No Answer From Nowhere: Causal Ancestry, Witness Testing, and a Conditional Route to P ≠ NP

Prepared as a companion research note to the entropy-dispersion manuscripts [1, 2].

Central claim. A reliably correct answer cannot come from nowhere. Whether a solver combines many small clues or appears to seize on a single decisive marker, that marker must itself have been produced from information contained in the problem. There is therefore a finite, traceable causal path connecting the problem presented to the solver with the answer it returns.

Terminology. Throughout, the *input* means the complete problem instance: the question together with all of its data, rules, and constraints.

Status of the argument. The existence of finite causal ancestry is *proved* for the finite sequential model defined below. The further claim — that every such ancestry can be *compressed* into a mathematical footprint reachable by known lower bounds — remains *conditional*, and is isolated as the principal open step (Conjecture 8.1).

General reader summary

The question behind this paper is the P versus NP question: when a proposed answer can be checked quickly, must there also be a quick way to find it? We approach that question from a deliberately elementary starting point — *a reliably correct answer cannot come from nowhere.*

By *input* we mean the complete problem given to the computer: the question together with its data, rules, and constraints. If a solver repeatedly returns answers that fit the real problem better than answers generated from an *unrelated* problem, then information from the real problem must have reached the output and shaped it. There is no other place for the advantage to come from.

That information may arrive as many small clues or as what looks like one decisive marker. But even a single marker must have been produced from the problem. The computation must therefore contain a **finite causal ancestry**: a traceable chain linking information in the problem to the answer eventually produced.

The first part of the paper makes this precise. Any finite advantage over an input-disconnected baseline forces at least one input-access event to have a measurable causal effect on the final usefulness test, and at least one event to carry a measurable amount of information about the

problem. This effect must moreover be visible on a non-negligible fraction of computation histories — it cannot hide on a vanishingly thin set.

The harder part is **compression**. A traceable ancestry need not be simple. Parity of n bits depends on every input bit, yet it can be computed with essentially no memory; so we do not assume that finite ancestry is automatically local or low-degree. Instead we build a *ladder* of analyzable forms — influential coordinates, juntas, low-degree approximations, branching programs, shallow circuits, and low-treewidth representations — and ask which rung a given ancestry can be pushed onto.

For planted and other mixture distributions we prove an **observed-or-latent dichotomy**: the useful correlation must arise either through an influential observed part of the problem or through a low-dimensional signal about the hidden planted solution. This sharply narrows the kinds of computation that could evade current lower-bound methods.

The result is a **conditional** argument, not a finished proof that P ≠ NP. What is proved is that useful answers have finite, traceable origins in the problem. What remains is to show that *every* witness-relevant causal ancestry can be compressed into an analyzable footprint on the same hard distribution — together with a separate bridge from distributional hardness to worst-case P versus NP.

Abstract

A reliably correct computational answer cannot come from nowhere: it must be causally connected to information in the problem instance. We formalize this principle for finite sequential randomized computations. Let X be a finite encoded instance, let A make at most m input-access events, let $Y = A(X; R)$, and let $\tau(X, Y) \in [0, 1]$ be a bounded usefulness test. Comparing the real output with the same algorithm run on an independent *ghost* instance, an operational distinguishing gap ε forces a finite traceable ancestry: some input event has interventional effect at least ε/m; some event carries at least $2\varepsilon^2/m$ nats of conditional information about X; and a non-negligible transcript region exhibits a same-sign effect.

Finite ancestry alone does not imply a simple representation — parity is the standing counterexample. To attempt structural compression we introduce a *robust local witness-testing interface*. A normalized solver that emits a locally testable proof only on genuine success produces a *local* real-versus-decoupled gap. An oblivious tester converts that gap into a local instance–output rectangle, and fixing the solver randomness isolates a deterministic witness-relevant causal cone.

Under product measures, OSSS-type inequalities convert the extracted covariance into a fixed influential observed coordinate. Under mixture-of-product distributions, a law-of-total-covariance argument yields an *observed-or-latent* dichotomy: either an observed block is influential inside a component, or a low-dimensional function of latent planted-witness coordinates carries the signal. We then give parameterized compression results keyed to ancestor count, decision-tree depth, memory, circuit depth, total influence, treewidth, and noise stability. The unrestricted compression of high-space, high-synergy cones remains conjectural; if

established on a common hard distribution and combined with a distributional-to-worst-case bridge, it would complete the proposed conditional route to P ≠ NP.

Contents

---

1. Introduction and contribution

The P versus NP question asks whether every problem whose proposed solutions can be checked quickly can also be solved quickly. This paper begins one level below that question, with a claim that is hard to deny: **a reliably correct answer cannot come from nowhere.** It must be connected, through the computation, to information contained in the problem actually presented.

Here *input* means the full problem instance — the question together with all data, rules, and constraints. The organizing contrast is between a real execution $A(X; R)$ and an *input-disconnected* execution $A(X'; R)$ generated from an independent problem. If the real output is more useful for X than the disconnected output, that difference must be carried by a finite chain of input observations, state changes, and downstream consequences. The output need not reveal that chain to the user, but the chain must exist inside the computation.

The program divides cleanly into three layers:

- **Causal.** Prove that useful output has bounded, traceable ancestry in the problem.
- **Relational.** Turn successful witness production into a *local* correlation between a part of the problem and a part of the output.

- **Representational.** Compress the resulting causal cone into a structure to which proof-complexity, influence, Fourier, branching-program, or knowledge-compilation lower bounds apply.

The causal layer is proved unconditionally for the model below. The representational layer is where the difficulty concentrates, and we do not pretend otherwise: universal compression is isolated as a conjecture (§8), not smuggled in as a lemma.

1.1 Central claim and formal scope

A reliably correct answer cannot come from nowhere. Whether the computation accumulates many small clues or appears to discover one decisive marker, the information that makes the answer fit the *actual* problem must have entered from the problem and reached the output through a finite causal path.

The theorem proved in §§2–3 is deliberately narrower than the final P versus NP claim: *operational distinguishability forces finite causal ancestry and quantitative information flow.* The paper then attempts to compress that ancestry into analyzable structure. Universal compression is not assumed as already established; it is stated explicitly as the witness-relevant causal-cone compression conjecture (C3) in §8.

