A Witness Logic for Semantic Composition

John Komkov, *Independent Researcher*

8 May 2026

Abstract. Classical program logics make local correctness compositional under a stable semantic frame. Open autonomous systems break that assumption: components satisfying their local contracts can produce globally incoherent behavior because the semantic frame itself fails to glue across interfaces. We identify the obstruction invariant forced by this failure. Working over the category of *semantic interface complexes* — finite diagrams of local semantic carriers with observable projections and latent seam dimensions — we state six natural axioms for an obstruction invariant and prove that, in the exact regime, the unique invariant satisfying them is the witness rank, equal to the dimension of the first cohomology of the seam complex. As consequences we obtain (i) a *repair duality* identifying minimum disclosure cardinality with the obstruction invariant; (ii) a *receipt normal form* exhibiting every coherence certificate as a typed disclosure derivation; (iii) a *communication lower bound* showing that protocols whose transcript carries fewer independent declarations than the witness rank cannot soundly certify coherence; and (iv) a *Shannon-rate restatement* of the lower bound, with the canonical perturbation family centered at a complex as the "source," its cardinality — the coefficient-group size raised to the witness rank — as the entropy, and the disclosure normal form receipt as the constructive saturating code, converse and achievability closing to within polynomial overhead. These results recast the existing local–global obstruction work, the sheaf-cohomological predicate-invention framing, and the operational receipt protocol as instances of one mathematical object.

§1. The Missing Composition Rule

Hoare's rule for sequential composition [1]

$$\frac{\{P\}\,C_1\,\{R\} \quad \{R\}\,C_2\,\{Q\}}{\{P\}\,C_1\,;\,C_2\,\{Q\}}$$

works because the predicate R has a stable meaning at the seam between C_1 and C_2. The same logical content is asserted on both sides of the join. This is so deeply assumed that classical program logics rarely state it as an axiom.

Throughout, we use *component* for the formal objects in a semantic interface complex; *agent* for the deployed, autonomous entity that acts on commitments after the principal who configured it is no longer in the loop; and *tool* for the operational, wire-protocol realization — typically an MCP server — by which an agent's interface is exposed to another. The doctrinal claim of this paper is at the agent level; tools and services are particular operational instantiations. The problem we address is **agentic composition** — failures that arise from convention mismatch at the boundaries between agents, not from any individual agent's unreliability. Single-agent reliability mechanisms — formal verification, deterministic execution, constrained decoding — do not eliminate this class of failure: deterministic agents with mismatched conventions still compose into globally incoherent workflows.

Open autonomous systems do not satisfy the Hoare assumption. A calendar tool returns a date in ISO-8601 with implicit UTC. An invoicing tool ingests it as a local-timezone date. Each tool's interface is type-correct; each tool's local contract is satisfied; the composition produces a wrong-by-twelve-hours invoice that is undetectable by either tool's local audit. The same predicate "this is the meeting date" carries one meaning on the left and another on the right. The seam silently changes the semantic frame.

This phenomenon is not a bug of one tool or one interface design. It is the structural feature of compositions whose components were authored by different parties under different conventions. Heterogeneous schemas, units, time zones, currencies, partition orders, null-handling discipline, identifier conventions: each is a degree of freedom along which the local meaning of a value can shift across a boundary. The empirical literature on multi-agent tool composition documents this at scale — bilateral validity does not imply compositional coherence — and provides calibrated diagnostics that detect it where local audits fail.

What is missing is not another diagnostic. What is missing is the *composition rule itself*: the axiom that says how local correctness combines into global correctness when the semantic frame is mobile across the seam. Existing local-reasoning systems — most prominently the frame rule of separation logic [2] — formalize when properties survive composition under a *stable* framing condition; they do not address the regime in which the framing itself is what fails to glue.

We call this missing rule the **witness logic**, and the obstruction it tracks the **witness rank**. The contribution of this paper is to identify the mathematical object that witness rank is, and to show that — in a regime made precise below — it is forced by axioms that any reasonable theory of compositional semantic obstruction would adopt.

The shape of the argument has precedent. Shannon characterized entropy by axioms on a function of probability distributions; the axioms are continuity, symmetry, additivity, and a normalization condition. Once stated, the axioms force entropy to be $-\sum p \log p$ uniquely, up to a constant. Eilenberg and Steenrod [10] characterized ordinary homology by five axioms — homotopy invariance, exactness, dimension, additivity, excision — and proved that any homology theory satisfying them is naturally isomorphic to singular homology. Tutte [7], Whitney [6], and others characterized matroid rank by exchange and submodularity (the modern reference is Oxley [8]). In each case the axiomatic move turned a useful construction into the canonical mathematical object: the unique answer to a structural question, not one answer among many.

The bid here is the same in shape, and we want to be precise about where the originality lies. We define a category $\mathsf{SemComp}$ of semantic interface complexes; we identify *seam-independent disclosure* as the primitive operation of compositional repair; we prove that any numerical invariant respecting this primitive equals the dimension of the first cohomology of the seam complex (the witness rank). The operative claim is that disclosure independence — a structural property of the cochain complex of a semantic interface complex — is the right primitive for the theory. Once that primitive is accepted, the obstruction invariant is forced.

This is a useful way to read the theorem. It does not say "we wrote down a few axioms and were astonished to discover cohomology." It says: identify the right structural primitive (disclosure independence on the seam complex), and the obstruction invariant is uniquely determined. Witness rank is then not one possible diagnostic among many. It is the obstruction invariant respecting the primitive operation of repair.

The consequences are immediate. *Minimum repair* equals the obstruction by the same axioms (§6). *Coherence certificates* normalize to typed disclosure derivations (§7). *Protocols whose transcripts cannot carry an obstruction-class-many independent declarations* cannot soundly certify coherence (§8); equivalently, the *rate* at which a sound coherence protocol must transmit semantic-frame information is forced to be $\log_2 |\mathbb{E}|$ bits per disclosed cocycle generator, with converse and constructive achievability closing to within a polynomial-in-$|G|$ correction (§8.6). The local–global impossibility that appears as a separate theorem in earlier work becomes a corollary: any invariant satisfying A1–A6 must give the same impossibility, because the invariant is forced.

The analogy to distributed systems is worth stating once. Lamport's happens-before relation [5] made causality in distributed systems a structural invariant: a partial order forced by what messages can carry, not chosen by the protocol designer. The role we propose for witness rank is the analogous one for compositional semantics: a structural invariant forced by what interfaces can carry, not chosen by the verifier designer. Whether the analogy holds at the level of adoption is a matter of subsequent decades; whether it holds at the level of mathematical structure is what this paper sets out to establish.

Scope of the contribution

We owe the reader an explicit table separating *modeling thesis* (what we propose as the right abstraction) from *verified theorem* (what is mathematically forced) from *open conjecture* (what is named but not yet proved).

Layer	Claim	Status
1	Semantic interface complex (§2) is the right formal object for compositional semantic obstruction.	Modeling thesis
2	Disclosure independence on the cochain complex (§3) is the structural primitive of compositional repair.	Modeling thesis
3	The exact regime (§3.5) is a natural subcategory closed under the basic operations.	Modeling thesis (with structural support)
4	Witness rank is the unique numerical invariant respecting the primitive (Theorem 5.1).	**Verified theorem** (Lean: `disclosure_characterization`, Aristotle run `ad67beb2`)
5	Repair-cardinality equals obstruction count (Theorem 6.2).	Verified theorem (mathematical, follows from Layer 4)
6	Disclosure normal form is canonical (Theorem 7.1) and receipts are sound + complete (Propositions 7.3–7.4).	Verified theorem (mathematical, follows from Layers 4–5)
7	Communication lower bound: protocols below witness rank cannot certify coherence (Theorem 8.4; four-form refactor §§8.2–8.6).	Verified theorem (pigeonhole; Lean: `pigeonhole_perturbation_family` + 9 supporting theorems in lean/CompositionDoctrine/CoherenceRate.lean, Aristotle runs `b84c8f33` and `3fd02ee6`)
8	Bulla, the seam protocol, BABEL, the Coherence Cliff are realizations of the abstract framework (§10).	Modeling claim with empirical support

Layer	Claim	Status
9	Surrogate-regime extension: parallel characterization with explicit correction term (Conjecture 9.1).	Open
10	Verifier universality: any sound certification factors through the witness functor (Conjecture 9.2).	Open
11	Semantic FLP-style impossibility (Conjecture 9.3).	Open
12	Eval-correctness inseparability formalized as theorem (Conjecture 9.4).	Open (empirically supported)
13	Coherence-preserving interface evolution (Conjecture 9.5).	Open
14	Quantum / classical sheaf-cohomological capacity correspondence (§9.6).	Frontier — verified theorem on quantum side (Kurisummoottil Thomas and Chen [18]); this paper supplies the canonical classical lift

The paper's *theorems* are layers 4–7. The paper's *modeling theses* are layers 1–3 and 8. The paper's *open conjectures* are layers 9–13. Layer 14 is a non-conjectural frontier note recording an observed cross-substrate correspondence.

The Lean formalization closes layer 4 (numerical uniqueness in the abstract exact regime). Layers 5–7 follow from layer 4 by short mathematical arguments not separately formalized. Layers 1–3 are not amenable to Lean verification — they are claims about whether a mathematical structure is the right model of a phenomenon, and external review (not machine proof) is the appropriate verification mechanism.

What this paper does, and what it does not

We prove the characterization theorem in the *exact regime*, defined in §3 as the class of semantic interface complexes whose relative seam complex truncates to two terms — equivalently, whose witness Gram has full row rank, equivalently, the regime in which the rank-counting decomposition of the obstruction is exact rather than a lower bound. In this regime the consequences in §6, §7, §8 are theorems. Outside this regime — in the surrogate regime, where the same machinery yields a sound lower bound rather than an exact value — the analogous characterization is conjectured and named as open in §9.

We prove that *minimum independent disclosure* equals the obstruction invariant. We do not prove that *every* sound certification scheme factors through an equivalent witness object; that would be a verifier-universality theorem, which is named as open in §9 and pursued in subsequent work.

We state and prove a communication lower bound: protocols that cannot carry an obstruction-class-many independent semantic declarations cannot soundly certify coherence. This is a clean pigeonhole result and not a claim about asynchrony, fault tolerance, or liveness.

We do not claim that the axioms are minimal in the sense of Eilenberg–Steenrod's careful proof of axiom independence. We give natural arguments for each axiom and verify non-circularity, but a fully formal independence analysis is left to future work.

We work in the finite-dimensional setting. The composition is finite; the seam complex is finite-dimensional; the cohomology is computable. Extensions to infinite or limit-shaped compositions are not pursued.

What the Lean verification closes

The companion Lean module (papers/composition-doctrine/lean/) formalizes the abstract setting at the level of `DoctrineCarrier` and proves Theorem 5.1 in that setting. Aristotle (Harmonic) verified the proof on 2026-04-26 (run `ad67beb2-9f8e-48d6-9e5e-4cce51520afa`); three theorems compile sorry-free, depending only on `propext` and `Quot.sound`.

A second module, lean/CompositionDoctrine/CoherenceRate.lean, formalizes the cardinality form of the communication lower bound (§§8.2–8.6) at theorem-statement precision: 10 sorry-free theorems covering the pigeonhole reduction (`pigeonhole_perturbation_family`), the abelian rank-times-alphabet form (`coherence_rate_abelian_finite`), the variable-alphabet form (`coherence_rate_variable_alphabet`), the uniform-alphabet specialization (`coherence_rate_variable_specializes_uniform`), the tight-bound corollary (`coherence_rate_bound_is_tight`, `coherence_rate_tight_abelian`), and four boundary cases (`coherence_rate_rank_zero`, `coherence_rate_rank_one`, `coherence_rate_binary_rank_one`, `coherence_rate_binary_rank_two`). Aristotle verified the module across two runs: `b84c8f33` (initial 6 theorems) and `3fd02ee6` (extension to 10). The non-abelian / pointed-set form of Theorem 8.4 is stated but not Lean-verified — it lies outside mathlib's current scope on non-abelian cohomology.

What this verification *does* close: the numerical-uniqueness layer (layer 4 of the scope table) and the cardinality form of the communication lower bound (layer 7). Given the abstract structure and the axioms, the Lean proofs confirm that the witness rank is the unique invariant and that the rate $\log_2 |\Omega_G|$ is forced by pigeonhole over the perturbation family.

What the verification *does not* close: that the abstract structure is the right model of compositional semantic obstruction (layers 1–3). Whether `DoctrineCarrier` faithfully captures what semantic interface complexes should be, whether disclosure independence is the natural primitive of repair, and whether the exact regime is a natural subcategory — these are modeling claims about the *adequacy* of the formalization, and machine verification cannot adjudicate them. External review by applied-topology, formal-methods, and distributed-systems readers is the appropriate verification mechanism for these layers.

The honest summary: *Lean closes the numerical uniqueness theorem in the abstract exact regime; external review must test whether the structural primitive is the right one.* This separation between mathematical content (verified by Lean) and modeling thesis (subject to external review) is preserved throughout the paper and made explicit in the scope table above.

Outline

§2 defines semantic interface complexes and the category SemComp. §3 introduces the witness complex, the witness rank, and the matroid characterization of disclosure independence. §4 states the doctrinal axioms (A1, A2, A3, A4a, A4b, A5, A6) with motivation. §5 states and proves the Disclosure Characterization Theorem in the exact regime, observing that A1–A4b alone suffice (Proposition 5.4). §6 derives repair duality. §7 establishes the disclosure normal form theorem and the soundness/completeness of receipts. §8 proves the communication lower bound, sharpens it into four explicit forms (cardinality, abelian rank-times-alphabet, variable-alphabet, non-abelian / pointed-set) over the canonical perturbation family centered at G (§8.2), and gives the Shannon-rate restatement showing converse and achievability close to within a polynomial-in-$|G|$ correction (§8.6). §9 names five open conjectures — surrogate-regime characterization with explicit correction term, verifier universality, semantic FLP-style impossibility, eval-correctness inseparability, and coherence-preserving interface evolution — and records one non-conjectural frontier note (§9.6): the canonical-lift coincidence between this paper's classical lower bound and Kurisummoottil Thomas and Chen's quantum capacity theorem [18]. §10 returns to the existing program — Bulla, the seam protocol, BABEL, the Coherence Cliff scaling family — and identifies each as a model of the abstract framework.

---

§2. Semantic Interface Complexes

We work over $\mathbb{Q}$ throughout. Field-independence of the rank invariant between $\mathbb{Z}/2$ and $\mathbb{Q}$ is established elsewhere; we may choose the rational presentation freely.

2.1 The basic object

Definition 2.1 (Semantic interface complex). A *semantic interface complex* is a tuple

$$G = (P, \mathcal{F}, \rho, \mathcal{O})$$

where:

(a) $P = (P, \leq)$ is a finite poset, the *context poset*. Its elements are *contexts*; covering relations $p \lessdot q$ represent refinement of context.

(b) $\mathcal{F} : P^{\mathrm{op}} \to \mathsf{Vect}_{\mathbb{Q}}$ is a finite-dimensional presheaf — the *convention presheaf* — assigning to each context $p \in P$ a finite-dimensional rational vector space $\mathcal{F}(p)$ of *conventions* at p.

(c) $\rho_{p,q} : \mathcal{F}(q) \to \mathcal{F}(p)$ for $p \leq q$ are the *restriction maps* of $\mathcal{F}$, satisfying $\rho_{p,p} = \mathrm{id}$ and $\rho_{p,r} = \rho_{p,q} \circ \rho_{q,r}$ for $p \leq q \leq r$.

(d) $\mathcal{O} \subseteq \mathcal{F}$ is a subpresheaf, the *observable subpresheaf*. For each p, $\mathcal{O}(p) \subseteq \mathcal{F}(p)$ is a designated subspace, and the restriction maps satisfy $\rho_{p,q}(\mathcal{O}(q)) \subseteq \mathcal{O}(p)$.

The quotient presheaf $\mathcal{L} := \mathcal{F}/\mathcal{O}$ is the *latent presheaf*. A class $[v] \in \mathcal{L}(p)$ is a *seam dimension at* p. The restriction maps of $\mathcal{L}$ are induced from those of $\mathcal{F}$.

We write $|G| := \dim_{\mathbb{Q}} \bigoplus_{p \in P} \mathcal{F}(p)$ for the size of the complex.

2.2 Reading the definition

The object encodes three pieces of data: *where* meanings live (the contexts P), *what* meanings are available there (the conventions $\mathcal{F}$), and *which of those meanings are exposed* across boundaries (the observables $\mathcal{O}$). The seam information — meanings that two contexts could disagree on without either being able to detect the disagreement locally — is precisely what is *not* in $\mathcal{O}$, and that residue is the latent presheaf $\mathcal{L}$.

