

Local Validity Does Not Compose

A Theorem on the Limits of Bounded AI Audit

John Komkov

April 2026

Contents

Abstract	2
1 Introduction	2
2 Setup: The Rank-1 Signed Model	4
Example 2.9 (One-day-early settlement in a three-tool payment pipeline)	6
3 Main Theorem	8
4 Proof	8
Lemma 1 (Tree Trivialization)	8
Lemma 2 (Local Gauge Indistinguishability on High-Girth Graphs)	8
Lemma 3 (Holonomy Classification)	9
Lemma 4 (Single-Cycle Obstruction)	9
Lemma 5 (Cycle Additivity)	9
Lemma 6 (Spectral Moment Blindness)	10
Proof of Theorem A	10
5 Corollaries	10
6 The Abstract Theorem	11
7 Lift to the Full Bulla Model	11
7.1 A Native Bulla Impossibility Construction (Theorem A')	12
7.2 What Theorem A' does and does not prove	15
8 The Constructive Complement	16
9 Empirical Confirmation	17
9.1 Spectral gap convergence	17
9.2 Same fee, different witness geometry, different repair cost	17
9.3 Minimum-cost repair is a matroid optimization problem	19
10 Implications for AI Assurance	21

11 Discussion	22
12 Conclusion	23
13 References	23

Abstract

Contemporary AI assurance is overwhelmingly local: components are evaluated individually, interfaces are tested pairwise, agent behaviors are sampled within bounded horizons, and spectral or structural summaries are computed from local neighborhoods. We prove that such evidence does not, in general, compose into a certificate of global semantic coherence. Specifically, we construct families of multi-agent compositions that are identical to every bounded-radius local audit yet differ arbitrarily in a computable global obstruction (the coherence fee). We show that low-order spectral summaries fail for the same structural reason: they depend on walk data that cannot traverse a cycle before the cycle closes. The proof uses high-girth graph families and classifies the obstruction as an element of $H^1(G; \mathbb{Z}/2)$ in a rank-1 signed model. A companion signed-incidence note (Komkov, 2026; [papers/signed-incidence/paper/signed-incidence.pdf](#)) supplies the middle-layer claim needed for the implementation-facing story: Bulla’s Q -valued fee is field-independent and therefore agrees numerically with the $\mathbb{Z}/2$ obstruction count, without identifying the two coefficient-field presentations as literally the same object. We then prove a native Bulla impossibility construction (Theorem A’) in the single-field-per-dimension (graphic matroid) regime by varying composition graph topology — a different mechanism (connectivity vs holonomy) yielding the same conclusion. We complement the impossibility results with two constructive contributions: an online maintenance algorithm ($O(n^2)$ per update via Schur complement) and a canonical benchmark family (semantic twins) that exposes the failure by construction. Empirical validation on 500 synthetic compositions shows that the connection Laplacian’s spectral gap converges to the random-noise baseline at scale, consistent with the theorem’s mechanism. A second empirical study on 240 real MCP server compositions demonstrates that compositions with identical scalar fee exhibit up to $4\times$ variation in repair cost — variation predicted by the witness Gram but invisible to the fee alone — and that leverage-guided repair consistently reduces total cost over geometry-naive strategies across four independent cost models, including an adversarial model designed to penalize geometry’s preferred ordering.

1 Introduction

As AI systems become composed — agents calling tools, tools calling other agents, retrieval-augmented pipelines orchestrating sub-pipelines, model-router stacks handing off between specialized models, multi-agent workflows coordinating across protocolized LLM systems — the

dominant failure mode shifts from component error to compositional semantic mismatch. Each interface passes its local check. Each tool produces valid output in its own terms. The composition fails silently because the terms do not agree globally.

The central mistake of current AI assurance is not insufficiently fine local analysis, but the assumption that sufficiently fine local analysis must compose. Current methods address compositions through local strategies:

- Component evaluation: benchmark each model, tool, or agent in isolation
- Pairwise interface testing: verify that each producer-consumer pair agrees on types, formats, and protocols
- Bounded-horizon rollouts: sample agent trajectories up to a depth budget and check for errors
- Mechanistic decomposition: inspect internal circuits, features, or activation patterns of individual components
- Spectral and structural summaries: compute graph metrics, Laplacian eigenvalues, or random-walk statistics over the composition topology

These all share a methodological assumption: that sufficiently detailed local certification composes into global system assurance. This paper gives a precise counterexample class and an exact alternative.

The issue is not merely theoretical. Consider a three-tool settlement pipeline. A capture service records `settlement_date = "2026-09-08"`, intending the next New York banking day after a holiday weekend. A risk engine converts that date into `days_to_settlement = 4`, using ordinary civil-day arithmetic. A treasury scheduler interprets that integer as four 24-hour intervals and emits `release_at = 2026-09-08T00:00:00Z`. Each interface passes its local checks: the date string is well formed, the horizon is plausible, the timestamp is schema-valid. Yet when the timestamp is translated back into the capture service's own semantic frame, it denotes the evening of 2026-09-07 in New York. The cycle does not close. Nothing is malformed locally; the failure is global.

The question here is not whether a sufficiently capable model may often guess the missing convention correctly. It is whether bounded local evidence can certify that a composition is semantically coherent before failure occurs. Stronger models may improve average-case semantic inference; they do not, by themselves, transform hidden conventions into explicit certificates.

Main contributions.

1. Impossibility. No bounded-radius local audit, and no spectral statistic of order below the composition's girth, can distinguish a globally coherent composition from one carrying k independent semantic obstructions (Theorem A). The result holds natively in the full \mathbb{Q} -valued Bulla model via a different construction mechanism — varying composition graph topology rather than connection values (Theorem A').
2. Observable. The coherence fee — the rank difference between full and observable

coboundary matrices — is an exact, computable invariant that captures the global obstruction. The underlying obstruction class is an element of $H^1(G; \mathbb{Z}/2)$; the fee is the number of independently twisted fundamental cycles in the rank-1 model (Theorem B).

3. Constructive tractability. The IncrementalDiagnostic maintains the exact fee under progressive field disclosures via Schur complement rank-1 updates, at $O(n^2)$ per disclosure. Minimum-cost full repair is an exact matroid optimization (min-weight basis of the contraction M/O), solvable in $O(\phi \cdot n^2)$ by the matroid greedy algorithm of Edmonds [5].
4. Benchmark family. The semantic twins construction provides canonical adversarial pairs: compositions with identical local views, identical low-order spectra, and arbitrarily different fees.
5. Implication. AI reliability must include compositional semantics — the study of whether individually legible components compose into a globally coherent system — not only component-level analysis.

