

Information-Theoretic Structure as a Necessary Consequence of the VERSF Framework

Keith Taylor — VERSF Theoretical Physics Programme

For the General Reader

Information theory — the mathematics of uncertainty, originally developed in the 1940s for telecommunications — has become indispensable across modern physics. It tells us how to quantify what is unknown about a system, how to measure the entropy of a thermodynamic state, even how to make sense of the probabilities in quantum mechanics. But there is a question that is rarely asked: where does it come from? Why is Shannon's particular formula the right way to measure physical uncertainty, rather than some other mathematical structure?

The standard answer is that information theory is a piece of pure mathematics that physics happens to find useful. We import it into physical theories the way we import calculus or linear algebra: as a tool, not as a consequence of physics itself.

This paper takes a different view. Within the VERSF framework, information theory is not imported into physics — it is a structural consequence of more basic physical principles. The argument runs roughly like this. Suppose you accept three things: that "physical content" means whatever distinctions observers can invariantly agree about; that any real, bounded system has a finite number of admissible alternatives it could commit to; and that the formation of a "fact" — a quantum measurement, a decoherence event, any recorded outcome — is a thermodynamically irreversible process with a real entropy cost. Then the mathematical structure that must describe such a system *before* it commits to a fact turns out to be Shannon entropy, uniquely. There is no other consistent choice that respects what observers can invariantly agree about and how independent systems combine.

This matters for two reasons. First, it changes the status of information theory in physics: not a useful import, but a structural feature of how reality must be described before facts are made. Second, it closes a loop in the broader VERSF programme. Earlier work in the programme has derived both quantum mechanics and thermodynamics from VERSF principles, but those derivations took an information-theoretic starting point as given. This paper grounds that starting point — the information layer used as input to those derivations is itself supplied by the framework, not assumed from outside.

We are honest about which parts of the argument are strictly forced and which are physically motivated. Three of the conditions that uniquely pin down Shannon entropy follow rigorously from the VERSF axioms. Two are natural physical desiderata — reasonable assumptions about

how independent systems combine and how partial information aggregates — for which the framework supplies principled but not strictly forcing reasons. The honest summary is: given the VERSF axioms together with these two natural desiderata, Shannon entropy is the unique admissible description of pre-commitment uncertainty. That is weaker than claiming the result is forced from VERSF alone, and stronger than claiming Shannon entropy is just convenient. It is, we think, the right middle ground — and it is the most that any honest derivation in this area can claim.

Table of Contents

1. Introduction
2. Foundational Structure
3. The Pre-Commitment Description Problem
4. Constraints on the Uncertainty Measure
 - 4.1 Permutation Invariance
 - 4.2 Continuity
 - 4.3 Monotonicity for the Uniform Distribution
 - 4.4 Compositional Additivity (motivated, not strictly forced)
 - 4.5 Coarse-Graining Consistency (motivated, not strictly forced)
5. Uniqueness of Shannon Entropy
6. Thermodynamic Connection
7. Reframing of Information Theory
8. Scope and the Representation Question
9. Scope and Limitations
10. Conclusion
11. References

Abstract

Information theory is conventionally introduced as an independent mathematical framework for quantifying uncertainty. We show that within the Void Energy–Regulated Space Framework (VERSF), information-theoretic structure arises necessarily — structurally entailed by, rather than imported into, the framework — given more primitive principles: observer-invariant distinguishability, finite admissibility, and irreversible fact formation.

The argument is structural. Any consistent description of pre-commitment physical reality must admit a measure of uncertainty satisfying permutation invariance, continuity, monotonicity, compositional additivity, and coarse-graining consistency. Each of these is grounded in a corresponding VERSF axiom or structural feature; the first three are forced, while the last two are honestly characterised as physical desiderata that VERSF motivates rather than strictly forces (we are explicit in §4 about which is which). These constraints match a standard formulation of

the Khinchin axioms, and they uniquely select Shannon entropy as the measure of unresolved physical alternatives prior to commitment. The thermodynamic cost of fact formation is then proportional to the information resolved, with the constant of proportionality fixed by the binary-resolution unit asserted by A1', which is a structural postulate of the framework rather than a derivation from more primitive thermodynamic principles.

This work isolates the **information-theoretic layer** of VERSF — the structural object that any admissible physical representation must encode. Prior work in the VERSF programme has shown that, given this information structure together with the discrete commitment substrate (TPB), quantum mechanics emerges as the minimal admissible representation. The present result completes that picture from below: the information layer used as input to the representation derivation is grounded in VERSF's foundational principles rather than imported externally, with the partly-forced/partly-motivated character of that grounding made explicit.