The classical case where there is no seam problem corresponds to $\mathscr{F} = \mathscr{O}$: every convention is observable across every refinement. In that case the obstruction we will define vanishes by construction. The interesting case — the open-systems case — is when $\mathscr{O}$ is strictly smaller than $\mathscr{F}$ at some contexts.

2.3 Operations and morphisms

Three operations on semantic interface complexes are needed for the axioms to have content. We define each explicitly; their categorical packaging is routine and we omit it.

(I) Isomorphism. An *isomorphism* $\phi : G \to G'$ is an order-isomorphism $\phi_P : P \to P'$ together with a family of linear isomorphisms $\phi_p : \mathscr{F}(p) \to \mathscr{F}'(\phi_P(p))$ commuting with restriction and carrying $\mathscr{O}(p)$ onto $\mathscr{O}'(\phi_P(p))$.

(II) Disjoint sum. Given G and H with disjoint context posets, the *disjoint sum* $G \sqcup H$ has context poset $P \sqcup P'$, convention presheaf $\mathscr{F} \oplus \mathscr{F}'$, and observables $\mathscr{O} \oplus \mathscr{O}'$.

(III) Admissible disclosure. Given G, a context $p \in P$, and a one-dimensional subspace $\bar{W} \subseteq \mathscr{L}(p)$ (equivalently, a line in the latent quotient at p), let $W \subseteq \mathscr{F}(p)$ be any lift mapping isomorphically onto $\bar{W}$. The disclosure (p, W) is *admissible* if it satisfies the **downward-closure condition**:

$$\rho_{a,p}(W) \subseteq \mathscr{O}(a) \quad \text{for every } a < p.$$

Equivalently, adding W to the observable subpresheaf at p extends $\mathscr{O}$ to a valid subpresheaf — the enlargement does not violate the presheaf restriction property at any context below p.

Given an admissible disclosure (p, W), the *disclosure of W at p*, denoted $G \setminus\!\!\setminus W$, is the complex $(P, \mathscr{F}, \rho, \mathscr{O}_W)$ where

$$\mathscr{O}_W(q) := \begin{cases} \mathscr{O}(p) + W & \text{if } q = p \\ \mathscr{O}(q) & \text{otherwise.} \end{cases}$$

Well-definedness. $\mathscr{O}_W$ is a subpresheaf of $\mathscr{F}$. For $q \geq p$: $\rho_{p,q}(\mathscr{O}(q)) \subseteq \mathscr{O}(p) \subseteq \mathscr{O}(p) + W = \mathscr{O}_W(p)$, so restrictions into p from above land in $\mathscr{O}_W(p)$. For $a < p$: $\rho_{a,p}(\mathscr{O}_W(p)) = \rho_{a,p}(\mathscr{O}(p) + W) \subseteq \mathscr{O}(a) + \rho_{a,p}(W) \subseteq \mathscr{O}(a) = \mathscr{O}_W(a)$, where the last inclusion is the admissibility condition. So disclosure modifies the observable subpresheaf at exactly one context, removing one dimension from $\mathscr{L}(p)$ while leaving $\mathscr{L}(q)$ unchanged for $q \neq p$.

Remark. If the downward-closure condition fails, the appropriate operation is the subpresheaf closure of $\mathscr{O}$ generated by W — a multi-context enlargement whose effect on the witness rank may exceed one. Such operations lie outside the one-generator disclosure calculus of §§4–7.

The *underlying disclosure* of $G \setminus\!\!\setminus W$ is the data (p, W). Two disclosures (p_1, W_1) and (p_2, W_2) on the same G are *cochain-independent* if the corresponding 1-cochains in the seam complex (§3.1 below) are linearly independent modulo coboundaries. Independence is a property of the seam complex of G, definable without reference to any invariant on G.

2.4 Covers

Definition 2.2 (Cover). A *cover* of G is a pair of full subcomplexes (U, V) — that is, restrictions to subposets $P_U, P_V \subseteq P$ inheriting the convention data — such that:

(i) $P_U \cup P_V = P$.

(ii) The intersection $U \cap V$ is the full subcomplex on $P_U \cap P_V$.

(iii) The convention and observable data on $U \cap V$ agree with the restriction from both U and from V.

A cover is *compatible* if the restriction maps from the union to U and to V exhibit $\mathcal{F}_G$ as the pullback of $\mathcal{F}_U \to \mathcal{F}_{U \cap V} \leftarrow \mathcal{F}_V$, and similarly for $\mathcal{O}$. Compatibility ensures that G is recoverable from its restrictions U, V, and $U \cap V$ as a presheaf — this is the condition under which Mayer–Vietoris exact sequences exist.

2.5 The category SemComp

Let SemComp be the category whose objects are finite semantic interface complexes and whose morphisms are generated under composition by isomorphisms, disjoint-sum injections, disclosure morphisms, and cover restrictions. The categorical scaffolding (presheaves, pullbacks, monoidal structure on disjoint sums) is standard; Mac Lane [9] is the canonical reference. We do not need a complete description of all morphisms in SemComp; we need only that the operations (I)–(III) and Definition 2.2 are well-defined and that an invariant on SemComp — a function $I : \mathrm{Ob}(\mathsf{SemComp}) \to \mathbb{Z}_{\geq 0}$ that is constant on isomorphism classes — has a well-defined value on each operation.

2.6 The exact regime

The axiomatic characterization in §5 is proved in a restricted class of objects, the *exact regime*. We give the abstract definition here; §3.5 shows that it agrees with the operationally-defined exact regime in the existing implementation literature.

Definition 2.3 (Exact regime, informal). A semantic interface complex $G = (P, \mathcal{F}, \rho, \mathcal{O})$ is in the *exact regime* if its seam cochain complex (Definitions 3.1–3.3) truncates to two terms: equivalently, if the rationally-defined rank decomposition of the seam coboundary satisfies

$$\dim_{\mathbb{Q}} H^1(G) = \operatorname{rank} \delta_1^{\mathcal{F}} - \operatorname{rank} \delta_1^{\mathcal{O}}$$

without correction terms from higher coboundaries. The precise form is given in Definition 3.5.

Let $\mathsf{SemComp}_{\mathrm{ex}}$ be the full subcategory of SemComp consisting of complexes in the exact regime. The category is closed under isomorphism, disjoint sum, disclosure, and compatible covers — these closure properties are immediate from the definitions and we record them without proof here.

2.7 What this section achieves

We have introduced a finite combinatorial-algebraic object — a presheaf of finite-dimensional vector spaces on a finite poset with a designated subpresheaf of "observables" — together with three operations (isomorphism, disjoint sum, disclosure) and a notion of cover. Nothing about this object refers to any specific implementation, programming language, protocol, or system. The object is general enough to carry instances ranging from finite-dimensional simplicial sheaves with observable structure to operational agentic compositions over MCP tool surfaces; it is specific enough that all axioms in §4 have content. The next section equips it with the cochain complex on which the obstruction invariant lives.

§3. The Witness Complex and the Witness Rank

This section defines the seam cochain complex of $\mathcal{F}/\mathcal{O}$, the obstruction module, and the witness rank. The construction is *relative seam cohomology*: the cohomology of latent semantic degrees that survive local observability and live specifically at composition seams. The pedigree is cellular-sheaf cohomology (Curry [14], Hatcher [15]) specialized to a finite presheaf on a finite poset, with the spectral framework on the resulting Laplacian operator developed in Hansen and Ghrist [13]. The construction is included in full because subsequent sections refer to specific cochain-level objects, and because the relationship between the abstract seam complex and the computationally equivalent *witness Gram* presentation (§3.4) is central to the existing program.

3.1 The seam complex

Fix a semantic interface complex $G = (P, \mathcal{F}, \rho, \mathcal{O})$. Let $\mathcal{L} := \mathcal{F}/\mathcal{O}$ be the latent presheaf with induced restrictions $\bar{\rho}_{p,q} : \mathcal{L}(q) \to \mathcal{L}(p)$ for $p \leq q$.

The obstruction invariant is not ordinary cohomology of the full context poset. It is the cohomology of a *relative seam complex* — the cochain complex that records how latent convention degrees at composition interfaces fail to glue under restriction.

Definition 3.1 (Seam partition). The context poset P decomposes into two classes by *structural role* in the observable-interface incidence graph:

- $\mathsf{Loc}(G)$ — the *locally resolved* contexts: leaf (minimal) elements of P, representing individual components whose observable interface is self-contained.
- $\mathsf{Seam}(G)$ — the *seam contexts*: non-leaf elements of P, representing composition interfaces where multiple local branches meet.

The partition is determined by the incidence structure of P, not by whether $\mathcal{L}$ vanishes. Under DFD+CHP (the operational sufficient conditions of §3.5), the class-homogeneous partition condition assigns convention dimensions to *interfaces*, so leaf contexts are necessarily fully observable: $\mathcal{L}(u) = 0$ for $u \in \mathsf{Loc}(G)$. More generally, a leaf context may carry latent data ($\mathcal{L}(u) \neq 0$); when it does, the latent data represents tool-internal hidden state that may propagate into seam contexts via the restriction structure (see Remark 3.2.1 below).

Definition 3.2 (Seam cochain complex). The *seam cochain complex* of G is the cochain complex

$$C^0_{\mathrm{seam}}(G) \xrightarrow{\delta^0} C^1_{\mathrm{seam}}(G) \xrightarrow{\delta^1} C^2_{\mathrm{seam}}(G) \xrightarrow{\delta^2} \cdots$$

where:

- $C^0_{\mathrm{seam}}(G) := \bigoplus_{u \in \mathsf{Loc}(G)} \mathcal{L}(u)$ — the latent data at leaf contexts. This is zero when every leaf is fully observable ($\mathcal{L}(u) = 0$ for all $u \in \mathsf{Loc}$); in the DFD+CHP regime it is always zero (see Definition 3.1).

- $C^1_{\mathrm{seam}}(G) := \bigoplus_{s \in \mathsf{Seam}(G)} \mathcal{L}(s)$ — the latent data at seam contexts.

- For $n \geq 2$, $C^n_{\mathrm{seam}}(G)$ is the degree-n cellular cochain group of the observable-interface cell complex $\Gamma(G)$ with cellular sheaf $\mathcal{L}$. In $\Gamma(G)$, the 0-cells are the leaf contexts $\mathsf{Loc}(G)$, the 1-cells are the seam contexts $\mathsf{Seam}(G)$ (each incident to the leaves below it), and higher cells are determined by the incidence structure of P: a k-cell exists for each $(k+1)$-element chain in P whose non-minimal elements are all seam contexts. In the exact regime, $\Gamma(G)$ is 1-dimensional (a graph), so $C^n_{\text{seam}} = 0$ for $n \geq 2$.

- The coboundary $\delta^0 : C^0_{\text{seam}} \to C^1_{\text{seam}}$ records how latent data at leaf contexts propagates into seam contexts via the convention-field incidence structure. For each covering relation $u \lessdot s$ ($u \in \mathsf{Loc}$, $s \in \mathsf{Seam}$) sharing a convention field d, the latent component of d at u embeds into $\mathcal{L}(s)$ along the corresponding dimension; under DFD+CHP, this is the transpose of the signed-incidence row (Paper IV-D, Komkov 2026). In the general case:

$$(\delta^0 c)(s) = \sum_{u \lessdot s} \iota_{u,s}(c(u)),$$

where $\iota_{u,s} : \mathcal{L}(u) \to \mathcal{L}(s)$ is the convention-field inclusion dual to the restriction $\bar{\rho}_{u,s}$.

In the *DFD+CHP* operational regime, leaf contexts are fully observable ($C^0_{\text{seam}} = 0$), so $\delta^0 = 0$ and the seam complex reduces to $0 \to C^1_{\text{seam}} \to C^2_{\text{seam}} \to \cdots$. In the *exact regime* (Definition 3.5), the complex truncates to two terms ($C^k = 0$ for $k \geq 2$), giving the two-term complex $C^0_{\text{seam}} \xrightarrow{\delta^0} C^1_{\text{seam}}$.

The pair $(C^\bullet_{\text{seam}}(G), \delta^\bullet)$ is the *seam cochain complex of* G.

Remark 3.2.1 (When $C^0_{\text{seam}} \neq 0$). Beyond DFD+CHP, a leaf context u may carry latent data — a hidden convention (e.g., an implicit timezone offset or a default currency) internal to the tool. When such latent data at u propagates into a seam context $s > u$ via $\iota_{u,s}$, the image $\iota_{u,s}(\mathcal{L}(u))$ lies in $\operatorname{im}\delta^0$: it is a *coboundary* in the seam complex. A seam dimension that is entirely explained by propagation from leaf latent data represents a convention mismatch that is *inherited* from a tool's internal state, not generated at the interface. Such a dimension does not contribute to H^1_{seam}, and disclosing it at the seam does not reduce the obstruction (axiom A4b; see Remark 4.2.1). This is the structural content of coboundary blindness: the invariant sees only *interface-generated* obstruction, not inherited tool state.

Definition 3.3 (Seam cohomology). The cohomology of the seam complex in degree n is

$$H^n_{\text{seam}}(G) := \ker \delta^n / \operatorname{im} \delta^{n-1}.$$

The first cohomology $H^1_{\text{seam}}(G)$ is the *obstruction module* of G.

Remark (Relation to ambient cohomology). The seam complex is the cellular sheaf cochain complex of $\Gamma(G)$ with sheaf $\mathcal{L}$ — not the ordered cochain complex of the full poset P with coefficients in $\mathcal{L}$. The full ordered-chain complex of P is a standard ambient construction (Curry [14]); the seam complex is the quotient that separates interface obstruction from tool-internal state, following the cellular-sheaf framework of Hansen and Ghrist [13]. When $C^0_{\text{seam}} = 0$ (the DFD+CHP case), the two formulations yield the same H^1; when $C^0_{\text{seam}} \neq 0$, the seam complex correctly quotients out propagated latent data, giving $H^1_{\text{seam}} = C^1 / \operatorname{im} \delta^0$ rather than C^1 itself.

3.2 The witness rank

Definition 3.4 (Witness rank). The *witness rank* of G is

$$r(G) := \dim_{\mathbb{Q}} H^1_{\text{seam}}(G).$$

Notational convention. From this point forward we write $H^1(G)$ for $H^1_{\text{seam}}(G)$ when unambiguous; the seam complex is the only cohomology theory applied to semantic interface complexes in this paper.

This is the central invariant. It is a non-negative integer; it vanishes when every seam dimension agrees on overlaps; it grows by one for each independent seam degree along which the latent data fails to glue. In the operational literature (Komkov 2026, Papers I–III), the witness rank is called the *coherence fee* $\varphi(G)$; the terms are interchangeable. We use "witness rank" in the doctrinal context to emphasize its characterization-theoretic status.

Lemma 3.5. $r(G \sqcup H) = r(G) + r(H)$.

Proof. The seam complex of a disjoint sum is the direct sum of the seam complexes; cohomology commutes with finite direct sums; dim is additive on direct sums of vector spaces. □

(This is the additivity property that enters axiom A3.)

3.3 Admissible disclosures act on the seam complex

An admissible disclosure of a one-dimensional latent line $\bar{W} \subseteq \mathcal{L}(p)$ changes the observable sub-presheaf and hence the latent presheaf: the dimension of $\mathcal{L}(p)$ drops by one while $\mathcal{L}(q)$ is unchanged for $q \neq p$ (by the admissibility condition of §2.3 III). This induces a chain map of seam complexes.

Lemma 3.6 (Disclosure chain map). Let $\sigma = (p, W)$ be an admissible disclosure on G, and let $G' = G \setminus\!\setminus W$. There is a chain map $\sigma_* : C^\bullet_{\mathrm{seam}}(G') \to C^\bullet_{\mathrm{seam}}(G)$ exhibiting the seam complex of G' as a subcomplex, with quotient concentrated in degrees where $\bar{W}$ contributes.

Proof sketch. Admissible disclosure gives $\mathcal{L}'(p) = \mathcal{L}(p)/\bar{W}$ and $\mathcal{L}'(q) = \mathcal{L}(q)$ for $q \neq p$. Choose any $\mathbb{Q}$-linear complement of $\mathcal{L}'(p)$ in $\mathcal{L}(p)$ — the choice is non-canonical, but any choice realizes $\mathcal{L}(p) \cong \bar{W} \oplus \mathcal{L}'(p)$ — and this splitting induces the inclusion $C^\bullet_{\mathrm{seam}}(G') \hookrightarrow C^\bullet_{\mathrm{seam}}(G)$. The quotient complex is the subcomplex generated by cochains valued in $\bar{W}$ at the seam context p. The cohomology computations of subsequent lemmas depend only on the isomorphism class of the splitting, not on the specific complement. □

The class $[\sigma] \in H^1_{\mathrm{seam}}(G)$ defined by an admissible disclosure σ is the cohomology class of the cocycle representing $\bar{W}$ in the seam complex. Two admissible disclosures σ_1 and σ_2 are *cochain-independent* (§2.3, restated cohomologically) precisely when $[\sigma_1]$ and $[\sigma_2]$ are linearly independent in $H^1_{\mathrm{seam}}(G)$.