Thesis. Local validity does not compose into global semantic coherence. We exhibit families of compositions that are identical to every bounded local audit yet differ arbitrarily in global obstruction, show that low-order spectral summaries fail for the same structural reason, and give an exact online diagnostic that remains tractable at scale.

For claim discipline: what is established here is the impossibility theorem in two forms (Theorem A in the rank-1 $\mathbb{Z}/2$ model, Theorem A' natively in the full \mathbb{Q} -valued Bulla model) and its constructive benchmark consequence. What failed productively is the hope that bounded local or low-order spectral evidence could certify global coherence. The companion signed-incidence note [4] ([papers/signed-incidence/paper/signed-incidence.pdf](#)) supplies the field-independence bridge between the two coefficient fields. Theorem A' bypasses this bridge entirely by constructing the impossibility directly in the Bulla regime [1, 2, 3].

2 Setup: The Rank-1 Signed Model

Definition 2.1 (Composition graph). A composition graph is a connected finite graph $G = (V, E)$ where V is the set of tools and E is the set of data-flow edges. For example, a retrieval-augmented generation pipeline has tools {retriever, ranker, generator} and edges for the document-to-query and query-to-response interfaces; a multi-agent workflow has one vertex per agent and one edge per handoff.

Definition 2.2 (Hidden semantic line). A rank-1 signed connection on G assigns to each oriented edge $e : u \rightarrow v$ a transport value $g_e \in \{\pm 1\}$, with the convention that reversing orientation preserves the value (since inversion in $\mathbb{Z}/2$ is the identity). A gauge transformation is a choice of signs $h_v \in \{\pm 1\}$ at vertices, acting by $g_e \mapsto h_v \cdot g_e \cdot h_u$. Gauge equivalence captures the

intuition that local relabeling — renaming “price” to “amount” or converting dollars to cents at a single tool — does not change the semantic content. What matters is whether such relabeling is globally consistent across cycles.

Definition 2.3 (Holonomy). For a directed cycle $\gamma = e_1 \cdot e_2 \cdot \dots \cdot e_m$, the holonomy is $\text{Hol}(\gamma) = \prod_i g_{e_i} \in \{\pm 1\}$. Holonomy is gauge-invariant: gauge transformations cancel at each internal vertex.

Definition 2.4 (Coherent and twisted instances). The coherent instance C_0 sets $g_e = +1$ for all edges. A twisted instance C_ω is defined by a cohomology class $\omega \in H^1(G; \mathbb{Z}/2)$: set transports so that $\text{Hol}(\gamma) = (-1)^{\langle \omega, [\gamma] \rangle}$ for every cycle γ . The cellular-cohomology framework adopted here is standard (Hatcher [11], Curry [10]); the spectral framework on the Laplacian operator that appears in Lemma 6 is developed by Hansen and Ghrist [9].

Definition 2.5 (Fee in the rank-1 model). A composition on G consists of a connection C (the full semantic structure, possibly hidden) and an observable projection that restricts what each vertex can see. The fee is the dimension of the obstruction that the hidden structure creates beyond what the observable structure can resolve:

$$\text{fee}(C) = \dim(\text{cycle obstructions under full connection}) - \dim(\text{cycle obstructions under observable projection})$$

In the rank-1 model, we make this concrete: the full connection assigns transports $g_e \in \{\pm 1\}$ to edges. The observable projection forgets the edge transports and retains only the vertex-local fiber values — each endpoint sees its own stalk but not the transition map to its neighbor. The fee counts the number of independent cycles whose holonomy is nontrivial — these are the obstructions that pairwise (observable) verification cannot detect but global (full) verification reveals.

Equivalently: choose a spanning tree T of G . Each non-tree edge e_i^* closes one fundamental cycle γ_i . After gauge-trivializing on T (Lemma 1), the residual transport on e_i^* equals $\text{Hol}(\gamma_i)$. The toy fee is the Hamming weight of the closure vector $(\text{Hol}(\gamma_1), \dots, \text{Hol}(\gamma_k)) \in \mathbb{F}_2^k$ — the number of fundamental cycles with nontrivial holonomy. (Note: this is the weight of the cohomology class ω expressed in the fundamental-cycle basis, not $\dim H^1(G; \mathbb{Z}/2)$ itself, which is always $\beta_1(G) = k$ regardless of the connection.)

For the coherent instance C_0 ($g_e = +1$ everywhere), every holonomy is $+1$, so $\text{fee}(C_0) = 0$. For a twisted instance C_ω with k independently twisted fundamental cycles, $\text{fee}(C_\omega) = k$.

This directly mirrors Bulla’s $\text{fee} = \text{rank}(\delta_{\text{full}}) - \text{rank}(\delta_{\text{obs}})$, where the rank difference counts the independent cycle-closing conditions that the full coboundary resolves but the observable coboundary does not.

Definition 2.6 (r-local audit). An r -local audit is any gauge-invariant estimator whose output depends only on the collection of rooted isomorphism classes of radius- r neighborhoods $\{B_r(v) : v \in V\}$, possibly aggregated across roots. This captures: pairwise interface checks ($r = 1$),

bounded-depth rollouts ($r = \text{depth}$), local feature decompositions ($r = \text{receptive field}$), and any method that inspects each component's neighborhood up to radius r .

Definition 2.7 (Low-order spectral summary). A spectral summary of order m is any function of the first m spectral moments $\{\text{tr}(A^j) : j = 1, \dots, m\}$ of a connection-dependent operator A . This captures: spectral gap estimation, Cheeger-type conductance bounds, random-walk mixing times, and any statistic that depends on closed walks of length $\leq m$.

Definition 2.8 (Graph family $G_{\{r,k\}}$ — Semantic Twins). For parameters $r \geq 1$ and $k \geq 1$, the graph $G_{\{r,k\}}$ is a bounded-degree connected graph with girth $g > 2r$ and first Betti number $\beta_1(G) = k$. Concretely: a tree backbone with k disjoint cycles attached, each of length at least $2r + 3$. The semantic twins on $G_{\{r,k\}}$ are the pair (C_0, C_ω) : identical local views, identical low-order spectrum, fee differing by k .

Example 2.9 (One-day-early settlement in a three-tool payment pipeline)

Consider a composed payment workflow with three tools:

- CaptureService, which records the intended settlement date of a trade
- RiskEngine, which converts that date into a settlement horizon
- TreasuryScheduler, which turns the horizon into a release timestamp

The composition graph has three edges:

1. CaptureService \rightarrow RiskEngine
2. RiskEngine \rightarrow TreasuryScheduler
3. TreasuryScheduler \rightarrow CaptureService

The third edge is a reconciliation edge: after scheduling release, the scheduler returns the release time to the capture service for hold reconciliation and ledger closure.