1.2 Main contributions

1. **No answer from nowhere.** An ε real-versus-decoupled distinguishing gap forces a finite causal ancestry, including one input event with interventional effect at least ε/m and one event carrying at least $2\varepsilon^2/m$ nats of conditional information (Theorems 3.1, 3.2).
2. **Transcript localization.** A non-negligible set of transcript prefixes supports a same-sign causal effect, preventing the advantage from being spread over infinitely fine or measure-zero distinctions (Lemma 3.3).
3. **Robust local witness distinguishability.** A normalized successful solver together with a sound local assignment tester produces a gap at least $\gamma \cdot s \cdot (1 - \beta)$ against an independent-instance output baseline (Theorem 5.3).
4. **Local rectangle extraction.** An oblivious test reading b bits yields a single instance-cylinder × output-cylinder rectangle of discrepancy at least $\gamma \cdot s \cdot (1 - \beta) \cdot 2^{-b}$ (Theorem 6.1).
5. **Fixed-seed causal cone.** One deterministic random seed preserves the rectangle covariance, isolating a definite subcomputation rather than an average over random runs (Theorem 6.2).
6. **Observed-or-latent mixture theorem.** For mixture-of-product distributions, the covariance yields either an influential observed block or a low-dimensional latent recovery signal (Theorem 6.4).
7. **Parameterized compression.** Exact or approximate compression guarantees keyed to ancestor count, decision-tree depth, memory, circuit depth, total influence, treewidth, and noise stability (§7).
8. **A sharper open conjecture.** The unrestricted problem is reduced to witness-relevant causal cones that evade *every* representation on the target hard distribution (§8).

1.3 What is and is not being claimed

The paper does **not** claim that finite causal ancestry is automatically low-degree, nor that bounded space implies low-degree approximation. It does **not** apply product-measure influence inequalities to conditioned or arbitrarily dependent distributions without an additional argument. It does **not** identify distributional hardness with worst-case P versus NP. These distinctions are load-bearing.

The central correction the paper makes to loose intuition is this: in an ordinary finite sequential machine, a useful answer cannot be causally disconnected from the problem and still be reliably correct. Input dependence is mediated by finite physical or logical events. The difficult question is *not* whether some causal ancestry exists — it always does — but whether its history-conditioned effects can always be *deconditioned* and compressed into a lower-bound-compatible representation.

---

2. Finite computation, distinguishability, and causal ancestry

2.1 Sequential input-access model

Definition 2.1 (Finite sequential computation). An instance X is drawn from a distribution D_n on a finite encoding space $\mathcal{X}_n$. A randomized algorithm A samples a complete random tape R independently of X and performs at most $m = m(n)$ input-access events.

Before event j, the operational history $H\{j-1\}$ *contains the internal state, previous query addresses, previous observations, and the random bits revealed so far. The algorithm chooses a query* Q_j *from* $H\{j-1\}$ and receives an observation $O_j = \psi_{Q_j}(X)$. Each observation reads at most $q = q(n)$ encoding coordinates, or one bounded-size block. For information accounting we use the *augmented* history $\bar{H}\{j-1\} = (R, Q_1, O_1, \ldots, Q\{j-1\}, O_{j-1})$, which includes the full random tape.

The output Y is a deterministic function of R and the observation transcript. Variable-length runs are padded with null events.

This model includes Turing machines, RAM computations, randomized decision trees, and circuit evaluations, provided their input dependencies are exposed as finite read events. A machine may read the entire polynomial-length input, but it must do so through finitely many events. Crucially, the model does **not** assume that the output depends on a fixed small subset of coordinates across all runs.

2.2 Operational distinguishability

Definition 2.2 (Input-decoupled baseline). Let X' be an independent copy of X and set $\tilde{Y} = A(X'; R)$. The pair $(X, \tilde{Y})$ has law $P_X \times P_Y$: the output keeps the same marginal distribution as the real output but is causally disconnected from the current instance.

Definition 2.3 (Operational distinguishing gap). For a bounded test $\tau : \mathcal{X}_n \times \mathcal{Y}_n \to [0, 1]$, define

$$\Delta\tau = | \mathbb{E}[\tau(X, A(X; R))] - \mathbb{E}[\tau(X, A(X'; R))] |. \quad (2.1)$$

For NP search, τ may be the verifier $V(X, Y)$, or a locally testable encoded verifier introduced in §5.

Because tests valued in [0, 1] characterize total variation distance,

$$\Delta\tau \le \| P_{XY} - P_X P_Y \|_{TV} .$$

A non-negligible gap is therefore a *finite statistical distinction*, not a limiting or infinitesimal one.

2.3 Causal ancestry and causal cones

Definition 2.4 (Realized causal ancestry). Represent one execution as a directed acyclic graph whose nodes are input-coordinate values, random bits, observations, state updates, output bits, and the final test result. A node lies in the *realized ancestry* of the test if a directed path connects it to the test node.

Definition 2.5 (Output causal cone). For a fixed random seed r and a selected set J of output bits, the *causal cone* $C_{r, J}$ is the union of all gates, states, observations, and input coordinates that can affect those bits in the deterministic computation $A(\cdot; r)$.

On a single run, at most $m \cdot q$ input-coordinate occurrences enter the transcript, and every causal path has length at most the running time. Across *different* inputs, however, the union cone may touch every coordinate. This distinction — between **execution-relative boundedness** and **globally fixed locality** — is the pivot of the entire argument.

2.4 Structural parameters of a cone

For a deterministic Boolean output event $f(X)$ extracted from a cone, we track: the number K of distinct input ancestors; deterministic decision-tree depth d; workspace s; circuit size S and depth h; total influence I under a product measure; underlying circuit treewidth w; and noise sensitivity under a specified perturbation kernel. **None** of these is assumed small in the general theorem.

---

3. Universal causal-ancestry theorems

3.1 Ghost-instance hybrid theorem

Theorem 3.1 (Finite causal-ancestry theorem). Suppose $\Delta\tau \geq \varepsilon$. Form hybrids $A^{(k)}$, $k = 0, \ldots, m$, that answer the first k input-access events from X and all remaining events from the independent ghost instance X'. Then there exists an event j such that replacing only the source of the j-th observation — together with all downstream consequences — changes the expected test value by at least ε/m:

$$\mathbf{|\, \mathbb{E}[\tau(X, A^{(j)})] - \mathbb{E}[\tau(X, A^{(j-1)})]\, | \geq \varepsilon/m.} \quad (3.1)$$

Proof. Let $a_k = \mathbb{E}[\tau(X, A^{(k)})]$. The endpoint $k = m$ is the real execution and $k = 0$ is the input-decoupled execution, so $|a_m - a_0| \geq \varepsilon$. The telescoping identity

$$a_m - a_0 = \Sigma_{k=1}^{m} (a_k - a_{k-1})$$

together with the triangle inequality forces some summand to have magnitude at least ε/m.

