GCT-STL Serie Position Paper · Phase I Results

Three-Dimensional Entropic Domains for TSP Complexity via 3-SAT: Foundations of the GCT-STL Research Program on P versus NP

Renato R. González

Instituto Tecnológico de Santo Domingo (INTEC), Santo Domingo, Dominican Republic

ORIGINAL DRAFT: JUNE 2026 | SCIENTIFIC VERSION: v4.9.3 · 8 NOVEMBER 2026

hexagon:2611.01230v1

Abstract

In the context of the P versus NP problem, this paper establishes the mathematical foundations of the GCT-STL research program using TSP_{dec} as a canonical NP-complete representative and its polynomial reduction to 3-SAT with an exact materialized 3-CNF artifact. The framework is organized into four interacting systems: a Fundamental Mathematical Representation System, a Logical and Computational-Complexity Formal System, a Computational-Heuristic-Experimental and Simulation System, and a transversal Mathematical and Symbolic Formalization and Certification System, complemented by an AI-Assisted Neuro-Symbolic Discovery and Formalization Layer.

Within the Fundamental Mathematical Representation System, nine provenance-controlled descriptors D1–D9 are organized across three computational-information domains: semantic/geometric (S), temporal/computational (T), and logical/syntactic (L). They are assembled into the canonical raw Valiant–González tensorial signature $R0 := T^{raw}_{VG}$, supported by explicit SourceMap and Provenance Bundle contracts and by three foundational entropy results. Paper I-A situates R0 within the branched four-representation architecture $R0 \to R1 = L_{VG}$ and $R0 \to R2 = T^{H}_{VG} \to R3 = R_{VG}$, where R1 is a componentwise logarithmic scale complement, R2 the canonical entropic/information representation, and R3 the exact-envelope normalized representation. The Full STL Tensorial Coupling Principle remains an open and falsifiable hypothesis concerning a possible obstruction in the complete coupled S–T–L structure rather than in any individual domain.

The logical-complexity system defines $Cost_{STL}, P_{STL}, NP_{STL}$, and the minimal route $F1 : P \subseteq P_{STL}$; F2: $TSP_{dec} \notin P_{STL}$; $F3_{std} : TSP_{dec} \in NP$; and the conditional bridge $F4 : F1 \wedge F2 \wedge F3_{std} \Rightarrow P \neq NP$. F1 and F2 remain open; empirical robustness and candidate obstructions remain discovery-level, while universality, representation transfer, and lower-bound transfer remain theorem-level obligations. Accordingly, Paper I-A does not prove P $\neq$ NP; it establishes the representational, logical, and methodological foundations for the subsequent GCT-STL program, whose companion papers develop the full four-representation mathematics, the logical-computational and heuristic-experimental architectures, formalization and certification, and the later experimental and obstruction-theoretic stages.

Keywords: P versus NP · TSP_{dec}· 3-SAT · 3-CNF · computational complexity · GCT-STL · Valiant–González tensor · provenance · entropy · four-representation architecture · tensorial coupling · $P_{STL} \cdot NP_{STL}$ · universal robustness · formal verification · neuro-symbolic AI

Thematic Index

1. Introduction

2. Conceptual and Methodological Foundations

    2.0 The Six Fundamental Disciplines of GCT-STL

    2.1 Classical Complexity, NP-Completeness, and TSP_{dec}

    2.2 Shannon Information and Complexity

    2.3 González (1995): Genealogy of GCT-STL

    2.4 Statistical-Mechanics Antecedents

    2.5 Geometric and Information-Geometric Antecedents

    2.6 Methodological Inheritance from Geometric Complexity Theory

    2.7 Classical Barriers to P vs NP and Their Implications for GCT-STL

3. The GCT-STL Research Architecture

    3.0 Four Systems and Two Support Layers

    3.1 Primary and Secondary Classification, Antecedents, and Outputs

    3.2 A_{AI} — AI-Assisted Neuro-Symbolic Discovery and Formalization Layer

    3.3 $G_{V\&V}$ — Program-Wide Verification, Validation, Statement-Audit and Scientific Governance

    3.4 Inter-System Dependencies and Claim Boundaries

    3.5 Roadmap and Scope

4. System I: Fundamental Mathematical Representation System

    4.0 The Concept of the System I

    4.1 Decision TSP, Canonical Reduction, and Computational/Structural Provenance Sources

    4.2 Canonical Descriptors and Foundational Theorems

    4.3 R0: The Canonical Valiant–González Tensorial Signature and Provenance Architecture

    4.4 Full STL Tensorial Coupling Principle

    4.5 The Four-Representation Mathematical Architecture and Exploratory Diagnostics

5. System II: Logical and Computational-Complexity Formal System

    5.0 Scientific Role and Logical Authority

    5.1 $Cost_{STL}$ and the Computational-Representational Core

    5.2 Internal Coherence of P_{STL} and NP_{STL}

    5.3 F1-DTM and F1 — Universal Standard-to-STL Representation

    5.4 F2 — Universality, Obstruction and Transfer

    5.5 $F3_{std}$ — Independent Standard NP Membership

    5.6 F4 — Conditional Bridge to P $\neq$ NP

    5.7 Program-Wide V&V and Classical-Barrier Governance

5.8 F_{comp} — Metatheoretical Interface

6. System III: Computational-Heuristic-Experimental and Simulation System

6.0 Scientific Role, Authority, and Research Questions

6.1 Canonical Inputs and the Four Representation Views

6.2 Experimental Universes and Comparative Design

6.3 Statistical Physics, Sampling, Uncertainty, and Finite-Size Analysis

6.4 Exploratory Algebraic, Tensorial, Orbital, and Information-Geometric Diagnostics

6.5 Negative Controls, Ablations, Sensitivity, and Falsification

6.6 Evidence Grading, UR1/UR2, and the System-II Handoff

6.7 Reproducibility, Program-Wide $G_{V\&V}, A_{AI}$, and Publication Handoff

7. System IV: Mathematical and Symbolic Formalization and Certification System

7.0 Scientific Role and Metalogical Boundary

7.1 Formal Specifications, Statements, and Encoding Contracts

7.2 Formalization Targets and Interfaces by Scientific System

7.3 Lean 4 and the Dependency Architecture

7.4 Proof Scope, Axiom Footprint, and Adversarial Audit

7.5 Independent Reproduction and Controlled Release

7.6 Boundaries with $F_{comp}, G_{V\&V}, and\ A_{AI}$

8. Research Programme and Companion-Paper Architecture

8.0 Research Programme

8.1 Canonical Publication Crosswalk and Release Boundaries

9. Open Obligations and Discussion

9.0 Antecedents and Open Targets

9.1 Scope Boundary of the González Homologation

9.2 Methodological and Epistemological Position Relative to Classical GCT

9.3 F_{comp} and Axiomatic Scope

10. Conclusions

References

Appendix A: From González's (1995) Unified Software-Complexity Metric to the GCT-STL Tensorial Architecture

Appendix B: Proof of Theorem 1 — González (1995) Canonical Syntactic-Entropy Bound

Appendix C: Proof of Theorem 2 — Shannon–Boltzmann Semantic Entropy Bounds

Appendix D: Proof of Theorem 3 — Canonical and Operational Computational Entropy Bounds

1. Introduction

The P versus NP problem asks whether every decision problem whose positive instances can be verified in polynomial time can also be decided in polynomial time. Paper I-A approaches this question through the decision version of the Traveling Salesman Problem, TSP_{dec}, used here as a canonical classical NP-complete representative and linked to an exact logical artifact through the polynomial pipeline $TSP_{dec} -> 3SAT -> 3-CNF$. The detailed justification for this choice, including the distinction between TSP optimization and TSP_{dec} and the role of the standard NP verifier, is given in §2.1.3. The present paper does not infer hardness from factorial tour growth, geometric richness, logical expansion, or any isolated observable; its objective is to construct a mathematically controlled representation in which these structures can be distinguished, related, and subjected to explicit proof obligations.

The mature GCT-STL program is organized into four interacting systems whose separation is both technical and philosophical. System I, the Fundamental Mathematical Representation System, is primarily ontological-mathematical: it defines the objects, relations, descriptors, representations, and provenance structures admitted by the model. System II, the Logical and Computational-Complexity Formal System, is primarily logical, with a metalogical and foundational component: it defines the complexity-theoretic meaning of the program and the theorem obligations required for any separation claim. System III, the Computational-Heuristic-Experimental and Simulation System, is primarily methodological and epistemological: it governs controlled computation, simulation, robustness testing, falsification, and discovery of candidate invariants or obstructions. System IV, the Mathematical and Symbolic Formalization and Certification System, is primarily metalogical and epistemic: it formalizes explicit statements, tracks axioms and dependencies, and certifies derivations in a trusted proof environment. Across them operates an AI-Assisted Neuro-Symbolic Discovery and Formalization Layer, primarily gnoseological and methodological, which assists knowledge generation and formalization without becoming a source of theorem authority. The governing distinction is therefore that knowledge generation, empirical justification, mathematical proof, and formal certification are related but non-interchangeable functions.

Paper I-A begins materially inside System I. For a canonical instance I and input x=enc(I), a deterministic Turing machine M supplies the standard computational artifact, while Trace(M,x) records a concrete execution-provenance channel rather than serving as a universal source for the model. Together with the exact materialized 3-CNF artifact, theorem-origin results, and structural/coupling evidence, these sources populate a multi-source Provenance Bundle and descriptor-level SourceMap for nine canonical descriptors D1-D9. The descriptors are organized across three computational-information domains: S, the semantic/geometric domain; T, the temporal/computational domain; and L, the logical/syntactic domain, and across the three resolution rows Y1-Y3. Only after the provenance obligations are explicit are they assembled into the canonical raw Valiant-González tensorial signature $R0 := T^{raw}_{VG}$ in V_Y tensor $V_{STL} = R^{(3x3)}$. R0 is the central mathematical object of Paper I-A: a provenance-controlled heterogeneous representation, not a Turing machine, not a complexity class, and not by itself an obstruction or computational lower bound. Accordingly, Domain $\neq$ Row $\neq$ CausalInfluence $\neq$ ProvenanceSource remains a non-negotiable distinction throughout the paper.

GCT-STL also has an explicit intellectual genealogy, but genealogy is not mathematical provenance. González (1995) supplies the structural antecedent of the three-domain by three-magnitude architecture and, most directly, the decisional-information lineage underlying the bounded syntactic-entropy construction. Statistical mechanics and information geometry provide methodological precedents for studying ensembles, collective behavior, and global structure, while Geometric Complexity Theory contributes an obstruction-search philosophy based on structured representations and certifiable separation mechanisms. These antecedents motivate the framework but do not transfer their theorems automatically to TSP_{dec} , do not numerically determine D1-D9, and do not substitute for computational lower-bound proofs. Their detailed roles and limits are developed in §2 and Appendix A.

Within System I, §4.2 establishes three foundational entropy results at descriptor level. Theorem 1 gives

the bounded canonical syntactic entropy $D9^{CAN} = H_{sint}^{CAN}$, with the inherited two-bit asymptotic envelope; Theorem 2 defines the semantic Shannon–Boltzmann entropy $D7^{CAN}$ and its maximum-envelope growth of order Θ(N log N); and Theorem 3 separates the canonical computational entropy $D5^{CAN}$ of the complete Boolean assignment space from the procedure-dependent $D5^{OP}$ and gives the stated bounds for the specified polynomial bounded family. These results characterize marginal information scales and their provenance; they are not deterministic-time lower bounds and do not establish that S, T, or L is individually the source of NP-hardness. Their joint placement forms the canonical entropic row $Y3 = \left(D7^{CAN}, D5^{CAN}, D9^{CAN}\right)$ within, but not identical to, the full raw signature R0.

System I is consolidated under §4. Subsections §§4.3–4.5 complete the representational sequence after the descriptor basis is established: §4.3 develops the provenance architecture and canonical construction of R0; §4.4 formulates the Full STL Tensorial Coupling Principle as an OPEN and falsifiable structural hypothesis concerning the complete coupled S–T–L organization; and §4.5 is the designated integration layer for the branched four-representation architecture $R0 \to R1 = L_{VG}$ and $R0 \to R2 = T_{VG}^{H} \to R3 = R_{VG}$ together with exploratory diagnostics. Representation-level diagnostics may guide discovery, but determinant/permanent, Gram, cyclic-tensor, orbital, Fisher–Ruppeiner, or related observables do not become computational lower bounds merely by being defined on these representations.

The remaining systems assign distinct authority to theorem-level reasoning, empirical discovery, and certification. System II contains $Cost_{STL}, P_{STL}, NP_{STL}$, the minimal $F1/F2/F3_{std}/F4$ route, theorem-level UR3, the Obstruction Theorem, Representation Transfer, F2c, and the F_comp metatheoretical interface. System III contains simulation, MCMC, finite-size analysis, negative controls, sensitivity analysis, UR1/UR2 empirical robustness, and candidate invariants or candidate obstructions under System-II definitions and claim boundaries. Thus UR1 and UR2 are explicitly System-III empirical constructs; they do not become UR3 or a universal theorem by accumulation of finite evidence. The transverse System IV provides mathematical and symbolic formalization and certification, with Lean 4 as the principal current proof environment rather than as the definition or foundation of the system itself. The A_{AI} layer supports all four systems through pattern discovery, conjecture generation, symbolic analysis, code and proof assistance, and formal translation, but its outputs remain candidates until validated by the appropriate empirical, mathematical, or certification process.

Paper I-A is therefore a foundation paper rather than a complete separation theorem. In the minimal logical route, $F1 : P \subseteq P_{STL}$ and $F2 : TSP_{dec} \notin P_{STL}$ remain open; $F3_{std} : TSP_{dec} \in NP$ is the standard independent antecedent; and F4 is the conditional bridge $F1 \wedge F2 \wedge F3_{std} \Rightarrow P \neq NP$. Candidate obstructions and UR1/UR2 remain System-III discovery-level results, whereas UR3, an Obstruction Theorem, Representation Transfer, and the independent obstruction-to-$Cost_{STL}$ transfer F2c remain System-II theorem obligations. The Research Programme and Companion-Paper Architecture later allocates these obligations across the publication series, while the final Conclusions/Open Obligations/Discussion synthesis states the scope, limitations, established results, and open obligations. Accordingly, Paper I-A does not prove P $\neq$ NP; it establishes the ontological-representational, logical, methodological-epistemic, and formal-certification boundaries within which any later GCT-STL separation claim would have to be proved. The distinct System-IV certification and I-D/I-E roles are specified in §7 and the definitive companion-publication crosswalk in §8.

2. Conceptual and Methodological Foundations

2.0 The Six Fundamental Disciplines of GCT-STL

The conceptual basis of GCT-STL is organized through six complementary research strands: classical computational complexity and polynomial reductions; Shannon information theory; the multidomain software-complexity genealogy of González (1995); statistical-mechanics methods for combinatorial ensembles; geometric and information-geometric analysis; and the obstruction philosophy of Geometric Complexity Theory.

Section 2.7 separately identifies classical complexity barriers as constraints on any mature proof strategy. These antecedents motivate and constrain the programme but do not themselves establish F1, F2, a computational lower bound, or the Full STL Tensorial Coupling hypothesis.

2.1 Classical Complexity, NP-Completeness, and $\mathbf{TSP_{dec}}$

2.1.1 Classes P and NP

Throughout this document we adopt the formal definitions of Cook (2000) as presented in the official CMI problem statement, according to which the class P consists of all languages accepted by some deterministic Turing machine running in polynomial time.

$$T_M(n) \leq \ n^k + \ k$$

for some fixed k; the class NP consists of all languages L for which there exists a polynomial-time verification relation R such that

$$w \ \in \ L \iff \exists y \left(|y| \leq \ |w|^k \ \wedge \ \ R(w, \ y) \right)$$

and the problem is to determine whether P = NP (Cook, 2000).

For general textbook treatments of deterministic machines, polynomial-time languages and verification, see also Papadimitriou (1994) and Sipser (2012). These references supply background definitions, not proof of the particular GCT-STL compiler or its representation costs.

2.1.2 NP-Completeness and the Cook–Levin Theorem

A decision problem L is NP-complete when L $\in$ NP and every language L$'$ $\in$ NP is polynomial-time many-one reducible to L. NP-completeness therefore identifies representative problems for the class NP: a polynomial-time algorithm for any NP-complete problem would imply polynomial-time algorithms for all problems in NP. The foundational Cook–Levin theorem establishes Boolean satisfiability (SAT) as NP-complete, and the standard reduction SAT$\leq_{\mathrm{m}}{}^{\mathrm{p}}$3SAT gives 3SAT the same class-level status (Arora & Barak, 2009; Cook, 1971).

The importance of Cook–Levin for GCT-STL is representational as well as complexity-theoretic. It establishes the classical precedent by which polynomial-time computation and verification can be encoded into Boolean structure of polynomially controlled size. This motivates the use of a formally controlled logical artifact as one layer of the GCT-STL representation, while preserving the distinction between the computational process and its Boolean encoding.

NP-completeness is not itself a lower bound. In particular, the fact that TSP_{dec} and 3SAT are NP-complete does not imply $TSP_{dec} \notin P$ and does not prove P$\neq$NP. The GCT-STL argument must still establish its independent representation and obstruction obligations, later isolated as F1 and F2.

2.1.3 $\mathbf{TSP_{dec}}$ as the Canonical NP-Complete Representative

Decision TSP is used in Paper I-A as the canonical NP-complete witness problem because it exposes, within one decision framework, a metric/geometric instance structure, a Boolean reduction interface, and a deterministic computational realization. For a well-formed instance I=(N,D,K), the program uses the canonical pipeline $TSP_{dec}\rightarrow$3SAT$\rightarrow$3-CNF to materialize a logical artifact φ_I. The role of the reduction is not to re-prove hardness but to provide a polynomially controlled interface through which the semantic-geometric, computational-temporal, and logical-syntactic dimensions can be represented together.

Accordingly, the classical sequence governing the front end of the paper is P/NP → NP-completeness → Cook–Levin → $TSP_{dec} \rightarrow 3SAT/3 - CNF$. This sequence fixes the standard complexity-theoretic status of the witness problem while leaving the GCT-STL representation and lower-bound obligations separate (Cook, 1971; Karp, 1972; Garey & Johnson, 1979; Arora & Barak, 2009). Its logical-formal role reappears in §5.5, where $F3_{std}$ records the standard and independent fact $TSP_{dec} \in NP$, whereas F1 and F2 remain distinct theorem obligations internal to the GCT-STL route.

The canonical-witness choice is also complexity-theoretic. NP-complete means membership in NP together with NP-hardness under polynomial-time many-one reductions; hence TSP_{dec} is simultaneously an NP language with polynomially verifiable certificates and an NP-hard representative to which every language in NP reduces polynomially. TSP_{dec} is selected not because TSP optimization is irrelevant, but because the decision formulation retains the TSP tour/cost substrate and the geometric realizations available for representation while supplying the exact language-theoretic interface required by NP-completeness, Cook–Levin/3SAT, the standard verifier formalized in $F3_{std}$, and the F1–F4 separation route. TSP_{opt} remains algorithmically and scientifically relevant, but it is not required as the canonical witness for the minimal P versus NP argument.

Claim boundary. NP-completeness and NP-hardness do not establish $TSP_{dec} \notin P$. They establish class-wide reduction status. The lower-bound burden remains F2, together with the required representation and transfer obligations.

2.2 Shannon Information and Complexity

2.2.1 Shannon Information Theory

This theory constitutes one of the most profound conceptual pillars of modern science, providing a mathematical framework for quantifying uncertainty, surprise, and the capacity for information transmission in symbolic systems. In this context, Shannon entropy measures the average amount of information contained in a random source, understood as the degree of unpredictability of its emissions (Shannon & Weaver, 1949). Shannon defined the entropy of a discrete random variable distribution as follows: $p = (p_1, \ldots, p_n)$

$$\mathbf{H(p)} = -\sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{p_i} \log_{\mathbf{2}} \mathbf{p_i}.$$

In this expression, each term represents the probability of occurrence of symbol i, while the logarithmic term measures the "surprise" or information associated with that symbol; the negative sign ensures that entropy is non-negative. The weighted sum captures the average uncertainty of the source measured in bits as a decision measure (0 or 1): the more uniform the distribution, the greater the entropy, and the more concentrated the distribution, the lower the entropy. Its significance transcends telecommunications, where it was originally formulated, and it has become a fundamental tool for analyzing computer systems, algorithms, and decision structures (González, 1995).

2.2.2 Complexity and Information Content

In the present framework, entropy is used as an information-theoretic language for characterizing uncertainty, distributional organization, decision structure, and ensemble structure (Shannon & Weaver, 1949; Cover & Thomas, 2006). An entropy associated with a solution space, logical artifact, or computationally derived object can therefore serve as a canonical descriptor of how information is distributed within that object. This role is descriptive and mathematical: it does not identify entropy with deterministic running time, membership in a classical complexity class, or a computational lower bound (Arora & Barak, 2009).

GCT-STL accordingly treats entropy as one component of a broader representation architecture. Canonical entropic coordinates may reveal regularities, scaling envelopes, or couplings relevant to later complexity

analysis, but any inference from such structure to non-polynomial computational cost requires an additional theorem-level bridge.

information structure ≠ deterministic-time complexity ≠ computational lower bound

2.3 González (1995): Genealogy of GCT-STL

González (1995) provides a structural genealogy for the GCT-STL three-domain/three-magnitude organization. Its syntactic, functional and computational domains motivate the homologation $E_F \to S, E_C \to T$ and $E_S \to L$; this mapping is not an isomorphism and does not supply numerical values or provenance for contemporary D1–D9. Those definitions remain governed by the present SourceMap and Provenance Bundle.

The direct information-theoretic inheritance is the decisional level-entropy construction in Appendix 3 of González (1995): a complete binary decision structure yields a Shannon level-entropy bounded by 2 bits and converging to that limit with growing height. Theorem 1 reformulates this law for the syntactic envelope of the exact materialized 3-CNF artifact through $D9^{CAN}$. Theorems 2 and 3 are GCT-STL extensions for $D7^{CAN}$ and $D5^{CAN}$; they are not asserted to originate in the 1995 model.

The 1995 model also combines its three domain vectors through a trilinear form evaluated by the determinant of the corresponding 3×3 matrix. This joint-coupling genealogy motivates the OPEN Full STL Tensorial Coupling hypothesis, without transferring a theorem, a universal obstruction or a computational time lower bound. Tensor methods and entropy-based complexity measures predate 1995; the historical claim is therefore restricted to the particular integrated construction. Appendix A provides detailed genealogy, formulas, historical scope and limitations.

2.4 Statistical-Mechanics Antecedents

Statistical-mechanics studies of random K-SAT, including Monasson et al. (1999), provide a methodological precedent for analyzing ensembles of combinatorial instances through entropy, control parameters, and changes in solution-space organization. Within GCT-STL, this precedent informs System III, the Computational-Heuristic-Experimental and Simulation System: it motivates finite-ensemble experiments, scaling analysis, MCMC procedures, and the search for robust empirical patterns. No random-K-SAT threshold theorem is transferred to the canonical $TSP_{dec} \to 3SAT \to 3-CNF$ family, and no finite transition, by itself, is interpreted as a computational lower bound or as theorem-level universality.

2.5 Geometric and Information-Geometric Antecedents

Geometric work such as Perelman (2002) illustrates, in a different mathematical setting, how entropy-like functionals and global geometry can organize the analysis of complex spaces. GCT-STL retains only this methodological lesson; Ricci-flow results are not imported into computational complexity. Information geometry supplies a more direct mathematical vocabulary for statistical manifolds, Fisher information, and curvature-type observables (Amari, 2016; Ruppeiner, 1995). Within GCT-STL, representation-level geometric quantities may be defined as mathematical observables in System I, while their numerical or ensemble behavior is investigated in System III. Any promotion of such behavior to a universal obstruction statement belongs to System II and requires a separate theorem.

2.6 Methodological Inheritance from Geometric Complexity Theory

2.6.0 GCT Principles

Geometric Complexity Theory (GCT), initiated by Mulmuley and Sohoni, is a programme for proving algebraic-complexity lower bounds using tools from algebraic geometry, geometric invariant theory and representation theory (Mulmuley & Sohoni, 2001, 2008). Rather than attacking lower bounds only through explicit algorithmic counting, GCT represents computational objects as points in spaces on which algebraic

groups act, and studies whether one class variety can lie in or embed into the orbit closure associated with another.

A central setting compares determinant and permanent polynomial families in the algebraic-complexity landscape associated with VP and VNP. The determinant is efficiently computable, whereas the permanent is the paradigmatic harder family. The GCT strategy seeks explicit geometric or representation-theoretic obstructions that witness the non-containment of the relevant orbit closures or class varieties. In that setting, determinant/permanent objects and orbit closures are precise algebraic constructions, not generic numerical matrix diagnostics.

GCT-STL adopts this obstruction philosophy as methodological inspiration while retaining standard Turing computation and the $Cost_{STL}$-based classes defined later in this paper. Its finite 3×3 determinant/permanent diagnostics, orbital quantities, tensorial observables, and Fisher–Ruppeiner-type quantities therefore have different system roles: their mathematical definitions may belong to System I, their finite numerical investigation belongs to System III, and only a proved universal obstruction together with an explicit obstruction-to-$Cost_{STL}$ transfer can enter the System-II route to F2 (Mulmuley & Sohoni, 2001, 2008). Paper I-A asserts neither VP $\neq$ VNP nor an equivalence between classical GCT and the P versus NP problem addressed through F1–F4.

2.6.1 Why a Geometric–Tensorial Obstruction Strategy?

GCT-STL does not adopt geometric, tensorial and information-geometric representations merely because classical lower-bound methods face known barriers. The primary rationale is representational. The programme seeks a mathematical language capable of expressing global coupling, joint non-degeneracy, symmetry, orbit structure, collective response and candidate non-collapse properties of a provenance-controlled representation. These are structural questions that are difficult to reduce to isolated scalar growth or local combinatorial counts without an additional invariant or transfer principle.

Logic, combinatorics and classical complexity theory therefore remain indispensable. Combinatorics supplies the discrete substrate; Cook–Levin and polynomial reductions supply the class-theoretic interface; System II supplies the deductive lower-bound obligations; and System IV supplies formal certification. Geometry and tensorial structure are used as a representation-and-obstruction language, not as substitutes for those foundations. Information geometry adds a complementary ensemble-level language for Fisher information and curvature-type observables, whose empirical or numerical behaviour remains System-III evidence until a theorem establishes universal relevance.