Suppose a trade is captured on Friday, 2026-09-04, immediately before the U.S. Labor Day weekend. CaptureService emits

```
settlement_date = "2026-09-08"
```

meaning: settle on the next New York banking day, which is Tuesday.

RiskEngine converts this to

```
days_to_settlement = 4
```

because September 8 is four civil days after September 4. This is locally reasonable. The field is well typed, in range, and consistent with ordinary date arithmetic.

TreasuryScheduler interprets that horizon as four 24-hour intervals and produces

```
release_at = 2026-09-08T00:00:00Z
```

This also passes its local checks. The timestamp is valid RFC3339, monotone in the input horizon, and lies within the expected release window.

The problem appears only when the value returns to CaptureService. In New York, midnight UTC on 2026-09-08 is still the evening of 2026-09-07. So the round-tripped value decodes to

```
date_NY(release_at) = "2026-09-07"
```

and the cycle-closing condition fails:

```
"2026-09-08" != date_NY(2026-09-08T00:00:00Z)
```

Each pairwise interface passed:

- CaptureService → RiskEngine: valid ISO date string
- RiskEngine → TreasuryScheduler: valid integer settlement horizon
- TreasuryScheduler → CaptureService: valid UTC timestamp

Yet the composition releases funds one business day early.

The hidden disagreement is not a bug in any single component. It is a mismatch among three unexposed semantic conventions:

- CaptureService: banking-calendar date
- RiskEngine: civil-day count
- TreasuryScheduler: UTC-midnight instant

These conventions are locally invisible but globally inconsistent. In the language of this paper, the composition has fee 1: one independent cycle-closing condition that observable interfaces cannot resolve.

A single additional disclosure removes the obstruction. Any one of the following would suffice:

- `calendar_basis = NY_FED_BUSINESS`
- an explicit `settlement_instant` instead of `days_to_settlement`
- `release_timezone = America/New_York` together with the release cutoff rule

Once one of these hidden semantic fields is made observable, the cycle closes and the fee drops to zero.

In the rank-1 toy model, each tool carries its own local semantic frame for the same latent quantity, settlement timing. What differs across edges is not the existence of the quantity but the transport used to compare it across tools. Each edge relation is locally satisfiable. The obstruction appears only in the product around the cycle.

3 Main Theorem

Theorem A (Local Indistinguishability of Global Obstruction). For every $r \geq 1$ and $k \geq 1$, there exists a bounded-degree connected graph $G_{\{r,k\}}$ and two compositions C_0, C_ω on the same graph such that:

1. For every vertex v , the restrictions of C_0 and C_ω to the radius- r ball $B_r(v)$ are gauge-isomorphic.
2. $\text{fee}(C_0) = 0$.
3. $\text{fee}(C_\omega) = k$.
4. Every spectral moment of the connection operator of order less than g agrees between C_0 and C_ω .

Hence no bounded-radius local gauge-invariant tester, and no low-order spectral statistic, can distinguish the coherent instance from the twisted one, while the global fee differs by k .

4 Proof

Lemma 1 (Tree Trivialization)

Every rank-1 connection on a tree is gauge-trivial.

Proof. Let T be a tree with root o . Define $h_o = +1$. For each vertex v , let $P(o,v)$ be the unique path from o to v . Set $h_v = \prod_{e \in P(o,v)} g_e$. Then for each edge $e : u \rightarrow v$ where u is the parent of v :

$$h_v \cdot g_e \cdot h_u = (\prod_{e \in P(o,v)} g_e) \cdot g_e \cdot (\prod_{e \in P(o,u)} g_e)$$

Since $P(o,v) = P(o,u) \cdot e$, the product telescopes to $+1$. Trees have unique paths, so the construction is well-defined. \square

Punchline: every 1-cocycle is locally a coboundary on a contractible neighborhood.

Lemma 2 (Local Gauge Indistinguishability on High-Girth Graphs)

If $\text{girth}(G) > 2r$, then every radius- r ball $B_r(v)$ is a tree. Hence any two rank-1 connections on G restrict to gauge-equivalent connections on $B_r(v)$.

Proof. A cycle inside $B_r(v)$ has length at most $2r$, contradicting $\text{girth}(G) > 2r$. So $B_r(v)$ is acyclic. Being a connected subgraph of a finite graph, it is a tree. Apply Lemma 1. \square

Consequence: any algorithm whose input is the collection of rooted radius- r neighborhoods, modulo gauge, receives identical data on C_0 and C_ω .

Lemma 3 (Holonomy Classification)

Two rank-1 connections on a connected graph are gauge-equivalent iff they have the same holonomy on every cycle. Gauge classes are classified by $H^1(G; \mathbb{Z}/2)$.

Proof. A rank-1 connection is a 1-cochain $c \in C^1(G; \mathbb{Z}/2)$. A gauge transformation adds a coboundary δh for $h \in C^0(G; \mathbb{Z}/2)$. So gauge classes are elements of $H^1(G; \mathbb{Z}/2) = \ker(\delta_1)/\text{im}(\delta_0)$.

If two connections differ by a coboundary, their evaluation on any cycle (which is a 1-boundary) agrees. Conversely, if their holonomies agree on all cycles, their difference pairs trivially with every element of the cycle space, hence lies in $\text{im}(\delta_0)$. \square

Punchline: what local observers can remove by gauge is exactly the coboundary part. What survives globally is exactly the cohomology class.

Lemma 4 (Single-Cycle Obstruction)

Let G contain a single simple cycle γ , and let the connection have $\text{Hol}(\gamma) = -1$. Then $\text{fee}(C_\omega) - \text{fee}(C_0) = 1$.

Proof. Choose a spanning tree T by deleting one edge e^* of γ . On T , gauge-trivialize the connection (Lemma 1). After gauge-fixing, all tree edges carry $g = +1$. The deleted edge e^* carries the residual transport $g_{\{e^*\}} = \text{Hol}(\gamma) = -1$.

The consistency equation for the cycle closing is: the propagated value along T from one endpoint of e^* to the other must equal the transport across e^* . When $\text{Hol}(\gamma) = +1$, the observable (gauge-trivialized) propagation agrees with the full propagation — no hidden obstruction, fee contribution = 0. When $\text{Hol}(\gamma) = -1$, the observable propagation produces $+1$ at the gap but the full propagation requires -1 — one independent hidden obstruction that pairwise verification cannot detect, fee contribution = 1. \square

Lemma 5 (Cycle Additivity)

Let G have cycle rank k with fundamental cycles $\gamma_1, \dots, \gamma_k$. If $\text{Hol}(\gamma_i) = -1$ for each i , then $\text{fee}(C_\omega) = k$.