The two adjacent hybrids are identical before event j: same randomness, same first $j - 1$ observations, hence the same j-th query. They differ only in whether the j-th answer is drawn from X or from X'. Every later difference is downstream of that single intervention. Thus event j lies on a finite causal path to the operational distinction. ■

The theorem does **not** claim that event j has a fixed query address across all histories. It fixes a *time index* and a *traceable intervention channel*. Pinning down a fixed *coordinate* requires additional distributional structure, supplied in §6 by OSSS-type influence inequalities.

3.2 Information-flow lower bound

Theorem 3.2 (Information-bearing event). Under the setup of §3.1, the real joint law satisfies $I(X; Y) \geq 2\varepsilon^2$ nats. Moreover, with augmented pre-event history $\bar{H}_{j-1} = (R, Q_1, O_1, \ldots, Q_{j-1}, O_{j-1})$, *some event j satisfies* $I(X; O_j \mid \bar{H}_{j-1}) \geq 2\varepsilon^2/m$:

$$\mathbf{I(X; Y) \geq 2\varepsilon^2 \quad and \quad \max_j I(X; O_j \mid \bar{H}_{j-1}) \geq 2\varepsilon^2/m.} \quad (3.2)$$

Proof. The test gap ε forces total variation distance at least ε between P_{XY} and $P_X P_Y$. Pinsker's inequality gives $D_{\mathrm{KL}}(P_{XY} \,\|\, P_X P_Y) \geq 2\varepsilon^2$, and this divergence is exactly $I(X; Y)$.

Because R is independent of X and Y is a function of R and the observation transcript, data processing and the mutual-information chain rule give

$$I(X; Y) \leq I(X; R, O_1, \ldots, O_m) = \Sigma_j I(X; O_j \mid R, O_1, \ldots, O_{j-1}).$$

Query addresses and machine states are deterministic functions of the augmented history, so each summand equals $I(X; O_j \mid \bar{H}_{j-1})$. Averaging yields the stated event. ■

The hybrid theorem measures *causal effect*; the information theorem measures *information transfer*. The event witnessing one need not be the event witnessing the other, but both are finite and quantitatively non-negligible whenever ε is inverse-polynomial and m is polynomial.

3.3 A non-negligible transcript region

Lemma 3.3 (Signed high-effect transcript region). There exist an event j, a sign $\sigma \in \{+1, -1\}$, and a measurable set B of pre-event histories such that $\Pr[H_{j-1} \in B] \geq \varepsilon/(4m)$, and, on B, the conditional signed hybrid effect is at least $\varepsilon/(4m)$.

Proof. Take j from Theorem 3.1 and let $Z(H_{j-1})$ be the conditional expected difference between the two adjacent hybrids, so $\mathbb{E}|Z| \geq \varepsilon/m$. One of Z_+ or Z_- has expectation at least $\varepsilon/(2m)$; orient σ toward it and threshold that part at $\varepsilon/(4m)$. Since $Z \leq 1$ pointwise, the threshold set $B = \{Z \geq \varepsilon/(4m)\}$ has probability at least $\varepsilon/(4m)$, and the conditional effect on B has the required sign and magnitude. (See Appendix A.1 for the constant.) ■

This lemma prevents the effect from living only on a vanishingly thin collection of histories: it produces an *exact* transcript footprint. The predicate defining B may nevertheless be computationally complicated — a caveat carried forward, not resolved here.

3.4 Fixing randomness

Lemma 3.4 (Deterministic seed reduction). After orienting the sign of the test gap, there exists a fixed random seed $r\star$ such that the deterministic computation $A(\cdot; r\star)$ retains a real-versus-decoupled test gap at least ε.

Proof. For each seed r, put $d(r) = \mathbb{E}_{X, X'}[\tau(X, A(X; r)) - \tau(X, A(X'; r))]$. The average of $d(R)$ is the oriented gap ε, so some seed satisfies $d(r\star) \geq \varepsilon$. ■

Randomization can therefore be removed for structural extraction. The resulting deterministic cone may succeed on a *different* subset of instances, but it preserves the distributional gap the argument needs.

---

4. Why ancestry alone does not imply simple structure

4.1 The parity obstruction

Example 4.1 (Finite, traceable, but high-order). Let X be uniform on $\{0, 1\}^n$, let $Y = \text{parity}(X)$ be the XOR of all n bits, and let $\tau(X, Y) = 1$ exactly when $Y = \text{parity}(X)$. The real test accepts with probability 1; the decoupled test accepts with probability ½. Hence $\Delta\tau = ½$.

The ancestry here is completely finite: a width-2 branching program updates a single parity bit while reading the input sequentially. Yet no proper subset of the n input bits determines the

result, the exact multilinear degree is n, and *all* non-constant Fourier mass sits at degree n. Parity therefore refutes the unconditional implication

finite causal ancestry $\Longrightarrow$ junta or low-degree structure.

What parity *does* possess is a different analyzable structure — a very small-width branching program. This is exactly why we replace a single universal target class with a **compression ladder** (§7).

4.2 The real obstruction: deconditioning and synergy

Theorem 3.1 identifies a *history-conditioned* event. Compressing it into a *fixed coordinate* requires removing its dependence on prior history. XOR-type synergy shows why this can fail: an observation may have zero unconditional correlation with the output, yet become decisive once the other observations are known. Product-measure influence theorems, low-dimensional latent decompositions, and circuit-structure assumptions are precisely the mechanisms that can overcome this **deconditioning problem** — and each carries its own hypotheses.

---

5. Robust local witness testing

5.1 Interface definition

Definition 5.1 (Robust local witness-testing interface). Let $R_n(x, w)$ be an NP witness relation. A *proof encoder* $P(x, w)$ maps a valid pair to an expanded proof z. A public-coin tester $T_\rho(x, z)$ is:

- **complete** if valid encoded proofs are accepted for every ρ;
- **γ-sound** if every malformed proof, or proof decoding to an invalid witness, is accepted with probability at most $1 - \gamma$ over ρ;
- **obliviously *b*-local** if, for each ρ, the tester reads fixed locations totaling at most $b = b_X + b_Z$ encoded bits (or bounded-size blocks) from x and z before applying a Boolean predicate.

