of the forbidden union

The next procedure adds a new forbidden cube Q into an existing disjoint family $\mathcal{U}$ by disjointizing Q against all existing cubes.

Algorithm 2 AddCube($\mathcal{U}, Q$): add Q to a disjoint family

Require: disjoint family $\mathcal{U} \subseteq \{0, 1, \bullet\}^n$, pattern $Q \in \{0, 1, \bullet\}^n$
Ensure: disjoint family $\mathcal{U}'$ with $\mathrm{Cov}(\mathcal{U}') = \mathrm{Cov}(\mathcal{U}) \cup Q(Q)$ and the added volume Δ

$\mathcal{P} \leftarrow \{Q\}$ ▷ pieces of Q not yet subtracted
for each $R \in \mathcal{U}$ **do**
  $\mathcal{P}_{\text{new}} \leftarrow \emptyset$
  for each $p \in \mathcal{P}$ **do**
    $\mathcal{P}_{\text{new}} \leftarrow \mathcal{P}_{\text{new}} \cup \mathrm{CubeDiff}(p, R)$
  end for
  $\mathcal{P} \leftarrow \mathcal{P}_{\text{new}}$
  if $\mathcal{P} = \emptyset$ **then**
    return $(\mathcal{U}, 0)$
  end if
end for
$\Delta \leftarrow \sum_{p \in \mathcal{P}} 2^{n-k(p)}$
$\mathcal{U}' \leftarrow \mathcal{U} \cup \mathcal{P}$ ▷ disjoint union since $\mathcal{P}$ is outside $\mathrm{Cov}(\mathcal{U})$
return $(\mathcal{U}', \Delta)$

3.4 COVERTRACE-SAT

Algorithm 3 COVERTRACE-SAT (core, no heuristics)

Require: CNF $F = \bigwedge_{j=1}^{m} C_j$ over n variables
Ensure: UNSAT if F is unsatisfiable; otherwise SAT with a witness assignment

$\mathcal{U} \leftarrow \emptyset$; $y \leftarrow 2^n$ ▷ y tracks surviving assignments
for $j = 1, \ldots, m$ **do**
  $Q \leftarrow p(C_j)$
  $(\mathcal{U}, \Delta) \leftarrow \mathrm{AddCube}(\mathcal{U}, Q)$
  $y \leftarrow y - \Delta$
  if $y = 0$ **then**
    return UNSAT
  end if
end for
extract some $x \in \Omega_n \setminus \mathrm{Cov}(\mathcal{U})$
return SAT with witness x

4 Correctness and exact model counting

Theorem 4.1 (Invariant and correctness). *After processing clauses $C_1, \dots, C_t$, the family $\mathcal{U}_t$ is disjoint and satisfies*

$$\mathrm{Cov}(\mathcal{U}_t) \;=\; \bigcup_{j=1}^{t} Q(p(C_j)).$$

Moreover,

$$y_t \;=\; 2^n - |\mathrm{Cov}(\mathcal{U}_t)| \;=\; \#\mathrm{SAT}\Big(\bigwedge_{j=1}^{t} C_j\Big).$$

Hence COVERTRACE-SAT outputs UNSAT *iff F is unsatisfiable, and otherwise outputs* SAT *with at least one satisfying assignment.*

Proof. We proceed by induction on t. For $t = 0$, $\mathcal{U}_0 = \emptyset$ and the claim holds.

Assume the claim for $t-1$. At step t, AddCube$(\mathcal{U}_{t-1}, Q)$ computes $\mathcal{P}$, a disjoint family whose union is exactly

$$Q(Q) \setminus \mathrm{Cov}(\mathcal{U}_{t-1}).$$

Subtracting sequentially over all $R \in \mathcal{U}_{t-1}$ computes a disjoint decomposition of $Q(Q) \setminus \mathrm{Cov}(\mathcal{U}_{t-1})$ because each call to CubeDiff maintains a disjoint decomposition of a difference. Since $\mathcal{P}$ is disjoint from $\mathrm{Cov}(\mathcal{U}_{t-1})$, the union $\mathcal{U}_t = \mathcal{U}_{t-1} \cup \mathcal{P}$ is disjoint and covers

$$\mathrm{Cov}(\mathcal{U}_t) = \mathrm{Cov}(\mathcal{U}_{t-1}) \cup Q(Q) = \bigcup_{j=1}^{t} Q(p(C_j)).$$