Proof. Choose a spanning tree T . Each non-tree edge e_i^* closes exactly one fundamental cycle γ_i . After gauge-trivializing on T , the consistency matrix decomposes as:

[Tree propagation block | Cycle closure block]

The tree propagation block is invertible (Lemma 1). The cycle closure block is diagonal: entry (i,i) encodes the closing condition for γ_i . When $\text{Hol}(\gamma_i) = -1$, row i is inconsistent. When $\text{Hol}(\gamma_i) = +1$, row i is redundant.

Since the non-tree edges define a basis of the cycle space (standard graph theory), the k closing conditions are linearly independent over \mathbb{F}_2 . Each twisted cycle contributes one independent

obstruction dimension. Total obstruction: k . \square

Lemma 6 (Spectral Moment Blindness)

Let A_ω be any connection-dependent operator whose m -th spectral moment is $\text{tr}(A_\omega^m) = \sum_{\{\text{closed walks of length } m\}} (\text{product of edge weights})$. If m is less than $\text{girth}(G)$, then $\text{tr}(A_0^m) = \text{tr}(A_\omega^m)$.

Proof. Consider a closed walk W of length m . If W traverses some edge e more times in one direction than the other, the image subgraph of W (the set of edges with net nonzero traversal count) contains a topological cycle of length $\leq m$. If m is less than $\text{girth}(G)$, no such cycle exists in G , so every closed walk of length m must be backtracking: it traverses each edge equally often in both directions.

For a backtracking walk, each edge e contributes $g_e \cdot g_e^{-1} = 1$ in pairs. Hence the walk's total weight is $+1$ regardless of the connection. Therefore the m -th moment coincides for C_0 and C_ω . \square

Corollary: all spectral statistics depending on moments of order below g are blind to the cohomological obstruction.

Proof of Theorem A

Choose $G_{\{r,k\}}$ with $\text{girth } g > 2r$ and cycle rank k . Let C_0 be the trivial connection ($g_e = +1 \forall e$) and C_ω be the connection with $\text{Hol}(\gamma_i) = -1$ on each fundamental cycle.

1. By Lemma 2, every $B_r(v)$ in C_0 and C_ω is gauge-isomorphic (local indistinguishability).
2. $\text{fee}(C_0) = 0$: the trivial connection has $\text{Hol}(\gamma) = +1$ on every cycle, so every cycle-closing condition is automatically satisfied. No hidden obstruction exists.
3. By Lemmas 4 and 5, $\text{fee}(C_\omega) = k$ (each twisted cycle contributes one independent obstruction).
4. By Lemma 6, all spectral moments of order below $g = 2r + 3$ agree (moment blindness below girth).

Hence no bounded-radius local tester and no low-order spectral statistic can distinguish C_0 from C_ω , while the fees differ by k . \square

5 Corollaries

Corollary 1 (No bounded-depth estimator). Fix r . There is no radius- r local gauge-invariant estimator that approximates the fee on all bounded-degree compositions to additive error less than $k/2$.

Proof. Apply the estimator to C_0 and C_ω . By local indistinguishability it outputs the same value on both. But $|\text{fee}(C_0) - \text{fee}(C_\omega)| = k$. \square

Corollary 2 (No low-order spectral estimator). Any estimator depending only on spectral moments of order below $\text{girth}(G)$ fails on the pair (C_0, C_ω) .

Proof. The spectral data coincide by Lemma 6, but the fees differ. \square

Corollary 3 (Necessity of global cycle computation). Any exact coherence diagnostic must, in the worst case, access information spanning an entire nontrivial cycle.

Proof. If it did not, it would factor through local contractible neighborhoods and fail by Theorem A. \square

6 The Abstract Theorem

Theorem B (Local Triviality, Global Necessity). In the rank-1 composition model, the hidden semantic obstruction is represented by a cohomology class $[\omega] \in H^1(G; \mathbb{Z}/2)$. Its restriction to every contractible chart vanishes (Lemma 1), but its global class may be nonzero (Lemma 3). The coherence fee — the number of independently twisted fundamental cycles — is derived from $[\omega]$ and is therefore a genuinely nonlocal observable: every local chart sees $\text{fee} = 0$, but the global composition may have $\text{fee} = k$ for any $k \leq \beta_1(G)$.

Theorem A is the constructive separation corollary of Theorem B: it exhibits explicit graph families where the gap between local and global is arbitrarily large.

7 Lift to the Full Bulla Model

The rank-1 signed model is the distilled core of the full Bulla computation, not a loose analogy. Each lemma has a precise counterpart in the full \mathbb{Q} -valued model, and the correspondence is theorem-by-theorem:

Toy model ($\mathbb{Z}/2$)	Full model (\mathbb{Q})	Status
Lemma 1: tree trivialization	Hidden witness equations solvable by forward propagation on acyclic subcompositions. <code>diagnose()</code> returns $\text{fee} = 0$ on any tree.	Implemented; verified on 703 compositions.

Toy model ($Z/2$)	Full model (Q)	Status
Lemma 3: H^1 classifies obstruction	$\text{fee}(G) = \text{rank}(\delta_{\text{full}}) - \text{rank}(\delta_{\text{obs}}) = \text{rank}(K(G))$ where $K(G) = H^T(I - P_O)H$.	Lean-verified (Aristotle UUID 044a00b1). Backbone theorem.
Lemmas 4–5: cycle additivity	Each independent hidden mismatch around a fundamental cycle yields one independent witness vector in the quotient space. The matroid structure (column matroid of δ_{full} contracted by δ_{obs}) governs independence.	Empirically validated: 703/703 compositions match exactly.
Lemma 6: spectral blindness	Connection Laplacian spectral gap converges to null-baseline gap at scale (ratio 1.40 at $n=5 \rightarrow 1.01$ at $n=50$).	Empirically confirmed on 500 synthetic compositions (this paper, Section 9).

The toy model compresses the full model rather than caricaturing it. The $Z/2$ sign structure isolates the nonlocality cleanly: the transport values are $\{\pm 1\}$, the obstruction is a parity mismatch, and the cohomology is literally a yes/no per cycle. In the full Q -valued model, the transport values are rational matrices, the obstruction is a quotient-space dimension, and the cohomology is richer. But the local-to-global structure — trivial on trees, obstructed on cycles, invisible to bounded probes — is identical.

7.1 A Native Bulla Impossibility Construction (Theorem A')

The correspondence table above shows that the toy model's conclusions transfer to the full Bulla model. We now prove a complementary result: the impossibility holds natively in the Q -valued regime, via a construction that does not factor through the $Z/2$ model. The construction covers the single-field-per-dimension (graphic matroid) regime, where each tool contributes exactly one field per convention dimension. The multi-field regime — where tools contribute several fields per dimension and the carrier graph is richer than the composition graph — remains an open frontier (see §7.2).