Lemma 3.7 (Independent disclosure decreases r by one). If $\sigma = (p, W)$ is an admissible disclosure on G with $[\sigma] \neq 0$ in $H^1_{\mathrm{seam}}(G)$, then

$$r(G \setminus\!\setminus W) = r(G) - 1.$$

Proof. The long exact sequence of the pair $(C^\bullet_{\mathrm{seam}}(G), C^\bullet_{\mathrm{seam}}(G'))$ gives

$$\cdots \to H^0(Q^\bullet) \to H^1_{\mathrm{seam}}(G') \to H^1_{\mathrm{seam}}(G) \to H^1(Q^\bullet) \to \cdots$$

where $Q^\bullet$ is the quotient complex. This quotient is concentrated in low degree (by Lemma 3.6) and its first cohomology is at most one-dimensional, generated by $[\sigma]$. When $[\sigma] \neq 0$, the connecting map kills exactly that class, dropping the rank of H^1_{seam} by one. □

This is the cochain-level statement of axiom A4a.

Remark 3.7.1 (Matroid structure of disclosure independence). The set of admissible disclosures on G, ordered by cochain-independence, satisfies the standard matroid axioms (hereditary, exchange, augmentation; Whitney [6], Tutte [7], Oxley [8]), with rank function equal to r.

This *disclosure matroid* coincides with the column matroid of the observable-column corank (Paper II, Backbone Theorem, Komkov 2026): the ground set is the set of admissible one-generator seam disclosures modulo coboundary equivalence; the rank function is $\dim H^1_{\mathrm{seam}}$; the independent sets are cochain-independent disclosure families. The doctrine paper takes the numerical-invariant route for clarity; the matroid route is equivalent and connects directly to the structural identity $\varphi(G) = \mathrm{corank}_M(\mathcal{O})$ of Paper II.

3.4 The witness Gram (computational equivalent)

The cohomological definition of $r(G) = \dim H^1_{\mathrm{seam}}(G)$ is the primary characterization. The *witness Gram* is a computationally equivalent matrix presentation that computes the same quantity via Schur complement / Kron reduction — it is the form used in the operational implementation.

The computational route begins not with the seam complex directly, but with the *full convention coboundary* $D_{\mathcal{F}}$ — the matrix of the coboundary operator on the full presheaf $\mathcal{F}$ restricted to covering relations of P. This is a $|E| \times |V|$ rational matrix, where $|E|$ counts (edge, dimension) pairs and $|V|$ counts (context, dimension) pairs. The observable subpresheaf $\mathcal{O}$ determines a subset of columns.

Definition 3.8 (Witness Gram). Let $D_{\mathcal{F}}$ be the full convention coboundary matrix, and write $K_{\mathcal{F}} := D_{\mathcal{F}}^{\top} D_{\mathcal{F}}$, partitioned into observable and latent blocks:

$$K_{\mathcal{F}} = \begin{pmatrix} K_{\mathcal{O}\mathcal{O}} & K_{\mathcal{O}\mathcal{L}} \\ K_{\mathcal{L}\mathcal{O}} & K_{\mathcal{L}\mathcal{L}} \end{pmatrix}.$$

The *Kron-reduced witness Gram* is the Schur complement over observable directions:

$$K^{\mathrm{red}}(G) := K_{\mathcal{L}\mathcal{L}} - K_{\mathcal{L}\mathcal{O}} K^{+}_{\mathcal{O}\mathcal{O}} K_{\mathcal{O}\mathcal{L}},$$

where $(\cdot)^{+}$ denotes the Moore–Penrose pseudoinverse.

Lemma 3.9 (Rank identity). $\mathrm{rank}\, K^{\mathrm{red}}(G) = r(G)$ in the exact regime.

Proof sketch. The Kron reduction eliminates observable directions from the full Gram, leaving a quadratic form on residual latent seam directions. Its rank equals $\dim H^1_{\mathrm{seam}}(G)$ in the regime where the seam complex is concentrated in degrees 0 and 1 — that is, the exact regime. The full proof is Theorem 1 of Komkov (2026), "Witness Geometry Beyond Fee." The connection between cellular-sheaf cochain complexes and graph Laplacian-style operators is developed in Hansen and Ghrist [13]. □

The rank identity does *not* assert that $\mathrm{rank}(D_{\mathcal{F}}^{\top} D_{\mathcal{F}}) = r(G)$; the unreduced Gram has rank equal to $\mathrm{rank}(D_{\mathcal{F}})$, which counts the *observable coboundary rank* plus the obstruction. The witness rank lives in the *residual after Kron reduction.* The two-step procedure — form the full Gram, then eliminate observable contributions by Schur complement — is what the operational implementation computes: $r(G) = \mathrm{rank}(D_{\mathcal{F}}) - \mathrm{rank}(D_{\mathcal{O}})$, equivalently $\mathrm{rank}\, K^{\mathrm{red}}(G)$.

Under CHP (the class-homogeneous partition condition), each seam context s receives latent contributions from exactly two locally resolved contexts (the two tools sharing a convention dimension), yielding the signed-incidence row structure — at most one $+1$ and one -1 per row of $D_{\mathcal{F}}$ — proved in Paper IV-D (Komkov 2026, "Signed-Incidence Structure"). This structural property implies total unimodularity and field-independence of the rank.

(This Lemma supplies the empirical computation in the existing program: $r(G)$ is computable in polynomial time as a matrix rank over the rationals, with no exotic structure required.)

3.5 The exact regime, made precise

Definition 2.3 left the exact regime informally specified. We now give the precise statement.

Definition 3.5 (Exact regime). A semantic interface complex G is in the *exact regime* $\mathsf{SemComp}_{\text{ex}}$ if its seam cochain complex $(C^{\bullet}_{\text{seam}}(G), \delta^{\bullet})$ satisfies $C^n_{\text{seam}}(G) = 0$ for every $n \geq 2$ — equivalently, if the seam complex truncates to two terms.

The seam complex is always well-defined; the exact regime is the condition under which it truncates. Under this condition $\delta^1 : C^1_{\text{seam}} \to C^2_{\text{seam}} = 0$ is the zero map, so $\ker \delta^1 = C^1_{\text{seam}}$, and the two-term rank formula becomes a *theorem*:

$$r(G) \;=\; \dim H^1_{\text{seam}}(G) \;=\; \dim C^1_{\text{seam}}(G) - \operatorname{rank} \delta^0_{\text{seam}}.$$

By Lemma 3.9, this equals $\operatorname{rank} K^{\text{red}}(G)$, the Kron-reduced Gram rank.

Remark (Operational rank formula). The formula $r(G) = \operatorname{rank}(D_{\mathcal{F}}) - \operatorname{rank}(D_{\mathcal{O}})$ used in the implementation literature (Komkov 2026, Paper I) is the equivalent computation obtained by lifting the seam complex to the full convention coboundary and subtracting the observable sub-coboundary contribution. In the exact regime, this coincides with $\dim C^1_{\text{seam}} - \operatorname{rank} \delta^0_{\text{seam}}$; the two are algebraically identical.

Operational sufficient conditions. The exact regime is characterized by abstract properties of the seam complex; for compositions arising in practice, it is implied by structural conditions on the data. Two such conditions, used in the existing implementation literature, are:

(i) *Disjoint field decomposition (DFD).* The latent presheaf $\mathcal{L}$ decomposes as a direct sum $\mathcal{L} = \bigoplus_d \mathcal{L}_d$ over independent field-types, each acting on disjoint contexts.

(ii) *Class-homogeneous partition (CHP).* Within each field-type, the convention dimension at every context is one-dimensional.

Under (i)+(ii), every higher-degree seam cochain vanishes and $G \in \mathsf{SemComp}_{\text{ex}}$. This is Theorem 2 of Komkov (2026), "Hierarchical Decomposition of Coherence Fee," and supplies the bridge between the abstract definition above and the operational regime named in the implementation literature. Under DFD+CHP, seam contexts are exactly the interface overlaps of the schema incidence graph; locally resolved contexts are tools; and all tool (leaf) contexts are fully observable ($C^0_{\text{seam}} = 0$), since CHP assigns each convention dimension to an inter-tool interface, not to tool-internal state. The seam complex reduces to graph cohomology of the incidence graph with coefficients in the latent presheaf. (This fully-observable-tools property is what the proof of Theorem 5.1 uses: minimal $\mathcal{L} \neq 0$ contexts are necessarily seam contexts when $C^0 = 0$.)

Examples and non-examples. To prevent the exact regime from feeling like "the regime in which the theorem works," we record several structural cases.

Case	In exact regime?	Reason
Acyclic observable composition ($\mathcal{F} = \mathcal{O}$)	Yes	$H^1_{\text{seam}} = 0$, seam complex trivial

Case	In exact regime?	Reason
Elementary hidden cycle C_e (§3.6)	Yes	Latent presheaf concentrated at one seam context; $C^1_{\mathrm{seam}} = \mathbb{Q}$, higher degrees vanish
Disjoint sum of finitely many elementary cycles	Yes	Seam complex is direct sum; closure under disjoint sum (§2.6)
DFD + CHP composition with bounded depth	Yes	Theorem 2 of Komkov (2026); the operational sufficient condition
Composition with *coloop* — a latent dimension simultaneously a coboundary and a non-trivial cocycle in different bases	No	Seam complex fails to split rationally; surrogate-regime correction term required
Composition with non-trivial higher-degree seam cohomology ($H^2_{\mathrm{seam}} \neq 0$)	No	Truncation condition $C^n_{\mathrm{seam}} = 0$ for $n \geq 2$ fails
Class-heterogeneous partition (CHP violation) — a single field-type whose convention dimension varies across contexts	No	Higher extension data in the seam complex; rank-counting gives a lower bound, not equality

The exact regime is structurally characterized: it is the class of complexes whose seam complex truncates to two terms. Outside it, the surrogate-regime correction term of Conjecture 9.1 enters.

3.6 Worked example: an elementary cycle

The smallest non-trivial example. Let $P = \{p_1, p_2, p_3, \hat{p}\}$ where $p_i < \hat{p}$ for each i and the p_i are pairwise incomparable. Let $\mathcal{F}(p_i) = \mathbb{Q}\langle e_i \rangle$ for $i = 1, 2, 3$ (each one-dimensional), and $\mathcal{F}(\hat{p}) = \mathbb{Q}^3 = \mathbb{Q}\langle e_1, e_2, e_3 \rangle$ (the interface context). Restriction maps from $\hat{p}$ project onto the corresponding axis. Let $\mathcal{O}(p_i) = \mathcal{F}(p_i)$ for each i (each tool observes its own convention), but $\mathcal{O}(\hat{p}) = \mathbb{Q}\langle e_1 - e_2, e_2 - e_3 \rangle$ — only pairwise differences are observable at the interface.

The latent presheaf $\mathcal{L}$ is zero on each p_i (tools are fully observable) and one-dimensional on $\hat{p}$, generated by the "absolute mean" $e_1 + e_2 + e_3$ modulo observable differences.

The seam partition: $\mathsf{Loc}(G) = \{p_1, p_2, p_3\}$ (leaves); $\mathsf{Seam}(G) = \{\hat{p}\}$ (interface). Since each leaf is fully observable ($\mathcal{L}(p_i) = 0$), this is the *fully locally resolved* case: $C^0_{\mathrm{seam}} = 0$. The seam complex:

$$C^0_{\mathrm{seam}} = 0, \qquad C^1_{\mathrm{seam}} = \mathcal{L}(\hat{p}) = \mathbb{Q}.$$

With $\delta^0 = 0$ (no leaf latent data to propagate), the seam complex is concentrated in degrees 0 and 1, so G is in the exact regime, and:

$$r(G) = \dim C^1_{\mathrm{seam}} - \operatorname{rank} \delta^0_{\mathrm{seam}} = 1 - 0 = 1.$$

Disclosing $W = \mathbb{Q}\langle e_1 + e_2 + e_3 \rangle \subseteq \mathcal{F}(\hat{p})$ — declaring the absolute mean to be observable — is admissible: for each $p_i < \hat{p}$, $\rho_{p_i,\hat{p}}(e_1 + e_2 + e_3) = e_i \in \mathcal{O}(p_i) = \mathcal{F}(p_i)$. After disclosure, $\mathcal{L}(\hat{p}) = 0$, so $r(G_{\mathrm{disclosed}}) = 0$.

This is the *elementary hidden semantic cycle*: three local contexts whose pairwise differences are observable but whose absolute alignment is not. It will play the role of generator in axiom A5.

3.7 A two-field example

The elementary cycle has $r = 1$. To exhibit higher rank concretely — and to verify the additivity and repair-duality claims explicitly — consider its two-field generalization.

Take the same context poset $P = \{p_1, p_2, p_3, \hat{p}\}$ but enlarge the convention spaces. At each leaf p_i, $\mathcal{F}(p_i) = \mathbb{Q}\langle d_i, a_i \rangle$ — two fields, “date” and “amount.” At the interface, $\mathcal{F}(\hat{p}) = \mathbb{Q}^6$ spanned by all six. Let $\mathcal{O}(\hat{p})$ contain only the four pairwise differences within each field: $d_1 - d_2$, $d_2 - d_3$, $a_1 - a_2$, $a_2 - a_3$. The two absolute means $d_1 + d_2 + d_3$ and $a_1 + a_2 + a_3$ remain latent.

The seam partition: $\mathsf{Loc} = \{p_1, p_2, p_3\}$ (leaves, fully observable); $\mathsf{Seam} = \{\hat{p}\}$ (interface, two-dimensional latent). Again fully locally resolved ($C^0_{\mathrm{seam}} = 0$). The seam complex:

$$C^0_{\mathrm{seam}} = 0, \qquad C^1_{\mathrm{seam}} = \mathcal{L}(\hat{p}) = \mathbb{Q}^2.$$

With $\delta^0 = 0$, we have $H^1_{\mathrm{seam}}(G) = \mathbb{Q}^2$ and $r(G) = 2$.

Disclosing $W_1 = \mathbb{Q}\langle d_1 + d_2 + d_3 \rangle$ promotes the date mean to observable. Both disclosures are admissible: $\rho_{p_i,\hat{p}}(d_1 + d_2 + d_3) = d_i \in \mathcal{O}(p_i)$. The cohomology class $[\sigma_1]$ is non-zero in $H^1_{\mathrm{seam}}(G)$; by Lemma 3.7, $r(G \setminus\!\!\setminus W_1) = 1$. Disclosing $W_2 = \mathbb{Q}\langle a_1 + a_2 + a_3 \rangle$ next reduces r to zero. The disclosure normal form (Theorem 7.1) is the receipt with disclosure set $\{W_1, W_2\}$, of cardinality $2 = r(G)$, matching repair duality (Theorem 6.2).

In the Gram presentation: the full convention coboundary $D_{\mathcal{F}}$ is a 6×6 matrix (three tools, two fields each, three interface overlaps per field). After Kron reduction eliminating the four observable directions (the pairwise-difference components), the residual 2×2 Kron-reduced Gram K^{red} has rank $2 = r(G)$. Operationally, this models a composition where each tool reports a local date and amount, and the interface requires both the date-frame and amount-frame to be globally aligned. Two independent disclosures suffice; no fewer.

3.8 What this section achieves

We have a finite, computable invariant $r : \mathrm{Ob}(\mathsf{SemComp}) \to \mathbb{Z}_{\geq 0}$, defined as the dimension of the first seam cohomology $H^1_{\mathrm{seam}}(G)$, computable as the rank of the Kron-reduced witness Gram $K^{\mathrm{red}}(G)$. The invariant is additive over disjoint sums (Lemma 3.5) and decreases by one under cochain-independent admissible disclosure (Lemma 3.7). The exact regime $\mathsf{SemComp}_{\mathrm{ex}}$ — the class of complexes whose seam complex truncates to two terms (Definition 3.5) — has both an abstract characterization and operational sufficient conditions (DFD + CHP). The next section asks: is r the *unique* invariant with these clean properties? §4 states the axioms; §5 proves it is.