The volume update is exact by disjointness, so $y_t = 2^n - |\mathrm{Cov}(\mathcal{U}_t)|$ equals the number of satisfying assignments of the prefix CNF. Finally, F is unsatisfiable iff $y_m = 0$, and if $y_m > 0$ then any extracted $x \notin \mathrm{Cov}(\mathcal{U}_m)$ satisfies all clauses. $\square$

4.1 Witness extraction

If $y_m > 0$, one may extract a witness by finding any point outside $\mathrm{Cov}(\mathcal{U}_m)$. There are many ways; one constructive approach is a greedy coordinate-by-coordinate procedure maintaining an upper bound on how many points remain covered under partial assignments.

Proposition 4.2 (Witness extraction is polynomial in the representation size). *Given a disjoint family $\mathcal{U} \subseteq \{0, 1, \bullet\}^n$ with $\mathrm{Cov}(\mathcal{U}) \neq \Omega_n$, one can compute $x \in \Omega_n \setminus \mathrm{Cov}(\mathcal{U})$ in time $O(n \cdot |\mathcal{U}|)$ using only membership tests against the cubes.*

5 Complexity in terms of fragmentation

The running time of COVERTRACE-SAT is governed by how large $\mathcal{U}$ becomes.

Definition 5.1 (Fragmentation parameters). Let $\mathcal{U}_t$ be the disjoint family after t insertions. Define the *fragmentation* $S_t = |\mathcal{U}_t|$ and $S = \max_t S_t$. Let T be the total number of cube pieces produced by all calls to CubeDiff across the run (counting multiplicity).

Proposition 5.2 (Parameterized time and space bounds). *The core algorithm runs in time $O(T \cdot c(n))$ and uses space $O(S \cdot c(n))$, where $c(n)$ is the cost of primitive cube operations under the chosen representation (e.g., $c(n) = O(1)$ with bitmasks plus occasional $O(n)$ scans). In the worst case, T, S can be $2^{\Theta(n)}$.*

Proof sketch. Each produced cube piece is generated, stored, and later processed by primitive operations a constant number of times. Hence total time is linear in the total number of produced pieces times the cost of primitives. Space is dominated by the maximum family size. Exponentiality follows from the lower bound in Section 6. □

6 A tight exponential worst-case lower bound

We give an unconditional worst-case lower bound for any disjoint axis-aligned subcube cover of an explicit set, implying that COVERTRACE-SAT must use exponential space/time on some instances.

6.1 Parity as an extremal obstruction

For $x \in \Omega_n$, define parity $\mathrm{par}(x) = x_1 \oplus \cdots \oplus x_n$ and let $O_n = \{x : \mathrm{par}(x) = 1\}$.

Lemma 6.1 (No nontrivial axis-aligned cube is monochromatic for parity)**.** *If $p \in \{0, 1, \bullet\}^n$ has at least one free coordinate ($k(p) < n$), then $Q(p)$ contains both even- and odd-parity points.*

Proof. Choose a free coordinate i of p and any $x \in Q(p)$. Let x' be x with bit i flipped. Then $x' \in Q(p)$ but $\mathrm{par}(x') = \mathrm{par}(x) \oplus 1$. □

Theorem 6.2 (Tight DSOP lower bound for parity)**.** *Any disjoint family $\mathcal{U} \subseteq \{0, 1, \bullet\}^n$ with $\mathrm{Cov}(\mathcal{U}) = O_n$ must satisfy $|\mathcal{U}| = 2^{n-1}$, and every $u \in \mathcal{U}$ is a singleton point ($k(u) = n$).*