The classical barriers enter as a second, independent consideration. A geometric or tensorial formulation is not automatically non-relativizing, non-natural or non-algebrizing. Consequently, GCT-STL chooses these representations because of their structural expressive power and then separately audits any mature System-II proof route against Relativization, Natural Proofs and Algebrization.

representational power $\rightarrow$ candidate obstruction language $\rightarrow$ theorem-level transfer $\rightarrow$ complexity conclusion

2.6.2 Algorithmic Bottlenecks of Classical GCT and the GCT-STL Design Choice

The methodological inheritance from Geometric Complexity Theory must be distinguished from a literal adoption of all of its representation-theoretic machinery. Classical GCT seeks explicit obstructions through algebraic-group actions, orbit closures and representation-theoretic multiplicities (Mulmuley & Sohoni, 2001, 2008). Subsequent complexity results show that some associated recognition tasks are themselves difficult: deciding positivity of a general Kronecker coefficient is NP-hard (Ikenmeyer, Mulmuley, & Walter, 2017), while general orbit-closure containment is also NP-hard (Bläser et al., 2021). These results do not invalidate GCT; they show that an obstruction programme cannot assume that recognition of every geometric or representation-theoretic witness is efficiently computable.

The distinction is important. Membership in moment polytopes is in NP ∩ coNP, including the asymptotic/stretched Kronecker-polytope setting (Bürgisser, Christandl, Mulmuley, & Walter, 2017). Thus, "Kronecker-related problem" is not a single complexity class, and the NP-hardness result for general Kronecker-coefficient positivity must not be generalized to every representation-theoretic task.

GCT-STL makes a different operational design choice. System I constructs a finite, provenance-controlled family R0–R3 and defines algebraic, tensorial, spectral, Gram and orbit-inspired observables directly on these representations. System III may evaluate such observables numerically to discover candidate invariants or candidate obstructions. Neither exact positivity of general Kronecker coefficients nor solution of the general orbit-closure containment problem is adopted as a primitive requirement of this representation-and-discovery pipeline. In this restricted operational sense, GCT-STL bypasses these classical algorithmic bottlenecks at the System-I/System-III stages rather than solving them.

This distinction does not remove the theorem-level obligation. If a future System-II Obstruction Theorem or Representation-Orbit Transfer requires exact high-dimensional orbit-closure non-membership, a hard multiplicity computation, or an equivalent decision problem, that complexity must be confronted explicitly. A tractable finite diagnostic is therefore not equivalent to an orbit-closure theorem, and bypass at the exploratory level is not a proof-level escape from the classical GCT bottleneck.

finite GCT-STL diagnostic ≠ classical orbit-closure obstruction theorem
operational bypass ≠ solution of Kronecker positivity / orbit-closure containment

2.7 Classical Barriers to P vs NP and Their Implications for GCT-STL

Any research program directed toward P versus NP must be evaluated against the principal methodological barriers known in complexity theory. Baker, Gill, and Solovay (1975) showed that relativizing techniques cannot by themselves resolve the problem, because there exist recursive oracles A and B for which $P^A = NP^A$ and $P^B \neq NP^B$.Razborov and Rudich (1997) later identified the Natural Proofs barrier: under standard pseudorandomness assumptions, a broad class of constructive and large properties useful for general circuit lower bounds cannot establish the required superpolynomial separation. Aaronson and Wigderson (2009) further showed that even many techniques combining relativization with arithmetization fall within the stronger algebrization barrier.

GCT-STL does not claim that tensorial coupling, orbit-inspired diagnostics, information geometry, or the four-representation architecture automatically escapes any of these barriers. Instead, they constrain the theorem-level route. Any eventual System-II proof of F2 must identify which decisive step is non-relativizing where necessary, must audit any lower-bound property against the Natural Proofs conditions when applicable, and must determine whether the argument algebrizes. The use of algebraic geometry is not itself a proof of non-algebrization, and the use of geometric or tensorial property is not itself evidence that the property is non-natural.

Accordingly, GCT-STL introduces a Classical-Barrier Audit as a proof-governance obligation attached to any mature F2 route. The audit does not constitute a new theorem and does not replace F1, F2, $F3_{std}$, Representation Transfer or F2c. Its role is to require an explicit account of the relation between the proposed proof mechanism and the known meta-barriers before a claimed P ≠ NP consequence can be accepted.

F2 candidate → Classical-Barrier Audit → validated System-II proof route

Within the Program-Wide $G_{V\&V}$ framework introduced in §3.3, using its System-I representation-specific VV-G1…VV-G5 profile, the Classical-Barrier Audit is attached to the theorem-level validation stage rather than to the existence of the representation itself. It therefore constrains VV-G5 and any mature F2 route: passing specification, provenance, numerical robustness or formal compilation gates cannot by itself establish that the decisive lower-bound argument escapes Relativization, Natural Proofs or Algebrization.

3. The GCT-STL Research Architecture

3.0 Four Systems and Two Support Layers

The mature GCT-STL programme is best understood not as a single linear pipeline but as a coordinated scientific architecture composed of four systems with different functions. Three systems are substantive scientific systems: a mathematical representation system, a logical and computational-complexity system, and a computational-heuristic-experimental system. A fourth system is transversal: mathematical and symbolic formalization and certification. A separate AI-assisted neuro-symbolic layer operates across all four systems, with its principal discovery role in System III and its principal formalization-assistance role in System IV. The programme is also governed by $G_{V\&V}$, the transversal Program-Wide Verification, Validation, Statement-Audit and Scientific Governance Framework.$G_{V\&V}$ covers the complete research and artifact lifecycle across the four systems, A_{AI} and the programme process; it is not a fifth scientific system and does not create theorem authority.

- **System I ↔ System III ↔ System II**
- **System IV: transversal formalization and certification across Systems I, II and III**
- **AI Layer: transversal discovery/formalization assistance; never a certification authority**
- **G_V&V Program-Wide: Verification, Validation, Statement-Audit and Scientific Governance Framework**

GCT-STL separates representation, theorem-level complexity reasoning, empirical discovery, and formal certification so that objects with different epistemic status are not conflated. This separation is the principal safeguard against converting a numerical pattern, a finite orbit signal, a simulation result, or an AI-generated proof sketch into a theorem without the required mathematical transfers. The categories below are operational rather than decorative: ontology concerns what mathematical entities and relations are admitted; epistemology concerns the warrant of claims; logic concerns deductive consequence; metalogic concerns formal systems, assumptions and certification; gnoseology concerns the generation and transformation of candidate knowledge; and methodology concerns concrete procedures of investigation (Bunge, 1977, 1983; Kleene, 1952; Popper, 1959).

3.1 Primary and Secondary Classification, Antecedents, and Outputs

The conceptual and methodological antecedents established in §2 enter this four-system architecture according to the scientific and epistemic role they perform, rather than through a one-to-one historical assignment.

Component	Primary category	Secondary categories	Why this classification is appropriate	Primary outputs	Antecedents: Conceptual and Methodological Foundations
System I — Fundamental Mathematical Representation System	Ontological	Logical; epistemological	Defines the mathematical ontology of GCT-STL: canonical $TSP_{dec}/3-CNF$ artifacts, DTM M, x=enc(I), Trace(M,x), D1-D9, provenance, R0-R3 and representation-level relations. Provenance provides a secondary epistemological account of descriptor origin.	Canonical artifacts; DTM/Trace interfaces; D1-D9; SourceMap/ Provenance Bundle; R0-R3; representation-level observables; Theorems 1-3.	González (1995): three-domain/three-magnitude genealogy, 3×3 organization and direct decisional-information lineage of D9. Shannon & Weaver (1949) and Cover & Thomas (2006): entropy/information language for theorem-origin coordinates and R2. Amari (2016) and Ruppeiner (1995): information-geometric observables. Mulmuley & Sohoni (2001, 2008): GCT-inspired orbit/obstruction vocabulary at representation level.

Component	Primary category	Secondary categories	Why this classification is appropriate	Primary outputs	Antecedents: Conceptual and Methodological Foundations
System II — Logical and Computational-Complexity Formal System	Logical	Metalogical; ontological; epistemological	Defines the complexity-theoretic objects, theorem obligations and deductive dependencies required for the public target: $Cost_{STL}$, P_{STL}/NP_{STL}, $F1/F2/F3_{std}/F4$, UR3, obstruction/transfer and F2c. F_{comp} is a metalogical/foundational subarchitecture.	$Cost_{STL}$; P_{STL}/NP_{STL}; $F1/F2/F3_{std}/F4$; UR3; Obstruction Theorem; Representation Transfer; F2c; F_{comp} interface.	Cook (1971, 2000), Karp (1972), Garey & Johnson (1979), Arora & Barak (2009): P/NP, reductions, NP-completeness and TSP_{dec}reference frame. Mulmuley & Sohoni (2001, 2008): obstruction/orbit-closure paradigm for theorem-level routes. Kleene (1952): logical/metalogical discipline.
System III — Computational-Heuristic-Experimental and Simulation System	Methodological	Epistemological; gnoseological	Provides controlled procedures for simulation, MCMC, finite-size analysis, ablations, negative controls, multi-universe testing and discovery of robust patterns or candidate invariants/obstructions. Its outputs have empirical rather than theorem-level warrant.	Simulations; MCMC; finite-size scaling; ablations; negative controls; UR1/UR2; candidate invariants; candidate obstructions.	Monasson et al. (1999): statistical-mechanics/phase-transition methodology as experimental precedent. Amari (2016) and Ruppeiner (1995): Fisher/information-geometric tools. GCT-inspired observables from System I supply candidate signals for testing. Popper (1959): falsifiability, controls and refutation discipline.
System IV — Mathematical and Symbolic Formalization and Certification System	Metalogical	Logical; epistemological; methodological	Represents stable definitions and theorems in a formal language, records axioms/dependencies and certifies proof objects with a trusted kernel. It checks formal validity rather than supplying missing scientific premises.	Formal definitions; mechanized lemmas/theorems; axiom/dependency audits; certified proof artifacts; reproducible proof inventories.	Kleene (1952): logic/metamathematics. System IV formalizes and certifies the GCT-STL instantiations of the mathematical content inherited from González, information theory, GCT, information geometry and standard complexity theory; it does not create their scientific meaning.
A_{AI} — **AI-Assisted Neuro-Symbolic Discovery and Formalization Layer**	Gnoseological	Methodological; epistemological	Assists human-machine generation and transformation of candidate knowledge: pattern recognition, conjecture formation, symbolic synthesis, proof search, code generation and formalization. It is not a fifth scientific system or a certification authority.	Candidate patterns/invariants; symbolic conjectures; proof sketches; code/simulation candidates; formal statements; proof-search suggestions.	No single §2 mathematical antecedent. The layer operates transversally over antecedent-derived objects in Systems I-III. Its status is governed by the gnoseological/epistemological distinction: AI may generate candidates, but validation belongs to the receiving scientific system and formal certification to System IV.
$G_{V\&V}$ **Program-Wide: Verification, Validation, Statement-Audit and Scientific Governance Framework**	Methodological / epistemological governance	Logical; metalogical; gnoseological; process-quality governance	Controls artifact identity, scientific horizon, evidence synthesis, conceptual integrity, statement adequacy, construction, semantic validity, inferential transfer, formal correspondence, independent review and release across Systems I–IV, A_{AI} and Process. It governs claim promotion but does not itself prove open theorems.	SMS/SLR registries; CM dependency graph; Statement Audits; Artifact/Claim Registry; PVV-0. . . PVV-9 evidence; VTM/MDR/AI-LOG; NC/CAPA; independent-review reports; release manifests and status/impact registers.	IEEE Std 1012-2024 (IEEE, 2025); Oberkampf & Roy (2010); Petersen et al. (2015) for SMS; Kitchenham & Charters (2007) and Page et al. (2021, PRISMA 2020) for SLR; Novak & Cañas (2008) for CM; National Academies of Sciences, Engineering, and Medicine (2019), Tabassi (2023; NIST AI RMF 1.0) and W3C PROV-O (2013) as methodological/provenance precedents. Current authority: Program-Wide $G_{V\&V}$ Framework v1.1, PI APPROVED/FROZEN.

Interpretive note. The antecedents are not assigned one-to-one to systems. The same mathematical tradition may define an object in System I, motivate an experiment in System III, and support a theorem strategy in System II. What changes is the epistemic status of the use. In particular, genealogy $\neq$ provenance; entropy $\neq$ time lower bound; information-geometric curvature $\neq$ hardness; and a GCT-inspired proxy $\neq$ a classical GCT obstruction.

3.2 $\mathbf{A_{AI}}$: AI-Assisted Neuro-Symbolic Discovery and Formalization Layer

A_AI is a transversal neuro-symbolic assistance layer, not System V. It supports the generation, transformation and adversarial testing of candidate knowledge: symbolic hypotheses and counterexamples in Systems I–II; computational experiments, code and candidate invariants in System III; and formal statement construction, proof search, proof repair and translation in System IV. Its main scientific cycle is numerical/symbolic discovery → conjecture → BREAK(T)/PROVE(T) → auditable formal statement → human review and, when applicable, kernel-checked proof. AI-generated outputs remain candidate artifacts until the responsible scientific system validates them. A_AI may participate in critical N4 tasks, but cannot approve or certify its own results; AI-LOG, independent audit and release controls are governed by G_V&V.

3.3 $\mathbf{G_{V\&V}}$ Program-Wide: Verification, Validation, Statement-Audit and Scientific Governance

$G_{V\&V}$ is the programme-wide governance layer governing the full life cycle of GCT-STL claims, knowledge, mathematical/logical/metamathematical artifacts, code, experiments, formal statements, AI contributions and scientific releases. Verification asks whether an artifact conforms to its frozen specification and recorded assumptions; validation asks whether it adequately supports its intended scientific interpretation or inference. Statement Audit applies before proof search and again after formal encoding to ensure that the formal statement corresponds to the intended mathematical claim. These obligations are not delegated to System IV alone: machine-checked correctness is not synonymous with semantic validity, representation-to-complexity transfer, or scientific release approval. G_V&V is transversal to Systems I–IV, A_AI and Process and is not a fifth scientific system.

The Program-Wide lifecycle is PVV-0 Scientific Horizon Mapping (SMS); PVV-1 focused evidence synthesis (SLR); PVV-2 versioned Concept Mapping; PVV-3 Statement Audit; PVV-4 Artifact Verification; PVV-5 Semantic/Scientific Validation; PVV-6 Transfer/Inferential Validation; PVV-7 Formal Certification; PVV-8 Independent Audit; and PVV-9 Release & Evolution Control. These are risk-based governance stages, not theorem antecedents and not compulsory linear execution order.

SMS, SLR and CM remain distinct: SMS periodically maps the relevant literature and generates scientific-change triggers; SLR critically synthesizes evidence for focused questions; CM keeps a typed graph of concepts, statuses and dependencies. New external results trigger an Impact Review (no reopen / scoped reopen / full reopen), not automatic invalidation. Stable governance rules, mutable operational status and release evidence are kept separate.

System-I representation V&V specialization (VV-G1…VV-G5). The five gates below remain the local representational profile inherited from T-3C. They are nested within the wider PVV lifecycle and are not a replacement for the programme-wide framework.

Gate	Question	Primary system responsibility	Status rule
VV-G1 — Specification	Is $R0/T_{VG}$ well defined on the declared domain?	I; IV certifies formalized statements	Must be proved/audited; definition alone is not a pass.
VV-G2 — Provenance	Can each D_i be reconstructed from its frozen source contract?	I; IV certification support	SourceMap existence is necessary but does not by itself close the gate.
VV-G3 — Polynomial construction	Is the end-to-end representation map polynomial in encoded input size?	I/II; IV formalization support	OPEN until a uniform polynomial bound is proved.
VV-G4 — Semantic fidelity	Does the representation preserve the computationally relevant semantics needed by the intended theorem?	II, with I interface	OPEN theorem-level validation obligation.
VV-G5 — Transfer	Does a proved representation obstruction imply the required $Cost_{STL}$ non-membership/lower bound?	II; IV certification	OPEN; contains Representation Transfer/F2c and Classical-Barrier Audit obligations.

Status discipline. Paper I-A does not call T_{VG} a fully validated computational representation. At the current

stage it is a canonical, provenance-controlled representation with several verification ingredients defined, while the polynomial-universality, semantic-fidelity and transfer gates remain open.

3.4 Inter-System Dependencies and Claim Boundaries

System I supplies canonical representation and observables. System III investigates those objects empirically under System-II targets and claim constraints. Robust finite evidence such as UR1/UR2 or a candidate obstruction remains in System III; theorem-level universality, an Obstruction Theorem, Representation Transfer, F2c and F2 belong to System II. System IV acts transversally on formalizable content, while A_AI assists across all systems without inheriting theorem authority. Any mature F2 proof route is additionally subject to the Classical-Barrier Audit defined in §2.7; this is a proof-governance obligation and does not itself establish F2 or alter the OPEN status of UR3, the Obstruction Theorem, Representation Transfer or F2c. G_V&V separately controls evidence/claim status, source and statement correspondence, cross-system transfer and release; a V&V sign-off may prevent an invalid promotion but can never replace proof of an OPEN System-II obligation.

The logical architecture is therefore branched rather than linear: the F1 representation branch and the independent standard branch $F3_{std}$ do not depend on the experimental obstruction route; the F2 branch may consume System-III discoveries only after they are reformulated and proved in System II. Only $F1 \wedge F2 \wedge F3_{std}$ yields the conditional bridge F4: P $\neq$ NP.

3.5 Roadmap and Scope

The remainder of Paper I-A develops this architecture without repeating the global explanations of A_{AI} and $G_{V\&V}$ supplied in §§3.2–3.3. Section 4 consolidates System I through the canonical artifacts/provenance sources (§4.1), descriptor theorems (§4.2), R0 and provenance (§4.3), OPEN Full Coupling (§4.4), and the R0–R3 representation family (§4.5). Section 5 consolidates System II and its theorem-level gates; Section 6 presents the integrated System-III discovery architecture and its empirical claim boundaries. Section 7 describes the transverse System-IV formalization and certification architecture; Section 8 provides the canonical companion-paper publication crosswalk; Sections 9–10 synthesize open obligations and conclusions. Section location is editorial; system membership and theorem authority follow the frozen scientific architecture.

4. System I: Fundamental Mathematical Representation System

4.0 The Concept of System I

System I — Fundamental Mathematical Representation System. This section consolidates the complete mathematical representation layer of GCT-STL. It begins from the canonical TSP_{dec} instance and its exact logical/reduction artifacts; establishes the heterogeneous computational, structural, artifact and theorem-origin provenance required for D1–D9; assembles the canonical raw representation $R_0 = T_{VG}^{raw}$ together with its SourceMap and Provenance Bundle; states Full STL Tensorial Coupling as an OPEN structural principle; and then integrates the broader R0–R3 representation family and exploratory diagnostics. The concentration of these components in one section is editorial rather than reductive: artifact construction, descriptor theorem status, representation assembly, structural hypotheses and exploratory observables remain epistemically distinct. System I defines what mathematical objects and relations constitute the GCT-STL representation; System II later determines the complexity-theoretic consequences that may validly be inferred from them.

4.1 Decision TSP, Canonical Reduction, and Computational/Structural Provenance Sources

4.1.0 Reduction Target: 3-SAT and the Materialized 3-CNF Artifact

The starting point is the classical reduction:

$$TSP_{dec} \leq_m^p \ 3SAT$$

Here N denotes the number of cities. The reduction uses a Boolean encoding whose positional core contains ν_pos=N^2 city-position variables, together with edge and auxiliary variables introduced by the materialized construction; M denotes the resulting number of clauses.

Notation convention. From this point onward, the primary asymptotic input-size variable is n=|x|, where x=enc(I) is the encoded decision instance. N denotes only the number of TSP cities; ν denotes the number of Boolean variables in the materialized 3-CNF (with N^2 positional variables as a distinguished subset); and M denotes the number of resulting clauses. Standard-complexity claims never substitute N, ν or M for n without an explicit encoding bridge.

The polynomial reduction targets the decision problem 3-SAT. Its concrete output for an instance I is a Boolean formula φ_I in 3-CNF, which supplies the logical artifact analyzed by the syntactic and compiler-derived coordinates. The distance/cost matrix D supplies metric-semantic structure, while the deterministic machine acting on the encoded instance supplies the execution channel.

3-SAT and 3-CNF must be distinguished. 3-CNF is a syntactic normal form: a Boolean formula written as a conjunction of clauses of width at most three under the compiler convention used here. 3-SAT is the decision problem asking whether a 3-CNF formula is satisfiable. Accordingly, the complexity reduction is written $TSP_{dec} \leq_m^p 3SAT$, whereas compileDecision3SAT materializes the concrete canonical 3-CNF artifact φ_I. D8 and D9, and the variable count entering D_5^{CAN}, are properties of that exact artifact rather than of the abstract decision problem "3-SAT."

Prospective three-domain entropic tensor (research direction; not a constructed object). The expression S⊗T⊗L is retained here as schematic notation for a possible third-order tensorial coupling of the semantic (S), temporal (T), and logical (L) domains. Such a future construction would use the fully entropic 3×3 representation $R_2 = T_{VG}^H$, defined in §4.5.3 and analyzed in Step 12, as its starting mathematical object. In the present architecture, however, R2 is a second-order tensor in $V_Y \otimes V_{STL}$ and is not itself S⊗T⊗L. A genuine three-domain construction would require precisely specified factor spaces V_S, V_T, and V_L, a defined provenance-compatible multilinear lift from R2, and separate examination of its invariants and transfer properties. The existing cyclic 3×3×3 diagnostic (§4.5.5) does not by itself establish this proposed semantic identification, a universal obstruction, or a computational lower bound.

4.1.1 Decision TSP and the Canonical 3-CNF Encoding

The NP-complete version of TSP is its decision version, defined as follows:

Given an instance $I_N = \langle N, D, K \rangle$, where:

- N is the number of cities;
- $D \in \mathbb{N}^{N \times N}$ is the distance matrix;
- K is a cost bound;

the problem asks if there exists a Hamiltonian tour π such that:

$$C_D(\pi) = \sum_{k=1}^{N} D_{\pi(k),\pi(k+1)} \leq K$$

with the final edge closing the Hamiltonian cycle over the N cities.$\pi(N+1) = \pi(1)$

The reduction to 3-SAT produces a canonical 3-CNF formula φ_I such that satisfiability of φ_I is equivalent to the corresponding TSP_{dec} instance being a yes-instance:

$$\langle N, D, K \rangle \in TSP_{dec} \iff \Phi_{TSP}^{3CNF} \in 3SAT$$

In the GCT-STL model, this formula decomposes into three substructures:

$$\Phi_{TSP}^{3CNF} = \Phi_{HAM}(X) \wedge \Phi_{EDGE}(X, Y) \wedge \Phi_{DIST}(Y, Z; D, K)$$

4.1.1.1 Hamiltonian component: $\boldsymbol{\Phi_{HAM}}$ The formula Φ_{HAM} encodes the syntactic structure of the tour. It uses variables where $X_{i,k} = 1 \Leftrightarrow city$ i occupies position k.

This component imposes:

- Each city appears exactly once.
- Each position contains exactly one city.
- The assignment encodes a valid permutation of the cities.

Therefore, this component is classified primarily in the logical/syntactic domain L (a domain annotation, not exclusive descriptor provenance):$\Phi_{HAM} \mapsto L$

4.1.1.2 Edge component: $\boldsymbol{\Phi_{EDGE}}$

Cyclic-index convention. Positional successor k+1 is interpreted modulo N: in particular, X_{j,N+1} denotes X_{j,1}. This closes the Hamiltonian cycle consistently with the tour-cost expression; it does not introduce an additional variable into the canonical 3-CNF. The formula Φ_{EDGE} connects the positional encoding with the structure of the edges used in the tour. It introduces variables, where $Y_{i,j,k} = 1 \iff$ at step k the tour travels from i to j.

The central relation is $Y_{i,j,k} \leftrightarrow (X_{i,k} \wedge X_{j,k+1})$. *Then* Φ_{EDGE} transforms a positional structure into an edge structure. Therefore:

$$\Phi_{EDGE} : L \longleftrightarrow S$$

In the GCT-STL model, Φ_{EDGE} plays the role of a structural L–S interface in the representation. Positional variables belong to the logical encoding of the tour, while edge variables express the corresponding edge structure used by the metric semantics. The equivalence therefore couples a local Boolean representation to an edge-level realization. This makes Φ_{EDGE} scientifically important as a coupling operator, but not as an established causal location of hardness or an obstruction theorem. Any non-additive effect associated with this interface must be evaluated only inside the complete provenance-controlled S–T–L signature represented by R0.

4.1.1.3 Metric component: Φ_{DIST} The formula Φ_{DIST} encodes the cost constraint:

$$\sum_{k=1}^{N} \sum_{i=1}^{N} \sum_{j=1}^{N} D_{ij} Y_{i,j,k} \leq K$$

Since this constraint depends on the distance matrix D, it belongs to the semantic domain $\Phi_{DIST} \mapsto S$. The constraint can be encoded by Boolean addition and comparison circuits and then transformed into 3-CNF using auxiliary variables Z.

Therefore:

$$\Phi_{TSP}^{3CNF} = \Phi_{HAM} \wedge \Phi_{EDGE} \wedge \Phi_{DIST}$$

Reading of the three terms.Φ_{HAM} represents the syntactic/logical tour organization;Φ_{EDGE} is the logical–semantic coupling interface; and Φ_{DIST} represents the metric/semantic cost constraint. These are interpretive component labels, not an additional theorem or new SourceMap assignment.

4.1.2 Deterministic Turing-Machine Execution and Machine-Derived Provenance

This section formalizes the deterministic-machine provenance layer of the GCT-STL model. The deterministic Turing machine M whose execution trace supplies the machine-provenance channel is not, by definition, the provenance source of the canonical 3-CNF artifact φ_I. That artifact is produced by the canonical reduction/compiler pipeline described in §4.1.1. Likewise, M does not operate on the semantic objects N, D, and K as unencoded mathematical entities; it acts on a finite binary encoding of the decision instance. This separation between the mathematical decision instance, its canonical logical artifact, and deterministic execution is essential for the provenance-controlled assembly of D1–D9 and for the later tensorial construction of T_{VG}^{raw}.