---

§4. The Six Axioms

We now state the axioms that characterize the obstruction invariant. Each axiom is a property that any reasonable measure of compositional semantic obstruction should satisfy: invariance under renaming, vanishing on trivial objects, additivity over independent components, exactness under disclosure, normalization on the elementary obstruction generator, and locality under cover decomposition. The cumulative force of the six is what determines the invariant uniquely; the cumulative naturalness of the six is what makes the resulting characterization a *canonical* description rather than a parochial one.

Throughout this section, $I : \mathrm{Ob}(\mathsf{SemComp}_{\mathrm{ex}}) \to \mathbb{Z}_{\geq 0}$ is a function on the exact regime, constant on isomorphism classes (a *numerical invariant*).

4.1 The axioms

(A1) Isomorphism invariance. For all $G, G' \in \mathsf{SemComp}_{\mathrm{ex}}$:

$$G \cong G' \implies I(G) = I(G').$$

(A2) Obstruction triviality. If $H^1_{\mathrm{seam}}(G) = 0$ — equivalently, every seam latent dimension is in the image of δ^0 (is a coboundary), equivalently $r(G) = 0$ — then

$$I(G) = 0.$$

(The special case $\mathcal{L} = 0$ implies $H^1_{\mathrm{seam}} = 0$ trivially; A2 says the converse also forces $I = 0$.)

(A3) Additivity. For all $G, H \in \mathsf{SemComp}_{\mathrm{ex}}$:

$$I(G \sqcup H) = I(G) + I(H).$$

(A4a) Disclosure sensitivity. Let $\sigma = (p, W)$ be an admissible disclosure on G whose induced cohomology class $[\sigma] \in H^1(G)$ is non-zero. Then

$$I(G \setminus\!\!\setminus W) = I(G) - 1.$$

(A4b) Coboundary blindness. Let $\sigma = (p, W)$ be an admissible disclosure whose induced cohomology class is zero — i.e., $[\sigma] = 0$, equivalently, $\bar{W}$ corresponds to a coboundary class in $\mathrm{im}\, \delta^0_{\mathrm{seam}}$. Then

$$I(G \setminus\!\!\setminus W) = I(G).$$

Together A4a and A4b say: independent admissible disclosures reduce the invariant by one; redundant (coboundary) disclosures leave it unchanged. We refer to the pair as *disclosure exactness* when convenient. Axiom A4 has two clauses (sensitivity and blindness) that together constitute disclosure exactness.

(A5) Cycle normalization. Let $C_e \in \mathsf{SemComp}_{\mathrm{ex}}$ be the elementary hidden semantic cycle (the worked example of §3.6). Then

$$I(C_e) = 1.$$

(A6) Excision under compatible cover. Let (U, V) be a compatible cover of G (Definition 2.2), with $G \in \mathsf{SemComp}_{\mathrm{ex}}$. Then

$$I(G) = I(U) + I(V) - I(U \cap V).$$

Equivalently, I is the unique (up to scaling) Mayer–Vietoris-additive function on $\mathsf{SemComp}_{\mathrm{ex}}$ that vanishes on the empty complex and respects compatible covers.

Remark. For complexes outside the exact regime, the Mayer–Vietoris sequence at H^1 may not be exact, and the equality above must be modified by a non-negative correction term recording classes lost in the gluing. The surrogate-regime version of A6 is named in §9 and pursued elsewhere.

4.2 Motivation, axiom by axiom

(A1) — Why isomorphism invariance. Renaming a context, relabeling a basis, or applying any structure-preserving bijection should not change a measure of obstruction. If I depended on the labels rather than on the structure, it would not be an invariant in the mathematical sense; it would be a metric of the presentation. A1 is the minimum content of "invariant."

(A2) — Why obstruction triviality. When $H^1_{\text{seam}} = 0$, every latent seam dimension is explained by propagation from leaf context data (lies in $\operatorname{im} \delta^0$) — no interface-generated obstruction exists. The composition may still have latent dimensions (tool-internal hidden state), but these do not produce interface failures. In the fully locally resolved case ($\mathcal{L} = 0$), this is even simpler: there is nothing latent at all. In either case, the obstruction must vanish.

(A3) — Why additivity. Two compositions sharing no contexts cannot interact; their obstructions cannot interfere. The obstruction of the disjoint sum is the simple sum. This is the same axiom that appears in homology (additivity), in entropy (independence of sources), and in valuations on disjoint unions of measurable sets. A non-additive invariant would entangle non-interacting components.

(A4a) — Why disclosure sensitivity. When a single previously-latent dimension corresponding to a non-trivial cohomology class is promoted to observable, the seam information at that dimension becomes locally checkable. The class is now observable and contributes nothing further to the obstruction — the rank drops by one. The "-1" carries both the cancellation rule (one disclosure removes one class) and the scale (the unit is "one elementary obstruction").

(A4b) — Why coboundary blindness. A disclosure whose cochain element is already in $\operatorname{im} \delta^0$ is *redundant*: it contributes no information beyond what the observable subpresheaf already encodes. Such a disclosure should not change the invariant. Without A4b, the theory could not distinguish redundant disclosures from informative ones, and minimum-repair cardinality (§6) would be ill-defined.

Remark 4.2.1 (A4b is non-vacuous). In the fully locally resolved case (DFD+CHP, where $C^0_{\text{seam}} = 0$), every non-trivial seam disclosure has $[\sigma] \neq 0$, so A4b is vacuously satisfied. The axiom becomes non-vacuous in the natural generalization where leaf contexts carry latent data ($C^0_{\text{seam}} \neq 0$). Concrete example: $P = \{u_1, u_2, s\}$ with $u_1 < s$, $u_2 < s$; $\mathcal{L}(u_1) = \mathbb{Q}$ (one hidden timezone convention at tool 1), $\mathcal{L}(u_2) = 0$, $\mathcal{L}(s) = \mathbb{Q}^2$ (two latent dims at the interface — one propagated from u_1's timezone, one a genuinely new currency convention). The coboundary $\delta^0 : \mathbb{Q} \to \mathbb{Q}^2$ maps the single leaf latent dim into the propagated component: $\delta^0(1) = (1, 0)$. Then $H^1 = \mathbb{Q}^2/\mathbb{Q}\langle(1, 0)\rangle \cong \mathbb{Q}$, giving $r = 1$. A seam disclosure along the propagated direction $(1, 0)$ has class $[(1, 0)] = 0 \in H^1$ — it is a coboundary, and A4b mandates that disclosing it leaves I unchanged. A seam disclosure along the currency direction $(0, 1)$ has $[(0, 1)] \neq 0$, and A4a applies. The structural content: disclosing a seam dimension that merely re-exposes tool-internal state is operationally inert; only interface-generated obstruction matters.

Non-circularity note (key). The conditions $[\sigma] \neq 0$ (in A4a) and $[\sigma] = 0$ (in A4b) are checked *entirely in the seam cochain complex of G* — they do not refer to I. The disclosure determines a one-generator seam 1-cochain in $C^1_{\text{seam}}(G)$, whose cohomology class is evaluated modulo $\operatorname{im} \delta^0_{\text{seam}}$. The axiom imports the *cohomological structure* of the seam complex (which is part of the data of G), not the *numerical invariant* we are characterizing. The honest reading: A4a/A4b say the invariant respects the cochain structure of disclosure-independence — a property of the seam complex, not a self-reference.

This is the locus of non-tautology and is discussed adversarially in notes/anti-tautology.md.

(A5) — Why cycle normalization. The elementary hidden semantic cycle (§3.6) is the smallest non-trivial obstruction: three local contexts agreeing pairwise but failing globally. Since $r(C_e) = 1 > 0$ (the obstruction is non-zero by construction); by symmetry of the cycle and A1 it is invariant under permutation of the three legs; the natural normalization is $I(C_e) = 1$. Any other choice of normalization gives an invariant proportional to I rather than equal — so A5 fixes the scale.

This axiom is the one that turns I from a function determined "up to multiplicative constant" into a function determined exactly. It plays the role that "$I(\text{point}) = 1$" or "$\dim I(\mathbb{Z}) = 1$" plays in homology, or that "$\log_2 2 = 1$" plays in Shannon entropy.

(A6) — Why excision. Compositions can be decomposed into overlapping pieces: a cover. Obstruction information should be local in the sense that the obstruction of G is determined by the obstructions of U, V, and their intersection $U \cap V$. The standard topological account of this is the Mayer–Vietoris sequence; in the exact regime, the sequence is exact and the axiom states that obstruction satisfies plain inclusion-exclusion. (Outside the exact regime, an additional non-negative correction term records cohomology classes that gluing can lose or create; that surrogate-regime version of A6 is named in §9.)

This is the axiom that makes the invariant *local*: $I(G)$ is determined by I on small pieces. Without it, the invariant could not be computed from local data, and the entire program of "compositional obstruction" would lose meaning.

Non-circularity note. Compatible covers (Definition 2.2) are properties of the cochain complex of G and its restrictions to U, V, $U \cap V$ — definable without any reference to I. The axiom is a relation among four values of I $(I(G), I(U), I(V), I(U \cap V))$, not a self-referential condition.

4.3 Independence

The seven axioms (A1, A2, A3, A4a, A4b, A5, A6) exhibit a non-uniform redundancy structure. Some are independent of the others; some are implied within the exact regime, but become non-redundant outside.

Independence of A4b within the exact regime. A4b is independent of A1, A2, A3, A4a, A5, A6 — exhibited by the *latent dimension invariant* $I_L(G) := \dim_{\mathbb{Q}} \mathcal{L}_{\text{tot}}(G)$, which satisfies all six other axioms (iso-invariance, vanishing on observable-only complexes, additivity under disjoint sum, sensitivity to disclosures, normalization on the elementary cycle, and inclusion-exclusion under cover) but fails A4b: I_L drops by exactly one under *every* disclosure, including cochain-trivial ones. Without A4b, redundant disclosures would inflate the rank-counting and the invariant would not be the cohomological one. A4b is the axiom that distinguishes information-bearing disclosures from redundant ones.

Independence of A4a. A4a is independent of A1, A2, A3, A4b, A5, A6 — the constant invariant $I = 0$ satisfies all of those (vacuously, in some cases) but fails A4a (which requires non-zero drop on non-trivial disclosure if any exists).

A1 is independent in the trivial sense that no permutation-non-invariant function can satisfy the iso quotient.

Internal redundancy in the exact regime. Proposition 5.4 shows that A5 and A6 are implied by A1+A2+A3+A4a+A4b within $\mathsf{SemComp}_{\mathrm{ex}}$. Their inclusion in the doctrinal axiom system is justified outside the exact regime, where they become non-redundant constraints (§9 and the companion note `axioms-discipline.md`).

Independence of A2, A3, and A5 from each other is more subtle. The companion note records the partial verifications and identifies the open cases. A fully formal minimality proof in the Eilenberg–Steenrod style is left to future work.

4.4 What the axioms exclude

It is worth being explicit about what the axioms forbid:

- **They forbid scale-dependence**: by A1+A5, I is the rank, not a quadratic form.
- **They forbid cross-component coupling**: by A3, I is sum-additive across disjoint components.
- **They forbid hidden penalties**: by A2, I vanishes on classically composable systems.
- **They forbid disclosure overcounting**: by A4a, one independent disclosure removes exactly one obstruction class — not two, not zero, not "depends on context." By A4b, redundant disclosures cannot inflate the count.
- **They forbid pathological non-locality**: by A6, I on a whole is determined by inclusion-exclusion on a compatible cover. (Outside the exact regime, an explicit gluing correction enters; this is named in §9.)

What the axioms *do not* forbid: any specific computational realization of I on a particular complex. The axioms constrain the abstract function; many concrete cochain-level computations may compute the same function. The characterization theorem says that any abstract function satisfying A1–A6 in the exact regime is the witness rank; it does not say that any one method of computing the rank is canonical, only that the value is.

4.5 What this section achieves

We have stated the doctrinal axioms (A1, A2, A3, A4a, A4b, A5, A6) on numerical invariants of the exact regime, motivated each as a property any reasonable measure of compositional obstruction should satisfy, and verified that the critical non-circularity points (A4a, A4b, and A6) are satisfied: each axiom is a relation involving the cochain-level structure of the seam complex but not the invariant being characterized.

The next section proves that the witness rank r — already known to satisfy these axioms by the lemmas of §3 — is the unique numerical invariant satisfying A1, A2, A3, A4a, A4b in the exact regime, with A5 and A6 automatic.

---

§5. The Disclosure Characterization Theorem

5.1 Two layers of the result

The result of this section operates at two layers, which we name explicitly to keep the locus of originality clear.

Layer 1 — the structural primitive. §2 defined disclosure as a morphism on semantic interface complexes; §3 defined cochain-level independence of disclosures. *The non-trivial conceptual content of the framework is that seam-independent disclosure is the right primitive operation of compositional repair.* Once accepted as primitive, disclosure independence is purely structural: it is a property of the cochain complex of G, definable without reference to any numerical invariant.

Layer 2 — numerical uniqueness. Given the primitive, the question is which numerical invariants of $\mathsf{SemComp}_{\mathrm{ex}}$ respect it. The theorem of this section says: only one such invariant exists, the witness rank.

The theorem is honest about the locus of work: the originality is not in proving Layer 2 (the proof is a short induction). The originality is in identifying Layer 1 — that disclosure independence on the cochain complex is the structural primitive of the theory. Once that is granted, the rank is forced.

5.2 Statement

Theorem 5.1 (Disclosure Characterization of Witness Rank). Let $I : \mathrm{Ob}(\mathsf{SemComp}_{\mathrm{ex}}) \to \mathbb{Z}_{\geq 0}$ be a numerical invariant satisfying axioms A1, A2, A3, A4a, A4b of §4. Then for every $G \in \mathsf{SemComp}_{\mathrm{ex}}$,

$$I(G) = r(G) = \dim_{\mathbb{Q}} H^1(G).$$

In words: the witness rank is the unique non-negative integer-valued invariant on semantic interface complexes in the exact regime that is iso-invariant (A1), locally trivial on observable-only complexes (A2), additive over disjoint sums (A3), sensitive to non-trivial seam disclosures (A4a), and blind to cochain-trivial disclosures (A4b).

Axioms A5 and A6 are *consequences* of A1–A4b in the exact regime (Proposition 5.4); they are part of the doctrinal axiom system because they become non-redundant outside it.

5.3 Existence

The witness rank r satisfies the axioms. We verify each:

- **A1** (iso-invariance): $r(G)$ depends only on the isomorphism class of the cochain complex of G, which is iso-invariant.
- **A2** (obstruction triviality): if $H^1_{\mathrm{seam}}(G) = 0$ then $r(G) = 0$ by definition. (The special case $\mathcal{L} = 0$ forces all seam cochains to vanish.)
- **A3** (additivity): Lemma 3.5.
- **A4a** (disclosure sensitivity): if $[\sigma] \neq 0$, $r(G \setminus\!\setminus W) = r(G) - 1$ by Lemma 3.7.
- **A4b** (coboundary blindness): if $[\sigma] = 0$, the cocycle representing W is a coboundary, contributing nothing to H^1, so $r(G \setminus\!\setminus W) = r(G)$.
- **A5** (cycle normalization): the worked example of §3.6 computes $r(C_e) = 1$.

- **A6** (excision): in the exact regime with compatible cover, the Mayer–Vietoris sequence is exact at H^1, giving $\dim H^1(G) = \dim H^1(U) + \dim H^1(V) - \dim H^1(U \cap V)$.

So r is one invariant satisfying all six axioms.

5.4 Uniqueness — the main argument

We show that any I satisfying A1, A2, A3, A4a, A4b equals r. The strategy is induction on a well-founded measure of complexity, with reduction steps that decrease the measure while preserving the relation $I(G) - r(G)$.

Definition 5.2 (Complexity measure). For $G \in \mathsf{SemComp}_{\mathrm{ex}}$, define the *latent complexity*

$$\mu(G) := \dim_{\mathbb{Q}} \mathcal{L}_{\mathrm{tot}}(G), \quad \text{where} \quad \mathcal{L}_{\mathrm{tot}}(G) := \bigoplus_{p \in P} \mathcal{L}(p).$$

The measure μ takes values in $\mathbb{Z}_{\geq 0}$. We prove $I(G) = r(G)$ by induction on $\mu(G)$.