Proof. By the lemma, no cube with $k(u) < n$ can be contained in O_n. Hence every $u \in \mathcal{U}$ must be a singleton. Since the family is disjoint and covers O_n, it must contain exactly one cube for each of the $|O_n| = 2^{n-1}$ points. □

6.2 A CNF family forcing exponential fragmentation

Corollary 6.3 (Worst-case fragmentation for COVERTRACE-SAT)**.** *For each $n \geq 1$, there exists a satisfiable CNF F_n on n variables such that any disjoint subcube representation of $U(F_n)$ requires 2^{n-1} cubes. Consequently, any run of COVERTRACE-SAT on F_n must use $\Omega(2^{n-1})$ space and $\Omega(2^{n-1})$ time.*

Proof. For each $a \in O_n$, define a clause C_a falsified only by assignment a:

$$C_a \;=\; \bigvee_{i:a_i=0} x_i \;\;\vee\;\; \bigvee_{i:a_i=1} \neg x_i.$$

Then $Q(p(C_a)) = \{a\}$ and thus $U(F_n) = \bigcup_{a \in O_n} \{a\} = O_n$ for $F_n = \bigwedge_{a \in O_n} C_a$. Since $E_n = \Omega_n \setminus O_n$ is nonempty, F_n is satisfiable. Theorem 6.2 implies any disjoint subcube cover of $U(F_n)$ has size 2^{n-1}. □

Remark 6.4 (Input-size caveat and a stronger open problem)**.** The family above has 2^{n-1} clauses, so its input size is exponential in n. A substantially stronger goal is to exhibit *polynomial-size* CNFs F_n such that every DSOP for $\neg F_n$ has size $2^{\Omega(n)}$. This becomes an explicit lower-bound problem in knowledge compilation / proof complexity (Section 7 and Section 9).

7 COVERTRACE as knowledge compilation

7.1 DSOP / deterministic DNF

Definition 7.1 (Cubes as terms, DSOP)**.** A pattern $p \in \{0, 1, \bullet\}^n$ corresponds to a conjunction (term)

$$T(p) \;=\; \bigwedge_{i \in \mathrm{supp}(p)} \begin{cases} x_i & \text{if } p_i = 1, \\ \neg x_i & \text{if } p_i = 0. \end{cases}$$

A *disjoint sum of products* (DSOP) for a Boolean function g is a DNF $g = \bigvee_{j=1}^{s} T(p_j)$ such that the sets $Q(p_j)$ are pairwise disjoint. Equivalently, it is a *deterministic DNF* where no assignment satisfies two distinct terms.

Proposition 7.2 (Forbidden cubes are the terms of $\neg F$)**.** *For a CNF $F = \bigwedge_{j=1}^{m} C_j$, one has*

$$\neg F \;\equiv\; \bigvee_{j=1}^{m} T(p(C_j)).$$

Thus $U(F)$ is the set of satisfying assignments of $\neg F$.

Theorem 7.3 (COVERTRACE compiles $\neg F$ into a DSOP)**.** *The output family $\mathcal{U}$ of COVERTRACE-SAT yields a* DSOP *for $\neg F$:*

$$\neg F \;\equiv\; \bigvee_{u \in \mathcal{U}} T(u),$$

and the terms are mutually disjoint.

Proof. By correctness, $\mathrm{Cov}(\mathcal{U}) = U(F)$. Each cube u corresponds to term $T(u)$ whose satisfying assignments are exactly $Q(u)$. Disjointness implies determinism. □

7.2 Position in the Darwiche–Marquis map

The knowledge compilation map organizes target languages by (i) *succinctness* and (ii) polytime support for queries/transformations such as model counting (CT), satisfiability (SAT), clausal entailment (CE), and conditioning (CD), among others. Deterministic DNF / DSOP supports model counting in time linear in the representation size by additivity. However, deterministic DNF is generally less succinct than languages such as d-DNNF, OBDD, or SDD which also support polytime counting under additional structural constraints.