4.1.2.0 Encoded Input and Primary Size Variable Let I=(N,D,K) be a well-formed instance of TSP_{dec}, where N is the number of cities, D is the distance or cost matrix, and K is the decision threshold. The deterministic Turing machine receives a canonical finite binary encoding x=enc(I), with n:=|x|. The variable n is the primary asymptotic input-size parameter for all standard complexity-theoretic statements in Paper I-A. The parameters N, ν(I)=|Vars(φ_I)|, and M(I)=|Clauses(φ_I)| retain distinct model meanings and may be related to n only through explicit encoding or compiler bounds.

4.1.2.1 Deterministic Machine, Configurations, Run, and Trace Let M be a standard deterministic Turing machine with finite control, finite input and tape alphabets, distinguished start/accept/reject states, and a deterministic transition function. A configuration C_k records the instantaneous machine state, including the current control state, tape contents, and head position or positions. Because the transition relation is functional, each non-halting configuration has at most one successor.

For a terminating computation on x, the unique execution path is $C_0 \to C_1 \to \cdots \to C_t$, and the execution trace is $Trace(M, x) := (C_0, C_1, \ldots, C_t)$. The trace is the primary evidence source for the machine-derived provenance channel, but it is not the universal source of the complete STL representation. In particular, Trace(M,x) is distinct from the canonical logical artifact φ_I and is not sufficient by itself to reconstruct T_{VG}^{raw}.

4.1.2.2 Machine-Derived and Machine/Artifact Provenance: D1–D4 The first four descriptors are introduced here at the level of their provenance and interpretation contracts so that the SourceMap formalized in §§4.3.4–4.3.5 can be read before the complete 3×3 Valiant–González signature is assembled. Their source structure is not uniform: D1, D2, and D4 are primarily trace/control derived, whereas D3

combines machine/compiler evidence with the exact materialized 3-CNF artifact. Consequently, Trace(M,x) is a primary source for this channel but is not, by itself, sufficient to reconstruct every one of D1–D4.

D1 — State Density. A logical-state or configuration-density coordinate extracted from machine-state or configuration evidence along the deterministic run Trace(M,x). It characterizes realized state structure rather than a property inferred from row or column placement.

D2 — Tape Depth. A tape/memory-depth or access coordinate tied to the concrete representation and deterministic execution of the encoded instance.

D3 — AST Bandwidth. A structural connectivity coordinate associated with the abstract-syntax structure of the exact materialized 3-CNF artifact φ_I. Its provenance combines compiler/artifact structure with machine-level processing evidence. It is neither a pure tape statistic nor an AST property of the raw binary encoding x.

D4 — Deterministic Control/Search Structure. A coordinate describing the control or search structure realized by the deterministic computation. It belongs to deterministic execution provenance and must not be interpreted as nondeterministic Turing-machine branching.

These definitions fix the provenance and interpretation contracts of D1–D4 without imposing premature numerical extraction formulas. Trace(M,x) is therefore a primary source of machine-derived provenance, but it is not the source of D5–D9 and is not sufficient by itself to construct the complete Valiant–González signature. The multi-source Provenance Bundle, SourceMap, and formal construction of T_{VG}^{raw} are introduced in §§4.3.4–4.3.6.

4.1.3 Structural and Artifact Provenance: D6 and D8

The machine-derived provenance channel introduced in §4.1.2 accounts for D1–D4. Two additional descriptors, D6 and D8, originate instead from the structural/artifact channel of the canonical TSP_{dec}→3SAT→3-CNF construction. Although both belong to the same provenance family, they represent different kinds of information. D8 records the exact size of the materialized logical artifact, whereas D6 characterizes the structural coupling between the metric-semantic and logical-syntactic representations. Neither descriptor is derived from the deterministic execution trace.

4.1.3.1 D8: Exact Materialized 3-CNF Artifact Let φ_I=compileDecision3SAT(I).formula denote the exact canonical 3-CNF artifact produced from a well-formed decision instance I=(N,D,K). The canonical structural descriptor is $D_8^{CAN}(I) := |Clauses\,(\varphi_I)|$, the exact number of clauses materialized by the canonical compiler. Accordingly, φ_I→D_8^{CAN}(I). D_8^{CAN} is an artifact-derived structural descriptor; it is not a running-time quantity, a machine-trace statistic, or a Shannon entropy.

The same artifact supplies the instance-specific structural input required for the canonical syntactic entropy D_9^{CAN}, giving the dependency $\varphi_I \to D_8^{CAN} \to D_9^{CAN}$. This dependency does not make D8 an entropy coordinate.

4.1.3.2 D6: Canonical Semantic–Logical Coupling Coordinate D6 occupies a different structural role. It represents the interface between the semantic geometry of the TSP instance and the logical structure required to encode feasible Hamiltonian solutions. In the canonical 3-CNF construction $\varphi_I = \Phi_{HAM} \wedge \Phi_{EDGE} \wedge \Phi_{DIST}$, Φ_{EDGE} provides the most direct representation interface between positional logical organization and metric edge realization. The selected edges must also satisfy global Hamiltonian constraints and the distance/cost conditions represented through Φ_{DIST}. Thus Φ_{EDGE} is a primary provenance anchor for the S↔L interface, but Φ_{EDGE}≠D6 and does not numerically determine D6 by itself.

A manuscript-safe dependency is therefore $(\Phi_{HAM}, \Phi_{EDGE}, \Phi_{DIST}, D) \to \mathcal{C}_{SL}(I) \to D_6^{CAN}(I)$, where $\mathcal{C}_{SL}(I)$ denotes the canonical semantic–logical coupling structure associated with the instance. At the foundation

level, D_6^{CAN} is the canonical S↔L structural coordinate recorded by the SourceMap; its operational numerical realization is kept epistemically distinct.

4.1.3.3 Operational D6 and the Collision Functional For experimental phases, the semantic–logical friction is operationalized by a collision functional that combines three normalized effects: input/output projection friction, subtour fragmentation, and the relative cost of patching subtours into a Hamiltonian cycle. With preregistered equal weights, $\rho_{coll}(D) = \frac{\rho_{IO}(D)+\rho_{subtour}(D)+\rho_{patch}(D)}{3}$ and $D_6^{OP}(D) = 1 + \lambda\rho_{coll}(D)$. The equal weighting and λ convention are operational choices, not theorems.

The three epistemic levels must remain separate:D_6^{CAN}(I), the canonical structural coordinate; D_6^{OP}(D), the per-instance operational realization and its finite empirical calibration; and κ_SL*, a possible universe-independent universal coupling constant that would require a future theorem. Empirical robustness does not establish κ_{SL}*.

4.1.3.4 Robustness and Non-Circularity Boundary The operational calibration of D6 is deliberately upstream of D7, critical-temperature or MCMC quantities, determinant/permanent diagnostics, orbital distance, obstruction rates, and other tensor-derived observables. This prevents outcome-tuning of the coupling coordinate. Stability observed across Euclidean, i.i.d., and Cluster universes constitutes finite robustness evidence; universality requires a mathematically defined STL-admissible class and a uniform concentration/common-limit theorem.

4.1.3.5 D6 and D8 Share a Provenance Channel but Not a Meaning The structural/artifact channel may be summarized schematically as $\mathcal{P}_{STRUCT}(I) \to D6, D8$. Yet D_8^{CAN} is an artifact-magnitude coordinate, whereas D_6^{CAN} is a semantic–logical structural-coupling coordinate. Their common provenance family does not make them mathematically interchangeable: D6 is not defined as a function of D8, and D8 is not defined as a function of D6.

4.1.3.6 Structural Provenance Contract At manuscript level, the structural provenance bundle may therefore record the exact artifact φ_I, its Hamiltonian/edge/distance components, the distance matrix D, and the semantic–logical coupling evidence required by the current SourceMap. The resulting mappings are $D_8^{CAN} \leftarrow \varphi_I$ and $D_6^{CAN} \leftarrow \mathcal{C}_{SL}(I)$, while any collision-functional calibration remains a downstream operational realization rather than a redefinition of canonical provenance.

4.2 Canonical Descriptors and Foundational Theorems

4.2.1 Syntactic Entropy and Theorem 1

4.2.1.1 The Concept of Syntactic Entropy and Its 1995 Lineage Syntactic entropy H_sint measures the information content associated with the vertical organization of a binary logical structure: not the number of symbols themselves, but how the logical nodes are distributed across decision levels. This distinction is central to the present formulation. The size and depth of a logical artifact may increase, while the Shannon entropy of its level distribution remains uniformly bounded.

The construction is inherited directly from González (1995). In the syntactical complexity domain of that paper, the decisional level of a software chunk was represented by a binary tree associated with a Boolean function. Appendix 3, 'Decisional Information Content', defined the numbers of decisions in the different tree levels as entropy classes, treated as random variables, and applied Shannon entropy to their relative frequencies. The resulting decisional information content was shown to be an asymptotic function upper bounded by 2 bits, with limit 2 as the height of the complete binary tree tends to infinity.

GCT-STL preserves this mathematical law but changes the represented object. Instead of a Boolean predicate belonging to a software chunk, the canonical object is the exact materialized 3-CNF artifact φ_I produced by

compileDecision3SAT. Thus the 2-bit bound is not introduced here as an ad hoc property of TSP→3-SAT; it is the canonical continuation of the level-entropy principle established in the 1995 model.

The key idea can be stated compactly: the number of nodes and the height of a complete binary logical structure may grow without bound, yet the Shannon entropy of the node distribution by levels remains below 2 bits and converges to 2. Consequently, growth of syntactic size does not entail unbounded growth of level-based syntactic information content.

4.2.1.2 Canonical Construction from the Exact Materialized 3-CNF For a well-formed TSP_{dec} instance I=(N,D,K), define the exact canonical logical artifact

$$\varphi_I := compileDecision3SAT(I).formula$$

and let its exact materialized structural size be

$$M(I) := |Clauses(\varphi_I)| = D_8^{CAN}(I)$$

Counting convention (C03-07). Clauses(φ_I) is the sequence of clause occurrences in the exact compiler-emitted 3-CNF artifact; D8 counts those emitted occurrences, without any further deduplication in the counting operation. Whether the frozen compiler itself removes duplicates during construction remains an implementation obligation. The existing requirement M(I)≥1 is unchanged.

The formula φ_I is not changed. In particular, no tautological clauses are appended and no dummy literals are introduced. From the exact D8 artifact we construct only an auxiliary canonical complete binary syntactic envelope. Unoccupied positions in that envelope are structural null slots and are not logical clauses or variables of φ_I.

Define the canonical syntactic-envelope height by

Domain note. The logarithmic syntactic-envelope construction below is defined for M(I)=D_8^{CAN}(I)≥1. For a possible empty materialized 3-CNF artifact (M=0), $\log_2$ M is undefined; establishing that the frozen compiler never produces an empty artifact, or adopting a separately audited empty-artifact convention, remains an explicit compiler obligation.

$$h_{syn}(I) := \lceil \log_2 M(I) \rceil + 2$$

The additive constant 2 records the fixed binary depth needed to represent a three-literal clause, for example $l_1 \vee (l_2 \vee l_3)$. For a complete binary envelope of height h=h_syn(I), level k contains 2ˆk nodes and the total number of structural nodes is 2ˆ(h+1)−1. Therefore

$$p_k = \frac{2^k}{2^{h+1}-1}, \qquad k = 0, \ldots, h$$

Following the same level-class principle used in González (1995), the canonical syntactic entropy is

$$H_{sint}^{CAN}(I) := -\sum_{k=0}^{h_{syn}(I)} p_k \log_2 p_k$$

The matrix descriptor is identified with this quantity:

$$D_9^{CAN}(I) \equiv H_{sint}^{CAN}(I)$$

This makes the GCT-STL provenance chain explicit:

$$\varphi_I \to D_8^{CAN}(I) \to h_{syn}(I) \to H_{sint}^{CAN}(I) \equiv D_9^{CAN}(I)$$

4.2.1.3 Theorem 1 — Canonical Syntactic-Entropy Bound Theorem 1 — González (1995) Canonical Syntactic-Entropy Bound. Let I=(N,D,K) be a well-formed TSP_{dec} instance and let φ_I=compileDecision3SAT(I).formula. Let $M(I) := |Clauses\,(\varphi_I)| = D_8^{CAN}(I)$, and let $H_{sint}^{CAN}(I) := D_9^{CAN}(I)$ be the Shannon level entropy of the canonical complete binary syntactic envelope derived from this exact materialized 3-CNF artifact. Then

$$0 \leq H_{sint}^{CAN}(I) = D_9^{CAN}(I) < 2\ bits$$

and, for growing families with D_8^{CAN}(I)$\to\infty$,

$$\lim_{D_8^{CAN}(I)\to\infty} H_{sint}^{CAN}(I) = \lim_{D_8^{CAN}(I)\to\infty} D_9^{CAN}(I) = 2\ bits$$

Thus $H_{sint}^{CAN}(I) = D_9^{CAN}(I)$ is uniformly bounded by the same 2-bit asymptote introduced for decisional syntactic information in González (1995), while its finite-instance value remains a deterministic function of the exact materialized syntax size. The universal statement concerns the bound; the numerical value itself is not independent of $D_8^{CAN}(I)$. The complete proof and the precise inheritance from the 1995 construction are given in Appendix B.

4.2.2 Semantic Entropy and Theorem 2

4.2.2.1 The Concept of Semantic Entropy The semantic entropy H_{SEM} of the GCT STL model captures the geometric and combinatorial structure of the directed-tour ensemble associated with a TSP_{dec} instance $I = (N, D, K)$. For a fixed origin, the set of admissible tours contains $(N-1)!$ permutations. The semantic entropy descriptor $D_7^{CAN}\,(I, \beta)$ quantifies the informational complexity of this ensemble under a Boltzmann weighting determined by the cost matrix D and an inverse-temperature parameter β. This entropy is a structural observable of the tour space; it is not a deterministic-time lower bound.

Semantic entropy describes the uncertainty of the Boltzmann probability distribution over the geometric tour space induced by a TSP cost matrix. It is therefore a property of the metric structure of the instance, not of the local 3-SAT grammar. D7 is explicitly a Shannon entropy; it is not the expected tour cost, and its magnitude alone is not interpreted as a computational lower bound or as the source of NP-hardness.

The finite tour-space cardinality supplies a universal entropy envelope. The dependence on the particular instance enters through the cost matrix D and the inverse-temperature parameter β, which determines the Boltzmann weights within that fixed space.

4.2.2.2 Tour Space and Boltzmann Distribution

$$\Omega_N^{\rightarrow} := \{\text{fixed-origin directed Hamiltonian tours}\}, \quad |\Omega_N^{\rightarrow}| = (N-1)!$$

C_I(π) := directed tour cost induced by D

$$Z_I(\beta) := \sum_{\pi \in \Omega_N^{\rightarrow}} \exp\left(-\beta C_I(\pi)\right)$$

$$p_\beta(\pi) := \frac{\exp\left(-\beta C_I(\pi)\right)}{Z_I(\beta)}$$

$$D_7^{CAN}(I,\beta) := -\sum_{\pi \in \Omega_N^{\rightarrow}} p_\beta(\pi) \log_2 p_\beta(\pi)$$

The threshold K belongs to the TSP_{dec} instance but does not enter the Boltzmann measure itself; D_7^{CAN} depends on N, the cost matrix D, and β. The mean tour cost $\langle C \rangle_\beta$ is a separate thermodynamic observable.

Domain clarification (C03-17; no change to Theorem 2). For finite $\beta \geq 0$, the Boltzmann construction requires a nonempty directed-tour ensemble and finite tour costs, ensuring $Z_I(\beta) > 0$. Under the directed two-arc-return convention, N=2 gives one tour and $D_7^{CAN}(I,\beta) = 0$; whether N=1 is an admissible tour instance depends on a separate degenerate-tour convention not fixed here. The minimum well-formed N is therefore not newly asserted. The Θ(N log N) envelope concerns N$\to\infty$.

4.2.2.3 Theorem 2 — Shannon–Boltzmann Semantic Entropy Bounds Theorem 2 — Shannon–Boltzmann Semantic Entropy Bounds. For a well-formed directed TSP_{dec} instance I=(N,D,K), let $\Omega_N^{\rightarrow}$ denote the fixed-origin directed Hamiltonian-tour space, with $|\Omega_N^{\rightarrow}| = (N-1)!$. For $\beta \geq 0$, let $D_7^{CAN}(I,\beta)$ be the Shannon entropy of the Boltzmann distribution induced by the tour costs. Then:

$$0 \leq D_7^{CAN}(I,\beta) \leq \log_2\left((N-1)!\right)$$

the upper bound is attained in the uniform high-temperature regime,

Equality qualification. The maximal entropy occurs whenever the induced Boltzmann distribution is uniform. Besides β=0, this includes any instance whose directed-tour costs are all equal, for which equality holds at every finite $\beta \geq 0$. The tour ensemble is combinatorial and is not filtered by the decision threshold K.

$$D_7^{CAN}(I,0) = \lim_{\beta \to 0} D_7^{CAN}(I,\beta) = \log_2\left((N-1)!\right)$$

and the maximum-entropy envelope satisfies

$$D_7^{MAX}(N) := \log_2\left((N-1)!\right) = \Theta\left(N \log N\right)$$

Moreover,D_7^{CAN}(I,β) is non-increasing in β.

The Θ(N log N) scale belongs to the maximal entropy envelope, not to every finite-temperature instance. The exact thermodynamic identity, monotonicity proof, and low-temperature degeneracy limit are given in Appendix C. D7 is an entropy and must not be identified with the expected tour cost.

The exact canonical D7 definition, Theorem 2 and its proof remain self-contained in Paper I-A. Extended Fisher–Ruppeiner analysis, estimator design and finite-MCMC methodology are developed in companion statistical-physics work and do not alter the definition or theorem status of D7 used here.

4.2.3 Computational Entropy and Theorem 3

4.2.3.1 The Concept of Computational Entropy Theorem 3 separates two quantities that were conflated in the earlier formulation. The canonical quantity is attached to the complete Boolean assignment space of the exact 3-CNF artifact produced by the canonical $TSP_{dec} \to 3SAT$ compiler; the operational quantity is attached to the subset of complete assignments actually reached or evaluated by a specified deterministic search procedure. This distinction preserves an entropic representation of computational search while avoiding any assumption that every deterministic algorithm must enumerate the complete Boolean space.

Let φ_I:= compileDecision3SAT(I). formula and ν(I) := |Vars(φ_I)|.

$$D_5^{CAN}(I) := H_{COMP}^{CAN}(I) := \log_2 \left|\{0,1\}^{\nu(I)}\right| = \nu(I)$$

For deterministic procedure A, let E_A(I)⊆{0,1}^{ν(I)} be the set of complete assignments explicitly reached or evaluated by A. The operational descriptor is

$$D_5^{OP}(A,I) := H_{COMP}^{OP}(A,I) := \log_2 max\{1, |E_A(I)|\}$$

Only D_5^{CAN} enters the canonical raw matrix T_{VG}^{raw}. D_5^{OP} is an associated operational observable whose value depends on both the instance and the chosen deterministic procedure.

4.2.3.2 Theorem 3 Theorem 3 — Canonical and Operational Computational Entropy Bounds.

For a well-formed TSP_{dec} instance I=(N,D,K), let $D_5^{CAN}(I)$ denote the canonical computational entropy associated with the Boolean search space of the materialized canonical 3-CNF representation, and let $D_5^{OP}(A,I)$ denote the operational computational entropy effectively explored by a deterministic search procedure A. For the natural family of instances with polynomial bounded edge weights and threshold,

$$D_5^{CAN}(I) = \Omega\left(N^2\right) \quad \text{and} \quad D_5^{CAN}(I) = O\left(N^3 \log N\right)$$

and

$$0 \le D_5^{OP}(A,I) \le D_5^{CAN}(I)$$

Equality D_5^{OP}(A,I)=D_5^{CAN}(I) holds under exhaustive exploration.

The complete derivation of the canonical lower and upper bounds, the compiler-level variable count, the operational inequality, and the conditional relation to realized search time are given in Appendix D. The theorem does not assert an exponential lower bound for arbitrary deterministic algorithms.

4.2.4 Entropic Integration and the Canonical Entropy Vector

Sections 4.2.1–4.2.3 establish three canonical entropy descriptors associated respectively with the syntactic, semantic, and computational components of the GCT-STL construction:D_9^{CAN}(I),D_7^{CAN}(I,β), and D_5^{CAN}(I). Their ordered integration defines the canonical entropy vector $H^{CAN}(\mathrm{I}, \beta) := (D_7^{CAN}(\mathrm{I}, \beta), D_5^{CAN}(\mathrm{I}), D_9^{CAN}(\mathrm{I}))$, with domain ordering (S, T, L). This subsection integrates those theorem-origin coordinates before the complete nine-descriptor tensorial signature is defined in §4.3.

4.2.4.1 The Integrating Corollary of the Three Dimensions Corollary 1 (Three-dimensional entropic integration of the TSP). Let H^{CAN}(I,β)=(D_7^{CAN}(I,β),D_5^{CAN}(I),D_9^{CAN}(I)). The canonical envelope scales satisfy D_9^{CAN}=O(1), 0≤D_7^{CAN}(I,β)≤D7^MAX(N)=Θ(N log N), and D_5^{CAN}=Ω(N^2) ∩ O(N^3 log N) for the stated natural polynomially bounded family.

For the canonical vector, define $||H^{CAN}||$=$\sqrt{[}(D_7^{CAN})^2+(D_5^{CAN})^2+(D_9^{CAN})^2]$. Under these envelopes, the canonical computational coordinate is asymptotically larger in descriptor magnitude, so $||H^{CAN}||$=$\Theta(D_5^{CAN})$, hence Ω(N^2) and O(N^3 log N). This ordering concerns the magnitudes of the chosen canonical descriptors; it neither identifies D_5^{CAN} with actual running time nor implies that every deterministic procedure realizes D_5^{OP}≈D_5^{CAN}.

The envelope-scale comparison O(1)≪Θ(N log N)≪Ω(N^2) is not a pointwise inequality for every instance and β and is not a causal ranking among L, S, and T. Candidate obstruction mechanisms must therefore be tested on the complete coupled STL object, where marginal coordinates may contribute through pairwise or higher-order interactions.

4.2.4.2 Role of the Entropy Vector in the STL Representation The ordered integration of the three canonical entropy descriptors defines H^CAN(I,β):=(D_7^{CAN}(I,β),D_5^{CAN}(I),D_9^{CAN}(I)). This vector acquires its precise tensorial interpretation in §4.3. After the canonical Valiant–González signature T_{VG}^{raw} is formally defined, H^CAN is identified with its third resolution row Y3. At this stage, however, H^CAN is introduced independently through Theorems 1–3; no tensorial projection or contraction is required for its definition.

Accordingly, the notation Y3 is introduced here only as the canonical entropic-row designation: H^{CAN}(I,β)≡Y3. Its formal realization as a row projection and contraction of T_{VG}^{raw} is deferred to §4.3.2, after the ambient tensor space $V_Y \otimes V_{STL}$ has been defined.

This distinction is important because the three resolution rows do not have the same mathematical provenance. Y3 is intrinsically entropic:D_7^{CAN},D_5^{CAN}, and D_9^{CAN} are canonical information quantities expressed in bits. By contrast, Y1 and Y2 contain non-entropic structural, machine-derived, and artifact-derived descriptors. The complete nine-descriptor object is therefore not an entropy matrix.

The downstream R1 componentwise logarithmic scale transform does not alter this provenance distinction. The non-entropic coordinates may be mapped to $\log_2$(1+d) for numerical scale compatibility, while the three canonical entropy coordinates remain unchanged because they are already expressed in bits. Thus, entropic coordinate ≠ log-scaled coordinate. R3 is a separate exact-envelope normalization of the R2 information representation.

4.3 R0: The Canonical Valiant–González Tensorial Signature and Provenance Architecture

The purpose of this section is to define the complete canonical Valiant–González tensorial signature that integrates the nine GCT-STL descriptors into a single provenance-controlled representation. The construction extends the canonical entropic row Y3 introduced in §4.2.4 with the two non-entropic resolution rows Y1 and Y2, thereby producing the complete 3×3 STL representation. Once this tensor space has been defined, the entropic-row projection and contraction introduced conceptually in §4.2.4 can be stated formally in §4.3.2. The section then specifies how deterministic execution, theorem-origin results, and structural/artifact evidence populate the nine descriptor positions through the Provenance Bundle and SourceMap. The resulting raw signature remains distinct from the downstream representation layers integrated in §4.5: $R1 = L_{VG}$ is the componentwise logarithmic scale complement, $R2 = T_{VG}^{H}$ is the canonical information lift, and R3=R_VG is the exact-envelope normalized representation. R0 does not by itself constitute an obstruction or computational lower bound.

4.3.1 Canonical Tensor Space and Formal Definition of the Valiant–González Signature

The Valiant–González object is the canonical two-axis tensorial signature used by GCT-STL to co-locate the three informational domains and the three resolution rows after the nine descriptors have been independently sourced. Its tensorial status is defined from the two structural index spaces of the model; the existing 3×3 signature is not identified with a third-order 3×3×3 tensor.

Let the resolution space and the STL-domain space be the real vector spaces

$$V_Y := span_R e_{Y1}, e_{Y2}, e_{Y3}, V_{STL} := span_R e_S, e_T, e_L.$$

For a well-formed TSP_{dec} instance I, a terminating deterministic realization M on x=enc(I), and a fixed inverse-temperature parameter β whenever D7 is evaluated, assume that the canonical Provenance Bundle has been materialized and that all SourceMap-required canonical descriptors are defined. The canonical raw Valiant–González tensor is

$$T_{VG}^{raw}(M, x; \beta) := \sum_{r=1}^{3} \sum_{c \in \{S,T,L\}} d_{rc}\, e_{Y_r} \otimes e_c \in V_Y \otimes V_{STL}$$

With the ordered bases (e_{Y1}, e_{Y2}, e_{Y3}) and (e_S, e_T, e_L), its coordinate matrix is

$$[T_{VG}^{raw}]_{Y \times STL} = \begin{pmatrix} D_6^{CAN} & D_2 & D_1 \\ D_8^{CAN} & D_4 & D_3 \\ D_7^{CAN} & D_5^{CAN} & D_9^{CAN} \end{pmatrix}$$

Equivalently, the STL columns are S=(D_6^{CAN},D_8^{CAN},D_7^{CAN})^T, T=(D2, D4,D_5^{CAN})^T, and L=(D1, D3,D_9^{CAN})^T, while the Y-rows are Y1=(D_6^{CAN},D2, D1), Y2=(D_8^{CAN},D4, D3), and Y3=(D_7^{CAN},D_5^{CAN},D_9^{CAN}). This coordinate placement records the architecture of the signature; it does not determine descriptor provenance. When β is fixed by context, the abbreviated notation T_{VG}^{raw}(M,x) is used.