Base case: $\mu(G) = 0$. Then $\mathcal{L} = 0$, hence $H^1_{\mathrm{seam}} = 0$ and $r(G) = 0$. By A2,

$$I(G) = 0 = r(G).$$

Inductive step: $\mu(G) > 0$. Then $\mathcal{L} \neq 0$, so there exists a context p with $\mathcal{L}(p) \neq 0$. Two reduction tools, drawn from A3 and A4a/A4b, suffice:

(R1) Disjoint sum split. If G is disconnected — i.e., its cover poset P admits a non-trivial decomposition $P = P_1 \sqcup P_2$ closed under inheritance, with the convention data block-diagonal — then $G = G_1 \sqcup G_2$ with both non-empty. By A3 and Lemma 3.5,

$$I(G) = I(G_1) + I(G_2), \qquad r(G) = r(G_1) + r(G_2).$$

Both G_i have $\mu(G_i) < \mu(G)$. By induction $I(G_i) = r(G_i)$ for each, hence $I(G) = r(G)$.

(R2) Disclosure reduction. If G is connected, we must find an admissible disclosure. The following lemma guarantees one always exists:

Lemma 5.3 (Minimal-context admissibility). If p is minimal among contexts with $\mathcal{L}(p) \neq 0$, then every one-dimensional line $\bar{W} \subseteq \mathcal{L}(p)$ lifts to an admissible disclosure.

Proof. For every $a < p$, minimality gives $\mathcal{L}(a) = 0$, hence $\mathcal{O}(a) = \mathcal{F}(a)$. Any lift $W \subseteq \mathcal{F}(p)$ of $\bar{W}$ satisfies $\rho_{a,p}(W) \subseteq \mathcal{F}(a) = \mathcal{O}(a)$, so the downward-closure condition holds. □

Choose p minimal with $\mathcal{L}(p) \neq 0$ and any admissible disclosure $\sigma = (p, W)$ representing a line $\bar{W} \subseteq \mathcal{L}(p)$. Let $G' := G \mathbin{\backslash\backslash} W$. The cohomology class $[\sigma] \in H^1_{\mathrm{seam}}(G)$ is computable from the seam complex; either it is non-zero or it is a coboundary.

- If $[\sigma] \neq 0$ in $H^1_{\mathrm{seam}}(G)$: by A4a and Lemma 3.7, $I(G) - I(G') = 1 = r(G) - r(G')$.
- If $[\sigma] = 0$ in $H^1_{\mathrm{seam}}(G)$: by A4b and the cochain-level fact that coboundaries contribute nothing to seam cohomology, $I(G) = I(G')$ and $r(G) = r(G')$.

In either case, $\mu(G') = \mu(G) - 1$, since exactly one latent dimension is promoted to observable. By induction $I(G') = r(G')$, hence $I(G) = r(G)$.

Termination. The latent complexity μ is a non-negative integer; each reduction (R1 or R2) strictly decreases it; so the recursion terminates after finitely many steps, ending at the base case ($\mu = 0$) where A2 gives the conclusion.

This completes the proof. □

5.5 A5 and A6 are automatic in the exact regime

Notice that the proof of Theorem 5.1 used only A1, A2, A3, A4a, A4b. The base case invokes A2; the disjoint-sum reduction invokes A3; the disclosure reduction invokes A4a and A4b. Axioms A5 (cycle normalization) and A6 (excision under compatible cover) were not used in the argument.

Proposition 5.4 (A5 and A6 redundant in exact regime). *Any invariant* $I : \mathrm{Ob}(\mathsf{SemComp}_{\mathrm{ex}}) \to \mathbb{Z}_{\geq 0}$ *satisfying A1, A2, A3, A4a, A4b automatically satisfies A5 and A6.*

Proof. The reduction-and-induction argument of §5.4 uses only A1, A2, A3, A4a, A4b to show $I = r$ on $\mathsf{SemComp}_{\mathrm{ex}}$. By the existence calculation (§5.3), r satisfies all seven axioms; hence I satisfies A5 and A6. □

Specifically: A5 ($I(C_e) = 1$) is forced because C_e has $\mu(C_e) = 1$, the unique disclosure that reduces μ to zero corresponds to a non-trivial cohomology class, and A4a + A2 then give $I(C_e) = 0 + 1 = 1$. A6 (excision) is forced because the witness rank, which any A1–A4b-invariant must equal, satisfies excision via the standard Mayer–Vietoris calculation.

Two structural remarks. First, in the exact regime five axioms suffice — A1 (iso-invariance), A2 (obstruction triviality), A3 (additivity), A4a (disclosure sensitivity), A4b (coboundary blindness) — to force the witness rank uniquely. The "-1" embedded in A4a is doing double duty: it specifies both the cancellation rule and the scale. A4b ensures the theory is not over-counted by redundant disclosures.

Second, A5 and A6 are *not* redundant outside the exact regime. When the cochain complex fails to split rationally — the surrogate regime — the disclosure reduction may not converge cleanly: an "independent" disclosure outside the exact regime can fail to reduce the obstruction by exactly one, requiring the cycle normalization (A5) and the excision rule (A6) to constrain the residual structure. Proposition 5.4 is a feature of the regime, not of the axioms.

We therefore retain A5 and A6 as part of the doctrinal axiom system. They express, respectively, the *normalization* property (one elementary obstruction equals one rank-unit) and the *locality* property (the obstruction of a whole is recoverable from its parts up to a defect term) that any *theory* of compositional obstruction must satisfy. In the exact regime they are automatic; outside, they are non-trivial constraints. The full axiom system describes what is required of the invariant in general — not only what is enforceable in the cleanest case.

A more detailed version of the uniqueness argument is given in notes/proof-skeleton-exact.md. An adversarial discussion addressing the "smuggling" objection — whether A4a/A4b implicitly contain the conclusion via cohomological independence — is given in notes/anti-tautology.md.

5.6 Realization: connecting seam presentations to the Lean formalization

The Lean module formalizes Theorem 5.1 at the level of `DoctrineCarrier` — an abstract structure with a ground set, a dimension function, an independence relation, and a rank. The following lemma anchors the Lean verification to the concrete mathematics of this paper.

Lemma 5.5 (Realization). Every $G \in \mathsf{SemComp}_{\text{ex}}$ determines a `DoctrineCarrier` instance in which:

- the ground set is the set of admissible one-generator seam disclosures on G, modulo coboundary equivalence;
- the independence relation is linear independence in $H^1_{\text{seam}}(G)$;
- the rank is $\dim H^1_{\text{seam}}(G) = r(G)$.

Under this realization, the Lean theorem `disclosure_characterization` instantiates Theorem 5.1: the unique numerical invariant respecting the independence relation equals the rank. The Lean theorem `pigeonhole_perturbation_family` formalizes the cardinality/pigeonhole core of Theorem 8.4 — the abstract fact that a family of $|\mathbb{E}|^{r(G)}$ distinct perturbations forces a transcript of at least $r(G) \cdot \log_2 |\mathbb{E}|$ bits — without encoding the full concrete-seam presentation or the convention-field structure of the perturbation family.

Proof. The ground set is well-defined: admissible disclosures at seam contexts (Definition 3.1), modulo the coboundary equivalence of §3.3, form the ground set of the disclosure matroid (Remark 3.7.1). The independence relation is linear independence in $H^1_{\text{seam}}(G)$, which satisfies the matroid axioms (hereditary, exchange, augmentation) by Remark 3.7.1. The rank of this matroid equals $\dim H^1_{\text{seam}}(G)$ by definition. The dimension function is inherited from the latent quotient dimensions at seam contexts. □

Remark. The disclosure matroid of the `DoctrineCarrier` coincides with the column matroid of the observable-column corank (Paper II, Backbone Theorem): the uniqueness theorem of this paper thus also characterizes the matroid rank function of the column-matroid backbone. The diagnostic invariant (§3), the repair calculus (§6), the matroid rank (Paper II), and the Gram computation (Paper III) are four projections of one underlying object.

5.7 Discussion

The proof has the same shape as the uniqueness arguments in Eilenberg–Steenrod (homology), in the Tutte–Whitney characterization (matroid rank), and in Shannon (entropy): show existence, show the candidate satisfies the axioms, then show uniqueness by induction over a structural reduction in which the axioms force the invariant on smaller objects to determine its value on larger objects. A more recent parallel in the persistence literature is the universality of the ℓ^p-metric on merge trees (Cardona–Curry–Lam–Lesnick, SoCG 2022): a single invariant is forced by being the *largest* one with a stability property, rather than the smallest satisfying axioms — but the underlying logic is the same, and the exact-weights theory of Bubenik–Scott–Stanley (*J. Appl. Comput. Topol.* 7, 2023) makes the connection between numerical invariants and stability conditions explicit.

Three observations on the structure of the proof and the role of the individual axioms.

1. Where the originality lives. The proof itself is short — a finite induction with two reduction operations. The originality is not in the proof. It is in the choice of category SemComp, in the identification of disclosure as the primitive operation, and in the claim that disclosure independence is the structural feature any reasonable theory of compositional obstruction must respect. Once those choices are accepted as the right primitives — which is Layer 1 of §5.1 — the numerical uniqueness (Layer 2) follows almost automatically. The theorem is structurally analogous to Shannon's: not because the proof is hard, but because the axioms identify the object.

2. Five axioms suffice in the exact regime. Proposition 5.4 establishes that A1 (iso-invariance), A2 (obstruction triviality), A3 (additivity), A4a (disclosure sensitivity), A4b (coboundary blindness) — together — force the witness rank uniquely on $\mathsf{SemComp}_{\mathrm{ex}}$. A5 and A6 are automatic in this regime. This is the analogue of Milnor's observation, in algebraic topology, that the dimension axiom of Eilenberg–Steenrod can be replaced by additivity in suitable categories of spaces.

3. The doctrinal axiom system is non-minimal by design. In the exact regime, A5 and A6 are redundant. Outside, both become essential: A4a may not converge cleanly, and the residual structure requires the cycle-normalization axiom (A5) to fix the scale of an irreducible elementary obstruction and the excision axiom (A6) to control how obstruction decomposes under cover. Stating all six axioms describes what the obstruction theory must look like *in general*, including in regimes where the proof in §5.4 does not directly apply.

The same pattern holds in algebraic topology: the Eilenberg–Steenrod axioms include redundancies in particular categories (e.g., for finite CW-complexes, dimension is implied by exactness and additivity), but the full axiom system remains the right description of what a homology theory is, because the redundancies dissolve in larger categories (infinite spectra, non-CW spaces).

5.8 What this section achieves

Witness rank is the unique invariant on $\mathsf{SemComp}_{\mathrm{ex}}$ that respects seam-independent admissible disclosure (Layer 1) — equivalently, the unique invariant satisfying axioms A1–A6 of §4 (Layer 2). The choice of $r := \dim H^1_{\mathrm{seam}}$ as the program's central diagnostic is not one option among many; it is forced by accepting disclosure independence as the primitive operation of compositional repair. The Realization Lemma (§5.6) anchors this theorem to the Lean formalization and to the column-matroid backbone of Paper II.

This is the result on which the doctrine of subsequent sections rests. The rest of the paper derives consequences: the minimum-repair invariant equals r (§6), coherence certificates have a disclosure normal form (§7), bounded-channel protocols cannot certify high-rank obstruction (§8), and the existing program — Bulla, the seam protocol, BABEL — supplies models of the abstract framework whose computational results confirm the theorem (§10).

---

§6. Repair Duality

We now establish that the minimum-repair invariant — the smallest number of independent disclosures needed to reduce the obstruction to zero — equals the witness rank. This identifies the *cost of certifying coherence* with the *amount of obstruction*: the dual sides of the same number.

6.1 The minimum-repair invariant

Definition 6.1 (Repair invariant). For $G \in \mathsf{SemComp}_{\text{ex}}$, the *repair invariant* is

$$\rho(G) := \min\{k \in \mathbb{Z}_{\geq 0} : \text{there exist disclosures } \sigma_1, \dots, \sigma_k \text{ on successive complexes such that } r(G \mathbin{\backslash\backslash} W_1 \cdots \mathbin{\backslash\backslash} W_k) = 0\}.$$

Equivalently, $\rho(G)$ is the smallest number of disclosures needed to make the latent presheaf cohomologically trivial.

This is a non-negative integer invariant of G, depending only on the iso class of the cochain complex.

6.2 The duality theorem

Theorem 6.2 (Repair Duality). *For all $G \in \mathsf{SemComp}_{\text{ex}}$,*

$$\rho(G) = r(G).$$

Proof. We show both inequalities.

Upper bound: $\rho(G) \leq r(G)$. By the reduction-and-induction argument of §5.3, any G can be reduced to a base-case complex (with $\mathcal{L}$ cohomologically trivial) by exactly $r(G)$ non-trivial disclosures (sub-case D2b applied $r(G)$ times) plus arbitrarily many cochain-trivial disclosures (sub-case D2a) which preserve r. So a sequence of $r(G)$ non-trivial disclosures suffices.

Lower bound: $\rho(G) \geq r(G)$. By Lemma 3.7, each non-trivial disclosure reduces r by exactly 1; cochain-trivial disclosures preserve r. So k disclosures, in any combination, reduce r from $r(G)$ to at least $r(G) - k$. To reach $r = 0$ requires $k \geq r(G)$. □

6.3 Discussion

Repair duality says that the cost of certifying compositional coherence — measured in independent declarations that must be added to the observable interface — is exactly the obstruction count. The diagnostic invariant and the repair invariant are the same number; they are projections of one underlying cohomological quantity. The closest existing local-reasoning result is the frame rule of separation logic [2], which describes when a property local to one heap region survives composition with disjoint regions; repair duality describes the dual quantity in the *non-frameable* setting where the seam itself fails to glue, and quantifies the minimum disclosure required to recover frameability.

The duality has both descriptive and operational content. *Descriptively*: there is no certification scheme that achieves coherence with fewer than $r(G)$ disclosures, no matter how cleverly the disclosures are chosen. *Operationally*: any algorithm that computes $r(G)$ also gives a constructive route to a minimum-cost repair, by exhibiting a sequence of $r(G)$ non-trivial disclosures, each reducing the obstruction by one.

This second statement is implicit in §3.4: the witness Gram presentation of H^1 exhibits $r(G)$ basis cocycles, and disclosing each in turn produces a minimum-cardinality repair. The witness Gram is therefore not only a *diagnostic* tool — telling us how much obstruction is present — but also a *prescriptive* tool — telling us which disclosures to make. The two roles coincide because the underlying invariant is the same.

6.4 The forced/optimizable decomposition

Repair sequences of cardinality $r(G)$ are not unique: there are typically many bases of $H^1(G)$, each giving a different repair sequence. But the *cardinality* is fixed; what varies is the *choice* of which independent dimensions to disclose first.

This induces a natural decomposition of the repair calculus into two parts:

- **Forced.** The cardinality $r(G)$ — the *number* of disclosures required — is invariant across all repair strategies. This is the structural cost of coherence; it cannot be optimized away.
- **Optimizable.** The *choice* of disclosure basis — which $r(G)$ specific declarations to add — admits optimization. Different bases may have different practical costs (different fields are easier or harder to declare). The repair-invariant theorem fixes the cardinality but leaves the basis choice open.

This is the formal version of the "forced versus optimizable disclosure" structure observed in the existing program: the witness rank determines the floor of repair cost; below that, one optimizes over basis choices subject to a cardinality constraint. The two halves of the decomposition are made precise here as: (i) cardinality is forced by Theorem 6.2; (ii) basis is optimizable subject to maintaining cardinality.

6.5 What this section achieves

Minimum-repair cardinality equals the witness rank. The diagnostic invariant and the repair invariant coincide. The cost of certifying coherence is the obstruction count; one disclosure per obstruction class, no fewer. This duality is a corollary of Theorem 5.1 — both sides follow from A1, A2, A3, A4a, A4b — and it makes the cost structure of compositional verification *forced by the same axioms* that fix the witness rank.

---

§7. Disclosure Normal Form

The repair calculus of §6 says that $r(G)$ disclosures suffice to bring a complex to a coherent state; the characterization theorem of §5 says that the witness rank is forced by the axioms. Together they yield a corollary about *certificates* of coherence: any proof that a composition is coherent normalizes to a particular shape — a *disclosure normal form* — which is small, replayable, and verifiable independent of the original argument. We call its operational realization a *receipt*.

This section makes that statement precise.

7.1 What a coherence certificate must contain

The disclosure normal form sits in the proof-carrying-code lineage [12]: a coherence certificate carries with it a small, independently checkable proof-of-correctness that the verifier can replay without re-running the original argument. The novelty here is the specific *canonical form* the certificate takes — disclosed-cocycle-basis plus rank witness — which the characterization theorem (§5) and repair duality (§6) jointly force. Suppose we have a proof that $G \in \mathsf{SemComp}_{\text{ex}}$ is coherent — that its obstruction vanishes, $r(G) = 0$. What must such a proof contain?