7.3 A conditional collapse of PH from uniform DSOP compilation

We now formalize a complexity-theoretic consequence suggested by the COVERTRACE viewpoint.

Theorem 7.4 (Conditional PH collapse under uniform disjoint-subcube compilation)**.** *Assume there exists a deterministic algorithm A such that, for every CNF formula F on n variables with size $m = |F|$, the algorithm runs in time* $\mathrm{poly}(n, m)$ *and outputs a disjoint family $\mathcal{U} \subseteq \{0, 1, \bullet\}^n$ satisfying:*

1. $|\mathcal{U}| \le \mathrm{poly}(n, m)$ *and the encoding length of $\mathcal{U}$ is* $\mathrm{poly}(n, m)$*;*
2. $\mathrm{Cov}(\mathcal{U}) = U(F)$ *(exact forbidden region).*

Then #SAT $\in$ P *and the Polynomial-Time Hierarchy collapses:* PH = P.

Proof. Given F, run $A(F)$ to obtain $\mathcal{U}$ in time poly(n, m). By disjointness,

$$|U(F)| = \sum_{u \in \mathcal{U}} 2^{n-k(u)},$$

computable in polynomial time. Hence $\#\text{SAT}(F) = 2^n - |U(F)|$ is computable in polynomial time, so #SAT $\in$ P. Since #SAT is #P-complete, this implies #P $\subseteq$ P. By Toda's theorem, $\text{PH} \subseteq \text{P}^{\#\text{P}}$; substituting the oracle with a polynomial-time algorithm yields PH $\subseteq$ P. Trivially P $\subseteq$ PH, hence PH = P. $\square$

Remark 7.5 (Why this does not prove P $\neq$ NP)**.** Theorem 7.4 is conditional and rules out a strong uniform DSOP compilation hypothesis under standard beliefs. Exponential lower bounds for DSOP (or for COVERTRACE's fragmentation) only show that this compilation language is not universally succinct; they do not rule out other algorithms for SAT.

7.4 Disjoint covers as proof certificates (UNSAT)

A disjoint family $\mathcal{U}$ with $\text{Cov}(\mathcal{U}) = \Omega_n$ constitutes a certificate that F is unsatisfiable, verifiable by checking disjointness and total volume 2^n. This induces a Cook–Reckhow proof system where proofs are disjoint covers; if all UNSAT instances had polynomial-size proofs in this system, then UNSAT $\in$ NP and NP = coNP.

8 An affine extension: exponential compression for parity

Parity is a worst-case obstruction for axis-aligned cubes, but it is a single linear equation over $\mathbb{F}_2$. This motivates extending the representation language.

8.1 Affine sets over $\mathbb{F}_2$

Definition 8.1 (Affine subspaces and cosets)**.** Let $A \in \mathbb{F}_2^{r \times n}$ and $b \in \mathbb{F}_2^r$. The solution set

$$\mathcal{A}(A, b) = \{x \in \Omega_n : Ax = b\}$$

is an affine subspace (a coset of $\ker A$). Its size is either 0 (inconsistent) or $2^{n-\text{rank}(A)}$.

Proposition 8.2 (Counting affine solutions is polynomial)**.** *Given A, b, one can decide emptiness and compute $|\mathcal{A}(A, b)|$ in time $O(n^3)$ by Gaussian elimination over $\mathbb{F}_2$.*

8.2 Exponential separation: parity

Theorem 8.3 (Axis-aligned DSOP vs. affine representation for parity)**.** *The odd-parity set O_n equals the affine set $\{x : x_1 \oplus \cdots \oplus x_n = 1\}$, representable with one linear equation. However, any disjoint axis-aligned subcube cover of O_n requires 2^{n-1} cubes (Theorem 6.2). Thus allowing affine pieces yields an exponential compression for this family.*

8.3 Toward *COVERTRACE-Affine*

The preceding theorem suggests a principled extension: augment the disjoint-subcube language with affine pieces (or affine decision nodes), so that XOR structure can be represented compactly while retaining polytime model counting on the compiled representation. Affine compilation languages and extended affine decision trees provide relevant precedents.