Assignment testers and PCPs of proximity are natural implementations of this interface after a suitable encoding [8, 9]. We treat the interface *abstractly* because the exact query, proof-length, and obliviousness parameters must be re-checked for each target relation and representation — this is an interface assumption, flagged again in §11.

5.2 Cross-instance anti-collision

Definition 5.2 (Witness anti-collision). For a distribution D_n on instances, define $\beta_n = \sup_w \Pr_{X \sim D_n}[R_n(X, w) = 1]$. Thus any witness drawn independently of the current instance is valid with probability at most β_n.

For random CSP distributions with linearly many independent constraints, β_n is typically exponentially small for every fixed assignment. For conditioned or highly correlated distributions this must be *proved* rather than assumed.

5.3 Normalizing failure behavior

Given a candidate solver A, define a **normalized solver** $\hat{A}$: run A to obtain w, verify $R\_n(x, w)$, and output the encoded proof $P(x, w)$ when valid; otherwise output a distinguished null symbol $\perp$. The local tester is defined to reject $\perp$. This postprocessing is polynomial-time and leaves the original success probability unchanged.

Theorem 5.3 (Local witness distinguishability). Let $s = \Pr[R\_n(X, A(X; R)) = 1]$. If the tester is γ-sound and the distribution has anti-collision β, then the tester distinguishes the real pair $(X, \hat{A}(X; R))$ from the input-decoupled pair $(X, \hat{A}(X'; R))$ with gap at least $\gamma \cdot s \cdot (1 - \beta)$:

$$\mathbb{E}\_{P}[T] - \mathbb{E}\_{\{P\_X P\_Z\}}[T] \geq \gamma \cdot s \cdot (1 - \beta). \quad (5.1)$$

Proof. In the real execution, $\hat{A}$ emits a complete proof exactly on successful runs and $\perp$ otherwise, so the tester accepts with probability exactly s.

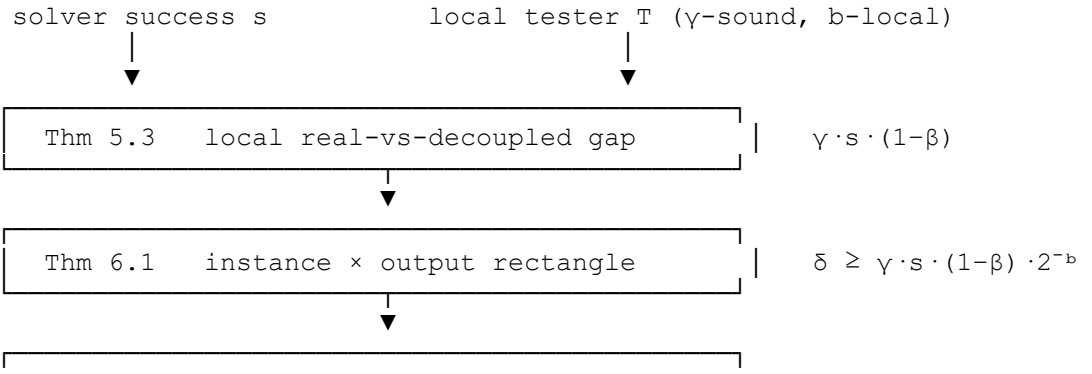
In the decoupled execution, the output marginal is $\perp$ with probability $1 - s$ and a proof from an independent successful instance with probability s. Fix any positive proof decoding to w. The current independent instance accepts w with probability at most β; on that event the tester accepts with probability at most 1, and otherwise γ-soundness bounds acceptance by $1 - \gamma$. Hence conditional acceptance of a positive independent proof is at most

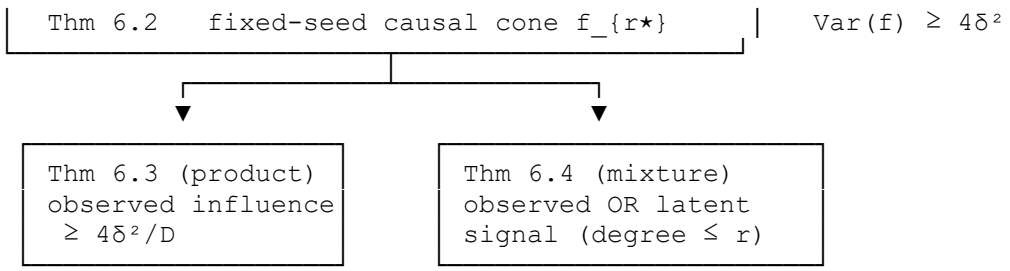
$$\beta + (1 - \beta)(1 - \gamma) = 1 - \gamma(1 - \beta).$$

Multiplying by s and subtracting from the real acceptance probability s gives $s - s \cdot [1 - \gamma(1 - \beta)] = \gamma \cdot s \cdot (1 - \beta)$. ■

This strengthens the basic "no camouflage" observation: it yields a *local*-test gap that scales directly with solver success, rather than merely detecting the final global verifier bit.

Figure 1 — the distinguishability-to-structure pipeline (§§5–7).





6. Rectangle extraction and observed-or-latent influence

6.1 Local rectangle footprint

Theorem 6.1 (Oblivious local rectangle extraction). Under Theorem 5.3, fix a tester random string ρ. If the tester reads at most b bits or block-bits, then there exist a local instance cylinder $U(X)$ and a local output cylinder $W(Z)$ such that

$$\mathbf{Cov}(U(X), W(Z)) \geq \delta, \text{ where } \delta \geq \gamma \cdot s \cdot (1-\beta) \cdot 2^{-b}. \quad (6.1)$$

Proof. Averaging the tester gap over ρ is at least $\gamma \cdot s \cdot (1-\beta)$, so some fixed $\rho\star$ achieves at least that gap. With $\rho\star$ fixed, the tester sees at most b bits, so its accepting set is a disjoint union of at most 2^b *atoms*. Each atom is a rectangle: one cylinder condition on the queried instance bits and one on the queried output bits.

The sum of the real-minus-product discrepancies over the accepting atoms equals the tester gap. Hence some atom has discrepancy at least the gap divided by 2^b. Rectangle discrepancy is exactly the covariance of the two cylinder indicators. (The atom count for general alphabets is handled in Appendix A.2.) ■

When $b = O(\log n)$, γ is constant, β is negligible, and s is inverse-polynomial, δ remains inverse-polynomial. This is the **first genuinely local structural object** in the chain.