By the characterization theorem, $r(G) = 0$ is equivalent to: G admits no non-trivial cohomology class in H^1. Equivalently, every latent dimension is in $\operatorname{im}\delta^0$ (every potential obstruction is already a coboundary). Equivalently, the latent presheaf $\mathcal{L}$ is fully observable up to coboundaries.

But this last statement is rarely the *original* form of the data. Compositions arise with non-trivial $\mathcal{L}$; coherence is achieved by *adding disclosures* to bring the cohomology to zero. So a certificate of coherence for an originally-incoherent G_0 takes the form:

$$G_0 \mathbin{\backslash\backslash} W_1 \mathbin{\backslash\backslash} W_2 \cdots \mathbin{\backslash\backslash} W_k = G_k, \quad r(G_k) = 0.$$

The certificate must therefore record:

(i) The *base complex* G_0 — the original interface data.
(ii) The *disclosure set* $\{W_1, \dots, W_k\}$ — the declarations added to reach coherence.
(iii) The *vanishing-obstruction proof* — a witness, computable cochain-level, that $r(G_k) = 0$.
(iv) The *regime tag* — confirmation that $G_0 \in \mathsf{SemComp}_{\text{ex}}$, so the calculations are exact.

We call a tuple $\mathcal{R} = (G_0, \{W_i\}, \pi, \tau)$ satisfying (i)–(iv) a *receipt*.

7.2 The disclosure-normal-form theorem

Theorem 7.1 (Disclosure Normal Form). *Let $G_0 \in \mathsf{SemComp}_{\text{ex}}$ be a semantic interface complex. Any proof that G_0 admits a coherent extension by some sequence of disclosures normalizes to a receipt $\mathcal{R} = (G_0, \{W_1, \dots, W_k\}, \pi, \tau)$ in which:*

- *(a) The disclosure set has cardinality $k = r(G_0)$ and is cochain-independent.*
- *(b) The vanishing-obstruction proof π is the linear-algebra certificate that* $\operatorname{rank}\delta^0_{G_k} = \dim_{\mathbb{Q}} C^1(G_k)$*, which is verifiable in $O(|C^1|^\omega)$ time using rational arithmetic (ω the matrix-multiplication exponent).*
- *(c) The regime tag τ is verifiable: a cochain-level check confirms that G_0 satisfies the exact-regime conditions of Definition 3.5.*

Two receipts $\mathcal{R}, \mathcal{R}'$ for the same base G_0 are equivalent if their disclosure sets are bases of the same subspace of $H^1(G_0)$.

Proof. By Theorem 5.1 and Theorem 6.2, any sequence of disclosures bringing r to zero must have non-trivial cardinality at least $r(G_0)$, and a cochain-independent sequence of cardinality exactly $r(G_0)$ suffices. So normalization to (a) is achieved by removing redundant or cochain-trivial disclosures from any longer sequence (each such removal does not increase the residual obstruction by Lemma 3.7).

For (b): once the disclosure set is fixed, the vanishing-obstruction proof reduces to a rank computation on the modified cochain complex, which is a polynomial-time linear-algebra problem over $\mathbb{Q}$.

For (c): the exact-regime conditions are checkable from the data of G_0 — operational sufficient conditions DFD + CHP (Definition 3.5) are direct structural properties of the convention presheaf.

The equivalence of receipts modulo the $H^1(G_0)$-basis ambiguity is the statement of §6.4: cardinality is forced, basis is optimizable. □

7.3 Compositionality of receipts

Receipts compose along compatible covers, by the standard Mayer–Vietoris exact sequence for the cochain complex.

Proposition 7.2 (Compositionality). *Let (U, V) be a compatible cover of $G_0 \in \mathsf{SemComp}_{\mathrm{ex}}$ in the exact regime. Let $\mathcal{R}_U, \mathcal{R}_V, \mathcal{R}_{U\cap V}$ be receipts for $U, V, U \cap V$ respectively. Then the disclosure set $\{W_i^U\} \cup \{W_j^V\} \setminus \{W_l^{U\cap V}\}$ — disclosures from U and V minus duplicates from the intersection — is a valid disclosure set for a receipt $\mathcal{R}_G$ for G_0, with cardinality $r(U) + r(V) - r(U \cap V) = r(G_0)$.*

Proof. Direct from A6 (or from its equivalent, Mayer–Vietoris for cochain cohomology). The cardinality formula is the inclusion-exclusion of the rank function, which is exactly axiom A6. □

This compositionality is the proof-theoretic counterpart of the cohomological locality recorded in A6: receipts can be assembled from receipts on covers, with redundancies subtracted.

7.4 Soundness

Proposition 7.3 (Soundness). *If $\mathcal{R} = (G_0, \{W_i\}, \pi, \tau)$ is a valid receipt — meaning π verifies and τ confirms the exact regime — then the disclosed complex $G_k = G_0 \setminus\setminus W_1 \cdots \setminus\setminus W_k$ has vanishing obstruction: $r(G_k) = 0$.*

Proof. Direct from the verifiable certificates: π verifies that the rank of the modified δ^0 matches $\dim C^1$, which by Definition 3.4 means $H^1(G_k) = 0$. □

In words: a receipt is *sound* if its certificates check; soundness implies that the certified composition has no hidden semantic obstruction relative to the axiomatized framework.

7.5 Completeness

The dual to soundness: every coherence-achievable complex admits a valid receipt.

Proposition 7.4 (Completeness). *For every $G_0 \in \mathsf{SemComp}_{\mathrm{ex}}$, there exists a valid receipt $\mathcal{R} = (G_0, \{W_1, \dots, W_{r(G_0)}\}, \pi, \tau)$ certifying that the disclosed complex has vanishing obstruction.*

Proof. By the construction in the proof of Theorem 6.2 (upper bound), choose any basis $\{[\sigma_1], \dots, [\sigma_{r(G_0)}]\}$ of $H^1(G_0)$, lift each to an underlying disclosure (p_i, W_i), and form the disclosure set $\{W_1, \dots, W_{r(G_0)}\}$. By repeated application of A4a, $r(G_0 \setminus\setminus W_1 \cdots \setminus\setminus W_{r(G_0)}) = 0$. The verification certificate π is the rank computation; the regime tag τ is the verification of exact-regime conditions on G_0. Both are computable in polynomial time. □

Together Propositions 7.3 and 7.4 say: receipts form a *sound and complete* certificate system for compositional coherence in the exact regime. Soundness says verifying the certificate suffices; completeness says certificates exist whenever they should.

7.6 Disclosure normal form is the canonical certificate form

Theorem 7.1 says that *any* proof of compositional coherence — however originally presented — normalizes to a receipt of cardinality $r(G_0)$. The argument is:

1. The proof must, in the end, certify $r(G_k) = 0$ for some sequence of disclosures bringing r to zero.
2. The cardinality is at least $r(G_0)$ by Theorem 6.2.

3. Any disclosures beyond cardinality $r(G_0)$ are redundant — cochain-trivial or basis-equivalent — and can be removed.
4. The minimum-cardinality disclosure set is unique up to basis choice in $H^1(G_0)$.

So every proof of coherence reduces, after normalization, to a receipt: a base complex, a minimum-cardinality disclosure basis, a rank-verification certificate, and a regime tag. There is no proof of coherence that does not, after this reduction, take this shape.

This is the proof-theoretic content of the framework: *coherence proofs have a disclosure normal form.* Receipts are not one possible certificate format among many; they are the unique form to which any proof reduces.

7.7 What this section achieves

Coherence certificates have a normal form: an independent disclosure set of cardinality $r(G_0)$ together with a verifiable rank-vanishing certificate. Receipts realize this disclosure normal form operationally. The proof system is *sound* (Proposition 7.3): valid receipts certify zero obstruction. The proof system is *complete* (Proposition 7.4): every coherent extension admits a receipt. Receipts compose along covers (Proposition 7.2).

This identifies disclosure normal form as the proof-theoretic counterpart of the witness rank: just as the rank is the unique invariant respecting seam-independent disclosure (Theorem 5.1), the disclosure set is the unique certificate form to which proofs reduce (Theorem 7.1). The diagnostic invariant, the repair calculus, and the certificate format are three projections of one underlying object.

---

§8. Communication Lower Bound

A *protocol* in this section is an information-exchange procedure between participants holding the local data of an interface complex. The protocol exchanges messages — a transcript — and at the end emits a verdict on whether the composition is coherent. We ask: how much must the transcript carry?

The answer is forced by the characterization theorem and a pigeonhole argument.

8.1 The setting

The result of this section is a clean pigeonhole bound in the communication-complexity tradition (Yao [3]; standard textbook treatment in Kushilevitz and Nisan [4]): we ask how much a protocol's transcript must carry to soundly *certify* the coherence-or-repair status of compositions in a family.

The lower bound applies to **receipt-producing certification**, not binary classification. The protocol receives the observable skeleton S of G as pre-shared input (the context poset, convention presheaf, and observable subpresheaf structure are known to the verifier). The transcript carries the *semantic-frame declarations* needed for repair — the cocycle-class information that identifies which obstruction-class assignment the input G carries over the fixed skeleton S. The protocol does not need to encode G itself in the transcript.

Let $\mathcal{C}$ be a family of semantic interface complexes sharing a fixed observable skeleton S — they differ only in the hidden obstruction-class assignment within $H^1_{\mathrm{seam}}(S;\mathbb{E})$.

A protocol Π on $\mathcal{C}$ takes $G \in \mathcal{C}$ as input and produces a transcript $T(G) \in \mathcal{T}$ together with a verdict $V(G)$ that is one of the following:

- *coherent*, asserting that $r(G) = 0$; or
- a *repair receipt* in the sense of Theorem 7.1 — namely a tuple $(\{W_1, \dots, W_k\}, \pi, \tau)$ recording a cochain-independent disclosure set (referencing the pre-shared skeleton S) whose application brings G to a coherent extension, together with the rank-vanishing certificate π and the regime tag τ.

We say Π is *sound on* $\mathcal{C}$ if for every $G \in \mathcal{C}$, $V(G) =$ coherent implies $r(G) = 0$, and every receipt verdict $V(G)$ is a *valid* receipt (its certificate π verifies and the disclosed extension has $r = 0$).

We say Π is *complete on* $\mathcal{C}$ if every coherent G receives the *coherent* verdict and every incoherent G receives a valid receipt.

A protocol is *correct* if it is both sound and complete. A correct protocol is *certifying* in the sense that the verdict carries enough information to either confirm coherence or reconstruct a minimum-cardinality repair — not merely to classify G as coherent or incoherent. We restrict the lower bound below to certifying protocols; for protocols that emit only the binary classifier $\{\text{coherent}, \text{incoherent}\}$, the bound collapses to the trivial $|\mathcal{T}| \geq 2$ (all incoherent inputs may share one transcript).

8.2 Perturbation-family centered at G

The pigeonhole bound below holds for any family $\mathcal{C}$ of pairwise non-isomorphic complexes containing both coherent and incoherent representatives. The natural such family — the one canonically associated with a single complex G — is the *perturbation family* obtained by varying G's cocycle representatives over the obstruction module while holding the underlying combinatorial structure fixed. Defining it here, before the lower bound, gives the four-form theorem of §8.3 a canonical source of cardinality.

Definition 8.1 (Perturbation family centered at G). *Fix $G \in \mathsf{SemComp}_{\mathrm{ex}}$ with finite coefficient group $\mathbb{E}$. Let $\Omega_G := H^1(G;\mathbb{E})$ be the obstruction-class set of G. In the abelian regime, fix a basis $\{[\sigma_1], \dots, [\sigma_{r(G)}]\}$ of $H^1(G;\mathbb{E})$; in the non-abelian regime, fix a pointed-set parameterization of Ω_G. The* perturbation family centered at G *is*

$$\mathcal{C}(G) \;:=\; \{G_{\vec{s}} : \vec{s} \in \Omega_G\},$$

where $G_{\vec{s}}$ is the complex obtained from G by setting the cocycle representative of basis class i to the equivalence-class element indexed by s_i, holding the underlying combinatorial structure of G — its context poset, its convention presheaf, its observable subpresheaf — fixed.

In the abelian regime, Ω_G is a finite abelian group of cardinality $|\mathbb{E}|^{r(G)}$, isomorphic abstractly to $\mathbb{E}^{r(G)}$. In the non-abelian regime, Ω_G is a pointed set with basepoint the trivial cocycle class, and $|\Omega_G| = |H^1(G;\mathbb{E})|$.

The point of the definition is that $\mathcal{C}(G)$ is *canonically* associated with G: no auxiliary worst-case family construction is needed, no choice of family is made for the bound. The family is determined by the obstruction module of G together with the chosen basis, and it captures exactly the information-content of that module.

Lemma 8.2 (Cocycle-class gauge invariance). *For $\vec{s}, \vec{s}' \in \Omega_G$ with $\vec{s} \neq \vec{s}'$, the complexes $G_{\vec{s}}$ and $G_{\vec{s}'}$ are non-isomorphic.*

Proof. The cohomology-class assignment is an invariant of the iso class: an isomorphism $\phi : G_{\vec{s}} \to G_{\vec{s}'}$ would induce, by functoriality of H^1, a bijection on cohomology classes carrying $\vec{s}$ to $\vec{s}'$. But the basis $\{[\sigma_1], \dots, [\sigma_{r(G)}]\}$ is fixed as part of the data defining $\mathcal{C}(G)$, so the induced bijection must fix the basis classes pointwise, forcing $\vec{s} = \vec{s}'$. □

Lemma 8.3 (Cardinality). *In the abelian regime, $|\mathcal{C}(G)| = |\mathbb{E}|^{r(G)}$. In the non-abelian or pointed-set regime, $|\mathcal{C}(G)| = |H^1(G; \mathbb{E})|$.*

Proof. By Lemma 8.2 the assignment $\vec{s} \mapsto G_{\vec{s}}$ is injective, and by construction it is surjective onto $\mathcal{C}(G)$, so $|\mathcal{C}(G)| = |\Omega_G|$. The abelian-regime equality $|\Omega_G| = |\mathbb{E}|^{r(G)}$ is the standard structure theorem for H^1 of a cellular sheaf with finite abelian coefficients on a finite poset (Curry [14]; equivalently the universal-coefficient theorem applied to the rationally split cochain complex of Definition 3.5). The non-abelian-regime equality is the cardinality of the pointed set $H^1(G; \mathbb{E})$ by definition. □

A protocol that claims to certify coherence on G must, by soundness on $\mathcal{C}(G)$, distinguish every non-trivial cocycle class from the trivial class — otherwise it accepts an incoherent perturbation as coherent. This turns the abstract pigeonhole bound recorded next into the explicit rank-times-alphabet form.

8.3 The lower bound

The lower bound takes four explicit forms, distinguished by what is assumed about the coefficient group $\mathbb{E}$ and the alphabet of disclosed cocycle classes. Form (a) is the most general — pigeonhole on an arbitrary mixed-coherence family. Forms (b)–(d) specialize to the canonical perturbation family $\mathcal{C}(G)$ of Definition 8.1 under different assumptions on $\mathbb{E}$.

Theorem 8.4 (Communication Lower Bound — four forms). *Let Π be a correct, certifying protocol (in the sense of §8.1) with transcript space $\mathcal{T}$. Then:*

(a) [Cardinality form.] If $\mathcal{C}$ is a family sharing a fixed observable skeleton whose members carry pairwise distinct obstruction-class assignments — equivalently, a family for which the cocycle-class tuples in $H^1_{\text{seam}}(S; \mathbb{E})$ differ pairwise (e.g. the perturbation family $\mathcal{C}(G)$ of Definition 8.1) — then Π on $\mathcal{C}$ requires $|\mathcal{T}| \geq |\mathcal{C}|$, equivalently $|T(G)| \geq \log_2 |\mathcal{C}|$ bits in the worst case.

(b) [Abelian rank-times-alphabet form.] If $\mathcal{C} = \mathcal{C}(G)$ is the perturbation family centered at $G \in \mathsf{SemComp}_{\text{ex}}$ (Definition 8.1) with finite abelian coefficient group $\mathbb{E}$, then $|\mathcal{T}| \geq |\mathbb{E}|^{r(G)}$, equivalently

$$|T(G)| \geq r(G) \cdot \log_2 |\mathbb{E}|$$

bits in the worst case over $\mathcal{C}(G)$.

(c) [Variable-alphabet form.] If the disclosed cocycle of class i in $\mathcal{C}(G)$ is drawn from a finite alphabet Σ_i — that is, the i-th cocycle generator carries a typed value with type-specific cardinality

— *then*

$$|T(G)| \geq \sum_{i=1}^{r(G)} \log_2 |\Sigma_i|$$

bits in the worst case.