Carrier graph dichotomy. Whether the composition graph's global topology is visible in the fee depends on how tool fields map to carrier-graph nodes:

- Single-field case: Each tool has one field per convention dimension. The carrier graph is isomorphic to the composition graph (nodes biject with tools, edges biject with composition edges). $\text{rank}(\delta_d) = n - c(G)$.
- Multi-field case: Tools have multiple fields per dimension. The carrier graph is richer; its component count depends on field routing, not just composition topology.

The single-field case makes the fee depend directly on global connectivity, which is the lever for the impossibility construction.

Lemma 7 (Carrier graph rank formula). Let $G = (V, E)$ be a composition on n tools, each with exactly one field in convention dimension d , where every edge connects that field to the corresponding field on the target tool. Then $\text{rank}(\delta_d) = n - c(G)$, where $c(G)$ is the number of connected components of G .

Proof. Under the single-field assumption, δ_d is the signed incidence matrix of G : column j corresponds to the unique (tool _{j} , field) pair, each row has one +1 (target) and one -1 (source). By the standard rank formula for signed incidence matrices (Schrijver, Theory of Linear and Integer Programming, Thm 19.3), $\text{rank} = n - c(G)$. \square

We now state the theorem with the full precision needed for a reviewer who will ask “what exactly does the local observer see?”

Definition 7.1 (r -local Bulla audit). An r -local Bulla audit is any function whose output depends only on, for each tool v , the following data restricted to the radius- r ball $B_r(v)$ in the composition graph:

- Tool schemas: For every tool $u \in B_r(v)$, the pair (internal_state(u), observable_schema(u)) — which fields exist, which are observable, and which are hidden.
- Edge structure: For every edge e within $B_r(v)$, the triple (source tool, target tool, dimension assignments) — which source field connects to which target field, in which convention dimension.
- Local coboundary restriction: The submatrix of δ consisting of rows (edges) and columns (tool-field pairs) supported within $B_r(v)$. This includes both the full coboundary (using internal_state) and the observable coboundary (using observable_schema) restricted to the ball.
- Derived local quantities: Any quantity computable from (a)–(c), including local rank, local leverage scores, local witness Gram restrictions, and local matroid structure.

The audit may aggregate this data across all roots $v \in V$ in any way. What it cannot do is access edges or tool-field pairs outside the r -ball of every root.

Theorem A' (Full-Bulla Local Impossibility). For every $r \geq 1$ and every $\Delta \geq 1$, there exist Bulla compositions G_0 and G_1 on the same tool set T such that:

1. Schema identity. Every tool has the same `internal_state` and `observable_schema` in both compositions.
2. Edge-type identity. Every edge uses the same convention dimension with the same field-to-field mapping in both compositions.
3. r -local indistinguishability. For every tool v , the data (a)–(d) of Definition 7.1 are identical in G_0 and G_1 . In particular, the local coboundary restrictions, local ranks, and local matroid structures agree on every r -ball.
4. Global fee gap. $\text{fee}(G_0) - \text{fee}(G_1) = \Delta$.

Hence no r -local Bulla audit (Definition 7.1) can distinguish G_0 from G_1 , despite the compositions differing by Δ units of coherence fee.

Construction. Set $n = 2(\Delta + 1)(r + 1)$. Define n tools $T_0, \dots, T_{\{n-1\}}$, each with: - `internal_state` = (d,) — one field in convention dimension d - `observable_schema` = () — no fields observable

All edges use dimension d with `from_field` = `to_field` = d .

- G_0 : single directed n -cycle ($T_0 \rightarrow T_1 \rightarrow \dots \rightarrow T_{\{n-1\}} \rightarrow T_0$)
- G_1 : $(\Delta + 1)$ directed cycles of equal length $n/(\Delta + 1) = 2(r + 1)$, on disjoint subsets of T

Proof.

Condition 1 (Schema identity). Both compositions use the same tool set T with identical schemas at each tool. ✓

Condition 2 (Edge-type identity). Every edge in both G_0 and G_1 connects field d on the source tool to field d on the target tool, in dimension d . The edge types are identical. ✓

Condition 3 (r -local indistinguishability). Both G_0 and G_1 are 2-regular directed graphs: each tool has in-degree 1 and out-degree 1. The girth of G_1 is $n/(\Delta + 1) = 2(r + 1) > 2r$. For any tool v and radius $r' \leq r$, the ball $B_{\{r'\}}(v)$ in a directed cycle of length $\geq 2(r + 1)$ is a directed path: the r' predecessors and r' successors of v , with r' incoming edges and r' outgoing edges. Since $\text{girth}(G_1) > 2r$, no ball of radius $\leq r$ contains a complete cycle in either composition. The r -balls are therefore isomorphic as rooted directed graphs.

Since all tools have identical schemas and all edges have identical dimension/field assignments, the data (a) and (b) are identical on corresponding r -balls. The local coboundary restriction (c) is the signed incidence matrix of the r -ball subgraph — a directed path in both cases, with the same dimensions. Its rank is (number of tools in ball) $- 1$, identical in both. All derived local quantities (d) — local leverage, local matroid structure, local witness Gram — are determined by (a)–(c) and therefore agree. ✓

Condition 4 (Global fee gap). By Lemma 7, $\text{rank}(\delta_d) = n - c(G)$. Since `observable_schema` is empty, $\text{rank}(\delta_{\text{obs}}) = 0$ in both compositions.

- G_0 has $c(G_0) = 1$: $\text{fee}(G_0) = (n - 1) - 0 = n - 1$.
- G_1 has $c(G_1) = \Delta + 1$: $\text{fee}(G_1) = (n - \Delta - 1) - 0 = n - \Delta - 1$.

- $\text{fee}(G_0) - \text{fee}(G_1) = \Delta$. ✓ □

Remark (Observable fields preserve the gap). If some tools' fields are made observable — with the same observable partition in both compositions — then $\text{rank}(\delta_{\text{obs}})$ depends only on data within the r -balls of observable tools. Since these r -balls are identical (Condition 3), $\text{rank}(\delta_{\text{obs}}, G_0) = \text{rank}(\delta_{\text{obs}}, G_1)$, and the fee gap is preserved.

7.2 What Theorem A' does and does not prove

What it proves. Any diagnostic that operates within the information class of Definition 7.1 — inspecting tool schemas, edge structures, coboundary restrictions, and derived local quantities within bounded-radius neighborhoods — cannot distinguish compositions with different fees. This rules out: pairwise interface checks ($r = 1$), bounded-depth rollouts ($r = \text{depth}$), local matroid analysis, local leverage computation, and any aggregation of such local data. The result sits in the locally-checkable-language tradition initiated by Naor and Stockmeyer [6] and Linial [7] (modern survey: Suomela [8]), which proves that distributed computation problems with non-trivial global structure cannot be decided in a constant number of communication rounds. The cohomological obstruction here plays the role that local indistinguishability plays in those classical results.