9 Toward an obstruction program: fragmentation lower bounds from structure

9.1 The minimal disjoint-subcube cover number

Definition 9.1 (Disjoint subcube cover number)**.** For a set $S \subseteq \Omega_n$, define

$$\chi_{\sqcup}(S) \;=\; \min\big\{|\mathcal{U}| : \mathcal{U} \subseteq \{0,1,\bullet\}^n \text{ disjoint and } \mathrm{Cov}(\mathcal{U}) = S\big\}.$$

Equivalently, $\chi_{\sqcup}(S)$ is the minimum DSOP size of the indicator function 1_S.

For a CNF F, the minimal representation size of the forbidden region is $\chi_{\sqcup}(U(F))$. Any exact COVERTRACE-style algorithm that stores a disjoint cover of $U(F)$ must use at least $\chi_{\sqcup}(U(F))$ pieces in the worst case (even with optimal clause ordering and maximal local merges).

Lemma 9.2 (Trivial volumetric lower bound)**.** *For any $S \subseteq \Omega_n$,*

$$\chi_{\sqcup}(S) \;\geq\; \frac{|S|}{\max\{|Q(p)| : p \in \{0,1,\bullet\}^n,\; Q(p) \subseteq S\}}.$$

Proof. In any disjoint cover of S by cubes, each cube has size at most the stated maximum. Since the cubes are disjoint and their sizes sum to $|S|$, the number of cubes is at least the ratio. □

The parity lower bound is an extremal case where the largest monochromatic cube inside O_n has size 1.

9.2 Influence as a fragmentation obstruction

For a Boolean function $f : \Omega_n \to \{0,1\}$, the (total) influence measures the expected boundary size of its level sets.

Definition 9.3 (Influence)**.** For $i \in [n]$, let $x^{\oplus i}$ denote x with bit i flipped. The *influence* of coordinate i on f is

$$\mathrm{Inf}_i(f) \;=\; \Pr_{x\sim\mathrm{Unif}(\Omega_n)}\big[f(x) \neq f(x^{\oplus i})\big].$$

The *total influence* is $\mathrm{Inf}(f) = \sum_{i=1}^{n} \mathrm{Inf}_i(f)$.

Total influence equals the normalized edge boundary of $S = f^{-1}(1)$: if $\partial_i S$ denotes the set of edges in direction i crossing between S and its complement, then $|\partial_i S| = 2^{n-1}\mathrm{Inf}_i(f)$ and $|\partial S| = \sum_i |\partial_i S| = 2^{n-1}\mathrm{Inf}(f)$.

Proposition 9.4 (Influence upper bound from a DSOP)**.** *Let $S \subseteq \Omega_n$ and $f = 1_S$. If $S = \bigsqcup_{u\in\mathcal{U}} Q(u)$ is a disjoint union of cubes, then*

$$\mathrm{Inf}(f) \;\leq\; \sum_{u\in\mathcal{U}} k(u)\, 2^{1-k(u)}.$$

Proof. For a cube $Q(u)$ of width $k = k(u)$, along each fixed coordinate $i \in \mathrm{supp}(u)$ every point in $Q(u)$ has a unique neighbor outside $Q(u)$, contributing $|Q(u)| = 2^{n-k}$ crossing edges in direction i. Thus $Q(u)$ alone would contribute $\mathrm{Inf}_i(1_{Q(u)}) = 2^{n-k}/2^{n-1} = 2^{1-k}$ for each $i \in \mathrm{supp}(u)$ and 0 otherwise, giving total influence $k2^{1-k}$. For a disjoint union, the boundary of S is a subset of the union of individual cube boundaries (adjacent cubes can only cancel boundary edges), hence the inequality after summing over cubes. □