Mathematically, T_{VG}^{raw} is therefore a second-order tensor (equivalently, a 3×3 bilinear-coordinate object after bases are fixed). Its nine entries are heterogeneous observables whose values remain governed by the Provenance Bundle and SourceMap. Tensor assembly co-locates the STL domains in one representation; it does not by itself prove non-additivity, irreducibility, an orbit obstruction, or a computational lower bound.

Full-coupling interpretation. The Full STL Tensorial Coupling principle remains an open and falsifiable research hypothesis about properties that may emerge from the complete source-audited signature rather than from any single marginal descriptor. It does not assert that all three domains are individually necessary in every instance, and it is not an obstruction theorem.

Ambient transformations and geometric analysis. The tensor space V_Y⊗V_STL admits the natural change-of-basis action of $GL(V_Y) \times GL(V_{STL})$. This is an ambient algebraic action on the representation space, not yet a theorem of computational equivalence. Any group later used for computational orbit or orbit-closure analysis must be separately defined and shown to preserve the relevant STL realizability, provenance, encoding, and $Cost_{STL}$ structure.

Higher-order coupling lifts. A genuine third-order coupling tensor in a space such as $V_S \otimes V_T \otimes V_L$ would have 27 ambient coordinates and is not identical to the canonical nine-coordinate T_{VG}^{raw}. If such an object is introduced downstream, it must be defined by an explicit lift $C_{VG} : TVGraw \mapsto C_{VG}(T_{VG}^{raw})$, together with a stated interpretation of its additional coordinates.

A lift constructed solely from nine coordinates of R0 cannot be assumed to generate 27 independent parameters; its image and multilinear properties require a separately defined map, provenance semantics and proof. The schematic third-order notation is not part of the established tensor signature.

Representation-stack boundary. The downstream matrix L_{VG} remains the component wise logarithmic scale complement of T^{raw}_{VG} defined in §4.5.2. The expression "tropical" is descriptive; Paper I-A does not claim that this transform constructs a tropical semiring, tropical variety, or formal tropical-algebra representation. The canonical exact-envelope normalized representation is $R3 = R_{VG}$ and is derived from R2, not from R1.

4.3.2 Canonical Entropic-Row Projection and Contraction

Having defined the canonical Valiant–González signature as $T^{raw}_{VG} \in V_Y \otimes V_{STL} \cong R^{3\times 3}$, with ordered bases (e_{Y1}, e_{Y2}, e_{Y3}) for V_Y and (e_S, e_T, e_L) for V_{STL}, it is now possible to formalize the role assigned in §4.2.4 to the canonical entropy vector. The third resolution row is Y3=(D_7^{CAN},D_5^{CAN},D_9^{CAN})=H^{CAN}.

4.3.2.1 Canonical Entropic Contraction Let e^{Y1}, e^{Y2}, e^{Y3} denote the dual basis of V_{Y*}. Define $\pi_{Y3} := e^{Y3} \otimes id_{VSTL}$, with $\pi_{Y3} : V_Y \otimes V_{STL} \to V_{STL}$. Applying this contraction to the canonical tensor gives $\pi_{Y3}\,(T^{raw}_{VG})$=$D_7^{CAN} e_S$+$D_5^{CAN} e_T$+$D_9^{CAN} e_L$=H^{CAN}.

4.3.2.2 Canonical Entropic-Row Projection Define $p_{Y3} : V_Y \to V_Y$ by $p_{Y3}(v) = e^{Y3}(v) e_{Y3}$ and $P_{Y3} := p_{Y3} \otimes id_{VSTL}$. Then $P_{Y3}\,(T^{raw}_{VG}) = e_{Y3} \otimes H^{CAN}$ and $P^2_{Y3} = P_{Y3}$. Thus P_{Y3} retains the entropy-bearing component inside the tensor space, whereas π_{Y3} extracts the corresponding vector in V_{STL}.

4.3.2.3 Mathematical Meaning of the Entropic Resolution The complete raw signature decomposes as $T^{raw}_{VG} = e_{Y1} \otimes Y1 + e_Y2 \otimes Y2 + e_Y3 \otimes H^{CAN}$. This is a resolution-space decomposition of the canonical tensorial signature, not a decomposition into three entropy vectors. Only Y3 is intrinsically entropic; therefore $H^{CAN} \neq T^{raw}_{VG}$.

4.3.2.4 Basis Dependence and Canonical Status P_Y3 and π_Y3 are canonical relative to the fixed ordered resolution basis and row semantics of the GCT-STL representation. A generic transformation in GL(V_Y) may mix the resolution rows and destroy the explicit interpretation of Y3 as the entropic row. Downstream admissible transformations must therefore state whether they preserve the canonical resolution structure.

4.3.2.5 Relation to the Downstream Representation Layers The projection and contraction operate on the raw canonical signature and precede the downstream representation layers integrated in §4.5. The Y3 coordinates are already information quantities expressed in bits; the remaining six coordinates are non-entropic raw quantities. The R1 transformation $T^{raw}_{VG} \to L_{VG}$ improves numerical scale compatibility but does not convert log-scaled coordinates into Shannon entropies. R2 is a separate information-theoretic lift, and R3 is its exact-envelope normalization.

4.3.2.6 Claim Boundary The identities $\pi_{Y3}(T^{raw}_{VG}) = H^{CAN}$ and $P_{Y3}(T^{raw}_{VG}) = e_{Y3} \otimes H^{CAN}$ are representation-theoretic statements internal to GCT-STL. They establish neither functional independence of the STL domains nor non-additive coupling, non-collapsibility, orbit-closure separation, or a computational lower bound. Y3 is the canonical entropic resolution, but Y3≠the complete obstruction object. The Full STL Tensorial Coupling Principle instead motivates analysis of the complete provenance-controlled T^{raw}_{VG}.

4.3.3 Relation between Machine Trace, Provenance Bundle, and $\mathbf{T^{raw}_{VG}}$

The canonical tensor is related to the deterministic machine trace, but the relation is mediated by provenance. Trace(M,x) is the primary evidence source for the machine-derived channel only; it is not the universal source

of D1–D9. The canonical construction is therefore

$$Trace(M,x) \to \mathcal{P}_{MTD}(M,x)$$

$$\mathcal{P}_{STL}(M,x) := (\mathcal{P}_{MTD}(M,x), \mathcal{P}_{THM}(M,x), \mathcal{P}_{STRUCT}(M,x), SourceMap)$$

$$\mathcal{P}_{STL}(M,x) \overset{\Phi_{STL}}{\to} STLSignature(M,x) \to T_{VG}^{raw}(M,x;\beta)$$

At descriptor level, D1–D4 are obtained from the machine/artifact execution channel, whereas D5, D7, and D9 retain theorem-origin authority and D6 and D8 retain structural/artifact provenance. Schematically,

$$(D_1, D_2, D_3, D_4) = E_{MTD}\left(Trace(M,x), \text{artifact evidence}\right)$$

$D_5^{CAN} \leftarrow \varphi_I$ and ν(I),$D_7^{CAN} \leftarrow$ (N,D,β),$D_8^{CAN} \leftarrow \varphi_I$,$D_9^{CAN} \leftarrow D_8^{CAN}$,$D_6^{CAN} \leftarrow$ S$\leftrightarrow$L structural coupling.

This separation preserves the ontology of Paper I-A: M computes; Trace(M,x) records the concrete deterministic execution; the Provenance Bundle integrates heterogeneous evidence; Φ_{STL} represents; STLSignature records; and T_{VG}^{raw} describes. Thus T_{VG}^{raw} is trace-informed but not trace-defined.

A trace-aware canonical formula may therefore be written as

$$T_{VG}^{raw}(M,x;\beta) = \sum_{r=1}^{3} \sum_{c \in \{S,T,L\}} d_{rc}\left(\mathcal{P}_{STL}(M,x),\beta\right)\, e_{Y_r} \otimes e_c$$

The β argument is present only because D_7^{CAN} is a family indexed by inverse temperature; it does not make the deterministic machine execution itself temperature dependent.

4.3.3.1 Prospective Tensorial Trace Trace_VG A distinct downstream object may be useful for studying the dynamics of the STL representation along a deterministic execution. This object must not be confused with the canonical static tensor T_{VG}^{raw}. Define provisionally a tensorial trajectory

$$Trace_{VG}(M,x) := \left(T_{VG}^{(0)}, T_{VG}^{(1)}, \ldots, T_{VG}^{(t)}\right)$$

where each $T_{VG}^{(k)}$ would be a state-dependent or prefix-dependent representation associated with the machine configuration C_k in $Trace(M,x) = (C_0, C_1, \ldots, C_t)$. Such a construction is not yet canonical, because not every descriptor has a natural configuration-wise value.

The first mathematically conservative dynamic object is therefore restricted to the descriptors whose provenance is genuinely trace-sensitive:

$$Trace_{VG}^{MTD}(M,x) := ((D_{1,k}, D_{2,k}, D_{3,k}, D_{4,k}))_{k=0}^{t}$$

The remaining canonical descriptors D5, D6, D7, D8, and D9 are global instance/theorem/artifact quantities unless a later theorem supplies a principled local or prefix-dependent refinement. No artificial time dependence is introduced merely to fill a tensorial trajectory.

If a full Trace_VG is later defined, the programme must specify: (i) the extraction map from C_k or a trace prefix to each dynamic coordinate; (ii) which canonical coordinates remain fixed along the path; (iii) the topology or metric on the trajectory space; (iv) admissible transformations; and (v) any theorem connecting trajectory geometry to $Cost_{STL}$ or F2. Until then, Trace_VG is a prospective derived object, not part of the canonical class definitions.

The resulting architecture is: $M \to Trace(M,x) \to \mathcal{P}_{MTD}$; $\mathcal{P}_{MTD} \oplus \mathcal{P}_{THM} \oplus \mathcal{P}_{STRUCT} \to \mathcal{P}_{STL} \to T_{VG}^{raw}$; and, prospectively, $Trace(M,x) \to Trace_{VG} \to$ geometric/orbital dynamics.

4.3.4 Descriptor-Source Architecture

The nine descriptors occupy fixed positions in the STL crystal, but their mathematical provenance is heterogeneous. Accordingly, the representation layer distinguishes three source channels: machine/trace provenance, theorem-origin provenance, and structural/artifact provenance. Matrix position is therefore not a provenance rule.

$$\mathcal{P}_{MTD} \to D1, D2, D3, D4,\ \mathcal{P}_{THM} \to D5, D7, D9,\ \mathcal{P}_{STRUCT} \to D6, D8.$$

For every descriptor, four classifications are kept separate: STL domain/column, Y-row, causal influence, and provenance source.

$$Domain \neq Row \neq CausalInfluence \neq ProvenanceSource$$

Descriptor	STL / Y	Primary provenance	Dependencies / evaluation inputs	Source contract
D1	L / Y1	Machine / Trace	Trace(M,x); logical-state observations	State-density/logical-state quantity is read from the machine-provenance channel, not from González genealogy or theorem-origin sources.
D2	T / Y1	Machine / Trace	Trace(M,x); tape/memory-access evidence	Basal tape/memory-access quantity tied to the concrete deterministic execution contract.
D3	L / Y2	Machine + artifact trace	Trace(M,x); compiler/AST connectivity evidence	AST/connectivity quantity exposed by machine/compiler refinement; it is not merely a tape statistic.
D4	T / Y2	Machine control / search trace	Deterministic control/search structure in Trace(M,x)	Search/tree structural quantity belongs to deterministic execution provenance; it is not nondeterministic DTM branching.
D5	T / Y3	Theorem 3 + exact 3-CNF dependency	φ_I; $\nu(I) = \lvert Vars(\varphi_I)\rvert$; optional $E_A(I)$ for D_5^{OP}	D_5^{CAN}(I)=ν(I) is the canonical assignment-space entropy;D_5^{OP}(A,I) is an operational explored-support observable and never replaces the canonical coordinate.
D6	S / Y1	Structural S–L coupling	Instance-level semantic/logical coupling evidence; optional finite calibration $\hat{\kappa}$_SL	D_6^{CAN}(I) is the canonical instance-level coupling descriptor when numerically defined.D_6^{OP}=$\hat{\kappa}$_SL is operational only; a universal κ^*_SL would require an independent theorem and is not assumed here.
D7	S / Y3	Theorem 2	N, D, β; Ω_N$\hat{}\to$ fixed-origin directed tours	D_7^{CAN}(I,β) is Shannon–Boltzmann entropy. Expected tour cost is distinct; downstream MCMC values are estimators, not a different descriptor.
D8	S / Y2	Exact materialized artifact	φ_I=compileDecision3SAT(I). formula	D_8^{CAN}(I)=|Clauses(φ_I)| is the exact structural size of the canonical materialized 3-CNF artifact.
D9	L / Y3	Theorem 1 + D8 artifact	Same exact φ_I and D_8^{CAN}(I); González (1995) as genealogy	D_9^{CAN}(I)$\equiv H_{sint}^{CAN}$(I) is evaluated on the same artifact as D8. Theorem 1 is immediate theorem-origin; González (1995), Appendix 3, supplies the level-entropy lineage.

4.3.5 STL Provenance Bundle

To prevent the representation map from visually or logically collapsing these heterogeneous sources into Trace(M,x), the model introduces a first-class STL Provenance Bundle. The bundle integrates the machine channel, theorem-origin source references, structural/artifact evidence, and the descriptor-level SourceMap.

$$\mathcal{P}_{STL}(M,x) := (\mathcal{P}_{MTD}(M,x), \mathcal{P}_{THM}(M,x), \mathcal{P}_{STRUCT}(M,x), SourceMap)\,.$$

For the canonical TSP path, $\mathcal{P}_{MTD}$ contains the machine code, x, Trace(M,x), and certified/refined machine-artifact evidence; $\mathcal{P}_{THM}$ records the static theorem-origin bridges Theorem 3→D5, Theorem 2→D7, and Theorem 1→D9; and $\mathcal{P}_{STRUCT}$ contains the materialized logical artifact required by D8 and the canonical semantic–logical coupling evidence required by D6. For D9 specifically, González (1995), Appendix 3, is recorded as genealogical/mathematical provenance of the level-entropy construction, while Theorem 1 remains the immediate theorem-origin authority and D8 supplies the exact instance artifact. The theorem references are provenance authorities, not theorem texts re-proved at run time.

4.3.6 Canonical Partial Representation Map and STL Signature

The canonical STL representation map is therefore defined over the explicit multi-source bundle. It is partial at the foundational level: the later representation theorem must establish the relevant totalization and polynomial-overhead properties rather than assuming them by definition.

$$\Phi_{STL}\left(\langle M\rangle, x, \mathcal{P}_{STL}(M,x)\right) \rightharpoonup STLSignature(M,x)$$

$$STLSignature := (D_1, \ldots, D_9, T_{VG}^{raw}, SourceMap, ProvenanceBundleRef, EncodingTag)$$

4.3.7 Provenance Materialization, compileSTL and $\mathrm{Cost}_{\mathrm{STL}}$

The manuscript-level materialization interface is factored through provenance construction. Instance-specific work required to extract, materialize, validate, and encode provenance evidence is part of the representation cost; static theorem-source references are not re-proved for every input.

$$\text{buildProvenanceBundle}(M,x) \rightharpoonup \mathcal{P}_{STL}(M,x).$$

$$\text{compileSTL}\left(\langle M\rangle, x\right) = \Phi_{STL}\left(\langle M\rangle, x, \text{buildProvenanceBundle}(M,x)\right).$$

The displayed `compileSTL` identity holds whenever both partial constructions are defined.

$$Cost_{STL}(M,x) := T_M(x) + T_{buildProv}(M,x) + T_\Phi(M,x) + Size\left(STLSignature(M,x)\right)$$

Equivalently, an implementation may fold $T_{buildProv} + T_\Phi$ into a single $T_{compileSTL}$ term. No determinant, permanent, orbital distance, Fisher–Ruppeiner observable, universal-robustness predicate, or obstruction criterion is part of this foundational cost definition.

System-interface boundary. The formula above specifies the representation-side accounting interface required to make STL construction computationally non-free. Its role here is to expose what must be charged when R0 is built. The class-theoretic semantics of $Cost_{STL}$, including its use in P_{STL}, NP_{STL} and the F1/F2 route, belong to System II; §4.3 does not derive those consequences.

4.3.8 Ontological and Claim Boundary and Definition of the Nine Components

The ontology of the representation layer is therefore explicit: M computes; Trace(M,x) records one machine-derived provenance channel; $\mathcal{P}_{STL}$ integrates heterogeneous provenance; Φ_{STL} represents; STLSignature records the descriptors and their source map; and T_{VG}^{raw} describes the resulting second-order 3×3 tensorial signature. None of these objects is identified with another, and no obstruction or lower-bound conclusion follows from the existence of the representation alone. The nine descriptor definitions are refined in the following table.

Component	Domain	Name	Analytical derivation formula	Interpretation
D1	L	State density	Trace-derived logical-state interface	Logical-state/state-density coordinates from the certified machine-provenance channel; not inferred from row or column placement.
D2	T	Tape depth	Trace-derived tape/memory interface	Basal DTM memory/tape-access coordinate tied to the concrete deterministic execution.
D3	L	AST bandwidth	Machine + compiler/AST artifact provenance	Structural complexity/connectivity of the exact materialized 3-CNF syntax tree.
D4	T	Deterministic search branching	Deterministic control/search provenance	Control/search structure of the concrete deterministic computation; not nondeterministic-machine branching.
D5	T	Canonical computational entropy	$D_5^{CAN}(I) = \nu(I) = \lvert Vars\,(\varphi_I)\rvert$	Entropy of the complete Boolean assignment space of the exact materialized 3-CNF.D_5^{OP}(A,I) is a separate explored-support observable.
D6	S	Semantic–logical coupling	$D_6^{CAN}(I)$: instance-level S–L coupling descriptor	Canonical cross-domain coupling coordinates.$D_6^{OP}{:=}\hat{\kappa}$_SL is an operational calibration only; κ^*_SL is a future universal candidate, not an established constant.
D7	S	Shannon–Boltzmann semantic entropy	$D_7^{CAN}(\mathrm{I},\beta){=}-\Sigma\pi\ \mathrm{p}\beta(\pi)\ \log_2\ \mathrm{p}\beta(\pi)$	$0{\le}D_7^{CAN}(\mathrm{I},\beta){\le}\log_2((\mathrm{N}{-}1)!)$. Equality is the $\beta{\to}0$ / uniform high-temperature envelope, not a general identity.
D8	S	Materialized 3-CNF size	$D_8^{CAN}(I) = \lvert Clauses\,(\varphi_I)\rvert$; $\varphi_I =$ `compileDecision3SAT(I).formula`	Exact structural size of the canonical materialized logical artifact under the fixed artifact-size convention.
D9	L	Canonical syntactic entropy	$D_9^{CAN}(I) \equiv H_{sint}^{CAN}(I) = -\sum_k p_k \log_2 p_k$	Computed from the same exact φ_I recorded by D8: $0{\le}D_9^{CAN}{<}2$ and $D_9^{CAN}{\to}2$ as $D_8^{CAN}{\to}\infty$. The 2-bit law descends from González (1995).

4.3.8.1 Representation Adequacy and V&V Status Representation Adequacy Principle (RAP). Let $\mathcal{R}_{VG}$ denote the canonical map from the declared computational/artifact sources to $R_0 = T_{VG}^{raw}$. No invariant, coupling measure, curvature response, orbit quantity, determinant/permanent diagnostic or other property $O\,(\mathcal{R}_{VG}(I))$ may be interpreted as a property of the underlying TSP_{dec} computation solely because it is well defined on R0. Such an interpretation requires a proved preservation/transfer relation whose hypotheses and conclusion match the computational claim being made.

Accordingly, representational completeness is not identified with informational exhaustiveness. R0 need not reconstruct every detail of the original instance or execution; it must preserve the information required by the theorem that later transfers a representation-level obstruction to the computational target. This is a falsifiable adequacy requirement, not an assumption granted by the 3×3 form of T_{VG}.

V&V property	Historical evidence at v4.3 (T-3C baseline; not current gate PASS)	T-3C classification
Well-definedness	Canonical descriptor definitions, R0 assembly and source contracts exist.	DEFINED / audit evidence present; do not declare VV-G1 formally closed without a dedicated completeness audit.

V&V property	Historical evidence at v4.3 (T-3C baseline; not current gate PASS)	T-3C classification
Semantic fidelity	TSP_{dec}, 3-CNF, DTM/Trace and representation interfaces are explicitly related.	OPEN as theorem-level adequacy; interface existence is not semantic-validation proof.
Polynomial constructibility	Some compiler/verifier components have polynomial bounds; $Cost_{STL}$ interface is defined.	OPEN end-to-end gate. No hidden superpolynomial representation preprocessing may be assumed free.
Provenance control	SourceMap and heterogeneous Provenance Bundle are explicit.	STRONGLY SPECIFIED; formal gate closure remains separate from existence of the metadata structure.
Transfer sufficiency	Representation Transfer and F2c are explicitly identified.	OPEN and critical; this is the decisive bridge from structural obstruction to F2.

Descriptor rule. The table records canonical descriptor meaning and source contracts; it does not collapse four distinct classifications. STL column membership identifies domain, Y1/Y2/Y3 identifies resolution stage, the ontological discussion in §4.3.8 records representational status and causal-influence boundaries, and SourceMap records provenance. These classifications may overlap without becoming identical.

D6 CAN/OP policy. D_6^{CAN}(I) is the canonical instance-level semantic–logical coupling coordinate recorded by the SourceMap. The operational collision-functional realization D_6^{OP}(D)=1+$\lambda\rho$_coll(D), together with any finite balanced calibration $\hat{\kappa}$_SL, must be explicitly tagged as operational/empirical and cannot replace D_6^{CAN} or be tuned to obtain a desired downstream result. A universal candidate κ_SL* would require an independent concentration/common-limit theorem over an explicitly defined STL-admissible class and is not assumed as an established constant in Paper I-A.

Exact-artifact identity for D8→D9. Let φ_I=compileDecision3SAT(I).formula. D_8^{CAN}(I)=|Clauses(φ_I)| records the exact number of clauses materialized by the canonical compiler, and $D_9^{CAN}(I) = H_{sint}^{CAN}(I)$ is evaluated from the canonical structural envelope derived from that same D8 artifact. No separately estimated clause count, padded formula, or alternate 3-CNF is permitted in the D8→D9 chain.

4.3.9 Location of the 3-CNF Components in the Valiant–González Crystal

The following annotation is a causal-influence view of the canonical raw signature T_{VG}^{raw}. It is not a replacement for SourceMap and does not redefine descriptor provenance:

$$T_{VG}^{\text{raw}}(N) = \begin{pmatrix} D_6^{(S)}[\Phi_{EDGE}, \Phi_{DIST}] & D_2^{(T)} & D_1^{(L)}[\Phi_{HAM}] \\ D_8^{(S)}[\Phi_{EDGE}, \Phi_{DIST}] & D_4^{(T)}[\Phi_{HAM} \to Search] & D_3^{(L)}[\Phi_{HAM}, \Phi_{EDGE}] \\ D_7^{(S)}[\Phi_{DIST}] & D_5^{(T)}[Search] & D_9^{(L)}[\Phi_{HAM}, 3CNF] \end{pmatrix}$$

The interpretation in terms of the three 3-CNF subcomponents is:

3-CNF component	Primary domain	Main cells
Φ_{HAM}	Syntactic (L)	D1, D3, D9
Φ_{EDGE}	Coupling (L↔S)	D3, D4, D8
Φ_{DIST}	Semantic (S)	D6, D8, D7
DTM execution / Boolean assignment space	Computational (T)	D2, D4, D5

Table 1 must be read on three separate axes. Column membership in T_{VG}^{raw} fixes the STL domain; row membership Y1/Y2/Y3 fixes the resolution stage; and the "Main cells" column records which 3-CNF subcomponent or machine process materially influences a descriptor. For example, D3, D4, and D8 can each have more than one causal influence without changing their canonical domain or provenance source. D2 remains machine/tape-derived even though it occupies the T column and is listed under DTM search.

D5 provenance qualification. The D5 entry in the comparison table denotes $D_5^{CAN}(I) = \nu(I) = |Vars(\varphi_I)|$, a property of the canonical compiler artifact. Search realized by a deterministic procedure A affects only the separate $D_5^{OP}(A, I)$ observable through E_A(I); the table is a causal-influence annotation and does not override the descriptor SourceMap.

The coupling entries require the same discipline. Φ_{EDGE} and Φ_{DIST} may influence D6 or other S/L coordinates, but causal influence does not numerically define D6. D_6^{CAN} remains the canonical S↔L coupling coordinate, while any finite $\hat{\kappa}$_SL used in experiments is D_6^{OP} and must be explicitly tagged as operational. Accordingly, statements about "structural pressure," computational expansion, or hardness are model interpretations to be tested by later bridge/obstruction theorems, not consequences of Table 1 itself.

4.4 Full STL Tensorial Coupling Principle

System-I placement and scope. Full STL Tensorial Coupling is an OPEN structural principle of System I. Its genealogy lies in the González (1995) three-domain × three-magnitude multilinear architecture, while the contemporary SourceMap and Provenance Bundle determine the realized D_i values. The principle is stated first on the canonical R0 signature defined in §4.3 and functions as an organizing interpretation across the later R1, R2 and R3 representations consolidated in §4.5; it neither defines those downstream representations here nor constitutes an obstruction theorem.

Having defined T_{VG}^{raw} and its provenance architecture in §4.3, the paper can now state the coupling question at the correct epistemic level. The following principle is a research hypothesis about the complete signature, not a theorem inferred from the existence of the tensor.