What it does not prove. The theorem does not rule out diagnostics that access global graph properties cheaply. In particular:

- Connected-component counting distinguishes G_0 from G_1 immediately (1 component vs $\Delta + 1$). The theorem's force is that component counting is itself a global operation — it requires traversing the entire graph, not just local neighborhoods. The construction does not claim that the fee gap is hidden from all efficient algorithms, only from bounded-local ones.
- Spectral methods with global access (computing the full Laplacian spectrum, not just low-order moments) can detect the component count. Theorem A' complements Lemma 6 (spectral blindness below girth) but does not extend it to full-spectrum methods.
- Adaptive message-passing that iterates beyond the girth may eventually detect the global topology. The theorem bounds what is achievable in a fixed number of rounds proportional to r .

The strongest objection and the response. A skeptic may say: "You changed the number of connected components. Of course the fee changed. This is graph counting, not an assurance impossibility." The response is precise: the construction demonstrates that the data available to any bounded-local audit (Definition 7.1) — which is exactly the data that component-level evaluation, pairwise testing, and bounded rollouts produce — is identical across compositions with different fees. The fact that a simple global computation (component counting) resolves the gap is not a weakness of the theorem; it is the theorem's point. The fee is a global invariant,

and the theorem proves that no local procedure can compute or approximate it. That global computation is easy does not make local computation sufficient.

Relationship to Theorem A. Theorem A varies connection values (holonomy $g_e \in \{\pm 1\}$) on a fixed graph; the mechanism is cohomological ($H^1(G; \mathbb{Z}/2)$). Theorem A' varies the composition graph while holding schemas fixed; the mechanism is connectivity ($n - c$). The two constructions produce the same conclusion — bounded-local audits are insufficient — through different algebraic routes. In the single-field-per-tool case, the signed-incidence bridge (total unimodularity, field-independence of rank) ensures the two rank formulas agree numerically, but the constructions themselves are logically independent.

Open frontier: the multi-field regime. Theorem A' covers compositions where each tool contributes one field per convention dimension (the graphic matroid regime). Real tool ecosystems often have tools with many fields per dimension — the MCP corpus includes tools with 12 path fields. In this regime, the carrier graph is richer than the composition graph, and $\text{rank}(\delta_d)$ depends on field routing, not just component count. Constructing r -locally indistinguishable pairs with different fees in this regime would require controlling carrier-graph topology through field assignments — a harder combinatorial problem that remains open.

8 The Constructive Complement

The impossibility result (Theorem A) is complemented by a constructive algorithm. The `IncrementalDiagnostic` maintains the exact fee under progressive field disclosures via Schur complement rank-1 updates:

$$K_{\text{new}} = K_{\text{old}} - (K_{\text{old}}[:,j] \cdot K_{\text{old}}[j,:]) / K_{\text{old}}[j,j]$$

Each disclosure costs $O(n^2)$ instead of $O(n^3)$ for full recomputation. The full repair trajectory ($\text{fee} = k \rightarrow \text{fee} = 0$) costs $O(kn^2)$. Correctness is validated by exact rational arithmetic: every intermediate K is bitwise identical to full recomputation.

The three contributions are tightly interlocked. The semantic twins (Definition 2.8) defeat every bounded-local audit (Theorem A). The same twins are exactly tracked by `IncrementalDiagnostic`: for an appropriate repair sequence of independent disclosures (each targeting a leverage-positive hidden field), the fee drops by one at each step until it reaches zero and the composition is certified coherent. Redundant disclosures leave the fee unchanged; the algorithm distinguishes the two cases in $O(1)$ via leverage lookup.

Beyond detection, the witness Gram supports exact repair optimization. The minimum-cost full repair is a matroid optimization problem — finding the minimum-weight basis of the contraction matroid M/O — solvable by the classical greedy algorithm using the same Schur complement infrastructure (Section 9.3). The scalar fee says how many fields must be disclosed; the

witness matroid says which, at minimum cost. The paper gives both the adversarial family that proves local methods must fail and the exact constructive algorithm that succeeds where they cannot.

9 Empirical Confirmation

9.1 Spectral gap convergence

The spectral gap experiment on 500 compositions across 7 scales provides empirical evidence consistent with the theorem’s worst-case mechanism:

Scale	$\lambda_2(\text{semantic})$	$\lambda_2(\text{null})$	Ratio	$R^2(\text{frustration})$	$R^2(\text{depth-8})$
5	0.0164	0.0117	1.40	0.986	0.497
20	0.0086	0.0067	1.29	0.976	0.328
40	0.0055	0.0053	1.04	0.966	0.233
50	0.0053	0.0053	1.01	0.972	0.270

The gap ratio converging to 1.0 is consistent with Lemma 6’s mechanism: spectral statistics that depend on walk data below the cycle scale lose discriminative power as compositions grow, while exact cycle computation ($R^2(\text{frustration}) > 0.96$ at all scales) remains stable. The theorem provides the worst-case separation; the experiment suggests the same local-to-global failure mode arises in practice on broader synthetic families.

9.2 Same fee, different witness geometry, different repair cost

The impossibility theorem establishes that global computation is necessary. A natural follow-up: once the fee is known, is the scalar fee sufficient for repair planning? Or does the witness geometry (leverage scores, effective resistance, loop/coloop structure) carry additional actionable information that the scalar fee does not?

Headline result. Within fee-matched groups of real MCP server compositions, repair cost varies by up to $4\times$ — compositions with identical fee have materially different repair difficulty, and the witness Gram predicts which is which.

Setup. We test on all 240 nonzero-fee pairwise compositions in the 703-composition MCP server corpus. For each composition, we simulate budgeted repair under five disclosure strategies:

- Geometry-guided (unit cost): greedy by leverage score — at each step, disclose the hidden field with highest leverage. Under unit costs, the matroid greedy theorem guarantees this reaches $\text{fee} = 0$ in exactly fee steps.

- Cost-weighted geometry: greedy by leverage-to-cost ratio, where costs are assigned by field semantics (pagination fields = 1, domain-state fields = 3, path/credential fields = 9).
- Cheapest-first: disclose fields in ascending cost order, ignoring witness geometry entirely.
- Random: mean of 50 uniformly random disclosure orderings.
- Worst-case: disclose loops first, then lowest-leverage fields — the ordering that wastes the most budget.

Repair step counts.