Corollary 9.5 (A width-aware influence lower bound on $\chi_\sqcup$)**.** *Let $S \subseteq \Omega_n$ and suppose every cube contained in S has width at least k_0 (equivalently, S contains no axis-aligned cube of dimension $> n - k_0$). Then*

$$\chi_\sqcup(S) \;\geq\; \frac{\mathrm{Inf}(1_S)}{k_0\, 2^{1-k_0}}.$$

Proof. In any disjoint cover of S, each cube has width $k(u) \geq k_0$, hence $k(u)2^{1-k(u)} \leq k_0 2^{1-k_0}$. Apply Proposition 9.4 and rearrange. □

For parity, $k_0 = n$ and $\mathrm{Inf}(1_{O_n}) = n$, so Corollary 9.5 yields $\chi_\sqcup(O_n) \geq 2^{n-1}$, matching Theorem 6.2.

9.3 Subcube partition complexity and recent structural lower bounds

Subcube partitions are a well-studied computation model: a *subcube partition* for a Boolean function f partitions Ω_n into disjoint subcubes on which f is constant, and its complexity is the number of parts. This measure is closely related to deterministic decision trees, but can be strictly smaller, and separations have been proved.

Recent work relates subcube partition complexity to additive structure and influence of the support of f. Such results suggest a path to *a priori* fragmentation lower bounds for COVERTRACE on structured formulas: if $U(F)$ has strong additive structure or large influence in an appropriate sense, then $\chi_\sqcup(U(F))$ (and hence worst-case fragmentation) must be large.

Problem 9.6 (Explicit CNF→DSOP lower bounds)**.** Exhibit an explicit family of CNFs F_n of size $\mathrm{poly}(n)$ such that $\chi_\sqcup(U(F_n)) = 2^{\Omega(n)}$.

Proving such a lower bound would separate polynomial-size CNF representations from polynomial-size DSOP representations for $\neg F_n$ and would provide a concrete obstruction to uniform disjoint-subcube compilation.

10 A GCT-inspired viewpoint: disjoint subcubes as restricted rank-one tensors

GCT studies lower bounds in algebraic complexity by relating explicit functions to geometric and representation-theoretic obstructions. While COVERTRACE operates in a discrete setting, it induces a natural tensor decomposition viewpoint.

10.1 Truth tables as tensors

Identify a Boolean function $f : \{0,1\}^n \to \{0,1\}$ with its order-n tensor $T_f \in \{0,1\}^{2\times\cdots\times 2}$, where

$$(T_f)_{x_1,\ldots,x_n} = f(x_1,\ldots,x_n).$$

Let $e_0 = (1,0)$, $e_1 = (0,1)$, and $e_\star = (1,1)$ in $\mathbb{R}^2$. A pattern $p \in \{0,1,\bullet\}^n$ induces a rank-one tensor

$$T_p = v_1 \otimes \cdots \otimes v_n, \qquad v_i = \begin{cases} e_0 & p_i = 0, \\ e_1 & p_i = 1, \\ e_\star & p_i = \bullet. \end{cases}$$

One checks that T_p is the indicator tensor of the cube $Q(p)$.

Proposition 10.1 (DSOP as restricted nonnegative tensor rank)**.** *Let $S \subseteq \Omega_n$ and $f = 1_S$. Then $\chi_\sqcup(S)$ equals the minimum r such that*

$$T_f = \sum_{j=1}^{r} T_{p_j}$$

with patterns $p_j \in \{0, 1, \bullet\}^n$ whose cubes are pairwise disjoint.

Remark 10.2. Ignoring the disjointness constraint yields a restricted *nonnegative* tensor rank measure (since all tensors have nonnegative entries). Allowing cancellations (signed sums) corresponds to inclusion–exclusion and can dramatically reduce representation size (as for parity), paralleling the gap between explicit sums and compact arithmetic circuits.

10.2 Obstructions and open problems

This tensor viewpoint suggests a discrete "GCT-like" program: identify invariants of $T_{1_{U(F)}}$ (under the natural action of the hypercube automorphism group) that lower bound $\chi_\sqcup(U(F))$. Even partial progress—e.g., proving superpolynomial $\chi_\sqcup$ for an explicit polynomial-size CNF family—would be a meaningful lower bound in a concrete representation model.