4.4.1 Full STL Tensorial Coupling Hypothesis

Primary structural hypothesis (open and falsifiable). GCT-STL does not assign a candidate obstruction a priori to S, T or L in isolation. Instead, if a genuine GCT-STL obstruction exists, the model tests whether it is a property of the complete, provenance-controlled coupled signature formed after the nine descriptors have been independently sourced and assembled. The working structural hypothesis is therefore

$$O_{STL}(I) = F_N\left(S(I), T(I), L(I); I_{ST}, I_{SL}, I_{TL}, I_{STL}\right)$$

The interaction symbols I_ST, I_SL, I_TL, and I_STL are schematic placeholders for pairwise and higher-order non-additivity. They do not constitute a proved decomposition, and the hypothesis does not assert that all three domains are individually necessary in every instance. Its methodological force is narrower: marginal scale alone must not be converted into causal primacy before the coupled object has been analyzed and subjected to controlled intervention, ablation, robustness, and theorem-level reconstruction.

$$StructuralGenealogy\left(T_{VG}\right) \neq ProvenanceSource\left(D_i\right) \neq ObstructionTheorem$$

González (1995) supplies the structural genealogy of a three-domain × three-magnitude multilinear architecture; the contemporary Provenance Bundle and SourceMap determine the realized source of each D_i; and the Full STL Tensorial Coupling principle concerns a research hypothesis about the complete coupled object. None of these layers entail the other. Likewise,D_6^{CAN} is a specific canonical S↔L coupling coordinate and D_6^{OP}=$\hat{\kappa}$_SL is an operational calibration; neither is identical to the hypothesized full coupled S–T–L interaction represented in R0 or to a proved obstruction theorem.

4.4.2 Geometric-Entropic Interpretation and Coupled-Obstruction Research Hypothesis

The concrete TSP→3-CNF construction considered in this paper is organized as the conjunction of three canonical substructures:

$$\Phi_I = \Phi_{HAM} \wedge \Phi_{EDGE} \wedge \Phi_{DIST}.$$

Φ_{HAM} establishes the admissible Hamiltonian/permutation structure; Φ_{EDGE} provides the representation interface between positional logic and edge structure; and Φ_{DIST} introduces the metric/cost constraint derived from D and K. The deterministic Turing computation supplies the execution/provenance channel acting on the encoded instance. None of these components, individually, is identified in this Paper I-A as the source of NP-hardness, and no theorem in this paper derives a deterministic lower bound from one component alone.

The geometric-entropic interpretation is governed by the Full STL Tensorial Coupling principle stated in §4.4.1 and positioned in the programme roadmap of §3.5. Its working statement is an open and falsifiable structural hypothesis:

$$O_{STL}(I) = F_N\left(S(I), T(I), L(I); I_{ST}, I_{SL}, I_{TL}, I_{STL}\right).$$

The interaction terms are schematic placeholders for pairwise and higher-order non-additivity, not a proved decomposition. Operationally, "coupled obstruction" means that the nine independently sourced descriptors are co-located in the provenance-controlled T_{VG}^{raw} so that later analyses can test properties that are not reducible to any single marginal coordinate. D_6^{CAN} is only one canonical S↔L coupling coordinate; it is neither the hypothesized full coupled S–T–L interaction represented in R0 nor an obstruction theorem.

This interpretation also fixes the complexity-theoretic claim boundary. A coherent coupled pattern, finite phase-transition signal, nonzero determinant, orbit distance, Fisher–Ruppeiner response, entropy contrast or MCMC regularity is not a computational lower bound by itself. The admissible downstream route is:

coupled hypothesis/evidence → mathematically universal obstruction candidate → independent transfer to $Cost_{STL}$ → $F2$.

Accordingly, the exploratory phase-transition and statistical-physics diagnostics developed downstream in §4.5 may supply finite discovery evidence for the coupled system, but they do not establish the hypothesized mechanism and do not replace the universality or transfer theorems required before a computational lower-bound consequence could follow.

Under the Representation Adequacy Principle, Full STL Tensorial Coupling is a representation-level structural hypothesis and candidate organizing principle. Even if a coupled invariant is empirically robust or mathematically non-degenerate on R0–R3, no computational obstruction follows until VV-G4 establishes the relevant semantic adequacy and VV-G5 establishes the required $obstruction - to - Cost_{STL}$ transfer. Full Coupling may guide discovery and theorem formulation; it is not itself F2.

4.5 The Four-Representation Mathematical Architecture and Exploratory Diagnostics

4.5.0 Architectural Overview

Sections §4.1–§4.4 have already fixed the canonical TSP_{dec}/3-CNF sources, the descriptor family D1–D9, the provenance-controlled raw signature R0, and the OPEN Full STL Tensorial Coupling principle. This subsection completes System I by situating R0 inside the four-representation architecture closed by Step 12A–L. The four layers have distinct mathematical roles and must not be collapsed into a single normalization chain.

- $R0 := T_{VG}^{raw}$ [Canonical provenance/descriptor layer]
- $R1 := L_{VG}$ [derived componentwise logarithmic scale layer]
- $R2 := T_{VG}^{H}$ [canonical entropic/information layer]

- $R3 := R_{VG}$ [canonical exact-envelope normalized layer]
- R0 → R1 and R0 → R2 → R3

The architecture is branched, not linear: R0→R1→R2→R3 is not the canonical relation. R1 and R2 are distinct transformations of the R0 foundation, and R3 is obtained from R2. No representation is declared universally superior. Inference across layers requires representation-specific transfer arguments; empirical agreement across layers is evidence of robustness, not logical equivalence.

4.5.1 R0 — Canonical Provenance/Descriptor Representation

R0 is the canonical raw representation constructed in detail in §4.3. It preserves the heterogeneous meanings and provenance of the nine descriptor coordinates and is assembled directly from the frozen SourceMap and Provenance Bundle. In matrix order,

$$T_{VG}^{raw} := \begin{pmatrix} D_6 & D_2 & D_1 \\ D_8 & D_4 & D_3 \\ D_7 & D_5 & D_9 \end{pmatrix}$$

At the R0 layer only Y3=(D7, D5, D9) is intrinsically entropic; Y1 and Y2 remain non-entropic raw-resolution rows. Section §4.5 does not redefine R0 or reproduce the D1–D9/provenance construction; it recalls R0 only to make the representation roles explicit.

4.5.2 R1 — Componentwise Logarithmic Scale Complement

$R_1 = L_{VG}$ is the derived numerical layer retained by the final Step-12L role decision as COMPLEMENT, DO NOT REPLACE. Its purposes are scale compression, finite-hardware stability, historical comparability with earlier GCT-STL experiments, and representation-sensitivity analysis. For numerical analyses in which every required coordinate has been materialized, retain the frozen componentwise scale policy

$$g_i(d) = d \;\; (i \in \{5,7,9\})\,; \qquad g_i(d) = \log_2(1+d) \;\; (i \in \{1,2,3,4,6,8\})$$

$$L_{VG}(M,x) := g\,[T_{VG}^{raw}(M,x)] \quad \text{(componentwise)}$$

The term "tropical-inspired" is retained only as a historical descriptor for this componentwise logarithmic scaling policy; it does not assert that Paper I-A has constructed a tropical semiring, tropical variety, or formal tropical-algebra representation. A logarithmically compressed coordinate is not thereby an entropy. R1 therefore does not replace R0, R2, or R3.

Numerical-domain qualification (M-19). For each real-valued non-entropic coordinate to which g_i(d)=$\log_2$(1+d) is applied, its realized value must satisfy d>−1 (normally d≥0 under the specified nonnegative coordinate contract). Inputs outside the domain remain undefined for R1; any actual computation must charge numerical representation and precision costs. Scalar invertibility alone is not a complexity-transfer theorem.

D6 handling remains provenance-sensitive. If D_6^{CAN}(I) has a numerical realization, g_6 is applied to that canonical value. If an experiment instead uses $D_6^{OP} = \widehat{\kappa}_{SL}$, the resulting object must carry an explicit operational/experimental tag; $L_{VG}^{O}P$ cannot be silently substituted into a theorem whose hypotheses require D_6^{CAN}.

The map Λ defining R1 is explicit and deterministic, but a property first discovered or formulated in L_{VG} does not automatically transfer back to T_{VG}^{raw}. Such a move requires an appropriate representation-transfer

result RT_Λ. Any downstream computation of R1 actually used by a decider or verifier must also be charged rather than treated as free preprocessing.

4.5.3 R2 — Canonical Entropic/Information Representation

R2 is the canonical information-theoretic lift closed in Step 12G. It replaces the heterogeneous raw magnitudes of R0 by provenance-specific information functionals H1–H9 while preserving the same Y×STL slot occupancy. For a well-formed instance I and, where applicable, $\beta \geq 0$,

$$T_{VG}^{H}(I,\beta) := \begin{pmatrix} H_6 & H_2 & H_1 \\ H_8 & H_4 & H_3 \\ H_7 & H_5 & H_9 \end{pmatrix}$$

Thus $Y_1^H = (H_6, H_2, H_1)$, $Y_2^H = (H_8, H_4, H_3)$, $Y_3^H = (H_7, H_5, H_9)$, with columns $S_H = (H_6, H_8, H_7)^T$, $T_H = (H_2, H_4, H_5)^T$, and $L_H = (H_1, H_3, H_9)^T$. The lift changes the coordinate functionals, not the tensor slots.

The label entropic/information representation is intentionally broader than Shannon entropy: each H_i is tied to its frozen carrier, probability law, state space, or code-length construction, and not every H_i is a Shannon entropy. The complete H1–H9 definitions, proofs, contexts, and envelope mathematics belong to the companion System-I treatment in Paper I-B.

4.5.4 R3 — Exact-Envelope Normalized Representation

R3 controls coordinate-scale anisotropy in R2 by normalizing each information coordinate with its exact canonical coordinatewise maximum over the frozen finite bit-width-controlled instance family $I_{N,B}$. For fixed $N \geq 4$ and $B \geq 1$,

$$H_i^{MAX}(N, B \mid context_i) := \max_{I \in \mathcal{I}_{N,B}} H_i(I \mid context_i)$$

$$\rho_i(I, \beta; N, B) := \frac{H_i(I,\beta)}{H_i^{MAX}(N,B)}$$

$$R_{VG}(I, \beta; N, B) := \begin{pmatrix} \rho_6 & \rho_2 & \rho_1 \\ \rho_8 & \rho_4 & \rho_3 \\ \rho_7 & \rho_5 & \rho_9 \end{pmatrix}$$

Whenever the exact denominator is positive, $0 \leq \rho_i \leq 1$. The H7 envelope also maximizes over its entropy parameter and is attained at $\beta = 0$ under the frozen Step-12 contract. The coordinate-wise envelope matrix need not be jointly realizable by one common instance; it is an envelope object, not automatically an instance tensor.

Envelope clarification (M-21). For the H7 coordinate, the exact envelope is taken jointly over the frozen admissible instance family and the allowed β range, preserving the fixed context conditions; the displayed max over-I notation abbreviates this context and does not silently fix β. A normalized entry is defined only when the corresponding exact envelope is finite and strictly positive.

Canonical R_{VG} uses H_i^{MAX} itself. A Θ-order fitted asymptotic curve, certified upper envelope, sample maximum, or universe-specific empirical maximum may not silently replace the canonical denominator. If an exact denominator is unavailable outside the certified domain, the corresponding canonical normalized entry is undefined rather than replaced by an empirical surrogate.

4.5.5 Exploratory Diagnostics Across the Representation Stack

System I also defines mathematical observables that can be evaluated on one or more representation layers. They are diagnostics, not automatic complexity statements. For a numerically instantiated nonnegative 3×3 matrix A, determinant and permanent are finite matrix functionals, and the Step-12 canonical cancellation ratio is

$$R_{pd}(A) := \frac{|\det(A)|}{perm(A)}, \quad perm(A) > 0$$

For such A, |det(A)|≤perm(A), so 0≤R_pd(A)≤1 whenever perm(A)>0; if perm(A)=0, the Step-12 state is DEGENERATE-PERM-ZERO and R_pd is undefined. The earlier exploratory quantity perm(A)/(|det(A)|+ϵ) is LEGACY/NONCANONICAL and is not used as the canonical R_pd in the present architecture.

The same System-I diagnostic family includes normalized row/domain Gram geometry, the canonical cyclic 3×3×3 lift and multilinear rank, the CMSW singular-mass redistribution proxy, and the conservative orthogonal/projective orbital baseline d_{O3}. Their detailed identities, invariance properties, reference objects, contraction/projection calculus, and anti-artifact controls are developed in Paper I-B rather than reproduced here.

These finite diagnostics remain sharply bounded in interpretation. det/perm quantities of a 3×3 signature are not Valiant's determinant/permanent polynomial families; CMSW is an $O(3)^3$ spectral proxy rather than a general $GL(3)^3$ invariant; d_{O3} is an orthogonal/projective baseline rather than a GL(N) orbit-closure obstruction distance. A fixed-dimensional synthetic orbit exclusion is therefore not a computational obstruction theorem.

Statistical-physics, Fisher–Ruppeiner, and phase-transition observables may also be used to probe collective organization of TSP-derived ensembles. Random-k-SAT phase-transition results provide methodological precedent, not a transfer theorem for the canonical TSP→3-CNF family. Finite-size transitions, susceptibilities, curvature response, or cross-representation persistence belong to discovery and falsification; their controlled simulation, MCMC estimation, sensitivity analysis, finite-size scaling, and multi-universe robustness tests belong to System III.

4.5.6 Representation-Transfer and Claim Boundary

The Step-12 closure fixes the representations but does not close the theorem obligations required to convert a representation-level signal into a computational lower bound. Let E denote the frozen entropic lift from canonical provenance to T^H_{VG} and N_MAX the exact-envelope normalization from T^H_{VG} to R_{VG}. A future transfer theorem need not make these maps globally invertible, but it must prove that the particular property used by the obstruction argument is preserved or reflected under explicit hypotheses.

$$O\left(R_k(I)\right) + \text{transfer hypotheses } \Rightarrow O_{target}(I)$$

Positive coordinate wise normalization preserves the coordinate decomposition and zero/nonzero status of positive-denominator entries, but it does not generically preserve determinant, permanent, singular spectra, or GL-orbits. Likewise, a property in R1 does not imply the corresponding property in R0 without RT_Λ, and a fixed-dimensional orbital signal does not imply GL(N) orbit-closure non-membership without a Representation-Orbit Transfer theorem.

Accordingly, cross-representation persistence is at most a stronger empirical candidate for representation-robust structure. UR1/UR2 remain empirical System-III notions; UR3, Representation-Orbit Transfer, an

obstruction theorem, and the independent $obstruction-to-Cost_{STL}$ transfer F2c remain System-II theorem obligations. None is closed merely by the Step-12 theoretical/conceptual freeze.

System-I observable $\neq$ System-III robustness $\neq$ System-II obstruction theorem $\neq$ F2.

This closes the System-I representation stack at foundation level. Paper I-A remains R0-centered and claim-complete for its foundational objects; Paper I-B develops the Step-12 A–L four-representation mathematics in full. No result in this subsection proves $TSP_{dec} \notin P_{STL}$ or P$\neq$NP.

5. System II: Logical and Computational-Complexity Formal System

5.0 Scientific Role and Logical Authority

System II fixes the complexity-theoretic objects and the deductive obligations required to connect the canonical STL representation to classical P versus NP. Its theorem authority is distinct from System-I representation, System-III finite evidence, System-IV kernel certification and Program-Wide V&V governance. Mathematical closure depends on proofs of the relevant quantified claims, not on empirical robustness, literature votes, formal compilation alone or administrative release gates.

5.1 $\mathbf{Cost_{STL}}$ and the Computational-Representational Core

5.1.1 Internal Classes and Upstream Accounting Interface

The GCT-STL framework uses P_{STL} and NP_{STL} as internal computational–representational classes. They are classes of decision languages equipped with a canonical STL representation and an explicit representation-cost discipline; they are not defined by tensorial volume, determinant/permanent behavior, orbital distance, Fisher/Ruppeiner geometry, entropy profiles or the absence of a future obstruction. The definitions in this section consume the bundle-aware compileSTL/$Cost_{STL}$ interface fixed in §4.3.7 and do not introduce a second representation map. Standard-complexity bounds use n=|x|; N remains the TSP city-count parameter.

The upstream interface used here is:

$$\mathcal{P}_{STL}(M,x) = (\mathcal{P}_{MTD}(M,x), \mathcal{P}_{THM}(M,x), \mathcal{P}_{STRUCT}(M,x), SourceMap)$$

buildProvenanceBundle(M,x) $\rightharpoonup$ $\mathcal{P}$_STL(M,x),

$$compileSTL\left(\langle M\rangle, x\right) = \Phi_{STL}\left(\langle M\rangle, x, buildProvenanceBundle(M,x)\right)$$

$$Cost_{STL}(M,x) := T_M(x) + T_{buildProv}(M,x) + T_\Phi(M,x) + Size\left(STLSignature(M,x)\right)$$

As required by the System-II accounting rule, any additional downstream computation actually used by a decider or verifier must also be accounted for and cannot be hidden as free preprocessing. This accounting rule is distinct from class-membership criteria: the geometric and diagnostic observables themselves do not define P_{STL} or NP_{STL}.

5.1.2 The Class $\mathbf{P_{STL}}$

Define totality of the canonical STL representation compiler on a deterministic machine M by

$$STLTotalOn(M) :\Leftrightarrow \forall x, compileSTL\left(\langle M\rangle, x\right)\downarrow$$

A decision language L belongs to P_{STL} precisely when it has a deterministic decider whose canonical STL representation is total and whose total STL cost is polynomially bounded in the encoded input length:

$$L \in P_{STL} :\Leftrightarrow \exists DTMM, Decides(M, L) \wedge STLTotalOn(M) \wedge \exists p \in Poly, \forall x, Cost_{STL}(M, x) \leq p(|x|).$$

Thus P_{STL} is a deterministic computational–representational class. It does not mean "no obstruction exists," "determinant is small," "orbital distance vanishes," or "tensorial volume is polynomial." A later obstruction theorem may be used to prove non-membership in P_{STL}, but the property intended to establish non-membership cannot be built into the definition itself. Because $Cost_{STL}$ includes ordinary machine running time T_M, forgetting the STL layer gives $P_{STL} \subseteq$P at the level of this cost definition. The inverse inclusion P$\subseteq P_{STL}$ is not definitional; it is precisely the open statement F1.

5.1.3 The Class NP_{STL}

NP_{STL} mirrors the certificate/verifier definition of classical NP while imposing the same canonical STL representation discipline on the verifier. Let ⟨x,w⟩ denote the canonical pair encoding. Then

$$L \in NP_{STL} :\Leftrightarrow \exists V, p, q \text{ such that}$$

$$x \in L \Leftrightarrow \exists w \ \ (|w| \leq p(|x|) \wedge V(x, w) = 1)$$

and ∀x,w with |w|≤p(|x|): compileSTL(⟨V⟩,⟨x,w⟩)↓

$$\wedge \ Cost_{STL}(V, \langle x, w \rangle) \leq q(|x| + |w|)$$

The universal requirement over all candidate certificates within the polynomial length bound is deliberate: the STL representation discipline belongs to the verifier as an algorithmic object, not only to a selected accepting witness. Historical C1/C2 entropy profiles and related geometric properties may later be studied as invariants, diagnostics, or characterization candidates, but they are not necessary-and-sufficient conditions in the logical definition of NP_{STL}.

5.1.4 Anti-Circularity and Excluded Observables

The following quantities are explicitly excluded as membership criteria for both internal classes:

- determinant or permanent thresholds of $T_{VG}^{raw} or L_{VG}$;
- L_{VG} magnitude, tensorial-volume bounds, orbital distance, orbit-closure conditions, or Fisher/Ruppeiner curvature;
- D_6^{OP} calibration, $\hat{\kappa}$_SL, empirically tuned thresholds, or a candidate universal κ^*_SL;
- universal-robustness predicates or any "absence of obstruction" condition;
- historical C1/C2 entropy profiles as complete class definitions; and
- substitution of N (number of cities) for n=|x| in standard complexity bounds without a proved encoding bridge.

These exclusions do not forbid later algorithms from computing such observables or later theorems from using them as proof ingredients. They only prevent the class definitions from assuming, by construction, the same obstruction properties that later work is intended to establish.

5.2 Internal Coherence of P_{STL} and NP_{STL}

A natural internal coherence statement is

$$P_{STL} \subseteq NP_{STL}$$

The standard empty-certificate construction gives the intended proof strategy: from a deterministic STL decider M, define a verifier V that ignores the certificate and simulates M. The STL version of the argument must additionally show that the canonical verifier representation remains total and that its compileSTL/ $Cost_{STL}$ overhead is polynomially preserved. Hence $P_{STL} \subseteq NP_{STL}$ is recorded as a theorem obligation rather than assumed by definition.

This internal inclusion is distinct from the public F1–F4 route. F1 is P$\subseteq P_{STL}$; $F3_{std}$ is the standard statement $TSP_{dec}\in$NP; and F2 is $TSP_{dec} \notin P_{STL}$. Membership of TSP_{dec} in NP_{STL} would be an additional internal result, but it is not required by the minimal F1–F4 bridge.

5.3 F1-DTM and F1 — Universal Standard-to-STL Representation

5.3.0 The Minimal F1–F4 Logical Route

The GCT-STL foundation connects to the classical P versus NP problem through four public statements. The internal classes P_{STL} and NP_{STL} are defined in §5.1; the present section states the minimal logical route required for Paper I-A.

- F1: P $\subseteq P_{STL}$.
- F2: $TSP_{dec} \notin P_{STL}$.
- $F3_{std}$: $TSP_{dec}\in$ NP.
- F4: F1 $\wedge$ F2 $\wedge$ $F3_{std} \Rightarrow$ P $\neq$ NP.

F1 and F2 are open theorem obligations. $F3_{std}$ is the standard NP-membership result for decision TSP. F4 is a conditional bridge: it becomes a proof of P$\neq$NP only if the open antecedents F1 and F2 are established over compatible encodings and definitions.

Public status of the F1–F4 system (System II; current freeze)

Statement	Canonical form	Status in Paper I-A
F1	$P \subseteq P_{STL}$.	Open representation theorem obligation.
F2	$TSP_{dec} \notin P_{STL}$.	Open computational lower-bound target.
$F_{3,std}$	$TSP_{dec} \in NP$.	PASS/FROZEN — Independent standard NP-membership branch (Paper I-D).
F4	$F1 \wedge F2 \wedge F_{3,std} \Rightarrow P \neq NP$.	VALID CONDITIONAL LOGIC; antecedents F1/F2 remain OPEN.

5.3.1 F1 — Standard-to-STL Representation Bridge

F1 requires every language in P to admit a deterministic polynomial-time decider whose canonical STL representation is total and whose representation overhead remains polynomial in the encoded input length n=|x|. This is a genuine representation theorem, not a consequence of defining P_{STL}. Demonstrating the STL representation of selected algorithms or selected TSP instances is insufficient to establish the universal inclusion $P \subseteq P_{STL}$.

F1-DTM — Uniform standard-to-STL representation obligation

The constructive theorem target quantifies over every standard polynomial-time deterministic Turing machine M and requires one effective uniform STL transformation (not one bespoke encoding for each instance), preservation of accept/reject/termination semantics, canonical provenance/R0 realization, representation size polynomial in the declared input and runtime parameters, and polynomial construction/evaluation overhead charged to $Cost_{STL}$. The end-to-end bounds are measured in the standard encoded input size n=|x|, with explicit encoding bridges. Selected examples, finite simulations, and syntactically correct representations do not prove the universal statement.

$$F1\text{-DTM} \Rightarrow F1: \ P \subseteq P_{STL}$$

STATUS: F1-DTM OPEN; F1 OPEN. VV-G4 and PVV-3/PVV-5 require statement and semantic adequacy, not just compiler existence.

5.4 F2 — Universality, Obstruction and Transfer

F2 states that decision TSP does not belong to the internal polynomial $Cost_{STL}$ class:

$$F2: \ TSP_{dec} \notin P_{STL}$$

Using the definition of P_{STL} in §5.1, F2 means that no deterministic Turing machine deciding TSP_{dec} can simultaneously have a total canonical STL representation and polynomial total STL cost on all encoded inputs. Equivalently, the intended quantified statement is

$$\neg\exists M, k, c \ \left[Decides\,(M, TSP_{dec}) \wedge STLTotalOn(M) \wedge \forall x, Cost_{STL}(M, x) \leq c\ (1 + |x|)^k\right]$$

This endpoint is stated in the primary input-size variable n=|x|. A determinant/permanent ratio, orbit distance, Fisher–Ruppeiner response, tensor or entropy diagnostic, or finite empirical regularity does not establish F2 by itself. Any such structural phenomenon must first be proved under the required universality conditions and then transferred independently to a $Cost_{STL}$ non-membership result.

Bound convention. The factor (1+|x|)^k includes the empty input; equivalently, an eventual polynomial bound for sufficiently long inputs may be combined with a finite-small-input adjustment. The correction is about the quantifier domain, not a proof of F2.

5.4.1 Required theorem-level F2 chain

A candidate empirical pattern may motivate UR3, but a mathematically universal obstruction (when distributional universality is required) must be established by an independent theorem over a defined STL-admissible class. The Obstruction Theorem must then identify the relevant structural impossibility; Representation Transfer (including any representation/orbit reflection hypotheses) must connect the property back to the computational object; and F2c must independently transfer that obstruction to a $Cost_{STL}$ lower bound/non-membership statement. Route-specific exceptions to UR3 must be justified by alternate universal premises, not by finite robustness.

candidate pattern → UR3 (if required) → Obstruction Theorem → Representation Transfer → F2c → F2

STATUS: UR3 OPEN; Obstruction Theorem OPEN; Representation Transfer OPEN; F2c OPEN; F2 OPEN. These are distinct theorems/obligations, not successive administrative approvals. A 3×3 determinant, orbit signal, curvature, entropy or Full Coupling hypothesis cannot bypass the transfer gates.

5.5 $F3_{std}$ — Independent Standard NP Membership

$F3_{std}$ is the classical statement $TSP_{dec} \in$NP. A certificate is a proposed Hamiltonian tour; a verifier checks that each city occurs exactly once, that the cycle is valid, and that its total cost does not exceed K, all in time polynomial in the standard encoded input length. $F3_{std}$ is independent of the GCT-STL entropy, coupling and geometric-obstruction hypotheses.

The standard $F3_{std}$ theorem is PASS/FROZEN in Paper I-D through the canonical TSP_{dec} certificate/verifier and its standard encoding. Its proof does not depend on D1–D9, R0–R3, Full Coupling, geometric observables, UR1–UR3, F1 or F2. NP-completeness does not itself establish a lower bound.

5.6 F4 — Conditional Bridge to P $\neq$ NP

Once F1, F2 and $F3_{std}$ are stated over compatible encodings and class definitions, the minimal bridge is

$$F1 \wedge F2 \wedge F3_{std} \Rightarrow F4 : P \neq NP.$$

Proof schema. Assume for contradiction that P=NP. By $F3_{std}$, $TSP_{dec} \in NP$, hence $TSP_{dec} \in$P under that assumption. By F1, every language in P belongs to P_{STL}, so $TSP_{dec} \in P_{STL}$. This contradicts F2. Therefore P$\neq$NP. The implication is logically valid as a bridge, but Paper I-A does not complete the proof because F1 and F2 remain open.