(d) [Non-abelian / pointed-set form.] If $\mathbb{E}$ *is non-abelian, or merely a pointed set with basepoint the trivial cocycle class, then* $|T(G)| \geq \log_2 |H^1(G;\mathbb{E})|$ *bits in the worst case, recovering form (b) in the abelian special case where* $|H^1(G;\mathbb{E})| = |\mathbb{E}|^{r(G)}$*.*

Proof. Form (a). Suppose $|\mathcal{T}| < |\mathcal{C}|$. By pigeonhole there exist $G, G' \in \mathcal{C}$ with $G \ncong G'$ and $T(G) = T(G')$. The verdict V is a function of the transcript, so $V(G) = V(G')$. By the certifying-protocol model of §8.1, $V(G)$ is either *coherent* or a receipt encoding the disclosed cocycle classes of G. The members of $\mathcal{C}$ have pairwise-distinct cocycle-class assignments by hypothesis, so $V(G) \neq V(G')$ on every pair: a coherent verdict would have to be issued for at most one of $\{G, G'\}$ (the one with $r = 0$, if any), and a receipt verdict encodes a specific cocycle-class tuple that distinguishes its input from every other member. Either way, the equality $V(G) = V(G')$ forced by transcript collision contradicts soundness on $\mathcal{C}$. Therefore $|\mathcal{T}| \geq |\mathcal{C}|$.

Form (b) is form (a) applied to $\mathcal{C}(G)$, whose pairwise cocycle-distinctness is Lemma 8.2 and whose cardinality is $|\mathbb{E}|^{r(G)}$ by Lemma 8.3. Form (c) is form (a) applied with variable-alphabet cardinality $|\mathcal{C}(G)| = \prod_{i=1}^{r(G)} |\Sigma_i|$, equivalent to form (b) under the alphabet substitution $|\mathbb{E}| \to |\Sigma_i|$ at each generator. Form (d) is form (a) applied with the non-abelian cardinality $|H^1(G;\mathbb{E})|$ from Lemma 8.3. □

The four forms differ only in what is plugged in for $|\mathcal{C}|$ at the pigeonhole step: family cardinality (a), abelian rank-power (b), product-of-alphabets (c), or non-abelian cohomology cardinality (d). The proof of (a) is the load-bearing argument; (b)–(d) follow by substitution. The Lean module lean/CompositionDoctrine/CoherenceRate.lean formalizes (a)–(c) at the cardinality form (Aristotle runs `b84c8f33` and `3fd02ee6`); form (d) is stated for completeness but not Lean-verified, since non-abelian cohomology lies outside mathlib's current scope.

The two subsections below elaborate two important special cases of the four-form theorem: §8.4 is the obstruction-structure refinement (a special case of (a) at the perturbation family); §8.5 is the fixed-structure case (a special case of (b) with uniform binary alphabet).

8.4 The obstruction-structure refinement

The bound of Theorem 8.4 is in terms of family cardinality. A finer bound expresses the cost in terms of the structural variation in obstruction class.

Corollary 8.5 (Obstruction-structure lower bound). *Let* $\mathcal{C}$ *be a family in which the obstruction modules* $H^1(G)$ *for* $G \in \mathcal{C}$ *collectively span a vector space of dimension* D*. Then any correct protocol on* $\mathcal{C}$ *transmits at least* D *bits in the worst case.*

Proof sketch. Construct a sub-family $\mathcal{C}' \subseteq \mathcal{C}$ realizing 2^D distinct subsets of obstruction classes — for instance, by varying which subset of a fixed D-dimensional obstruction space is "actively obstructed" in each member. Apply Theorem 8.4 to $\mathcal{C}'$ with $|\mathcal{C}'| = 2^D$. □

This sharpens the lower bound: the *number* of bits is the *dimension* of the obstruction variation, not just the family cardinality.

8.5 The fixed-structure case

Let $\mathcal{C}_*$ denote the family of complexes with a fixed underlying context poset and convention dimensions, where what varies is which dimensions are observable and which are latent. The witness rank ranges over $\{0, 1, \dots, R\}$ for some R determined by the structure.

Corollary 8.6 (Fixed-structure lower bound). *Any correct protocol on $\mathcal{C}_*$ must produce transcripts whose underlying message space has cardinality at least 2^R, equivalently transmit at least R bits.*

Proof. Identical pigeonhole, with $|\mathcal{C}_*| \geq 2^R$ from the family construction. □

8.6 Rate-theoretic restatement

The lower bound of Theorem 8.4 — and its sharper four-form refinements above — admit a Shannon-rate restatement. The perturbation family $\mathcal{C}(G)$ of Definition 8.1 plays the role of a *source* in the source-coding analogy of Shannon [16] and Cover and Thomas [17]: each element of $\mathcal{C}(G)$ is one realization of the obstruction state, and a sound coherence-certifying protocol must distinguish every realization. The cardinality $|\mathcal{C}(G)|$ plays the role of the message space's entropy, and $\log_2 |\mathcal{C}(G)|$ is the per-instance bit cost.

Converse, restated. By Lemma 8.3, $|\mathcal{C}(G)| = |\mathbb{E}|^{r(G)}$ in the abelian regime, so any sound and complete protocol on $\mathcal{C}(G)$ transmits at least $r(G) \cdot \log_2 |\mathbb{E}|$ bits in the worst case. Equivalently: the *rate*, in bits per disclosed cocycle generator, is at least $\log_2 |\mathbb{E}|$. This is the Shannon-style converse: a source with H bits of entropy per symbol cannot be compressed below H bits per symbol [16].

Achievability, restated. The disclosure normal form of Theorem 7.1 produces a receipt $\mathscr{R} = (G_0, \{W_1, \dots, W_{r(G_0)}\}, \pi, \tau)$ that saturates this bound up to polynomial overhead. The cocycle-disclosure portion encodes exactly $r(G_0)$ generators, with the i-th generator carrying $\lceil \log_2 |\mathbb{E}| \rceil$ bits in the uniform-alphabet case (or $\lceil \log_2 |\Sigma_i| \rceil$ bits in the variable-alphabet case). The vanishing-obstruction proof π and the regime tag τ each cost $O(\mathrm{poly}(|G_0|))$ bits, sublinear in $r(G_0)$ for fixed regime. Proposition 7.4 (completeness) constructs the receipt in polynomial time, so the achievability is constructive.

Combining. The converse and achievability close to within a polynomial-in-$|G|$ correction: at least $r(G) \cdot \log_2 |\mathbb{E}|$ bits, at most $r(G) \cdot \lceil \log_2 |\mathbb{E}| \rceil + \mathrm{poly}(|G|)$ bits. *The rate at which a sound coherence protocol must transmit semantic-frame information is, in the exact regime, the entropy of the obstruction module's coefficient group.*

This is the structural counterpart of Shannon's source-coding theorem [16; 17, Ch. 5]: a memoryless source with entropy H requires H bits per symbol (asymptotically tight); a coordination graph with witness rank r requires $r \cdot \log_2 |\mathbb{E}|$ bits per instance to certify (tight up to polynomial overhead). The source is the perturbation family; the channel is the protocol's transcript; the alphabet is the coefficient group $\mathbb{E}$. The witness rank — the obstruction-counting invariant of §3 — is the *length* of the source-coding analog: diagnostic rank (§3), repair cardinality (§6), and certification rate (§8) are forced by Theorems 5.1, 6.2, and 8.4 to be the same number, projected onto three different operational axes.

8.7 The semantic communication interpretation

The lower bound has a clean interpretation in the language of communication complexity.

Corollary 8.7 (Witness rank as a communication lower bound). *For any family of semantic interface complexes whose witness ranks range over* $\{0, \dots, R\}$*, any sound protocol that certifies coherence-or-repair status must transmit at least* R *bits of semantic information about the obstruction structure.*

In other words: the witness rank lower-bounds the *semantic communication complexity* of distinguishing coherent from incoherent compositions in a family.

This is a clean formulation of the principle that "channel capacity must exceed obstruction count": a verifier whose protocol cannot transmit enough independent semantic declarations cannot soundly certify coherence over a family of complexes whose obstruction structure spans more than the channel allows.

8.8 Relation to distributed systems

The communication lower bound has formal kinship with classical lower bounds in distributed systems — for instance, with the $\Omega(\log n)$ bandwidth requirement for leader election under partial information, or with the bandwidth requirements for Byzantine agreement. We do not claim it is itself a distributed-systems result; it is a structural lower bound in a different setting. The structural-invariant framing follows the spirit of Lamport's happens-before [5]: a partial-order constraint on what a transcript can encode, forced by what messages can carry rather than by protocol design choices.

The analogy is with FLP-style impossibility arguments: just as FLP shows that a particular kind of agreement is impossible without certain structural primitives (atomic commit, randomization, partial synchrony), Theorem 8.4 shows that a particular kind of certification is impossible without certain structural transmission primitives (witness-rank-many independent semantic declarations). The two are structurally analogous; whether the analogy can be tightened into a formal "semantic FLP" theorem with liveness conditions is left to subsequent work.

We emphasize: §8 is *not* a semantic FLP theorem. It is a clean pigeonhole communication bound. The richer FLP-style analysis with liveness, asynchrony, and fault tolerance is named as open in §9.

8.9 What this section achieves

The witness rank lower-bounds the communication complexity of correct semantic-coherence verification: any sound, complete protocol on a family of complexes whose obstruction structure spans R bits must transmit at least R bits of independent semantic information. The bound is structurally simple — pigeonhole — but its consequences are wide: any verifier with bounded message bandwidth fails to certify coherence on a sufficiently complex family. The cost of certification is, again, the witness rank; this time as a transmission requirement rather than a disclosure cardinality (§6) or a certificate length (§7).

---

§9. Boundaries

The doctrine established in §§4–8 holds in the exact regime $\mathsf{SemComp}_{\mathrm{ex}}$: complexes whose relative seam complex truncates to two terms. Outside this regime — and even within, in directions the present paper does not pursue — natural and important questions remain. We name the most central as conjectures, with the hope that some are provable, some are refinable, and some will turn out to require fundamentally new machinery.

9.1 Surrogate-regime characterization with explicit correction term

Conjecture 9.1 (Surrogate regime). *Let* $\mathsf{SemComp}_{\mathrm{sur}}$ *be the class of finite semantic interface complexes whose relative seam complex does not truncate to two terms — i.e., whose obstruction module exhibits non-trivial higher-cohomology contributions. There exists a numerical invariant* $\rho^* : \mathrm{Ob}(\mathsf{SemComp}_{\mathrm{sur}}) \to \mathbb{Z}_{\geq 0}$ *— the* surrogate witness rank *— and an explicit correction term* $\kappa(G) \geq 0$ *— the* coloop burden *— such that:*

(i) $\rho^*(G) = r(G) + \kappa(G)$ *for all* $G \in \mathsf{SemComp}_{\mathrm{sur}}$*, where* $r(G) = \dim H^1(G)$ *is the standard cohomological rank.*

(ii) ρ^* *is the unique invariant satisfying axioms A1–A4b plus a modified A5* (cycle normalization with multiplicities) *and a modified A6* (excision with explicit correction term κ).

(iii) κ *is computable from the cochain complex in polynomial time and bounds the gap between operational rank-counting and true cohomological rank.*

The conjecture asserts that the surrogate regime admits a parallel characterization theorem, but with two new axioms (modified A5 and A6) that become non-trivial outside the exact regime, and a new invariant κ tracking the excess of operational diagnostics over the true cohomological obstruction.

Status. Open. Partial results in the existing program (the leverage-conservation identity in [witness-geometry-beyond-fee], the additive decomposition in [hierarchical-fee]) suggest the structure is correct in special cases (DFD-violating compositions with bounded coloop count). Full characterization is a multi-quarter research program.

9.2 Verifier universality

Conjecture 9.2 (Verifier universality). *Let* V *be any verifier — i.e., a functor from* $\mathsf{SemComp}_{\mathrm{ex}}$ *to a category of certificates and verifications — that is sound and complete in the technical senses defined elsewhere. Then* V *factors through an equivalent of the witness functor: there exists a natural transformation from* V *to the witness-rank-cum-receipt construction, and this transformation is unique up to natural isomorphism.*

In words: any sound verifier of compositional coherence reduces, up to natural transformation, to one that computes the witness rank and emits the corresponding receipt.

Status. Open. The characterization theorem (§5) and the receipt normal form theorem (§7) suggest the result, but they do not directly prove it. A proof would likely use representation-theoretic or Yoneda-flavored arguments on the verification category.

This is a substantially harder result than the characterization theorem of §5. Theorem 5.1 says: among invariants satisfying A1–A6, witness rank is the unique one. Conjecture 9.2 says: among all verifiers (an a-priori broader category), witness-rank-style verification is essentially the only one. The two are related but not identical.

9.3 Semantic FLP-style impossibility

Conjecture 9.3 (Semantic FLP). *Consider a distributed protocol exchanging messages between heterogeneous components, each holding a piece of a semantic interface complex, in the presence of asynchrony and bounded fault tolerance. There exist regimes in which: (i) every protocol either fails to terminate (loss of liveness) or fails to certify semantic coherence (loss of safety); and (ii) the impossibility is parametrized by the witness rank in a way analogous to the impossibility for state agreement in FLP.*

Status. Open. The communication lower bound of §8 is a clean precursor — a pigeonhole result on transcript size — but does not yet engage with the asynchrony, fault, and termination conditions that distinguish FLP from a static lower bound.

A full semantic-FLP would relate witness-rank-bounded distributed protocols to the classical impossibility tradition in distributed systems. We do not pursue it here; we name it as the natural meeting point between the doctrine and the distributed-systems literature.

9.4 The eval/correctness gap as a clean theorem

Conjecture 9.4 (Eval-correctness inseparability). *There exists a parameterized family of semantic interface complexes* $\{G_n\}_{n\in\mathbb{Z}_{\geq 0}}$ *such that:*

(i) Every output-evaluation metric $E : \mathrm{Ob}(\mathsf{SemComp}) \to [0,1]$ *that is locally computable (in a precisely-definable sense — depending only on bounded-radius local data) is asymptotically uncorrelated with* r*:*

$$\lim_{n\to\infty} \mathrm{Cor}(E(G_n), r(G_n)) = 0.$$

(ii) The family $\{G_n\}$ *realizes arbitrary witness rank* $r(G_n) \to \infty$ *while every locally-computable eval metric remains bounded.*

In words: there is no locally-computable evaluation metric that tracks the true compositional obstruction at scale.

Status. Empirically confirmed in the existing program (Coherence Cliff scaling experiment achieves $R^2 > 0.96$ for sheaf-cohomological diagnostics while best non-sheaf baselines degrade from $R^2 = 0.83$ at small scale to $R^2 = 0.52$ at large scale). The empirical pattern suggests the conjecture holds; the formal theorem requires precise specification of "locally computable" and the construction of an explicit family realizing the asymptotic separation.

This conjecture is the cleanest of the four for near-term formalization: the empirical confirmation is in hand, and the theoretical statement is a finite-dimensional combinatorial-asymptotic claim. We name it here as the natural follow-up paper to the doctrine.

9.5 Coherence-preserving interface evolution

Conjecture 9.5 (Coherence-preserving update). *Let* G *and* G' *be semantic interface complexes related by an interface-update map* $f : G \to G'$ *— for instance, a tool whose manifest is rewritten to add fields, rename conventions, or refine restrictions. If the induced map* $f^\bullet : C^\bullet(G) \to C^\bullet(G')$ *on seam cochain complexes is a chain homotopy equivalence, then witness rank, obstruction class, and disclosure-normal-form receipts are invariant:* $r(G) = r(G')$*, and any valid receipt for* G *transports to a valid receipt for* G' *via the explicit chain-homotopy data.*

In words: an interface update preserves all certificate validity exactly when the cochain-level update is homotopic to the identity. The conjecture identifies the precise condition under which *incremental* recertification is sufficient — versus the conditions under which a global re-derivation is required.

Status. Open. Mathematically the statement is a routine specialization of standard cochain-homotopy invariance; the work is in defining the right notion of *interface-update map* on SemComp and identifying the operational sufficient conditions (analogous to DFD + CHP for the exact regime) that practical tool updates can satisfy. The closest existing template is the persistence-diagram stability theorem of Cohen-Steiner, Edelsbrunner and Harer [11]: small perturbations to the underlying space induce small perturbations to the cohomological invariant, in a quantitatively controlled way. Conjecture 9.5 asks for the analogous statement for our seam-cochain-complex-valued invariant under interface-update perturbation.

This is the most product-adjacent of the open problems. A proof would license a *coherence-preserving versioning discipline* in which manifest updates are accompanied by chain-homotopy data, and existing receipts remain valid without re-verification — solving the operational problem of incremental recertification under continuous interface evolution.