6.2 Fixed-seed causal-cone extraction

Theorem 6.2 (Fixed-seed cone covariance). Let U be the extracted instance cylinder and W the output cylinder. Define $f\_r(x) = W(\hat{A}(x; r))$. There exists a fixed seed $r\star$ with

$$\mathrm{Cov}(U(X), f\_\{r\star\}(X)) \geq \delta,$$

and $f\_\{r\star\}$ depends only on the causal cone of the output bits read by W.

Proof. The rectangle covariance equals $\mathbb{E}\_R \mathrm{Cov}(U(X), f\_R(X))$: averaging $\mathbb{E}[U f\_R]$ recovers the real rectangle probability, and averaging $\mathbb{E}[f\_R]$ recovers the output marginal of W. A signed average of at least δ has some summand at least δ, giving $r\star$.

With $r\star$ fixed, f is deterministic. Only the output bits inspected by W matter, so every gate outside their union causal cone can be deleted without changing f. ■

Let $a(X) = U(X) - \mathbb{E}[U]$. Since U and f are indicators, Cauchy–Schwarz gives

$$\mathbf{Var}(f) \geq 4\delta^2. \quad (6.2)$$

So the extracted cone computes a genuinely non-constant event with inverse-polynomial variance whenever δ is inverse-polynomial.

6.3 Product-measure influence extraction

Theorem 6.3 (Observed influence under a product measure). Assume the instance encoding is a product of N independent coordinates or blocks, and that a deterministic decision tree computes f with expected query count D. Then some observed coordinate i has resampling influence

$$\max_i \mathbf{Inf}_i(f) \geq \mathbf{Var}(f)/D \geq 4\delta^2/D. \quad (6.3)$$

Proof. The OSSS inequality gives $\mathrm{Var}(f) \leq \Sigma_i \mathrm{reveal}_i \cdot \mathrm{Inf}_i(f)$, where reveal_i is the probability that the determining decision tree reads coordinate i [3]. The reveal probabilities sum to the expected query count D. Combining OSSS with (6.2) yields the bound. ■

A polynomial-time bit-access computation has $D = \mathrm{poly}(n)$, so inverse-polynomial rectangle covariance yields inverse-polynomial fixed-coordinate influence. Under the uniform cube, KKL supplies an alternative bound of order $\mathrm{Var}(f)\cdot\log N / N$ independent of the algorithm [4].

This is the point at which an explicit sequential model dissolves part of the global-to-local ambiguity: under a product encoding, a successful solver-derived output event *cannot* remain dependent only through diffuse history-conditioned effects — **one fixed coordinate must matter.** The caveat is equally sharp: conditioned random SAT and many planted distributions are *not* product measures in their observed encoding.

6.4 Mixture-of-product distributions

Planted models often become product distributions *after conditioning on a latent signal.* Let Θ be a latent variable, and suppose $X \mid \Theta = \theta$ is a product of independent observed blocks. The rectangle covariance then splits into within-component and between-component terms:

$$\mathbf{Cov}(U, f) = \mathbb{E}_\Theta[\, \mathbf{Cov}(U, f \mid \Theta)\,] + \mathbf{Cov}(\, \mathbb{E}[U \mid \Theta], \mathbb{E}[f \mid \Theta]\,). \quad (6.4)$$

Theorem 6.4 (Observed-or-latent influence dichotomy). Let $\mathrm{Cov}(U, f) \geq \delta$. Suppose $X \mid \Theta = \theta$ is product and f is computable with expected observed-block query count at most D in every component. Let $a(\theta) = \mathbb{E}[U \mid \Theta = \theta]$, and assume a depends only on a latent coordinate set J of size r whose coordinates are independent under the latent law. Then at least one of the following holds:

- **(A) Observed influence.** For some θ and observed block i, $\mathrm{Inf}_i^{\wedge}\theta(f) \geq \delta^2/D$.
- **(B) Latent signal.** The r-local function $g\_J(\Theta\_J) = \mathbb{E}[f \mid \Theta\_J]$ has variance at least δ^2, and some latent coordinate $j \in J$ has L^2 resampling influence at least δ^2/r.

Proof. Apply the law of total covariance (6.4).

Within-component case. If the absolute within-component term is at least $\delta/2$, some component θ has $|\mathrm{Cov}(U, f \mid \theta)| \geq \delta/2$. Since $\mathrm{Var}(U \mid \theta) \leq ¼$, Cauchy–Schwarz gives $\mathrm{Var}(f \mid \theta) \geq \delta^2$. OSSS within that product component yields observed influence at least δ^2/D.

Between-component case. Otherwise the between-component covariance has magnitude at least $\delta/2$. Because a depends only on Θ_J, it has the same covariance with $g\_J = \mathbb{E}[f \mid \Theta\_J]$. Again $\mathrm{Var}(a) \leq ¼$, so $\mathrm{Var}(g\_J) \geq \delta^2$. Efron–Stein over the r independent latent coordinates gives $\Sigma\_\{j \in J\}\ \mathrm{Inf\_j}^{\wedge}\{(2)\}(g\_J) \geq \mathrm{Var}(g\_J)$, so some latent coordinate has influence at least δ^2/r. (See Appendix A.4 for the L^2 normalization.) ■

Case (B) is *already* a low-dimensional footprint: when the latent coordinates are Boolean, g_J is a function of only r latent bits and therefore has exact degree at most r in the multilinear basis. The theorem does **not** require the algorithm to reveal the latent variable explicitly. For *dependent* latent laws, the same conclusion needs an appropriate Poincaré or Efron–Stein inequality, and the constants change accordingly.

6.5 Planted CSP interpretation

For planted k-SAT, conditioning on the planted assignment makes the clauses independent. A cylinder fixing b local clause blocks depends, through the sign-generation rule, on at most $k \cdot b$ planted-assignment coordinates once the variable indices are fixed. So Theorem 6.4 says an extracted solver footprint must appear *either* as influence of an observed clause block, *or* as a degree-$O(k \cdot b)$ signal about a small set of planted-assignment bits.

This is far sharper than a blanket assertion that every global algorithm is local. The remaining work is to show that the hard regime rules out *both* alternatives at inverse-polynomial strength, or to transfer the extracted footprint to the low-degree-recovery and proof-complexity lower bounds already available for the distribution.