5.7 Program-Wide V&V and Classical-Barrier Governance

The Program-Wide $G_{V\&V}$ framework (introduced in §3.3) governs all System-II claim promotion. PVV-3 audits formal and human theorem statements; PVV-5 validates semantic adequacy; PVV-6 audits the actual inferential bridge, including VV-G4/VV-G5 and Representation Transfer/F2c where relevant. A governance gate may stop or reopen a claim but cannot supply a missing proof. Finite results from System III and closed mathematical results from System I are consumed only with their exact epistemic status.

Any mature route to F2 requires a Classical-Barrier Audit covering Relativization, Natural Proofs and Algebrization where applicable, and must account for any general Kronecker positivity/orbit-closure containment or equivalent hard recognition requirement reintroduced by its proof mechanism. Operational bypass in finite System-I/III diagnostics is not theorem-level barrier escape. SMS (PVV-0) and SLR (PVV-1) can detect literature and frame evidence but cannot close an OPEN mathematical theorem.

System-IV formalization is assessed under PVV-7 for statement correspondence, dependencies/axioms, no-sorry footprint and independent clean builds. Critical N4 releases require PVV-8 independent mathematical/formal audit and PVV-9 provenance/release manifests; AI is permitted as assistant but never sole certifier.

5.8 F_{comp} — Metatheoretical Interface

Beyond the mathematical representation layer developed in Paper I-A, the GCT-STL program includes a metatheoretical framework denoted F_comp. In the present paper, F_comp is an interface declaration rather than a premise: T_{VG}^{raw}, the Provenance Bundle and SourceMap, the $Cost_{STL}$-based definitions of P_{STL} and NP_{STL}, and the conditional F1–F4 bridge are stated independently of F_comp. Paper I-A therefore does not use F_comp to strengthen, replace, or justify F1 or F2.

The detailed axiomatic status of F_comp, its intended relation to ZFC, any Kolmogorov-visibility or related metatheoretical axioms, and any independence-oriented interpretation of P versus NP belong to Paper I-C (System II) and dedicated metatheoretical work. Cross-reference to F_comp in Paper I-A serves only to locate the foundation presented here within the broader program and to prevent later metatheoretical developments from being read as hidden assumptions of the present tensorial construction.

6. System III: Computational-Heuristic-Experimental and Simulation System

6.0 Scientific Role, Authority, and Research Questions

System III is the computational and empirical discovery system of GCT-STL. It explores the mathematical objects defined by System I under the formal questions set by System II, aiming to discover, challenge, calibrate, and falsify candidate structural mechanisms. Simulation, Monte Carlo sampling, information-geometric response and cross-representation diagnostics may support a research conjecture, but never supply a universal theorem merely through repeated numerical success. In particular, the OPEN Full STL Tensorial Coupling Principle motivates experimental questions without becoming an additional antecedent of the F1–F4 conditional bridge.

Its scientific questions are operational: Which observables are stable under changing the data-generating universe? Which regularities survive changes of N, ensemble parameters, encodings, estimators and representation? Which apparently strong signals disappear under negative controls or adversarial instances? Which surviving patterns warrant a mathematically precise conjecture for System II? The right output of System III is a reproducible evidence and falsification dossier, not a claim that TSP_{dec} has acquired a proven non-polynomial lower bound.

6.1 Canonical Inputs and the Four Representation Views

Experiments begin from the frozen canonical decision instance I=(N,D,K), its explicit encoding x=enc(I), the materialized Boolean/reduction artifacts where applicable, canonical D1–D9 source contracts, and the provenance-controlled representation R0. The cost matrix D is an explicit nonnegative cost object in the canonical formulation; symmetry, triangle inequality or Euclidean realizability are not generally imposed. Geometric assumptions enter only when an experimental generator expressly supplies them. CAN and OP observables must retain their separate definitions and source tags, and no calibrated descriptor may be substituted for its frozen canonical counterpart.

$$R0 = T_{VG}^{raw}; \quad R1 = L_{VG}; \quad R2 = T_{VG}^{H}; \quad R3 = R_{VG}$$

$$R0 \to R1; \qquad R0 \to R2 \to R3$$

The architecture is branched rather than a linear four-step transformation. R1 provides complementary componentwise logarithmic scaling; R2 is the information-functional representation; R3 is the exact-envelope normalization of R2. They are not automatically equivalent. Each test must identify the representation, coordinate basis, normalizer, source/version, relevant parameter values and estimator. A signal observed in R1 cannot be silently attributed to R0; an R3 quantity must use its declared exact coordinatewise H_iˆMAX(N,B) denominator rather than a dataset maximum or fitted scale. The complete mathematics remains in Paper I-B; Step 12J supplies the closed comparative design, not executed Stage-II results.

6.2 Experimental Universes and Comparative Design

The planned comparative programme uses three deliberately different generator families: Euclidean instances (spatial/metric structure), i.i.d. cost instances (independently generated entries under the declared law), and Cluster instances (multimodal or community-like structure). Each family requires its own documented generator, constraints, distributional parameters and reproducible seeds. Comparisons must control N, input encoding, admissible cost scales, threshold K selection and numerical precision; a change in apparent obstruction may otherwise be a generator, scale or serialization artifact rather than a property of computation.

Experiments should preregister the relevant strata of N, generator regimes, sample allocation, independent replications and analysis comparisons in their separate Stage-II protocol. These are requirements for future execution, not parameter settings retroactively claimed by Paper I-A. Exact small-N enumeration, where computationally feasible, supplies reference cases for validating instance construction, descriptor evaluators and sampling routines. Failed conditions, null responses, contradictory universes and removed instances must be reported with their provenance rather than selectively suppressed.

Agreement across Euclidean, i.i.d. and Cluster cases can substantiate only finite multi-universe robustness under the operationalized sampling contract. The non-circular class of STL-admissible universes (UR0), and proofs that these generators lie within it or appropriately approximate it, require independent mathematical specification. Three benchmark families cannot be treated as universal quantification over all relevant computational instances or machines.

6.3 Statistical Physics, Sampling, Uncertainty, and Finite-Size Analysis

The information-theoretic pathway includes the fixed-origin directed-tour Shannon–Boltzmann ensemble underlying canonical D7 and, where justified, inverse-temperature β, partition-function or response diagnostics. Proposed Monte Carlo/MCMC and tempering procedures must specify their target distributions, kernels, stationary-measure assumptions, burn-in/initialization policy and estimator definitions; they must verify balance or other invariance requirements and report autocorrelation, effective sample size, independent-chain agreement, mixing concerns and uncertainty quantification as appropriate. An attractive trace alone is not evidence that a sampler has explored the target distribution.

Finite-size analyses should compare a declared N grid and parameter regime, show estimator errors and sensitivity bands, and distinguish a finite scaling fit from a proved asymptotic law. Fisher- or Ruppeiner-type metrics/curvatures must be attached to a specified ensemble and parameter space, with numerical-derivative choices and degeneracies disclosed. Apparent critical response, sharp crossover or singular-looking finite behavior may motivate a candidate mechanism; it does not imply a deterministic time lower bound. Detailed kernels, tuning, sample sizes, β/N schedules and preregistrations belong to Paper II-A rather than being invented in Paper I-A.

6.4 Exploratory Algebraic, Tensorial, Orbital, and Information-Geometric Diagnostics

The diagnostic portfolio may include determinant and permanent proxies, normalized Gram/angular structure, cyclic tensor constructions and multilinear rank, CMSW spectral summaries, restricted orthogonal/projective orbital baselines such as d_{O3}, and information-geometric quantities. Each diagnostic must name its input matrix/tensor and representation Rk, dimensions, hypotheses, algorithm, uncertainty treatment and symmetry/invariance group; comparing values from different representations requires a declared mapping or controlled paired experiment.

$$R_{pd}(A) = \frac{|\det(A)|}{perm(A)}, \qquad \text{when } perm(A) > 0$$

For the Step-12H canonical ratio R_pd, record zero-denominator cases explicitly; with a nonnegative 3×3 matrix A, $0\leq R_{pd}\leq 1$ whenever defined. The formerly used inverse ϵ-regularized permanent-to-determinant expression is LEGACY/NONCANONICAL and cannot be renamed R_{pd}. Likewise CMSW and finite d_O3 diagnostics operate under their stated restricted groups and are not general GL(N) orbit-closure obstructions. Persistence of an algebraic or geometric proxy does not establish classical GCT separation, Full Coupling, UR3, F2c or F2.

6.5 Negative Controls, Ablations, Sensitivity, and Falsification

Before promoting a pattern, System III must actively attempt to break it. Required controls include generator-matched nulls, permuted/reshuffled structures where mathematically meaningful, encoding and cost-scale perturbations, normalization sensitivity, adversarial counterexamples, exact small-N identity checks and alternative independent estimators. Each ablation must declare the operator, retained dimension and diagnostic that remains meaningful after the operation; reporting a lower-dimensional invariant as though it were a full 3×3 determinant is prohibited.

Step 12I supplies a decisive anti-artifact constraint. Deleting or zeroing an entire row or column forces det(A)=0 by elementary linear algebra; this collapse has no evidentiary value for a conjectured computational obstruction. A single-cell ablation (I−E_(i, α)) is the preferred minimal test of determinant/permanent sensitivity when the retained full matrix remains meaningful. Whole-row/column removal may still support suitable lower-dimensional Gram/spectral analyses, but not a 3×3 determinant-obstruction argument.

A candidate must be revised, downgraded or rejected when a generator change reverses it, an exact counterexample appears, it depends on a sample maximum mislabelled as an exact envelope, or its apparent strength is explained by finite precision, chain nonconvergence, rank-forcing ablation, or uncontrolled encoding. Failed replications and null discoveries are valid outputs of System III and must survive into its released record.

6.6 Evidence Grading, UR1/UR2, and the System-II Handoff

Evidence status is tracked explicitly: OBSERVATION denotes a finite measurement; CANDIDATE denotes a stated, still-falsifiable proposed regularity; UR1 denotes empirical multi-universe robustness under a documented protocol; UR2 denotes a strong candidate-universality dossier that organizes reproducible evidence and hypotheses but is not itself a universal theorem. The progression from UR1 to UR2 is methodological and conjectural, not a logical implication. No number of finite confirmations promotes either state automatically to UR3.

$$\text{finite evidence / UR1 / UR2} \not\Rightarrow \text{UR3} \not\Rightarrow \text{F2}$$

When a candidate is sufficiently precise, System III may hand System II its quantifiers, tested universes, counterexamples, invariance qualifications, estimated uncertainties and unresolved assumptions for a new theorem statement. System II alone owns the independent UR0/UR3 specification and proof (where required), Obstruction Theorem, Representation Transfer and F2c, leading, if proved, to F2 (TSP_{dec} is not in P_{STL}). RAP requires semantic fidelity and an explicit property-specific inference from representation-level structure to computational resources. Experimental evidence, even across all four representation views and all three generators, never bypasses these gates.

6.7 Reproducibility, Program-Wide $\mathbf{G_{V\&V}}$, $\mathbf{A_{AI}}$, and Publication Handoff

The Program-Wide $G_{V\&V}$ framework of §3.3 governs the System-III lifecycle without constituting another scientific system. PVV-0 SMS and PVV-1 SLR monitor external evidence and scoped methodological questions; PVV-2 Concept Mapping and PVV-3 Statement Audit record concepts and predeclared experimental claims; PVV-4 verifies generators, evaluators, scripts and sampling artifacts; PVV-5 assesses scientific/measurement validity; PVV-6 blocks unsupported transfers; PVV-7 applies when code/specifications receive formal certification; PVV-8 requires proportionate independent replay of high-risk results; and PVV-9 governs release, impact review and revision.

The A_{AI} layer of §3.2 may assist simulation design, code creation, data checking, adversarial attacks, diagnostics and symbolic hypothesis generation. Its contributions require AI-LOG and human review; high-risk results may not be self-certified by the same AI process that produced them. A valid evidence package records canonical instance IDs, generator and parameter definitions, seeds, source hashes, software/toolchain,

normalization choices, raw and processed data, uncertainty, null and negative outcomes, preregistration deviations and independent-replay status.

Step-12J is CLOSED as a theoretical/conceptual four-layer comparative protocol; its production execution remains DEFERRED to Stage II. Paper II-A is responsible for full statistical-physics/MCMC methodology; Paper II-B+ will report preregistered comparative data, uncertainty and any UR1/UR2 evidence actually obtained. Paper I-C supplies the System-II/III conceptual bridge, while later Paper III work concerns UR3, universal obstruction and transfer theorems. Those editorial allocations do not change mathematical truth conditions or imply that experiments have already been executed.

7. System IV: Mathematical and Symbolic Formalization and Certification System

System IV provides the transverse mathematical and symbolic formalization/certification interfaces of GCT-STL. It formalizes precisely declared objects, algorithms and theorems arising in Systems I–III, checks derivations under declared formal assumptions, and records the proof artifacts that support those derivations. System IV is a scientific subsystem, not an implementation synonym for Lean 4, not the Program-Wide $G_{V\&V}$ governance layer of §3.3, and not the A_AI assistance layer of §3.2. These interfaces are intentionally presented at architecture level; Paper I-E develops their general formal details.

7.0 Scientific Role and Metalogical Boundary

Formal theorem proving can establish that a statement follows from precisely encoded definitions and axioms. It does not, merely by accepting a proof term, establish that the encoded proposition is the intended human claim or that the claim suffices for a computational lower bound. System IV therefore certifies formal derivations; $G_{V\&V}$ assesses claim identity, validity, transfer, risk and publication readiness. Mathematical truth conditions remain those of the corresponding Systems I and II, and experimental findings remain System-III evidence.

7.1 Formal Specifications, Statements, and Encoding Contracts

The formalization process begins with an explicit mathematical target: quantified domain, hypotheses, types, intended conclusion, and dependency footprint. For canonical TSP_{dec}, the F0 instance I=(N,D,K) uses an explicitly encoded nonnegative cost matrix D and threshold K; no symmetry, triangle inequality, metric structure or Euclidean realization is silently assumed. Encoded input size is n=|x| for x=enc(I), and size/cost, certificate, verifier and language semantics must match their frozen standard or STL definitions.

The Double Statement Audit separates a pre-proof Mathematical Statement Audit from the post-encoding Formal Statement Correspondence Audit. The first checks that the target proposition and its premises are precisely those needed; the second checks that the Lean theorem, its implicit arguments and local definitions actually correspond to that target. A valid proof of a related or weakened statement cannot be silently promoted to the intended scientific claim.

7.2 Formalization Targets and Interfaces by Scientific System

System-I targets include well-formed canonical artifacts, D1–D9/source contracts, provenance identities, the R0 construction, appropriate R0–R3 and exact-envelope lemmas, and representation-specific semantics. The existence of a definition or Step-12 conceptual derivation does not assert that every associated uniformity or adequacy theorem is formally closed.

System-II targets include the constructive universal F1-DTM theorem and F1, theorem-level UR3 where needed, the Obstruction Theorem, Representation Transfer, F2c, and F2. These remain OPEN. The valid

conditional implication $F1 \wedge F2 \wedge F3_{std} \Rightarrow F4{:}P \neq NP$ does not close its two open antecedents. The separate $F3_{std} : TSP_{dec} \in NP$ standard branch has the programme status PASS/FROZEN and is assigned to the specialized Paper I-D companion, not to the general architecture of Paper I-E.

System-III formalization may cover exact ensemble/descriptor evaluators, sampling-kernel specifications, generator contracts, deterministic test oracles, bounded numerical properties and reproducible build steps. Such certification does not transform a finite MCMC result, a robust empirical pattern, or a candidate obstruction into UR3 or a complexity lower bound. The Step-12J comparative protocol is CLOSED at conceptual level but its Stage-II production execution is DEFERRED.

7.3 Lean 4 and the Dependency Architecture

Lean 4 and its mathematical libraries are principal implementation environments, not the definition of System IV. A formal release should maintain canonical theorem and module identifiers, import graph, declaration dependencies, source/encoding references, an exact theorem-statement inventory and a reproducible toolchain. The actual release must audit its current module tree, rather than treating an obsolete historical file list as normative.

The formal dependency graph must distinguish standard complexity foundations and the independent $F3_{std}$ branch from GCT-STL representation definitions and open lower-bound obligations. A Lean namespace, proof search result or accepted file does not establish logical independence, a missing representation transfer, or any stronger claim than its checked theorem.

7.4 Proof Scope, Axiom Footprint, and Adversarial Audit

For a critical theorem, the proof scope includes explicit imports, custom axioms (if any), theorem-local assumptions, classical axioms invoked, compiler/toolchain identity, and the source-level dependency footprint. The absence of sorry/sorryAx is necessary for a clean critical proof package but insufficient for scientific validation: an unchecked assumption, altered encoding, vacuous implication or unintended theorem statement can survive a compilation-only review.

Adversarial checking should seek counterexamples, scope narrowing, circular definitions, silent changes of input size, hidden oracle assumptions, and false passage from geometric observables to computational claims. Barrier relevance (relativization, natural proofs, algebrization and relevant classical GCT recognition bottlenecks) must be addressed by any mature System-II F2 strategy; formalization alone does not constitute barrier escape.

7.5 Independent Reproduction and Controlled Release

Critical proof releases require PVV-7 formal certification, PVV-8 appropriately independent audit and PVV-9 controlled release under the PI-approved Program-Wide $G_{V\&V}$ framework. Release evidence comprises a mathematically pre-audited statement, a formal statement-correspondence record, axiom/import audit, pinned Lean/Mathlib and supporting tools, clean rebuilding, source and proof hashes, reviewed exceptions, and an auditable release manifest. An internal AI rerun of its own proof is not strongly independent audit.

Proof Provenance and Process/Release Provenance must record who or what contributed definitions, code, conjectures, counterexamples, proofs and review decisions. AI-assisted high-risk (N4) work is permitted for discovery, adversarial search and formalization, but AI cannot self-certify the critical claim it produced.

7.6 Boundaries with $\mathbf{F_{comp}}$, $\mathbf{G_{V\&V}}$, and $\mathbf{A_{AI}}$

F_comp belongs to the separate System-II metatheoretical interface and is not a premise of the ordinary F1–F4 separation route. The failure of proof search, an unsuccessful Lean compilation or an incomplete

formalization is not evidence of independence from ZFC. Proving a ZFC-related independence statement would require its own precisely formulated and appropriately justified metatheoretical argument.

System IV certifies encoded derivations; $G_{V\&V}$ governs statement correspondence, semantic adequacy, transfer validation, independent review and releases; $A_AI$contributes discovery and formalization assistance without acquiring proof authority. In editorial terms, Paper I-D records the specialized independent $F3_{std}$ certification branch, while Paper I-E is the general System-IV architecture companion. Neither paper substitutes for the open mathematical work on F1, UR3, Representation Transfer, F2c or F2.

8. Research Programme and Companion-Paper Architecture

8.0 Research Programme

Paper I-A is the mathematical-foundation article. The companion series separates the complete System-I representation mathematics, the Systems-II/III theoretical-conceptual architecture, the independent standard $F3_{std}$ formal branch, general System-IV architecture, future System-III empirical methodology/execution, and downstream System-II theorem work. The allocations below are editorial scope boundaries, not mathematical premises or evidence of completed theorems. This §8 is the single authoritative publication crosswalk; earlier sections reference it rather than duplicating the table.

8.1 Canonical Publication Crosswalk and Release Boundaries

Each destination has a specific scientific responsibility and an explicit non-duplication rule. In particular, I-D is distinct from I-E; I-C covers the theoretical-conceptual architecture of both Systems II and III; and the Stage-II empirical programme is divided between methodology and execution. Neither publication allocation nor a released companion is a substitute for proof of a still-open theorem.

Paper	Canonical function	Material retained / developed there	Boundary relative to Paper I-A
I-A	Mathematical foundation	D1–D9; canonical artifacts; SourceMap/Provenance; R0; Theorems 1–3; concise R0–R3, F1–F4 and systems map.	PRIMARY foundation. No P≠NP proof, no complete Step-12 derivation, and no unexecuted Stage-II evidence.
I-B	Complete System-I representation mathematics	Step 12A–L; branched R0–R3; H1–H9, canonical envelopes; Gram, cyclic, orbital and related diagnostics; contraction/projection calculus.	SYSTEM-I DETAIL. Representation statements are not automatically transferable complexity lower bounds.
I-C	Systems II + III theoretical-conceptual architecture	$Cost_{STL}, P_{STL}/NP_{STL}, F1-DTM/F1, F2, UR3$, Obstruction Theorem, Representation Transfer, F2c, F_{comp}; scientific discovery and inference governance.	THEORETICAL-CONCEPTUAL HOME. Not a report of Stage-II execution, and no theorem is closed by editorial assignment.
I-D	Specialized independent standard formal companion	$F3_{std} : TSP_{dec} \in NP$ under the canonical F0/G2 encoding; standard certificate/verifier; Lean and reproducibility package.	INDEPENDENT FORMAL BRANCH. Not the general System-IV architecture or a proof of F1/F2.
I-E	General System-IV architecture and certification	Formal specifications; interfaces across Systems I–III; Lean modules, axioms/imports, theorem dependencies, Statement Audit, release and certification governance.	SYSTEM-IV COMPANION. References I-D without duplicating its specialized $F3_{std}$ proof; certification ≠ scientific V&V.
II-A	System-III operational methodology	Shannon–Boltzmann, Fisher–Ruppeiner, MCMC, preregistration, negative controls, convergence, finite-size analysis and uncertainty protocols.	METHODS, NOT EXECUTION. Do not imply that Stage-II comparative production data exist.
II-B+	System-III Stage-II execution and empirical evidence	Preregistered multiverse experiments, R0–R3 comparisons, finite-size outcomes, UR1/UR2 and candidate obstruction diagnostics if/when executed.	EMPIRICAL LAYER. Step-12J production remains DEFERRED; no UR3 theorem follows from finite experiments.
III-A/B	Downstream System-II theorem programme	UR3 as required, universal obstruction theorem, Representation Transfer, F2c/F2 and possible later System-IV certification.	THEOREM OBLIGATIONS OPEN until proved; no implied resolution of P≠NP.

The eight destinations form a publication plan rather than a proof dependency graph. Before external release of the specialized Paper I-D, its historical §2 publication-role text must be aligned to this canonical I-A–I-E crosswalk; that editorial correction does not change the $F3_{std}$ proof package or its PASS/FROZEN programme status.

The Program-Wide G_V&V framework governs pre-publication Statement Audit, source/proof provenance, review independence and release controls for the relevant artifact. In particular, the Paper-II experimental series cannot be presented as an executed study before its methods, preregistration, data and independent audit have been produced.

9. Open Obligations and Discussion

9.0 Antecedents and Open Targets

The reorganized architecture makes explicit which components are established mathematical definitions or theorem-level results, which are methodological antecedents, and which remain open research targets. The following discussion preserves those boundaries while locating GCT-STL relative to its structural genealogy and to classical GCT.

9.1 Scope Boundary of the González Homologation

The González (1995) correspondence is a structural homologation, not an isomorphism. No map preserving all operations, distances, measures or volumes has been proved, and the homologation is not a source of D_i values. It is also not a premise for F1, F2, $F3_{std}$, any complexity-class separation or any lower-bound theorem. Such claims must be derived from the canonical descriptor definitions, the Provenance Bundle and independent bridge/obstruction theorems. This preserves the genealogical role of González (1995) without allowing analogy to substitute for mathematical provenance.

9.2 Methodological and Epistemological Position Relative to Classical GCT

GCT-STL is inspired by the obstruction philosophy of classical Geometric Complexity Theory, but the two programmes have different immediate objects and complexity interfaces. Classical GCT addresses algebraic complexity through representation-theoretic and orbit-closure methods. GCT-STL, by contrast, keeps standard Turing computation and the intermediate $Cost_{STL}$-based class P_{STL} as its computational interface. Consequently, no equivalence between the two settings is assumed or required in Paper I-A.

The Full STL Tensorial Coupling principle supplies the primary structural hypothesis guiding this diagnostic layer: a candidate obstruction is sought in the full coupled S–T–L signature represented by the provenance-controlled R0, not assigned a priori to semantic, computational or logical syntax alone. This hypothesis motivates joint observables and controlled interaction studies, but it is not an additional antecedent of F4 and is not itself an obstruction theorem.

The determinant and permanent of the finite 3×3 logarithmically transformed representation R1 = L_{VG} are therefore not Valiant's permanent and determinant, and their ratio is not a formal GCT obstruction. Likewise, orbit distance, Fisher–Ruppeiner observables, entropic contrasts, and related normalized/tensor diagnostics are discovery or diagnostic quantities until separate theorems establish both their required universality and their computational consequence.

Finite experiments have the same limited epistemic role. Monte Carlo and finite-size computations may provide empirical evidence, numerical checks, phase-transition signals or falsification tests for candidate mechanisms; they do not certify a uniform lower bound over an infinite problem family. If a candidate obstruction is discovered empirically, it must first be reconstructed as a mathematically universal statement under explicit hypotheses and then connected by an independent transfer theorem to non-membership in polynomial $Cost_{STL}$.

Full-coupling hypothesis/evidence $\rightarrow$ mathematically universal obstruction theorem $\rightarrow$ independent $obstruction - to - Cost_{STL} transfer \rightarrow F2 : TSP_{dec} \notin P_{STL}$.

The methodological sequence is therefore discovery → mathematical reconstruction → theorem with explicit hypotheses → formal verification. This sequence separates finite discovery evidence from the theorem-level universality and transfer results required for a complexity lower bound.

The central epistemic risk of GCT-STL is representational rather than merely numerical. T_{VG} is intentionally a compact structured signature of a much larger computational object, so distinct instances or executions may share identical or nearby descriptor configurations. Such compression is not a defect by itself; a scientific representation need not be informationally exhaustive. The decisive question is whether the information discarded by the representation is irrelevant to the theorem that is later used to transfer an obstruction to computational complexity. GCT-STL therefore treats failure of semantic adequacy or failure of representation-to-cost transfer as explicit falsification/stop conditions rather than as effects to be repaired by retuning downstream observables.