Strategy	Mean steps to fee = 0	Overhead vs geometry-guided
Geometry-guided	fee (exact, 240/240)	—
Random (mean of 50)	+12.7%	+12.7%
Worst-case	+43.7%	+43.7%

Geometry-guided disclosure reduces the fee by exactly 1 at every step on all 240 compositions, confirming the matroid greedy theorem empirically. Random disclosure wastes steps on loops (leverage-zero fields that do not reduce fee); worst-case disclosure does so deliberately.

Repair cost under the semantic cost model. Under the 1/3/9 semantic cost model, both strategies run until fee = 0 (full repair). Cost-weighted geometry (which uses leverage scores to avoid wasting expensive disclosures on redundant fields) never exceeds cheapest-first total cost on any of the 240 compositions, and saves 9.6% on average across compositions where the two differ (aggregate: 5582 vs 5998 cost units). The savings come entirely from loop-bearing compositions, where geometry avoids disclosing high-cost fields that contribute nothing to fee reduction.

The decisive finding: fee-matched variation.

Fee	N	Cost range (cost-weighted geometry)	Cost ratio
1	89	1–3	3×
2	67	6–18	3×
10	25	22–52	2.4×
11	36	23–93	4×

At fee = 11 (36 compositions), cost-weighted geometry repair ranges from 23 to 93 cost units — a 4× spread at the same scalar fee. The witness Gram captures which hidden fields are substitutable (low effective resistance, disclosing either suffices) vs irreplaceable (high effective resistance, only that specific field reduces the fee). This variation is invisible to any diagnostic that reports only the scalar fee.

Structural detail. Among the 240 compositions, 105 have multi-component witness Gram (block-diagonal hidden-field dependency), and 92 contain at least one loop (leverage-zero

hidden field). In loop-bearing compositions, random disclosure wastes steps on redundant fields that cannot reduce the fee — wasted work that leverage-guided disclosure avoids entirely.

Cost-model robustness. A natural objection is that the geometry advantage is an artifact of the specific semantic cost model (1/3/9). To test this, we repeat the comparison under four independent cost models: uniform (all fields cost 1), semantic (the 1/3/9 model), random (costs drawn from {1, 2, 3, 5, 8, 13} per field, seeded), and adversarial (cost = 1 + 12 × leverage, so the fields geometry most wants are deliberately made most expensive).

Cost model	Geometry wins	Cheapest wins	Ties	Mean savings
Uniform	68/240 (28.3%)	0/240	172	7.6%
Semantic	102/240 (42.5%)	0/240	138	9.6%
Random	56/240 (23.3%)	0/240	184	6.4%
Adversarial	105/240 (43.8%)	0/240	135	6.1%

Across all four models (960 total comparisons), cheapest-first never outperforms geometry-guided — not once. The ties are compositions where every hidden field is a coloop (leverage = 1), so every disclosure order yields the same total cost; there, geometry and cheapest-first agree by construction. The wins are compositions with loops, where geometry avoids wasting expensive disclosures on redundant fields. The result is not a cost-model artifact: it holds even when the cost model is designed to penalize geometry’s preferred ordering.

These results establish robust dominance of leverage-guided disclosure over cheapest-first across four cost models and confirm that scalar fee does not determine repair cost. Section 9.3 establishes that this dominance is not merely empirical: the geometry-guided strategy achieves the exact minimum-cost repair.

9.3 Minimum-cost repair is a matroid optimization problem

The full repair problem — disclose the cheapest set of hidden fields that drives the fee to zero — has an exact combinatorial characterization.

Theorem (Minimum-cost witness repair). Let G be a composition with fee $\phi > 0$ and hidden field set H . Let M/O denote the contraction of the column matroid M of δ_{full} by the observable column set O . Then:

1. The feasible full-repair sets (subsets $S \subseteq H$ such that disclosing S drives the fee to 0) are exactly the sets containing a basis of M/O .
2. The minimum-size full-repair set has cardinality $\phi = \text{rank}(M/O)$.
3. Under a linear cost function $c : H \rightarrow \mathbb{Q}_{>0}$, the minimum-cost full repair has cost equal to the minimum-weight basis of M/O under c .

4. The matroid greedy algorithm — process hidden fields in ascending cost order, disclose each field if it has positive leverage ($K[j][j] > 0$), skip it otherwise — computes the exact minimum in $O(\phi \cdot n^2)$ time.

Proof sketch. The witness matroid M/O has as its independent sets the subsets $S \subseteq H$ such that $\text{rank}(O \cup S) = \text{rank}(O) + |S|$, i.e., each element of S contributes one new rank dimension. The rank of M/O is $\text{fee}(G) = \text{rank}(M) - \text{rank}(O)$. A basis of M/O is a minimum-cardinality spanning set. The leverage condition $K[j][j] > 0$ is the independence oracle: it tests whether field j is independent of the already-disclosed fields in the contraction. By Rado’s theorem (1957), the greedy algorithm on any matroid with a linear weight function finds the minimum-weight basis. Each independence test costs $O(n^2)$ (one Schur complement update or leverage check), and there are at most n elements to process, giving $O(n^3)$ worst case or $O(\phi \cdot n^2)$ amortized since only ϕ elements are actually disclosed. \square

Empirical verification. We compare three strategies on all 240 nonzero-fee compositions under the semantic cost model (1/3/9):

Strategy	Aggregate cost	Overhead vs oracle
Oracle (matroid greedy by cost)	5582	—
Geometry-guided (greedy by leverage/cost)	5582	0.0%
Cheapest-first (greedy by cost, no loop skip)	5998	+7.5%

The geometry-guided strategy matches the oracle on all 240 compositions — not approximately, but exactly. Both achieve the provably minimum total cost. Cheapest-first pays 7.5% more because it wastes disclosures on loops (leverage-zero hidden fields that do not reduce the fee).

Why geometry-guided matches the oracle. The oracle and geometry-guided use different orderings (ascending cost vs descending leverage/cost ratio) but both skip loops. On this corpus, the matroid structures are simple enough — predominantly uniform or near-uniform contraction matroids — that all loop-skipping orderings produce the same minimum-weight basis. In principle, the two strategies can diverge on compositions with more complex matroid structure (non-uniform circuits, parallel classes); on the 703-composition MCP corpus, they do not.

Implication. Minimum-cost repair is not a heuristic problem requiring approximate methods or machine-learned priorities. It is an exact combinatorial optimization with a polynomial-time solution. The scalar fee tells you how many fields to disclose. The witness matroid tells you which fields to disclose and in what order to minimize cost. The matroid greedy algorithm, implemented via `IncrementalDiagnostic`’s Schur complement updates, provides the exact minimum at $O(\phi \cdot n^2)$ cost — the same complexity as the existing incremental diagnostic.