---

7. Parameterized causal-cone compression

7.1 General approximation lemma

Theorem 7.1 (Approximation preserves the footprint). Let $a(X) = U(X) - \mathbb{E}[U]$ and suppose $\mathbb{E}[a(X) \cdot f(X)] \geq \delta$. If a structured real-valued function p satisfies $\|f - p\|_2 \leq \eta$ under D_n, then $\mathbb{E}[a \cdot p] \geq \delta - \eta/2$. In particular, $\eta \leq \delta$ yields $\mathbb{E}[a \cdot p] \geq \delta/2$.

Proof. By Cauchy–Schwarz, $|\mathbb{E}[a(f-p)]| \le \|a\|_2 \cdot \|f-p\|_2$. Since U is an indicator, $\|a\|_2 = \sqrt{(\mathrm{Var}\ U)} \le \tfrac{1}{2}$. Subtract this error from $\mathbb{E}[a \cdot f] \ge \delta$. ■

This converts *any* representation or approximation of the fixed-seed output event into a representation that retains inverse-polynomial covariance with the local instance condition.

7.2 Compression ladder

Table 1 — structural conditions on the extracted causal cone and the resulting footprint.

Condition on f or its cone	Compressed structure	Quantitative statement	Status / source
K distinct input ancestors	Exact K-junta; degree $\le K$	No approximation loss	Immediate
Decision-tree depth d	Exact multilinear polynomial, degree $\le d$	No approximation loss	Standard query representation
Expected query count D (product measure)	Fixed influential coordinate	$\max_i \mathrm{Inf}_i \ge 4\delta^2/D$	OSSS [3]
Workspace s for T steps	Branching program	width $\le 2^s$, length $\le T$	Immediate state expansion
AC^0 cone, size S, depth h	Low-degree Fourier approximation	degree $(\log(S/\eta))^{O(h)}$ for L^2 error η	LMN [6]
Total influence I (product measure)	Approximate junta	$K \le \exp(O(I/\eta^2))$ for L^2 error η	Friedgut [5]
Circuit treewidth w, size S	Structured decomposable circuit	size $S \cdot 2^{O(w)}$	Knowledge compilation [13]
Small input-noise sensitivity	Low-degree Fourier truncation	tail bounded by Prop. 7.2	Fourier identity

7.3 Bounded memory gives a transcript structure, not low degree

If the fixed-seed cone uses s bits of workspace, its sequence of internal states is a branching program of width at most 2^s and length at most the running time. For $s = O(\log^k n)$ this is polynomial or quasipolynomial width depending on k. The statement is *exact*, but it does **not** imply low-degree approximation — parity has width 2 and degree n. The correct conclusion from bounded memory is a transcript or branching-program footprint, not a spectral one.

7.4 Low influence and shallow circuits

Friedgut's junta theorem: a Boolean function with low total influence is close to a function of few coordinates [5]. Linial–Mansour–Nisan: constant-depth polynomial-size circuits have Fourier mass concentrated at polylogarithmic degree [6]. Through Theorem 7.1, either property converts the extracted cone into a low-complexity solver footprint while preserving a controlled fraction of the rectangle covariance.

7.5 Noise stability as a compression bridge

Proposition 7.2 (Operational robustness implies spectral compression under product noise). Let X^{ρ} be obtained by independently resampling each product coordinate with probability ρ, and expand f in the orthonormal product basis. Then the Fourier weight above degree d satisfies

$$\mathbf{W}_{>d}(f) \le \mathbf{N}_{\rho}(f) / \left(2 \cdot [1 - (1 - \rho)^{d+1}] \right), \quad (7.1)$$

where $N_{\rho}(f) = \mathbb{E}[(f(X) - f(X^{\rho}))^2]$.

Proof. The product noise operator multiplies a degree-k Fourier coefficient by $(1 - \rho)^k$, so

$$\mathbb{E}[(f(X) - f(X^{\rho}))^2] = 2 \sum_S [1 - (1 - \rho)^{|S|}] \hat{f}(S)^2.$$

Every term of degree above d has multiplier at least $1 - (1 - \rho)^{d+1}$; rearranging gives the bound. ■

This is a precise version of the finite-resolution intuition: if a witness-relevant output event is *stable* under a specified finite perturbation kernel, its high-degree mass is controlled. Stability is an *additional* assumption — it is not implied by mere output distinguishability, and parity fails it sharply.

7.6 The compression-or-synergy dichotomy

Corollary 7.3 (Degree-d dichotomy). For any d, either the degree-d Fourier truncation of f has L^2 error at most δ (and therefore preserves at least $\delta/2$ covariance with the local cylinder U), or the causal cone has high-order spectral mass $W_{>d}(f) > \delta^2$.

So failure of low-degree compression is *not* structurelessness — it is a **positive certificate of high-order synergy.** Any counterexample to the proposed P versus NP route must sustain such synergy *while also* generating valid, locally testable witnesses with non-negligible probability. That is a demanding pair of requirements to meet simultaneously.

---

8. The causal-cone compression conjecture

8.1 Sharpened conjecture

Conjecture 8.1 (Witness-relevant causal-cone compression, C3). On a target hard CSP distribution equipped with a robust local witness encoding, every polynomial-time solver-derived fixed-seed cone f that has inverse-polynomial covariance with a local instance cylinder admits at least one lower-bound-compatible footprint:

1. a fixed observed or latent influence that translates to a low-degree recovery statistic;

2. a polynomial-size proof-search or resolution trace;
3. a polynomial-width transcript or decomposable representation covered by an appropriate lower bound; or
4. a degree-n^{α} approximation with inverse-polynomial retained covariance, for some fixed $\alpha < 1$.

C3 is strictly narrower than asserting that *every* polynomial-time function is low-degree or local. It applies only to cones extracted from **successful witness production**, only **after** local rectangle compression, and only on the **target distribution.** Parity is not a counterexample, because it already has a width-2 branching representation; a true counterexample must evade *every* listed footprint.

8.2 Equivalent escape-hatch description

A solver evading the program must produce a cone with *all* of the following at once:

- inverse-polynomial witness relevance, certified by the local rectangle covariance;
- no useful fixed-coordinate or latent-coordinate translation on the target dependent distribution;
- high Fourier degree or high noise sensitivity;
- large total influence, or no small-junta approximation;
- superlogarithmic effective memory, or no polynomial-width branching representation covered by lower bounds;
- large causal-cone treewidth, or no compact decomposable compilation;
- no polynomial-size resolution-style explanation of successful runs.

This is a far more specific object than a generic mysterious global algorithm. Constructing one would itself identify a new computational paradigm.