Methodological comparison

Characteristic	**Classical GCT**	**GCT-STL in Paper I-A**
Primary complexity interface	Algebraic complexity classes and orbit-closure questions.	Standard Turing classes connected through the intermediate $Cost_{STL}$-based class P_{STL}.
Role of determinant/permanent	Central algebraic objects in the classical programme.	3×3 L_{VG} determinant/permanent are numerical diagnostics/proxies only; not Valiant's permanent-vs-determinant problem.
Geometric observables	Representation-theoretic/orbit-closure obstruction machinery.	Orbit, Fisher, normalized tensor and entropic observables are candidate diagnostics until independent theorems justify them.
Experimental computation	Not the defining proof interface.	Finite simulation/MCMC provides empirical evidence or falsification tests, not uniform lower-bound certification.
Required transfer to a lower bound	Provided by the relevant algebraic obstruction theorem in its own setting.	Requires an independent theorem transferring an established STL obstruction into a $Cost_{STL}$ lower-bound/non-membership result.
Status in this paper	Background methodological reference.	Programmatic inspiration plus explicit claim boundaries; no equivalence theorem is asserted.
Primary obstruction hypothesis	Obstruction is formulated in the algebraic/representation-theoretic objects defined by the classical programme.	Open/falsifiable Full STL Tensorial Coupling thesis: a candidate obstruction is investigated on the existing provenance-controlled R0 (3×3) signature, rather than assigned to one marginal domain. S⊗T⊗L denotes only a prospective, as-yet-unconstructed third-order lift.

Classical complexity barriers and classical GCT recognition bottlenecks therefore remain theorem-level governance questions, not obstacles automatically overcome by choosing geometric, tensorial or information-geometric coordinates. In §5, the Classical-Barrier Audit is attached to any mature F2 strategy. Finite R0–R3 diagnostics, UR1/UR2 evidence and operational avoidance of general Kronecker/orbit-closure recognition tasks are not proofs of a non-relativizing, non-natural or non-algebrizing lower-bound method.

9.3 $\mathbf{F_{comp}}$ and Axiomatic Scope

F_comp is the System-II metatheoretical and axiomatic-scope workstream, developed conceptually in Paper I-C (see §8), not a hidden antecedent of the ordinary GCT-STL route. The validity of the canonical representation T_{VG}^{raw}, the descriptor theorems, the $Cost_{STL}$-based internal classes, and the conditional implication F1 ∧ F2 ∧ $F3_{std}$⇒ P ≠ NP does not require F_comp or a result on independence from ZFC. Conversely, a future metatheoretical analysis cannot retrospectively transform empirical diagnostics or an OPEN F2 target into a theorem. Failure of Lean compilation or proof search is not a proof of undecidability or independence from ZFC.

10. Conclusions

This work establishes a three-dimensional, provenance-controlled entropic representation of TSP_{dec} through reduction to 3-SAT and its canonical 3-CNF artifact. The model distinguishes logical syntax (L), geometric semantics (S) and deterministic computational dynamics (T), but treats their organization as a coupled representational system rather than as a pre-assigned causal decomposition of NP-hardness.

The theorem-level results are deliberately narrower. Theorem 1 canonically transports the González (1995) decisional-level entropy law to the exact materialized syntax: $H_{sint}^{CAN}(\mathrm{I}) \equiv D_9^{CAN}(\mathrm{I})$ satisfies $0 \leq H_{sint}^{CAN}(\mathrm{I}) < 2$ and approaches 2 bits as the D8 artifact grows. Theorem 2 defines $D_7^{CAN}(\mathrm{I},\beta)$ as Shannon–Boltzmann entropy on the fixed-origin directed-tour ensemble, with maximum envelope $\Theta(\mathrm{N}\log\mathrm{N})$. Theorem 3 separates D_5^{CAN} from procedure-dependent D_5^{OP} and, on the stated natural polynomially bounded family, gives $D_5^{CAN}(\mathrm{I}) = \Omega(\mathrm{N}^2)$ and $\mathrm{O}(\mathrm{N}^3 \log \mathrm{N})$, with $0 \leq D_5^{OP}(\mathrm{A},\mathrm{I}) \leq D_5^{CAN}(\mathrm{I})$. None of these results is a universal deterministic-time lower bound.

Taken together, these results establish distinct asymptotic descriptor-envelope scales: $\mathrm{O}(1) \ll \Theta(\mathrm{N}\log\mathrm{N}) \ll \Omega(\mathrm{N}^2)$. This is a statement of marginal informational anisotropy, not a pointwise inequality among all finite-instance entropies and not a causal hierarchy L<S<T or T>S>L. The bounded syntactic level entropy, the semantic maximum envelope and the canonical Boolean-search entropy can each participate in cross-domain interactions without any single coordinate being declared the source of hardness.

The nine descriptors are assembled only after their sources have been made explicit through the multi-source Provenance Bundle and SourceMap, yielding T_{VG}^{raw}. Any L_{VG} object is the downstream componentwise logarithmic scale complement R1; the exact-envelope normalized representation is R3 = R_VG. Determinant/permanent values, orbital distances, Fisher–Ruppeiner quantities, entropy contrasts and finite MCMC estimates are diagnostic or discovery objects unless independent theorems establish stronger roles. No tensor-volume, non-collapsibility or proxy criterion defines P_{STL} or NP_{STL}.

The primary structural interpretation is the open and falsifiable Full STL Tensorial Coupling hypothesis. It asks whether a genuine obstruction, if one exists in this framework, emerges from non-additive structure in the complete provenance-controlled R0 signature with coupled S–T–L domains rather than from S, T or L separately. The concrete decomposition $\Phi_\mathrm{I} = \Phi_{HAM} \wedge \Phi_{EDGE} \wedge \Phi_{DIST}$ motivates that question but does not prove it. D_6^{CAN} remains a specific S↔L coupling coordinate, not the complete coupling and not an obstruction theorem.

The complexity-theoretic endpoint is correspondingly explicit. P_STL and NP_{STL} are $Cost_{STL}$-based computational–representational classes. F1:P$\subseteq P_{STL}$ and F2: $TSP_{dec} \notin P_{STL}$ remain open. $F3_{std}$: $TSP_{dec} \in$NP is the standard antecedent. Thus F4:F1$\wedge$F2$\wedge F3_{std} \Rightarrow$P$\neq$NP is a valid conditional bridge, not a proof completed in this paper. Candidate coupled or geometric evidence can contribute to F2 only after it has been promoted to an appropriately universal theorem and independently connected to $Cost_{STL}$ non-membership.

The programme is subject to the PI-approved transverse G_V&V lifecycle (§3.3), not only the specialized System-I representation V&V profile. The canonical object and its provenance are defined, but definition is not full scientific validation. Open obligations include specification/edge-case completeness, polynomial realization and uniformity where required, semantic fidelity, and the independent Representation Transfer and F2c bridge necessary for F2. PVV Statement Audit, inference validation, formal certification, independent audit and release control assess these obligations without substituting for their mathematical proof. System-III experiments may support or falsify candidate mechanisms, but do not establish universal theorem claims.

The programme accordingly preserves discovery and falsification → empirical multi-universe robustness → mathematical reconstruction with explicit hypotheses → theorem-level universality where required → obstruction and $Cost_{STL}$ transfer → formal certification and independent review. The publication architecture is recorded once in §8: I-A provides the mathematical foundation, I-B develops complete System-I R0–R3

mathematics, I-C develops the Systems-II/III theoretical-conceptual architecture, I-D contains the specialized $F3_{std}$ formal companion, and I-E the general System-IV architecture. Paper II-A covers methodology, II-B+ future Stage-II empirical execution, and III-A/B the later open theorem/transfer route. This division does not change the logical dependency of F1–F4.

Accordingly, the contribution of Paper I-A is not a claimed resolution of P versus NP but a mathematically disciplined representation layer: precisely bounded entropy descriptors, explicit provenance, a canonical 3×3 STL signature, non-circular internal class definitions, an open coupled-obstruction hypothesis and a precise public map of the remaining F1–F4 obligations. The central complexity claims remain explicitly open for subsequent stages of the GCT-STL research programme.

About the Author

Renato R. González holds a bachelor's degree in mathematics and statistics from Pedro Henríquez Ureña University (UNPHU), Dominican Republic (1980), and a degree in Computer Programming and Statistics from the same institution (1975). He holds a master's degree in economics from the Pontifical Catholic University Madre y Maestra (PUCMM) (1999). With over 50 years of professional experience in teaching and research in Information Technology, software engineering, artificial intelligence, and data science, he is a Senior Lifetime Member of the IEEE Computer Society and the Computational Intelligence Society. He is Academic Coordinator of the Master's Program in Data Science and Principal Investigator of the CogniSofIA project at the Instituto Tecnológico de Santo Domingo (INTEC).

Email: renato.gonzalez@intec.edu.do

References

Aaronson, S., & Wigderson, A. (2009). Algebrization: A new barrier in complexity theory. ACM Transactions on Computation Theory, 1(1), Article 2. https://doi.org/10.1145/1490270.1490272

Alder, A., & Strassen, V. (1981). On the algorithmic complexity of associative algebras. Theoretical Computer Science, 15(2), 201–211.

Amari, S.-I. (1985). Differential-geometrical methods in statistics. Springer.

Amari, S.-I. (2016). *Information geometry and its applications*. Springer. https://doi.org/10.1007/978-4-431-55978-8

Arora, S., & Barak, B. (2009). *Computational complexity: A modern approach*. Cambridge University Press.

Baker, T., Gill, J., & Solovay, R. (1975). Relativizations of the P = ? NP question. SIAM Journal on Computing, 4(4), 431–442. https://doi.org/10.1137/0204037

Bläser, M., Ikenmeyer, C., Lysikov, V., Pandey, A., & Schreyer, F.-O. (2021). On the orbit closure containment problem and slice rank of tensors. In Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA) (pp. 2565–2584). Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611976465.152

Bunge, M. (1977). *Treatise on basic philosophy: Vol. 3. Ontology I: The furniture of the world*. D. Reidel Publishing Company. https://doi.org/10.1007/978-94-010-9924-0

Bunge, M. (1983). *Treatise on basic philosophy: Vol. 5. Epistemology & methodology I: Exploring the world*. D. Reidel Publishing Company. https://doi.org/10.1007/978-94-009-7027-4

Bürgisser, P., Christandl, M., Mulmuley, K. D., & Walter, M. (2017). Membership in moment polytopes is in NP and coNP. SIAM Journal on Computing, 46(3), 972–991. https://doi.org/10.1137/15M1048859

Cook, S. A. (1971). The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing* (pp. 151–158). Association for Computing Machinery. https://doi.org/10.1145/800157.805047

Cook, S. A. (2000). *The P versus NP problem.* Clay Mathematics Institute.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Wiley.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness.* W. H. Freeman.

González, R. R. (1995). A unified metric of software complexity: Measuring productivity, quality, and value. *Journal of Systems and Software, 29*(1), 17–37. https://doi.org/10.1016/0164-1212(94)00126-8

Harrison, W. (1992). An entropy-based measure of software complexity. IEEE Transactions on Software Engineering, 18(11), 1025–1029.

Håstad, J. (1990). Tensor rank is NP-complete. Journal of Algorithms, 11(4), 644–654.

Ikenmeyer, C., Mulmuley, K. D., & Walter, M. (2017). On vanishing of Kronecker coefficients. Computational Complexity, 26(4), 949–992. https://doi.org/10.1007/s00037-017-0158-y

Institute of Electrical and Electronics Engineers. (2025). IEEE standard for system, software, and hardware verification and validation (IEEE Std 1012-2024). IEEE. https://doi.org/10.1109/IEEESTD.2025.11134780

Karp, R. M. (1972). Reducibility among combinatorial problems. In R. E. Miller, J. W. Thatcher, & J. D. Bohlinger (Eds.), Complexity of computer computations (pp. 85–103). Plenum Press. https://doi.org/10.1007/978-1-4684-2001-2_9

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (Technical Report EBSE-2007-01). Evidence-Based Software Engineering. https://ebse.webspace.durham.ac.uk/ebse-bibliography/guidelines-for-performing-systematic-literature-reviews-in-software-engineering/

Kleene, S. C. (1952). Introduction to metamathematics. North-Holland Publishing Company.

Monasson, R., Zecchina, R., Kirkpatrick, S., Selman, B., & Troyansky, L. (1999). Determining computational complexity from characteristic 'phase transitions'. Nature, 400(6740), 133–137. https://doi.org/10.1038/22055

Mostow, G. D., & Sampson, J. H. (1969). Linear algebra. McGraw-Hill.

Mulmuley, K. D., & Sohoni, M. (2001). Geometric complexity theory I: An approach to the P vs. NP and related problems. SIAM Journal on Computing, 31(2), 496–526. https://doi.org/10.1137/S009753970038715X

Mulmuley, K. D., & Sohoni, M. (2008). Geometric complexity theory II: Towards explicit obstructions for embeddings among class varieties. SIAM Journal on Computing, 38(3), 1175–1206. https://doi.org/10.1137/080718115

National Academies of Sciences, Engineering, and Medicine. (2019). Reproducibility and replicability in science. The National Academies Press. https://doi.org/10.17226/25303

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them (Technical Report IHMC CmapTools 2006-01 Rev 01-2008). Florida Institute for Human and Machine Cognition. https://cmap.ihmc.us/docs/theory-of-concept-maps

Oberkampf, W. L., & Roy, C. J. (2010). Verification and validation in scientific computing. Cambridge University Press.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M.

M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Papadimitriou, C. H. (1994). Computational complexity. Addison-Wesley.

Perelman, G. (2002). The entropy formula for the Ricci flow and its geometric applications. arXiv. https://arxiv.org/abs/math/0211159

Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, 64, 1–18. https://doi.org/10.1016/j.infsof.2015.03.007

Popper, K. R. (1959). The logic of scientific discovery. Hutchinson.

Razborov, A. A., & Rudich, S. (1997). Natural proofs. Journal of Computer and System Sciences, 55(1), 24–35. https://doi.org/10.1006/jcss.1997.1494

Ruppeiner, G. (1979). Thermodynamics: A Riemannian geometric model. Physical Review A, 20(4), 1608–1613. https://doi.org/10.1103/PhysRevA.20.1608

Ruppeiner, G. (1995). Riemannian geometry in thermodynamic fluctuation theory. Reviews of Modern Physics, 67(3), 605–659. https://doi.org/10.1103/RevModPhys.67.605

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. University of Illinois Press.

Sipser, M. (2012). Introduction to the theory of computation (3rd ed.). Cengage Learning.

Strassen, V. (1983). Rank and optimal computation of generic tensors. Linear Algebra and its Applications, 52–53, 645–685.

Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

World Wide Web Consortium. (2013). PROV-O: The PROV ontology (W3C Recommendation, 30 April 2013). https://www.w3.org/TR/prov-o/

Appendix A: From González's (1995) Unified Software-Complexity Metric to the GCT-STL Tensorial Architecture

A.1 Purpose, Historical-Priority Status, and Epistemic Boundary

This appendix makes explicit the mathematical, historical, and methodological relationship between the unified software-complexity model of González (1995) and the contemporary GCT-STL architecture. The relationship is substantial but bounded. GCT-STL is neither a direct numerical continuation of the 1995 metric nor a claim that the earlier software-complexity model already contained TSP_{dec}, 3SAT, Turing-machine universality, P_{STL}/NP_{STL}, or a class-separation theorem. The appendix distinguishes structural genealogy, direct mathematical inheritance, reinterpretation/formalization, and genuinely new GCT-STL constructions.

The historical-priority claim is deliberately narrow. Tensor methods in algebraic complexity and entropy-based software-complexity measures both existed before 1995. The claim supported here is that, to the best of the programme's bibliographic audit, González (1995) appears historically distinctive in explicitly combining three heterogeneous software-complexity domains represented in information-valued vector spaces, a trilinear tensorial map, determinant evaluation, and an absolute determinant interpreted as a unified tensorial complexity volume. This is a PLAUSIBLE STRONG HISTORICAL PRIORITY claim, not an absolute "first ever" claim.

Bibliographic verification boundary. González (1995), including Appendix 3 and the pertinent pages on the three domains, entropy, and trilinear determinant construction, has now been directly inspected as a primary source in the PVV-9 audit. Its bibliographic identity and the presence of those historical constructions are verified. The broader historical-priority assessment nevertheless remains provisional until a scoped scientific mapping study/systematic or focused literature review (SMS/SLR), and the mathematical equivalence of historical and contemporary GCT-STL objects is not inferred from this inspection.

A.2 Pre-1995 Genealogies: Entropy, Algebraic Tensor Complexity, and Information Geometry

Three pre-1995 traditions must be separated. First, information-theoretic software metrics predate González (1995); Harrison (1992), cited in the 1995 paper itself, is one explicit antecedent. Second, algebraic complexity had already developed tensor rank, bilinear complexity, and optimal tensor computation through work including Alder and Strassen (1981), Strassen (1983), and Håstad (1990). Third, differential-geometric traditions in statistics and thermodynamics used metric tensors, curvature, and determinant-based volume elements, including Amari (1985) and Ruppeiner (1979). These traditions establish that neither "tensor + complexity" nor "entropy + software complexity" originates in 1995.

The 1995 contribution examined here is therefore not priority over those ingredients individually, but the specific integrated construction described below. This separation also prevents a category error: the covariant trilinear tensor of González (1995) is not a Riemannian metric tensor, and the tensor-rank programme of algebraic complexity is not the same mathematical object as a determinant-based software-complexity volume.

A.3 González (1995): Original Three-Domain × Three-Magnitude Taxonomy

González (1995) organizes software complexity into three domains—syntactical (E_S), functional (E_F), and computational (E_C)—and three characteristic magnitudes per domain. The resulting nine-factor taxonomy is the structural ancestor relevant to the 3×3 organization later used in GCT-STL.

1995 domain	1995 factor	Original magnitude role	GCT-STL domain genealogy	Boundary
Syntactical E_S	Code length	Length/position	L (logical/syntactic)	Structural genealogy only; no descriptor identity.
Syntactical E_S	Syntactical time	Time	L (logical/syntactic)	Structural genealogy only; no descriptor identity.
Syntactical E_S	Decisional level	Level/depth	L (logical/syntactic)	Direct information-theoretic lineage to D9 arises only after Appendix 3 entropy construction.
Functional E_F	Data coupling	Length/position	S (semantic/geometric)	Domain-level homologation; no descriptor identity.
Functional E_F	Cohesiveness	Time	S (semantic/geometric)	Domain-level homologation; no descriptor identity.
Functional E_F	Functional coupling	Level/depth	S (semantic/geometric)	Domain-level homologation; no descriptor identity.
Computational E_C	Data length	Length/position	T (temporal/computational)	Domain-level homologation; no descriptor identity.
Computational E_C	Algorithmic time	Time	T (temporal/computational)	Genealogical affinity; not equal to D2/D4/D5.
Computational E_C	Data hierarchical level	Level/depth	T (temporal/computational)	Domain-level homologation; no descriptor identity.

The domain homologation used in Paper I-A is therefore E_F→S, E_C→T, and E_S→L. The contemporary row axis Y1–Y3 is a GCT-STL construction and is not identified with the original length/time/level coordinates. Likewise, the numerical values of D1–D9 are determined by contemporary SourceMap and Provenance Bundle contracts, not by historical analogy.

A.4 Complete González (1995)–GCT-STL Correspondence and Status Matrix

The correspondence is classified as I = Inherited, R = Reinterpreted, C = Corrected/Formalized, and N = New in GCT-STL. A single item may carry more than one code because GCT-STL may inherit an organizing principle while changing its mathematical semantics.

1995 object	Original role	GCT-STL relation	Code	Claim boundary
Three heterogeneous domains	Syntactical / Functional / Computational	S / T / L computational-information domains	I+R	Structural genealogy; semantics reassigned.
Three magnitudes per domain	Length / Time / Level	Rows Y1 / Y2 / Y3	R+N	No one-to-one row identity.
Information-valued coordinates	Entropy/information measures across domains	Y3 intrinsically entropic in raw R0; R2 information lift	I+R+C	Modern architecture separates raw heterogeneous coordinates from explicitly entropic representation.
Trilinear map Φ	Three-linear form / covariant tensor	Historical tensorial genealogy of coupled representation	I+C	Tensor status retained; modern object differs.
det(M)	Scalar tensor evaluation	Determinant-based representation diagnostics	I+R	No lower bound follows by definition.
|det(M)|	Tensorial complexity volume	Historical antecedent for joint coupling/non-degeneracy	I+R	Volume functional ≠ obstruction theorem.
|det(M_j−M_i)|	Differential software-complexity volume	No direct orbital descendant	C	Explicitly separated from d_orb.
Permutation-group orbits in structural entropy	Equivalence classes in program graph	GCT-STL orbit-inspired geometry	R+N	Conceptual resonance only; not the same orbit construction.
Decisional entropy, 2-bit limit	Appendix 3	D_9^{CAN} / Theorem 1	I+C	Direct mathematical lineage plus canonical reformulation.
D1–D9, R0–R3, P_{STL}/NP_{STL}, F1–F4	Absent	Contemporary GCT-STL objects	N	No historical back-projection.

A.5 Information-Theoretic Lineage and the 2-Bit Decisional Bound

The strongest direct mathematical inheritance in Paper I-A is the decisional-information construction of Appendix 3 of González (1995). A Boolean function is represented by a binary decision tree, the populations at different levels are used as entropy classes, and the resulting Shannon entropy approaches an upper limit of 2 bits. Paper I-A transports that level-entropy law to the exact canonical 3-CNF artifact and defines D_9^{CAN}=H_{sint}^{CAN}. Appendix B supplies the contemporary proof.D_7^{CAN} and D_5^{CAN} are not inherited 1995 formulas; they are GCT-STL extensions of the information-theoretic principle to semantic and computational coordinates.

$$\lim_{h\to\infty} H_{decisional}(h) = 2 \text{ bits}$$

A.6 The 1995 Trilinear Tensor Construction

The 1995 framework does not terminate with nine disconnected scalar measures. Let E_syn, E_fun, and E_comp be the three complexity vector spaces. González defines a map Φ:E_syn×E_fun×E_comp→H and explicitly requires linearity in each argument. Under the algebraic p-linear convention cited through Mostow and Sampson (1969, pp. 259–260), and identifying H with the real-valued entropy range for the scalar-form statement, Φ is a covariant tensor of order three. Equivalently, by the universal property of the tensor product, Φ corresponds to a linear functional on E_syn⊗E_fun⊗E_comp.

$$\Phi : E_{syn} \times E_{fun} \times E_{comp} \to H$$

$$\Phi \in E_{syn}^* \otimes E_{fun}^* \otimes E_{comp}^* \quad \text{(scalar-valued case)}$$

For each software chunk i, the three domain vectors are arranged as the rows of a 3×3 matrix M_i. The matrix is therefore the coordinate arrangement of the three arguments used in the evaluation; it is not the definition of the tensor itself. The paper then sets:

$$\Phi\left(e_{syn,i}, e_{fun,i}, e_{comp,i}\right) = det\left(M_i\right)$$

This point matters historically: the tensorial language is explicit in the primary source and is not a retrospective label applied to a mere 3×3 array.

A.7 Tensorial Complexity Volume and Joint Non-Degeneracy

González (1995) takes the absolute magnitude of the scalar tensor evaluation and interprets it geometrically as the volume of the parallelepiped spanned by the three domain vectors:

$$V_i = \left|\Phi\left(e_{syn,i}, e_{fun,i}, e_{comp,i}\right)\right| = \left|\det\left(M_i\right)\right|$$

This may be described safely as a determinantal tensorial complexity-volume functional. It need not be interpreted as a matrix norm on the full matrix space, and it is not a Riemannian metric tensor. The relevant coupling consequence is immediate: three individually nonzero domain vectors can nevertheless have zero joint volume when they are linearly dependent.

$$e_1 \neq 0,\ e_2 \neq 0,\ e_3 \neq 0 \ \not\Rightarrow\ \left|\det\left(e_1, e_2, e_3\right)\right| > 0$$

$$e_3 = \alpha e_1 + \beta e_2 \ \Rightarrow\ e_1 \wedge e_2 \wedge e_3 = 0 \ \Rightarrow\ \det(M) = 0$$

This joint non-degeneracy property is the legitimate mathematical bridge to the modern Full STL Coupling research question. It does not establish a computational obstruction, uniformity theorem, $Cost_{STL}$ lower bound, or P≠NP consequence.

A.8 Determinantal Differential Volume

The 1995 paper also uses |det(M_j−M_i)| when comparing consecutive complexity configurations. In the present historical reconstruction, this is treated as a determinantal differential-volume construction internal to the original software-measurement framework. Its semantic role is a volumetric measure of the joint change matrix ΔM_ij=M_j−M_i, including the singular case in which the differential volume vanishes.

$$\Delta M_{ij} = M_j - M_i, \qquad \Delta V_{ij}^{\det} = \left|\det\left(\Delta M_{ij}\right)\right|$$

This quantity must not be conflated with ||det(M_j)|−|det(M_i)|| or |det(M_j)−det(M_i)|. Nor is it a historical precursor that is mathematically identified with the orbital distances later defined in GCT-STL. The two constructions belong to different mathematical roles:

$$V_{\det}^{(1995)} \neq d_{orb}^{(GCT\text{-}STL)}$$

A further formal clarification is important: for a general trilinear map, Φ(a−a′,b−b′,c−c′) is not equal in general to Φ(a,b,c)−Φ(a′,b′,c′), because multilinear expansion produces cross-terms. Accordingly, |det(M_j−M_i)| is treated here as the defined differential-volume construction of the software metric, not as a universal identity between differences of tensor evaluations.

A.9 Post-1995 Developments: Documented Influence and Independent Convergence

Post-1995 developments relevant to this genealogy fall into two categories. Documented influence requires citation or explicit adoption of the 1995 work; independent convergence denotes later research that develops related ideas without evidence of descent. The bibliographic review undertaken for this programme identified later entropy-based and semantic software metrics that cite or extend information-theoretic software measurement, while algebraic complexity, Geometric Complexity Theory, information geometry, tensor networks, tensor hardness, and geometric notions of computation largely constitute parallel or convergent traditions. Paper I-A does not convert thematic resemblance into a claim of historical derivation.

Tradition	**Historical relation**	**GCT-STL interpretation**
Entropy/semantic software metrics	Documented or near-direct software-metrics lineage	Supports influence of unified information-theoretic complexity; does not establish adoption of the determinant/tensor-volume construction.
Algebraic complexity / tensor rank	Independent antecedent and continuing parallel tradition	Shows tensors in complexity predate 1995.
Geometric Complexity Theory	Independent post-1995 obstruction programme	Provides orbit/obstruction methodology; no historical derivation from González is asserted.
Information geometry / thermodynamic geometry	Parallel geometric tradition	Provides Fisher/Ruppeiner tools later used as diagnostics; mathematically distinct from Φ.
Tensor networks / tensor hardness / multilinear computation	Independent later tensor-computation developments	Demonstrates growth of tensor methods after 1995; not evidence of direct influence.