9.6 Quantum / classical sheaf-cohomological capacity correspondence

The five conjectures of §9.1–§9.5 are open problems on the *classical* substrate: extensions or refinements of the framework as established in §§2–8. We close §9 with one *non-conjectural* note: a recently-published result on a *quantum* substrate that coincides with the classical lower bound of §8 under canonical lift. This is a frontier note — not a sixth conjecture — recording an observed correspondence rather than a problem to be solved.

Kurisummoottil Thomas and Chen [18] develop a sheaf-cohomological capacity theory for quantum semantic communication. Their Theorem 1 (cohomological semantic rate, [18, §III-A]) states: *for a quantum semantic sheaf $\mathcal{S}$ over a graph G, the minimum communication rate required for perfect semantic alignment is*

$$R^*_{\text{sem}} = \log_2 \dim H^1(G, \mathcal{S}),$$

measured in qubits (or bits for classical sheaves). The stalks $\mathcal{H}_v$ are finite-dimensional Hilbert spaces, restriction maps are completely positive trace-preserving, and dim refers to Hilbert-space dimension over $\mathbb{C}$.

Canonical-lift correspondence. When the classical obstruction module $H^1(G; \mathbb{E})$ is realized as the computational basis of a $|H^1|$-dimensional Hilbert space — the canonical embedding of a finite group into ℓ^2 over $\mathbb{C}$ — the two formulas coincide numerically: $\dim_{\mathbb{C}} H^1 = |H^1|$ in this lift, so $\log_2 \dim_{\mathbb{C}} H^1$ (KT-Chen [18], qubits) and $\log_2 |\Omega_G|$ (this paper, classical bits) are the same number. The classical-deterministic-protocol bound of Theorem 8.4 — and its sharpened forms in §§8.2–8.6 — is the classical-substrate specialization of [18, Thm 1].

What each side strictly adds. The two results are not equivalent; they extend in orthogonal directions. This paper adds the explicit perturbation family of Definition 8.1 and the polynomial-time constructive achievability of Theorem 7.1's receipt; KT-Chen [18] add entanglement-assisted reduction ([18, Thm 2]), contextuality bounds ([18, Thm 3]), and discord-integration duality ([18, Thm 4]) — quantum-resource extensions the classical setting does not address. The classical analogs of KT-Chen's quantum extensions would require a precisely-specified shared-randomness resource model, explicitly out of scope for this paper.

The frontier this opens. The correspondence raises a structural question that neither paper alone resolves: is the agreement between KT-Chen's quantum bound and the classical bound of §8 a numerical coincidence in the canonical lift, or does it reflect a deeper functorial relation — a quantum-classical adjunction at the level of sheaf-cohomological capacity? A clean adjunction would say that the perturbation-family construction $\mathcal{C}(G)$ and the entanglement-assisted reduction H^1_{EA} are dual structures on the same underlying object. We do not know whether such an adjunction exists; the question is the natural meeting point between the two frameworks.

We record the correspondence here without conjecturing its lift to an adjunction. The empirical observation — that two independently-derived lower bounds, on substrates as different as classical-deterministic protocols and quantum-entanglement-assisted channels, coincide numerically through the same cohomological invariant — is itself worth flagging. Whether the coincidence reduces to a structural theorem is left for subsequent work.

9.7 Discussion

The five conjectures of §9 are not equally hard. Conjecture 9.1 (surrogate-regime characterization) is the natural extension of the present work: a multi-quarter research program with concrete intermediate milestones. Conjecture 9.2 (verifier universality) is genuinely deeper: it requires proving that no fundamentally different verifier shape can soundly certify coherence — a stronger claim than uniqueness within an axiom system. Conjecture 9.3 (semantic FLP) is at the boundary with distributed systems: a full development requires formal apparatus the present paper does not provide. Conjecture 9.4 (eval-correctness inseparability) is the most empirically grounded and theoretically tractable. Conjecture 9.5 (coherence-preserving update) is the most product-adjacent: it directly addresses the operational problem of incremental recertification under interface evolution, and a proof would yield an immediate engineering discipline.

The doctrine paper does not require any of these to be true. Its main result — Theorem 5.1 — stands on its own in the exact regime. The conjectures describe the natural directions in which the doctrine generalizes; whether each direction yields a clean theorem is for subsequent work to decide.

The frontier note of §9.6 is of a different kind. KT-Chen's quantum capacity result [18] is a published theorem on a substrate this paper does not work on; we record its canonical-lift coincidence with the classical bound of §8 because the coincidence is genuine and worth flagging, not because we conjecture either result is downstream of the other. The structural question of whether the correspondence lifts to a quantum-classical adjunction is a frontier between two independent research programs; it is named here so that subsequent work in either program has a fixed reference point.

9.8 What this section achieves

We have named five open conjectures, each natural and concrete, each tied to an existing piece of the program. Surrogate-regime characterization extends the doctrine outside its current scope. Verifier universality strengthens the characterization theorem. Semantic-FLP connects the doctrine to distributed-systems impossibility. Eval-correctness inseparability formalizes an empirically-confirmed phenomenon. Coherence-preserving update gives the operational discipline of incremental recertification a precise mathematical condition. We have additionally recorded one non-conjectural frontier note (§9.6): the canonical-lift coincidence between this paper's classical lower bound and Kurisummoottil Thomas and Chen's quantum capacity theorem [18] — a published result on a different substrate, flagged because the numerical agreement is structural enough to warrant subsequent investigation. The doctrine paper is positioned as the canonical reference for the exact regime; the conjectures and the frontier note jointly define its frontier.

---

§10. Models of the Doctrine

The framework of §§2–8 is product-neutral by design: the category SemComp is a mathematical object, the witness rank is a cohomological invariant, and the receipt is a proof-theoretic normal form. The same machinery applies to any concrete instance of the abstract structure.

This section identifies four such instances — the *Bulla protocol*, the *seam manifest specification*, the *BABEL benchmark*, and the *Coherence Cliff scaling family* — and verifies that each is a model of the doctrine in a precise sense: the operational quantities computed in each instance correspond to the abstract invariants of the framework, and the empirical results in each instance are consistent with the theorems of §§5–8.

10.1 Bulla: a witness-receipt protocol

Bulla (Komkov 2026, [bulla]) is a Python implementation of a witness-receipt protocol for agentic compositions, with the MCP tool-use surface as its first operational target. A *Bulla composition* consists of a finite list of tools, each with a manifest declaring its observable schema and convention dimensions. From this data, Bulla constructs a coboundary operator over the field $\mathbb{Q}$ and computes the *coherence fee* — the rank of the Kron-reduced witness Gram — and a *witness receipt* recording the disclosed dimensions and the policy disposition.

Mapping to the framework. A Bulla composition G_{Bulla} corresponds to a semantic interface complex $G \in \mathsf{SemComp}_{\text{ex}}$ as follows:

Bulla artifact	Framework object
Tool list with manifests	Convention presheaf $\mathcal{F}$ on the context poset P
Declared (observable) field schema	Observable subpresheaf $\mathcal{O}$
Latent (undeclared) convention dimensions	Latent presheaf $\mathcal{L} = \mathcal{F}/\mathcal{O}$
Composition seam (the boundary between two tools)	Cover overlap in the context poset
Coherence fee	Witness rank $r(G) = \dim H^1(G)$
Witness Gram	Cochain complex coboundary δ^0 in matrix form
Bridge / disclosure	Disclosure morphism (§2.3 III)
Witness receipt	Receipt in the sense of Theorem 7.1
Operational exact regime (DFD + CHP)	$\mathsf{SemComp}_{\text{ex}}$ (§3.5)
Surrogate regime	$\mathsf{SemComp}_{\text{sur}}$ (Conjecture 9.1)

Under this correspondence, the coherence fee of a Bulla composition equals the witness rank of the corresponding semantic interface complex; the bridge calculus equals the repair calculus; the witness receipt equals the receipt normal form.

The empirical calibration of Bulla on 703 real schema compositions — coherence fee = 0 has zero false negatives, coherence fee ≥ 4 predicts mismatches with $\sim 100\%$ confidence — is consistent with the framework: the witness rank is the obstruction invariant; below threshold the composition is coherent (no obstruction to detect); above threshold the obstruction is detectable by independent means.

10.2 Seam manifest specification: a typed-interface model

The seam protocol (Komkov 2026, [seam]) specifies what gets published at composition time: the *seam manifest*, a typed declaration of observable schema and convention dimensions. The manifest is signed and content-addressed; semantic witnesses are exchanged at composition; fraud proofs are constructed when local audits and global manifests disagree.

The seam manifest is the operational realization of the *interface specification* of a semantic interface complex: it specifies which conventions are observable and which are latent at each tool boundary. The publication of a seam manifest is the operational analog of the disclosure morphism (§2.3 III): both add structure to the observable subpresheaf.

Mapping. Each seam manifest field added to the observable schema corresponds to an independent disclosure W_i in the receipt of Theorem 7.1. The minimum seam-manifest publication required for coherence is $r(G_{\text{Bulla}})$ many fields, by Theorem 6.2. Operational metrics in the seam paper — the number of published declarations per fee point — are consistent with this duality.

10.3 BABEL: a benchmark for diagnostic correlation

BABEL (Komkov 2026, [babel]) is a 932-instance benchmark spanning seven families (calendar, invoicing, etc.) with deterministic ground truth (mean holonomy of a composition family). The benchmark measures correlation between diagnostics and ground-truth coherence: structural sheaf-cohomological diagnostics achieve Spearman $\rho = 0.99$; schema-validation diagnostics achieve $\rho = 0.17$; output-evaluation baselines fall in between.

Mapping. Each BABEL instance is a finite semantic interface complex with a designated "ground truth" coherence label (mean holonomy $= 0$ or > 0). Structural diagnostics compute the witness rank or a close proxy; their high correlation with ground truth is consistent with the framework: the witness rank is the obstruction invariant, and ground truth coherence corresponds to obstruction $= 0$.

Empirical content of Conjecture 9.4. BABEL provides an explicit family realizing the conjecture: locally-computable evaluation metrics (schema validation, GPT-4o judgment) achieve correlation ≤ 0.82, while the cohomological diagnostic achieves $\rho \geq 0.99$. The empirical separation is consistent with the conjecture's claim of asymptotic decoupling, though the conjecture's full formalization (§9.4) requires a precise definition of "locally computable."

10.4 Coherence Cliff: a scaling family

The Coherence Cliff scaling experiment (Komkov 2026, [coherence-cliff]) constructs 500 compositions across seven scales ($n = 5$ to $n = 50$ agents). At each scale, structural sheaf-cohomological diagnostics achieve $R^2 > 0.96$; non-sheaf baselines (Random Forest on topological features) degrade from $R^2 = 0.83$ at small scale to $R^2 = 0.52$ at large scale.

Mapping. Each scale corresponds to a family of semantic interface complexes with growing context-poset size and growing witness-rank distributions. The structural diagnostic (witness rank) tracks ground truth at all scales because it computes the unique obstruction invariant. Non-structural baselines degrade because they approximate the obstruction without computing it directly; as the obstruction structure becomes higher-dimensional with scale, the local approximations lose information that the cohomological diagnostic preserves.

This empirical pattern is consistent with Theorem 8.4: any baseline whose information capacity is sub-witness-rank cannot match witness-rank diagnostics on a scaling family.

10.5 Summary of the program — doctrine view

The existing program — papers, implementation, calibration, benchmark — is a coordinated investigation of one mathematical object: the witness-rank invariant on semantic interface complexes. Each component contributes a different projection:

Program component	Doctrine view
[scpi] paper	Formalization of the underlying descent / sheaf-cohomology framework; verification of structural lemmas in Lean.
[hierarchical-fee] paper	The additive decomposition (Lemma 3.5) and the witness-Gram presentation (Lemma 3.9).
[witness-geometry-beyond-fee] paper	The Kron reduction and leverage-conservation identity used in Definition 3.8 and §10.1.
[signed-incidence] paper	Field-independence of the witness rank between $\mathbb{Z}/2$ and $\mathbb{Q}$, used implicitly throughout.
[local-global-obstruction] paper	The impossibility of bounded local audits, equivalently the non-vanishing of H^1 on a constructed family.
[bulla] software	Computational realization: $r(G)$ as polynomial-time matrix rank over $\mathbb{Q}$; receipts as the proof normal form.
[seam] paper	Operational protocol: seam manifest as the typed interface declaration.
[coherence-cliff] paper	Empirical confirmation at scale: structural diagnostics dominate non-structural at all sizes.
[babel] benchmark	Empirical correlation: $\rho = 0.99$ for structural, $\rho = 0.17$ for baselines.
[bridge] paper	Empirical foundation: bilateral validity does not imply compositional coherence.

Each is a model of one or more aspects of the doctrine. None alone gives the full structure; together they describe a single mathematical object — the witness-rank invariant — from formal, computational, operational, and empirical sides.

10.6 What this section achieves

The doctrine is not abstract. Each of its four components — characterization theorem, repair duality, receipt normal form, communication lower bound — has a concrete realization in the existing program, and the empirical results of those realizations are consistent with the theorems.

The relationship is *not* symmetric, and we do not claim that the program "validates" the doctrine. The existing program supplies examples, computations, calibrated benchmarks, and empirical tests of an abstract framework whose mathematical core was implicit in the work but never named as the unique invariant. The doctrine paper supplies that invariant — the obstruction quantity those artifacts were converging toward — and the axiomatic argument that no other invariant respects the structural primitive. Whether this is the *right* primitive is a modeling question (§1 scope table, layers 1–3) that this paper does not settle and that machine verification cannot adjudicate.

References

[1] Hoare, C. A. R. (1969). An axiomatic basis for computer programming. *Communications of the ACM* 12(10), 576–580. doi:10.1145/363235.363259.

[2] Reynolds, J. C. (2002). Separation logic: A logic for shared mutable data structures. *Proceedings of the 17th IEEE Symposium on Logic in Computer Science (LICS '02)*, 55–74. doi:10.1109/LICS.2002.1029817.

[3] Yao, A. C.-C. (1979). Some complexity questions related to distributive computing (preliminary report). *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC '79)*, 209–213. doi:10.1145/800135.804414.

[4] Kushilevitz, E., and Nisan, N. (1997). *Communication Complexity.* Cambridge University Press.

[5] Lamport, L. (1978). Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM* 21(7), 558–565. doi:10.1145/359545.359563.

[6] Whitney, H. (1935). On the abstract properties of linear dependence. *American Journal of Mathematics* 57(3), 509–533. doi:10.2307/2371182.

[7] Tutte, W. T. (1965). Lectures on matroids. *Journal of Research of the National Bureau of Standards, Section B* 69B, 1–47.

[8] Oxley, J. G. (2011). *Matroid Theory*, 2nd edition. Oxford Graduate Texts in Mathematics 21, Oxford University Press.

[9] Mac Lane, S. (1998). *Categories for the Working Mathematician*, 2nd edition. Graduate Texts in Mathematics 5, Springer.

[10] Eilenberg, S., and Steenrod, N. (1952). *Foundations of Algebraic Topology.* Princeton Mathematical Series 15, Princeton University Press.

[11] Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. (2007). Stability of persistence diagrams. *Discrete & Computational Geometry* 37(1), 103–120. doi:10.1007/s00454-006-1276-5.

[12] Necula, G. C. (1997). Proof-carrying code. *Proceedings of the 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL '97)*, 106–119. doi:10.1145/263699.263712.

[13] Hansen, J., and Ghrist, R. (2019). Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology* 3(4), 315–358. doi:10.1007/s41468-019-00038-7.

[14] Curry, J. (2014). *Sheaves, Cosheaves and Applications.* PhD thesis, University of Pennsylvania. arXiv:1303.3255.

[15] Hatcher, A. (2002). *Algebraic Topology.* Cambridge University Press. Available at https://pi.math.cornell.edu/~hatcher/AT/ATpage.html.

[16] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27(3), 379–423; 27(4), 623–656. doi:10.1002/j.1538-7305.1948.tb01338.x.

[17] Cover, T. M., and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd edition. Wiley-Interscience. doi:10.1002/047174882X.

[18] Kurisummoottil Thomas, C., and Chen, M. (2026). Fundamental limits of quantum semantic communication via sheaf cohomology. arXiv:2601.10958v2.

Companion papers in the Res Agentica program are referenced inline in §10 and elsewhere by bracketed name (e.g., [scpi], [bridge], [seam], [sheaf], [bulla], [hierarchical-fee], [witness-geometry-beyond-fee], [signed-incidence], [local-global-obstruction], [coherence-cliff], [babel]); see papers/CANON.md for the canonical reading list and version status.