8.3 Why proving C3 remains difficult

General polynomial-size circuits can compute high-degree functions, use polynomial memory, and have large treewidth. Compressing *all* such circuits would amount to major circuit-lower-bound progress. The leverage in C3 must therefore come from the **witness-relevant covariance** and the **target distribution** — not from generic circuit structure alone. That is the honest location of the remaining difficulty.

---

9. Consequences for the P versus NP framework

9.1 What this route can replace

In a standard finite sequential model, the claim that a useful answer must have a traceable origin in the problem is *not* conjectural: Theorems 3.1–3.3 establish finite causal effect, information

flow, and a non-negligible transcript region directly, and under product encodings Theorem 6.3 extracts a fixed influential coordinate. The remaining global-to-local issue should therefore be **reformulated as a compression problem**: can every witness-relevant causal ancestry be converted into an analyzable structure on the target hard distribution?

9.2 Conditional proof chain

1. Assume a polynomial-time solver produces valid witnesses with success $s \geq 1/\text{poly}(n)$.
2. Because its reliable output is more useful for the real problem than an output generated from an independent problem, that output has finite, traceable causal ancestry in the real input.
3. A robust local witness tester and an anti-collision bound convert solver success into a finite local real-versus-decoupled gap.
4. An oblivious tester yields a local instance–output rectangle of inverse-polynomial discrepancy, and a fixed seed isolates a deterministic witness-relevant causal cone.
5. Product or mixture structure yields either an influential observed coordinate or a low-dimensional latent recovery signal.
6. C3 — or one of the parameterized compression theorems of §7 — converts that signal into a low-degree, proof-search, branching-program, or decomposable footprint.
7. A matching lower bound on the *same* distribution rules out that footprint, and therefore rules out the assumed polynomial-time solver.
8. A separate distributional-to-worst-case bridge is still required to reach the traditional statement $P \neq NP$.

Steps 1–5 are proved for the stated model. Step 6 is conditional (C3). Steps 7–8 require distribution-specific lower bounds and a hardness-transfer argument, respectively. The paper's contribution is to make the boundary between the proved and the conditional parts *exact.*

9.3 Relation to the two accompanying manuscripts

The parent manuscript proves detectability of reliable success and develops the physical entropy-dispersion interpretation [1]. The companion manuscript formalizes total influence, martingale variation, and structure extraction for restricted models while acknowledging the general gap [2]. The present note refines that gap in three ways: a universal ghost-instance causal theorem; robust local witness testing to obtain a genuinely local joint footprint; and the observed-or-latent dichotomy for mixture distributions.

9.4 The distribution problem is now explicit

Product-measure OSSS and KKL results do not automatically apply to random k-SAT conditioned on satisfiability. Planted models carry a useful latent product structure, but lower bounds for planted and conditioned distributions must be aligned. Recent influence theorems for mixing and monotonic measures suggest possible extensions beyond products [14, 15], but their hypotheses must be verified for the CSP distributions of interest.

10. Falsification criteria and research agenda

10.1 Falsification criteria

The framework would be refuted or materially weakened by any of the following:

- a finite sequential solver with non-negligible real-versus-decoupled advantage but *no* event satisfying the hybrid or information bounds (contradicting Theorem 3.1 or 3.2);
- a target witness relation for which no robust local encoding preserves solver success at useful parameters;
- a hard distribution with large cross-instance witness collision β, making independent proofs locally useful;
- a solver-derived local rectangle whose fixed-seed cone provably has none of the proposed representations and defeats the observed-or-latent dichotomy assumptions;
- a concrete polynomial-time NP solver whose witness-relevant cone is simultaneously high-space, high-synergy, high-treewidth, and outside all current proof and statistical models.

10.2 Immediate technical targets

1. Construct an explicit oblivious robust witness tester for the exact SAT/CSP encoding used in the entropy-dispersion program, with $b = O(\log n)$ or better.
2. Prove a uniform anti-collision bound $\beta = 2^{-\Omega(n)}$ for the target random or planted satisfiable distribution.
3. Instantiate Theorem 6.4 for planted k-SAT and compute the exact latent locality r of every tester atom.
4. Develop low-degree recovery lower bounds for the specific local encoded-witness events produced by the rectangle theorem, using recovery-oriented low-degree methods [12].
5. Extend OSSS- or KKL-style influence extraction to the conditioned satisfiable distribution via contiguity, mixing, or a planted-to-conditioned transfer theorem.
6. Determine whether high spectral synergy in the extracted cone forces large space, large treewidth, or a proof-complexity trace on random CSPs.
7. Use Valiant–Vazirani isolation to study a unique-witness specialization, where local proof bits have a canonical semantic target [11].
8. Separate the average-case compression theorem from the final worst-case reduction step.

10.3 Suggested first paper-level lemma

Target Lemma 10.1 (Planted local-recovery compression). For planted k-SAT in a specified hard density regime, any inverse-polynomial local rectangle covariance produced by Theorem 6.1 implies *either* an inverse-polynomial degree-$O(\log n)$ estimator of a local planted-witness block, *or* an inverse-polynomial influence of a fixed observed clause block. Proving this with exact constants would connect the present causal framework directly to low-degree recovery lower bounds.

11. Limitations

- **The usefulness test matters.** A global verifier gives distinguishability but may not give a local rectangle without an additional robust encoding.
- **Oblivious local testing is an interface assumption.** Standard PCP soundness for no-instances is not by itself enough; the tester must control invalid witnesses on yes-instances.
- **The anti-collision parameter is distributional** and may fail when many instances share witnesses.
- **Influence theorems are measure-sensitive.** Product, mixture, mixing, and conditioned measures each require separate treatment.
- **A fixed influential coordinate is not automatically a low-degree assignment-space polynomial.** Basis translation and recovery lower bounds remain necessary.
- **Noise stability is a sufficient compression condition, not a consequence** of finite distinguishability.
- **The fixed-seed reduction isolates one deterministic seed** but does not claim the same seed succeeds uniformly on every instance.
- **Even a complete average-case separation on one distribution** does not by itself establish worst-case $P \neq NP$.

12. Conclusion

This paper begins with a simple claim: a reliably correct answer cannot come from nowhere. The input is the complete problem instance — the question together with its data, rules, and constraints. Whether a solver follows many small clues or appears to identify one decisive marker, the useful distinction must have been produced from information in that problem and must reach the answer through a finite causal path.