11 Conclusion

COVERTRACE-SAT gives a simple and exact geometric compilation of $\neg F$ into a disjoint subcube cover, enabling exact #SAT and witness extraction whenever fragmentation remains controlled. We proved a tight exponential worst-case lower bound using parity, and we formalized a strong conditional consequence: a uniform polynomial-time disjoint-subcube compiler for all CNFs would collapse PH. We proposed an affine extension that compresses parity exponentially and articulated GCT-inspired obstructions via a restricted tensor-rank interpretation.

The remaining frontier is explicit lower bounds for polynomial-size CNFs in the DSOP language and principled hybrid compilation strategies that combine axis-aligned and affine structure while retaining exactness.

A Implementation notes and heuristics (not part of the core mathematics)

This appendix collects optimizations that can significantly improve practical performance but do not change the formal core (Algorithms 1–3) and do not evade worst-case lower bounds.

A.1 Bitmask representation

A pattern $p \in \{0, 1, \bullet\}^n$ can be stored as two n-bit masks: a fixed-mask F indicating which coordinates are fixed and a value-mask V specifying their values. Intersection and containment can then be tested with constant-time bit operations.

A.2 Buddy merges

If two cubes have the same fixed coordinates and differ in exactly one fixed bit, they are "buddies" and can be merged by freeing that bit, reducing fragmentation without changing the covered set. Repeated merges correspond to local boundary cancellation.

A.3 Clause ordering

Processing clauses that forbid large-volume cubes early (e.g., short clauses) may reduce intermediate fragmentation on some instances. Such orderings are instance-dependent heuristics and are excluded from the core correctness analysis.

B A quantum-computing viewpoint (optional)

The hypercube Ω_n indexes the computational basis $\{|x\rangle : x \in \{0,1\}^n\}$ of n qubits. A cube $Q(p)$ corresponds to the span of basis states consistent with p; disjointness of cubes corresponds to orthogonality of these subspaces. Exact model counting $\#\mathrm{SAT}(F)$ equals the dimension of the satisfying subspace.

While Grover's algorithm provides a quadratic speedup for searching satisfying assignments given an oracle, it does not yield polynomial-time SAT in general. Nevertheless, COVERTRACE-style geometric structure could inform oracle construction or hybrid classical-quantum heuristics (e.g., amplitude amplification restricted to surviving regions), although this remains speculative.

References

[1] S. Cook and R. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44(1):36–50, 1979.

[2] S. Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991.

[3] A. Darwiche and P. Marquis. A knowledge compilation map. *Journal of Artificial Intelligence Research*, 17:229–264, 2002.

[4] F. Koriche, J.-M. Lagniez, P. Marquis, and P. Thomas. Knowledge compilation for efficient model counting. In *Proceedings of IJCAI*, 2013.

[5] R. Kothari, C. Meka, and P. Raghavendra. Separating decision tree complexity from subcube partition complexity. In *Proceedings of APPROX/RANDOM*, 2015.

[6] N. Hegyvári. The complexity of subcube partition relates to the additive structure of the support and influence. *European Journal of Combinatorics*, 2024.

[7] K. D. Mulmuley and M. Sohoni. Geometric complexity theory I: An approach to the P vs. NP and related problems. *SIAM Journal on Computing*, 31(2):496–526, 2001.

[8] K. D. Mulmuley and M. Sohoni. Geometric complexity theory II: Towards explicit obstructions for embeddings among class varieties. *SIAM Journal on Computing*, 38(3):1175–1206, 2008.

[9] M. Bläser and C. Ikenmeyer. Introduction to Geometric Complexity Theory. *Theory of Computing*, 2025.

[10] J. Dörfler, C. Ikenmeyer, and G. Panova. On geometric complexity theory: multiplicity obstructions are stronger than occurrence obstructions. In *ICALP*, 2019.

[11] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of STOC*, 1996.

[12] A. Lucas. Ising formulations of many NP problems. *Frontiers in Physics*, 2:5, 2014.

[13] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997.