10 Implications for AI Assurance

The theorem identifies a structural blind spot in current AI assurance methodology. Component-wise evaluation, pairwise interface testing, bounded-horizon rollouts, and local spectral summaries are all instances of r -local auditing (Definition 2.6) or low-order spectral summarization (Definition 2.7). By Theorem A, none of these — individually or in combination — can certify global semantic coherence in the worst case.

This is not a claim about the limitations of any particular tool or technique. It is a claim about the information content of bounded local views within the hypothesis class studied here (Definitions 2.6–2.7). The obstruction lives in cycle-closing conditions that are structurally inaccessible to any probe that does not span a complete cycle. Within the bounded-local regime, finer-grained analysis (more features, more circuits, higher resolution) does not help — resolution alone does not escape the bounded-local regime.

Scaling can improve average-case semantic inference. It does not, by itself, convert hidden conventions into explicit certificates, especially in open compositions assembled from independently specified and independently evolving components. End-to-end training may hide interface mismatches inside one jointly tuned policy; it does not solve the verification problem for systems composed across organizational or protocol boundaries.

The practical consequence is that AI reliability for composed systems requires a new layer of assurance: one that examines global compositional semantics, not only component-level behavior. The coherence fee is a candidate observable for this layer: exact, computable, incrementally maintainable, and equipped with a theorem explaining precisely when local alternatives must fail.

Relation to mechanistic interpretability. Mechanistic interpretability studies internal mechanism — what circuit or feature produced a behavior. Compositional semantics studies whether individually legible mechanisms compose into a globally coherent system. The theorem establishes the boundary: local mechanism recovery does not settle global semantic coherence. This positions the fee not as a competitor to mechanistic interpretability but as a missing global layer above it.

Relation to formal verification and distributed systems. The result belongs to a longer local-to-global tradition in verification and distributed computing. Some properties are checkable by local certificates (e.g., graph coloring can be verified by checking each edge). Others require genuinely global witnesses (e.g., non-3-colorability requires a global argument). Our contribution identifies global semantic coherence of AI compositions as belonging to the latter class. The closest structural analogue is the impossibility of local checkability for certain graph properties in distributed computing (Naor and Stockmeyer 1995), where the boundary between locally checkable and globally necessary properties is a central object of study.

11 Discussion

On negative results and self-honesty. The spectral gap experiment (Section 9) rejected the phase-transition hypothesis per the precommitted failure condition. This illustrates the kind of epistemic posture the theorem enables: sharp domain-of-validity statements backed by impossibility results, not hedged caveats. A diagnostic system equipped with this theorem can state “no local method can do better on this family” — transforming abstention from a reliability heuristic into a theorem-backed impossibility.

On the hypothesis class. Theorem A covers r -local gauge-invariant estimators and spectral statistics of order below the girth. Theorem A' provides a complementary construction in the single-field-per-dimension (graphic matroid) regime, covering r -local Bulla audits (Definition 7.1) that inspect tool schemas, edge structures, coboundary restrictions, and derived local quantities within bounded neighborhoods. The multi-field regime — where tools have several fields per dimension and the carrier graph has richer structure than the composition graph — is not yet covered by the impossibility construction, though the correspondence table (§7) provides strong evidence that analogous results hold there. Neither theorem covers all conceivable local methods. Message-passing procedures with adaptive depth, iterative refinement schemes, and methods that access global topology through sampling would require separate analysis. The semantic twins family provides a concrete benchmark for testing whether any proposed method exceeds the proven bounds.

On girth realism. The semantic twins construction requires high girth ($g > 2r$), which is a worst-case construction. Real composition graphs may have shorter cycles where bounded-depth testing succeeds. The theorem's practical import is that hardness scales with cycle length — and composed systems with long dependency chains (multi-stage settlement pipelines, hierarchical agent orchestration, supply-chain workflows with reconciliation loops) naturally produce cycles that exceed any fixed testing radius.

On dynamic compositions. The IncrementalDiagnostic tracks disclosures on a fixed composition graph. Real agent systems discover tools and edges dynamically. Extending exact maintenance to graph-growth operations (tool addition, edge creation) is a natural next step; the Schur complement framework extends to column and row additions, though the details require separate treatment.

On the field. The triad of impossibility theorem, canonical observable, and constructive algorithm suggests a research direction that could be called compositional AI verification or local-to-global AI assurance: the study of when and why local evidence composes (or fails to compose) into global reliability certificates. This paper provides the first impossibility result and the first exact alternative for one precisely defined class of compositional semantic failure.

12 Conclusion

The coherence fee occupies a precise position in the local-to-global hierarchy: it is derived from an obstruction class that vanishes on every contractible neighborhood and survives only on nontrivial cycles. This is not a contingent property of the algorithm that computes it — it is a structural consequence of the fee being derived from a cohomological obstruction whose local restrictions are trivial.

The practical consequence is sharp: no amount of local testing, no matter how sophisticated, can approximate the fee on compositions whose cycle structure exceeds the testing radius. The only route to exact diagnosis is global cycle computation. The `IncrementalDiagnostic` provides this at $O(n^2)$ per update, making the impossible-locally-but-tractable-globally gap practically navigable.

Semantic mismatch is invisible for the same reason curvature is invisible in a coordinate patch: locally exact, globally obstructed.

13 References

- [1] J. Komkov, “The Coherence Fee,” Res Agentica program, 2026. (Paper I.)
- [2] J. Komkov, “The Column-Matroid Backbone,” Res Agentica program, 2026. (Paper II.)
- [3] J. Komkov, “The Witness Gram as a Kron-Reduced Laplacian,” Res Agentica program, 2026. (Paper III.)
- [4] J. Komkov, “Signed-Incidence Structure in Compositional Verification,” Res Agentica program, 2026.
- [5] J. Edmonds, “Matroids and the greedy algorithm,” *Mathematical Programming* 1 (1971), 127–136.
- [6] M. Naor and L. Stockmeyer, “What can be computed locally?,” *SIAM Journal on Computing* 24(6) (1995), 1259–1277.
- [7] N. Linial, “Locality in distributed graph algorithms,” *SIAM Journal on Computing* 21(1) (1992), 193–201.
- [8] J. Suomela, “Survey of local algorithms,” *ACM Computing Surveys* 45(2) (2013), Article 24.
- [9] J. Hansen and R. Ghrist, “Toward a spectral theory of cellular sheaves,” *Journal of Applied and Computational Topology* 3(4) (2019), 315–358.
- [10] J. Curry, *Sheaves, Cosheaves and Applications*. PhD thesis, University of Pennsylvania, 2014. arXiv:1303.3255.

[11] A. Hatcher, Algebraic Topology. Cambridge University Press, 2002.