The finite causal-ancestry, information-flow, transcript-region, and fixed-seed results make that statement mathematical. Robust local witness testing then supplies a relational bridge from successful witness production to a local real-versus-decoupled gap, a local instance–output rectangle, and a deterministic witness-relevant causal cone. Product measures yield observed influence; mixture-of-product models yield an observed-or-latent alternative.

The remaining condition is **compression, not causal origin.** One must show that every witness-relevant cone on the target hard distribution admits a lower-bound-compatible representation, or else exhibits a precisely characterized high-order synergy that can itself be ruled out. If that compression statement, together with the required common-distribution and worst-case bridges, is established, the framework yields the proposed conditional proof that quickly verifiable NP solutions cannot always be found quickly — that is, **$P \neq NP$.**

Appendix A. Proof details and constants

A.1 Transcript-region constants

Let $a = \varepsilon/m$ and let Z be the conditional signed hybrid difference at the event supplied by Theorem 3.1, so $\mathbb{E}|Z| \geq a$. Then either $\mathbb{E}[Z_+] \geq a/2$ or $\mathbb{E}[Z_-] \geq a/2$. In the first case set $B = \{Z \geq a/4\}$; since $\mathbb{E}[Z_+] \leq \Pr(B) + a/4$, we get $\Pr(B) \geq a/4$. The negative case is identical after reversing sign.

A.2 Rectangle atom count

For fixed tester randomness, suppose the tester reads b binary coordinates in total. Each assignment to those coordinates determines one atom $A\_u \times B\_v$, and the acceptance set is a union of at most 2^b disjoint atoms. If blocks have alphabet sizes $L_1, \ldots, L\_t$, replace 2^b by $\prod_i L\_i$, or by 2 raised to the total binary description length.

A.3 Influence normalization

For Boolean f on a product space, define resampling influence $\mathrm{Inf}_i(f) = \Pr[f(X) \neq f(X^{(i)})]$, where coordinate i is independently resampled. Equivalent normalizations differ by at most constant factors on unbiased Boolean coordinates. The OSSS statement used in Theorem 6.3 should be read with a consistent resampling convention.

A.4 Latent L^2 influence

For real-valued g and independently resampled latent coordinate i, define $\mathrm{Inf\_i}^{(2)}(g) = \tfrac{1}{2} \cdot \mathbb{E}[(g(\Theta) - g(\Theta^{(i)}))^2]$. Efron–Stein gives $\mathrm{Var}(g) \leq \Sigma_i \mathrm{Inf\_i}^{(2)}(g)$. This is the normalization used in Theorem 6.4.

Appendix B. Planted *k*-SAT specialization

B.1 Conditional product structure

In planted k-SAT, a planted assignment Θ is drawn first and clauses are then sampled independently subject to being satisfied by Θ. Thus the observed formula is a mixture of product distributions indexed by Θ. Theorem 6.4 applies to any fixed-seed output event once the local rectangle has been extracted.

B.2 Local latent dependence of a clause cylinder

Fix a cylinder specifying a constant number b of clause blocks, including their variable indices and sign patterns. Conditional on the indices, its probability under a planted assignment depends only on the planted bits of the variables appearing in those clauses; the union has size at most $k \cdot b$. Consequently the between-component term in Theorem 6.4 produces a degree-at-most-$k \cdot b$ latent footprint.

B.3 Remaining checks for a full instantiation

- The robust tester must be implemented on the same encoded-instance distribution used by the lower bounds.
- The anti-collision bound must hold for witnesses output by an independent successful run, uniformly over their marginal distribution.
- The observed-influence case must be translated from clause-generation coordinates to the canonical formula encoding.
- The latent case must be matched to a recovery lower bound covering the extracted local proof or witness event.
- The planted distribution must be related to the conditioned satisfiable distribution if both proof-search and low-degree lower bounds are to be combined.

References

[1] K. Taylor. *Operational Separation of Discovery and Verification: An Entropy-Based Framework for NP-Hardness Under Physical Admissibility.* Unpublished manuscript, 2026.

[2] K. Taylor. *Entropy Dispersion and Total Influence: A Structural Route Toward Conditional $P \neq NP$.* Unpublished companion manuscript, 2026.

[3] R. O'Donnell, M. Saks, O. Schramm, and R. A. Servedio. Every decision tree has an influential variable. *Proc. FOCS 2005*, 31–39; arXiv:cs/0508071.

[4] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. *Proc. FOCS 1988*, 68–80. DOI:10.1109/SFCS.1988.21923.

[5] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica* 18(1):27–35, 1998. DOI:10.1007/PL00009809.

[6] N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, Fourier transform, and learnability. *J. ACM* 40(3):607–620, 1993. DOI:10.1145/174130.174138.

[7] A. C.-C. Yao. Some complexity questions related to distributive computing. *Proc. STOC 1979*, 209–213. DOI:10.1145/800135.804414.

[8] I. Dinur and O. Reingold. Assignment testers: towards a combinatorial proof of the PCP theorem. *SIAM J. Comput.* 36(4):975–1024, 2006. DOI:10.1137/S0097539705446962.

[9] E. Ben-Sasson, O. Goldreich, P. Harsha, M. Sudan, and S. Vadhan. Robust PCPs of proximity, shorter PCPs, and applications to coding. *SIAM J. Comput.* 36(4):889–974, 2006. DOI:10.1137/S0097539705446810.

[10] M. Braverman and A. Rao. Information equals amortized communication. *IEEE Trans. Inform. Theory* 60(10):6058–6069, 2014. DOI:10.1109/TIT.2014.2347282.

[11] L. G. Valiant and V. V. Vazirani. NP is as easy as detecting unique solutions. *Theoret. Comput. Sci.* 47:85–93, 1986. DOI:10.1016/0304-3975(86)90135-0.

[12] T. Schramm and A. S. Wein. Computational barriers to estimation from low-degree polynomials. *Ann. Statist.* 50(4), 2022; arXiv:2008.02269.

[13] A. Amarilli, F. Capelli, M. Monet, and P. Senellart. Connecting knowledge compilation classes and width parameters. *Theory Comput. Syst.* / extended version arXiv:1811.02944, 2018.

[14] F. Koehler, N. Lifshitz, D. Minzer, and E. Mossel. Influences in mixing measures. arXiv:2307.07625, 2023.

[15] H. Duminil-Copin, A. Raoufi, and V. Tassion. Sharp phase transition for the random-cluster and Potts models via decision trees. *Ann. of Math.* 189:75–99, 2019; arXiv:1705.03104.