A.10 Historical Priority Matrix

Claim	**Status**	**Paper I-A claim boundary**
González introduced entropy in software complexity	PREDATED	Not allowed.
González introduced tensors in computational complexity	PREDATED	Not allowed.
González constructed an explicit scalar-valued trilinear tensor across three software-complexity spaces	SUPPORTED BY PRIMARY SOURCE	Allowed.
The tensor evaluation is det(M) and |det(M)| is interpreted as tensorial complexity volume	SUPPORTED BY PRIMARY SOURCE	Allowed.
The integrated combination three domains + information spaces + trilinear tensor + determinant + complexity volume is historically distinctive	HISTORICALLY DISTINCTIVE / PLAUSIBLE STRONG PRIORITY	Allowed only with "to the best of our bibliographic audit".
González was absolutely the first to construct such an object	NOT ESTABLISHED	Not allowed without exhaustive priority proof.
The 1995 volume proves a complexity lower bound or P≠NP	FALSE / OUT OF SCOPE	Forbidden.

A.11 Genealogy, Provenance, System Membership, and Proof Status

Within the four-system architecture, the González (1995) inheritance belongs primarily to System I, the Fundamental Mathematical Representation System. It motivates the organization of representation objects and part of the information/tensorial vocabulary. It does not numerically determine D1–D9, which remain governed by current provenance contracts, and it does not supply System-II theorem authority. A System-I genealogy can motivate a System-III diagnostic or a System-II conjecture, but any universal obstruction, Representation Transfer, F2c, or F2 must be established independently at theorem level.

$$\text{genealogy} \neq \text{provenance} \neq \text{empirical evidence} \neq \text{theorem} \neq \text{formal certification}$$

A.12 From González (1995) to Full STL Tensorial Coupling

The controlled genealogical claim is that González (1995) supplies a structural ancestor for coupling heterogeneous complexity domains through a joint multilinear quantity. GCT-STL preserves that coupling intuition while changing the represented objects, the provenance contracts, the coordinate semantics, and the complexity-theoretic goal. The contemporary raw signature is R0=$T_{VG}^{raw} \in V_Y \otimes V_{STL} \cong R^{(3\times 3)}$, and the Full

STL Tensorial Coupling Principle asks whether a relevant obstruction is carried by the complete coupled S–T–L structure rather than by any isolated marginal domain.

CONTROLLED GENEALOGY CLAIM
GCT-STL preserves the three-by-three multidimensional architecture and the multilinear-coupling principle introduced in González (1995), but reassigns their semantics to the contemporary semantic/geometric, temporal/computational, logical/syntactic, and resolution-layer architecture. The determinant-based complexity volume is a conceptual/mathematical antecedent of the modern full-coupling principle; it is neither identical to nor sufficient to establish a GCT-STL obstruction.

COUPLING NON-DEGENERACY CLAIM
The determinant construction exhibits a basic coupling fact: nonzero components do not imply nonzero global complexity volume; joint non-degeneracy is required. GCT-STL generalizes this structural question by investigating whether computational obstruction is likewise carried by the complete S–T–L configuration. The first statement is a determinant consequence; the second remains an OPEN GCT-STL research hypothesis.

A.13 Final Claim Boundary

- Historical priority over a specific integrated construction ≠ priority over tensors or entropy individually.
- Tensorial complexity volume ≠ deterministic-time complexity lower bound.
- V_detˆ(1995) ≠ d_orbˆ(GCT-STL).
- Joint non-degeneracy of a determinant ≠ uniform computational obstruction.
- Full STL Tensorial Coupling hypothesis ≠ F2.
- Genealogical inheritance ≠ descriptor provenance.
- Historical or conceptual convergence ≠ proof authority.

Appendix B: Proof of Theorem 1 — González (1995) Canonical Syntactic-Entropy Bound

B.1 Genealogical Link to Appendix A

Appendix A documents the broader three-part contribution of González (1995) to GCT-STL. For Theorem 1, the relevant direct mathematical inheritance is specifically Appendix 3, "Decisional Information Content": binary decisional structure, level populations treated as entropy classes, Shannon information of the resulting distribution, and the asymptotic 2-bit bound. The present appendix does not repeat the full genealogy; it proves the canonical GCT-STL instantiation on the exact materialized 3-CNF artifact φ_I and records the $D_8^{CAN} \to H_{sint}^{CAN} \to D_9^{CAN}$ dependency.

B.2 Theorem Statement

Theorem 1 — González (1995) Canonical Syntactic-Entropy Bound. Let I=(N,D,K) be a well-formed TSP_{dec} instance and let φ_I=compileDecision3SAT(I).formula. Let M(I):=|Clauses(φ_I)|=D_8^{CAN}(I), and let H_{sint}^{CAN}(I):=D_9^{CAN}(I) be the Shannon level entropy of the canonical complete binary syntactic envelope derived from this exact materialized 3-CNF artifact. Then

$$0 \le H_{sint}^{CAN}(I) = D_9^{CAN}(I) < 2 \; bits$$

and, for any growing family for which D_8^{CAN}(I)→∞,

$$\lim_{D_8^{CAN}(I)\to\infty} H_{sint}^{CAN}(I) = 2 \text{ bits}$$

B.3 Exact Artifact and Canonical Syntactic Envelope

Domain qualification inherited from §4.2.1.2: the displayed formula for h_syn assumes M(I)=|Clauses(φ_I)|≥1. The empty-artifact case is not covered by this logarithmic construction unless a separate compiler-domain convention is established.

Fix a well-formed instance I and let φ_I be the exact canonical 3-CNF emitted by compileDecision3SAT. Define

$$M(I) := |Clauses(\varphi_I)| = D_8^{CAN}(I)$$

The construction used for D9 does not replace φ_I by another formula. In particular, it does not add tautological clauses to reach a power of two and does not add dummy literals to equalize parse depths. Those operations belong to an older normalization argument and are not required for the entropy bound.

Instead, construct an auxiliary complete binary syntactic envelope whose only purpose is to define the level distribution. Missing positions are structural null slots; they do not belong to the 3-CNF artifact and have no semantic effect. Let

$$h := h_{syn}(I) = \lceil \log_2 M(I) \rceil + 2$$

The term +2 accounts for the fixed binary template of a three-literal clause. For each level k=0,...,h, a complete binary envelope contains

$$N_k = 2^k$$

nodes, and the total number of structural nodes is

$$N_{tot} = \sum_{k=0}^{h} 2^k = 2^{h+1} - 1$$

Choose one structural node uniformly at random and let the random variable L record its level. Then

$$p_k := P(L = k) = \frac{2^k}{2^{h+1} - 1}$$

B.4 Shannon Level Entropy and Closed Form

The canonical syntactic entropy is the Shannon entropy of this level distribution:

$$H_{sint}(h) = -\sum_{k=0}^{h} p_k \log_2 p_k$$

Substituting p_k=2^k/(2^(h+1)−1) gives

$$H_{sint}(h) = \log_2\left(2^{h+1} - 1\right) - \frac{1}{2^{h+1} - 1} \sum_{k=0}^{h} k\, 2^k$$

Using the finite weighted geometric-series identity

$$\sum_{k=0}^{h} k\, 2^k = (h-1)2^{h+1} + 2$$

we obtain the exact closed form

$$H_{sint}(h) = \log_2\left(2^{h+1} - 1\right) - \frac{(h-1)2^{h+1} + 2}{2^{h+1} - 1}$$

A form that makes the limiting constant explicit is

$$H_{sint}(h) = 2 + \log_2\left(1 - 2^{-(h+1)}\right) - \frac{h+1}{2^{h+1} - 1}$$

B.5 Universal Bound and 2-Bit Asymptote

For every finite h, $\log_2(1-2\hat{}(-(h+1)))<0$ and $(h+1)/(2\hat{}(h+1)-1)>0$. Therefore both corrections to the constant 2 are negative and

$$H_{sint}(h) < 2$$

Shannon entropy is non-negative, so

$$0 \leq H_{sint}(h) < 2$$

As $h\to\infty$,

$$\log_2\left(1 - 2^{-(h+1)}\right) \to 0 \quad \text{and} \quad \frac{h+1}{2^{h+1} - 1} \to 0$$

Hence

$$\lim_{h\to\infty} H_{sint}(h) = 2$$

Because h_syn(I)$\to\infty$ whenever D_8^{CAN}(I)=M(I)$\to\infty$, the canonical GCT-STL descriptor satisfies

$$0 \leq H_{sint}^{CAN}(I) = D_9^{CAN}(I) < 2$$

$$\lim_{D_8^{CAN}(I)\to\infty} D_9^{CAN}(I) = 2$$

Thus D_9^{CAN}=O(1); for growing families with D_8^{CAN}(I)$\to\infty$,D_9^{CAN}=Θ(1). This is the exact sense in which the syntactic entropy is asymptotically bounded.

B.6 What Is Inherited and What Is New in GCT-STL

Inherited from González (1995) are four mathematical ideas: (1) binary representation of a logical decisional structure; (2) use of tree levels as entropy classes; (3) Shannon information content of the level distribution; and (4) the asymptotic upper bound of 2 bits. These form the genealogical and mathematical provenance of H_sint.

New in GCT-STL is the canonical instantiation of that law on φ_I=compileDecision3SAT(I).formula, the exact D8→H_sint→D9 dependency, the explicit separation between the logical artifact and the auxiliary structural envelope, and the placement of D9 as the L/Y3 descriptor of T_{VG}^{raw}.

Accordingly, three provenance levels must not be conflated. González (1995), Appendix 3, supplies the genealogical and mathematical lineage of the level-entropy law; Theorem 1 supplies the theorem-origin authority for D9 inside Paper I-A; and D_8^{CAN} supplies the instance-specific exact artifact on which H_{sint}^{CAN} is evaluated.

B.7 Claim Boundary

The 2-bit result is a bound on the Shannon entropy of a distribution over syntactic levels. It is not a bound on the number of clauses, the number of Boolean variables, the byte size of the 3-CNF artifact, proof complexity, or deterministic running time. The fact that H_sint remains bounded while the materialized syntax grows is an information-theoretic property of the level distribution and does not by itself imply a computational lower bound.

This corrected proof therefore preserves the substantive 1995 insight while removing formula-altering padding and making the GCT-STL artifact provenance exact.

Appendix C: Proof of Theorem 2 — Shannon–Boltzmann Semantic Entropy Bounds

C.1 Theorem Statement

Theorem 2 — Shannon–Boltzmann Semantic Entropy Bounds. For a well-formed directed TSP_{dec} instance I=(N,D,K), let Ω_Nˆ→ denote the fixed-origin directed Hamiltonian-tour space, with |Ω_Nˆ→|=(N−1)!. For $\beta \geq 0$, let D_7^{CAN}(I,β) be the Shannon entropy of the Boltzmann distribution induced by the tour costs. Then:

$$0 \leq D_7^{CAN}(I,\beta) \leq \log_2\left((N-1)!\right)$$

$$D_7^{CAN}(I,0) = \lim_{\beta\to 0} D_7^{CAN}(I,\beta) = \log_2\left((N-1)!\right)$$

$$D_7^{MAX}(N) := \log_2\left((N-1)!\right) = \Theta\left(N \log N\right)$$

and D_7^{CAN}(I,β) is non-increasing in β.

C.2 Proof — Geometric and Probabilistic Construction

Proof. Fix a well-formed directed TSP_{dec} instance I=(N,D,K). Let Ω_Nˆ→ be the fixed-origin directed Hamiltonian-tour space and let C_I:Ω_Nˆ→→R_≥0 be the tour-cost function induced by D. The canonical directed-tour convention gives |Ω_Nˆ→|=(N−1)!. The threshold K belongs to the decision instance but does not enter the Boltzmann weights.

Boundary-case cross-reference (C03-17). The probability normalization in this proof uses the nonempty finite-cost ensemble specified in §4.2.2.2. This clarification does not determine the compiler minimum admissible N or amend the statement of Theorem 2. The cases N=1 and N=2 retain the qualifications recorded there.

$$Z_I(\beta) = \sum_{\pi \in \Omega_N^{\rightarrow}} \exp\left(-\beta C_I(\pi)\right)$$

$$p_\beta(\pi) = \frac{\exp\left(-\beta C_I(\pi)\right)}{Z_I(\beta)}$$

$$D_7^{CAN}(I,\beta) = -\sum_{\pi \in \Omega_N^{\rightarrow}} p_\beta(\pi) \log_2 p_\beta(\pi)$$

C.2.1 Thermodynamic Entropy Identity

Because ln p_$\beta(\pi)$=$-\beta$C_I(π)−ln Z_I(β), substitution into Shannon entropy gives the exact identity

$$D_7^{CAN}(I,\beta) = \frac{\ln Z_I(\beta) + \beta \langle C \rangle_\beta}{\ln 2}$$

where ⟨C⟩_β=Σ_π p_$\beta(\pi)$C_I(π). Hence entropy and expected tour cost are related but distinct observables.

C.3 Universal Entropy Bound

For any probability distribution supported on m states, Shannon entropy satisfies 0≤H≤$\log_2$m, with equality at the upper bound exactly for the uniform distribution. Taking m=(N−1)! yields

$$0 \le D_7^{CAN}(I,\beta) \le \log_2\left((N-1)!\right)$$

C.4 Uniform High-Temperature Limit

At β=0, all Boltzmann weights are equal, so p_$0(\pi)$=1/(N−1)! for every π∈Ω_N^→. Therefore

$$D_7^{CAN}(I,0) = \log_2\left((N-1)!\right)$$

$$\lim_{\beta \to 0} D_7^{CAN}(I,\beta) = \log_2\left((N-1)!\right)$$

Define the maximum-entropy envelope

$$D_7^{MAX}(N) := \log_2\left((N-1)!\right)$$

C.5 Asymptotic Scale of the Maximum Envelope

Stirling's approximation gives

$$\log_2\left((N-1)!\right) = (N-1)\log_2(N-1) - \frac{N-1}{\ln 2} + O\left(\log N\right)$$

$$D_7^{MAX}(N) = \Theta\left(N \log N\right)$$

This Θ(N log N) statement applies to the maximum-entropy envelope. It is not a universal equality for D_7^{CAN}(I,β) at arbitrary β.

C.6 Monotonicity in Inverse Temperature

Differentiating the canonical finite-state entropy identity and using d⟨C⟩_β/dβ=−Var_β(C_I) yields

$$\frac{\partial D_7^{CAN}(I,\beta)}{\partial \beta} = -\frac{\beta}{\ln 2} Var_\beta\left(C_I\right) \leq 0$$

Thus D_7^{CAN} is non-increasing in β. Consequently, a generic interior maximum of D7 is not an appropriate finite-size pseudocritical locator. Downstream pseudocritical analysis should instead use fluctuation/susceptibility observables such as C_V=β^2Var_β(C_I), together with preregistered diagnostics.

C.7 Low-Temperature Limit and Optimal-Tour Degeneracy

Let C_min(I) be the minimum directed tour cost and let g_*(I) be the number of fixed-origin directed tours attaining C_min(I). As $\beta\to\infty$, all non-optimal weights vanish relative to the optimum set, and the limiting distribution is uniform on the g_*(I) minimizers. Hence

$$\lim_{\beta\to\infty} D_7^{CAN}(I,\beta) = \log_2 g_*(I)$$

In particular, if the optimum is unique under the adopted convention, then g_*(I)=1 and D_7^{CAN}(I,β)→0.

C.8 Relation to Theorem 3 / D5 and Theorem 1 / D9

Theorem 2 provides an upper envelope for the semantic descriptor, while Theorem 3 provides a structural lower envelope for the canonical computational descriptor. For N≥2,

$$D_7^{CAN}(I,\beta) \leq \log_2\left((N-1)!\right) < N^2 \leq D_5^{CAN}(I)$$

Therefore D_5^{CAN}(I)>D_7^{CAN}(I,β) in descriptor magnitude. This analytic inequality does not imply causal dominance, hardness, or a deterministic-time lower bound. Theorem 1 independently gives 0≤D_9^{CAN}(I)<2. No universal ordering D7>D9 is asserted because D7 may approach $\log_2$g_*(I), including zero for a unique optimum.

Full-coupling interpretation. The inequality above is a statement about marginal descriptor magnitude only. Theorem 2 therefore supplies a semantic coordinate and its exact thermodynamic behavior, not a causal ranking. Semantic dominance is retained only as a subordinate, falsifiable research hypothesis: depending on regime, S may dominate, amplify, uniformize or interact with T and L, but no such role is assumed from D7 alone. Testing causal contribution requires analysis of the complete coupled STL signature and, for lower-bound purposes, independent universality and obstruction-to- $Cost_{STL}$ transfer theorems.

C.9 Canonical versus Experimental D7

Unlike D5, D7 does not require a separate CAN/OP descriptor pair in Paper I-A. The mathematical object is the full canonical family D_7^{CAN}(I,β). Companion statistical-physics work develops estimator design, convergence and uncertainty methodology for finite MCMC approximations of this same quantity, denoted for example by D7_hat(I,β), and reports multi-universe finite-size evidence. These downstream activities do not define a different matrix coordinate.

$$D_7(I,\beta) \equiv D_7^{CAN}(I,\beta), \qquad \widehat{D}_7(I,\beta) \approx D_7^{CAN}(I,\beta)$$

C.10 Claim Boundary and Conclusion

Theorem 2 establishes the canonical Shannon–Boltzmann descriptor D7, its finite-support bounds, uniform high-temperature maximum, Θ(N log N) maximum envelope, monotonicity in β and low-temperature degeneracy limit. It does not establish semantic causality, a deterministic-time lower bound, hardness from Fisher curvature or a separation of P_{STL} from NP_{STL}. Under the full-coupling hypothesis, any causal or lower-bound claim must be reconstructed on the complete coupled object and satisfy the required universality and transfer conditions.

Appendix D: Proof of Theorem 3 — Canonical and Operational Computational Entropy Bounds

D.1 Theorem Statement

Theorem 3 — Canonical and Operational Computational Entropy Bounds.

For a well-formed TSP_{dec} instance I=(N,D,K), let D_5^{CAN}(I) denote the canonical computational entropy associated with the Boolean search space of the materialized canonical 3-CNF representation, and let D_5^{OP}(A,I) denote the operational computational entropy effectively explored by a deterministic search procedure A. For the natural family of instances with polynomially bounded edge weights and threshold,

$$D_5^{CAN}(I) = \Omega\left(N^2\right) \quad \text{and} \quad D_5^{CAN}(I) = O\left(N^3 \log N\right)$$

and

$$0 \le D_5^{OP}(A,I) \le D_5^{CAN}(I)$$

Equality D_5^{OP}(A,I)=D_5^{CAN}(I) holds under exhaustive exploration.

D.2 Canonical 3-CNF and the CAN/OP descriptors

Proof. Let φ_I=compileDecision3SAT(I).formula be the exact canonical 3-CNF materialized by the canonical compiler and let

$$\nu(I) := |Vars\,(\varphi_I)|$$

The complete Boolean assignment space is A_I={0,1}^{ν(I)}. Under the uniform distribution on this finite support, Shannon entropy equals Hartley entropy:

$$H_{comp}^{CAN}(I) = \log_2 |A_I| = \log_2\left(2^{\nu(I)}\right) = \nu(I)$$

Accordingly,D_5^{CAN}(I):=H_comp^CAN(I)=ν(I). For a deterministic search procedure A, let E_A(I)⊆A_I be the set of complete assignments explicitly reached or evaluated by A and define

$$D_5^{OP}(A,I) := H_{COMP}^{OP}(A,I) := \log_2 max\{1, |E_A(I)|\}$$

D.3 Structural lower bound

The Hamiltonian positional encoding contains N^2 distinct Boolean variables $x_{i,p}$, one for every city-position pair. These variables are part of the canonical decision artifact independently of all auxiliary variables introduced later by the structural and weighted-cost circuits. Therefore

$$\nu(I) \geq N^2$$

and hence

$$D_5^{CAN}(I) = \Omega\left(N^2\right)$$

This is a structural lower bound on the canonical descriptor. It does not state that a deterministic algorithm must inspect N^2 variables in a particular order or enumerate all $2^{\nu(I)}$ assignments.

D.4 Compiler-derived upper envelope

Define

$$B(I) := max\{bitlen(K), \max_{i,j} bitlen\,(D_{ij})\}$$

and let D_max(I)=$\max_{i,j}$D_ij. The canonical cost circuit chooses accumulator width

$$W(I) := bitlen\,(max\{N\,D_{\max}(I), K\})$$

The canonical compiler allocates fresh Boolean variables in the following groups. Here (N−3)_+ := max{N−3,0}.

Compiler component	**Fresh-variable contribution**	**Reason**
Positional variables	N^2	One $x_{i,p}$ per city-position pair.
Structural long-OR auxiliaries	2N(N−3)_+	N city-at-least-once and N position-at-least-once chains; pairwise at-most-one clauses add no fresh variables.
Edge selectors	N^3	One y_(p,i,j) for every triple (p,i,j).
Ripple-carry arithmetic	$5W(I)N^3$	Each of the N^3 selected weights is added through W full-adder positions; encodeFullAdder allocates five fresh Tseitin variables.
Comparator loop	2W(I)+2·popcount(K)	Two fresh variables for a 0-bit of K and four for a 1-bit across W positions.
Final comparator output	1	Fresh variable for accept = lt OR eq.
Boolean constants	2	constFalse and constTrue.

Consequently, the number of distinct variables in the materialized formula is bounded above by the compiler allocation count:

$$\nu(I) \leq N^2 + 2N(N-3)_+ + N^3 + 5W(I)N^3 + 2W(I) + 2popcount(K) + 3$$

This inequality is sufficient for Theorem 3. An exact equality with the final fresh-variable counter would additionally require showing that every allocated variable occurs in the final materialized artifact; that stronger artifact-level statement is not needed for the present upper bound.

Implementation-evidence qualification (M-07/M-08). These are bounds derived from the declared compiler allocation scheme. The present manuscript does not supply an independently replayed implementation-level proof that every materialized clause, adder/carry constraint and overflow rule agrees with that scheme. Such

source correspondence remains an OPEN verification obligation; the displayed algebra is not a certificate of an unprovided executable compiler.

D.5 Reduction to N and coefficient bit-width

Let ℓ_N=bitlen(N). By the definition of B(I), D_max<$2^{B(I)}$ and K<$2^{B(I)}$, while N<2^{ℓ_N}. Hence

$$W(I) \le B(I) + \ell_N, \qquad popcount(K) \le B(I)$$

Substitution gives the finer compiler-derived envelope

$$\nu(I) \le \widetilde{\widehat{p}_\nu}(N, B) := N^2 + 2N(N-3)_+ + N^3\left[1 + 5\left(B + \ell_N\right)\right] + 4B + 2\ell_N + 3$$

Thus, in general,

$$D_5^{CAN}(I) = O\left(N^3\left[B(I) + logN\right]\right)$$

For N≥2 one may also replace bitlen(N)≤N to obtain the purely algebraic coarse polynomial

$$p_\nu(N, B) = 5N^4 + 5BN^3 + N^3 + 3N^2 + 2N + 4B + 3$$

The N^4 term is only a coarse consequence of replacing the logarithmic bit-length of N by N; it is not the natural growth scale of the compiler.

D.6 Polynomially bounded instance family

For the natural family stated in Theorem 3, assume there exists a fixed constant c>0 such that D_ij≤Nˆc for every matrix entry and K≤Nˆc. Then B(I)=O(log N). Therefore

$$D_5^{CAN}(I) = O\left(N^3 \log N\right)$$

Combining this with the structural lower bound yields exactly the canonical part of Theorem 3:

$$D_5^{CAN}(I) = \Omega\left(N^2\right) \quad \text{and} \quad D_5^{CAN}(I) = O\left(N^3 \log N\right)$$

D.7 Operational entropy bound

Because E_A(I)⊆A_I and |A_I|=$2^{\nu(I)}$,

$$|E_A(I)| \le 2^{\nu(I)}$$

Taking logarithms and using D_5^{CAN}(I)=ν(I) gives

$$0 \le D_5^{OP}(A, I) \le D_5^{CAN}(I)$$

If A performs exhaustive exploration, E_A(I)=A_I and therefore |E_A(I)|=$2^{\nu(I)}$. Hence

$$D_5^{OP}(A, I) = D_5^{CAN}(I)$$

D.8 Conditional relation to realized search time

Domain qualification (M-09; conditional comparison only). In the following inference assume E_A(I)≠∅ and a cost model with c_A>0 independent of I over the compared instance family. Then D_5^{OP}(A,I)=$\log_2$|E_A(I)|, so T_A(I)≥c_A·2^{D_5^{OP}(A,I)}. If E_A(I)=∅, the definition D_5^{OP}=$\log_2$(max{1,|E_A|}) yields zero while |E_A|=0, and the displayed inference is not asserted. This qualification neither changes Theorem 3 nor establishes a lower bound for arbitrary deterministic deciders.

The relation with running time is a corollary of the operational quantity and is not part of the theorem statement. If an operational cost model charges at least one counted operation for every complete assignment reached by A, then for some constant c_A>0,

$$T_A(I) \geq c_A \, |E_A(I)|$$

and hence

$$T_A(I) = \Omega\left(2^{D_5^{OP}(A,I)}\right)$$

This is conditional on the realized exploration of the specified procedure A. It does not imply T_A(I)=Ω(2^{D_5^{CAN}(I)}) for arbitrary deterministic algorithms and therefore does not constitute an exponential lower bound for 3-SAT, TSP_{dec}, or NP-complete problems in general.

D.9 Conclusion and descriptor provenance

Theorem 3 establishes a canonical/operational distinction. The theorem-origin descriptor placed in the T coordinate of the canonical Valiant–González crystal is

$$D_5^{CAN}(I) = \nu(I) = |Vars\,(\varphi_I)|$$

with Ω(N^2) and O(N^3 log N) bounds for the stated polynomially bounded instance family.D_5^{OP}(A,I) measures realized explored-support entropy and satisfies $0 \leq D_5^{OP} \leq D_5^{CAN}$, reaching the canonical value only under exhaustive exploration. Theorem 3 therefore supplies a canonical computational-entropy scale without importing a universal deterministic-time lower bound into the definition of D5.