1. Introduction

Information theory is foundational across physics, computation, and communication, yet it is typically imported into physical theory rather than derived from physical principles. The status of Shannon entropy in particular — whether it is a mathematical convenience, an epistemic measure, or a structural feature of physical reality — is rarely settled at the level of foundational axioms.

The VERSF programme begins from a different starting point. Physical reality is described in terms of:

- finite distinguishable alternatives,
- irreversible commitment events ("facts"), and
- observer-invariant physical content.

Two questions then arise:

1. **What mathematical structure is required to describe physical systems prior to fact formation?**
2. **What representation realises that structure under the full set of VERSF constraints?**

Prior work in the VERSF programme has answered Question (2) (Taylor, *From Information to Quantum Mechanics*; see References): given an information-theoretic description of pre-commitment uncertainty together with the discrete commitment substrate (TPB), the minimal admissible representation is finite-dimensional complex quantum mechanics. That derivation, however, takes the information-theoretic layer as input. The present paper isolates and answers Question (1), grounding that input: Shannon entropy is the unique measure of pre-commitment uncertainty consistent with VERSF, and is therefore not an external assumption to the representation derivation but a structural consequence of the framework.

The separation matters. Conflating the two questions has been a recurring source of overclaiming in earlier VERSF drafts. The information-theoretic layer is grounded in VERSF's foundational axioms — partly forced, partly motivated as physical desiderata, with the distinction made explicit in §4 — in a clean and self-contained sense; the representation result depends on further structural input (TPB, and the additional postulates needed to select linearity and complex structure). Establishing the two layers separately makes the logical dependence explicit and closes a gap in the prior derivation.

2. Foundational Structure

We assume only the VERSF principles required for the present result. Additional axioms in the broader framework — notably the discrete commitment substrate (TPB) governing inter-event evolution — are not invoked here; they enter the representation result established in prior work.

A0 — Observer-Invariant Distinguishability. Physical content depends only on distinctions invariant under admissible observer transformations.

A1 — Finite Distinguishability. Any bounded physical system admits a finite set of admissible alternatives:

$$\mathcal{C} = \{c_1, c_2, \dots, c_N\}$$

A1' — Thermodynamic Fact Formation. A physical fact corresponds to an irreversible commitment event satisfying:

$$\text{Fact} \Rightarrow \Delta S \geq k_B \ln 2$$

The forward direction asserts that fact formation incurs at least the Landauer-bounded entropy cost of resolving a binary distinction. We do not assert the converse. Many thermodynamic processes produce $\Delta S \geq k_B \ln 2$ without resolving an admissible-set distinction in the VERSF sense (e.g. bulk dissipation, thermal mixing). The converse — that any sufficiently large entropy increment constitutes a fact — would require additional structural input identifying the increment with the resolution of a specific admissible-set distinction, and is not claimed here.

Pre-Commitment Structure. Prior to commitment, no alternative is selected and no entropy-producing event has occurred. The system must nevertheless be physically describable, since it is not empty and carries physically relevant structure — it determines what *can* be committed and with what weight.

3. The Pre-Commitment Description Problem

Let a system have admissible alternatives $\mathcal{C} = \{c_1, \dots, c_N\}$. Before commitment, the system is not in any single c_i , yet it is not without structure: it determines which c_i are admissible and what weight each carries upon commitment.

We therefore require a function

$$H : \Delta_{\{N-1\}} \rightarrow \mathbb{R}$$

assigning a real number to each admissible weighting $\mathbf{p} = (p_1, \dots, p_N)$ with $p_i \geq 0$ and $\sum p_i = 1$, where $H(\mathbf{p})$ measures the unresolved physical content of the pre-commitment configuration.

A note on the simplex. Taking the pre-commitment description to be a normalised non-negative weighting $\mathbf{p} \in \Delta_{\{N-1\}}$ is itself a substantive choice. A richer pre-commitment object — for instance, one that projects to a simplex only at the moment of commitment, the canonical example being a quantum state, which projects to a simplex over outcomes only once a measurement basis is fixed — is conceivable, and this is precisely where classical-versus-quantum representational ambiguity arises. We take $\Delta_{\{N-1\}}$ as the *minimal description sufficient for the present argument*: a measure over admissible alternatives is required by the structural demands of §3, and $\Delta_{\{N-1\}}$ is the smallest object that supplies one. The question of whether a richer pre-commitment object is required, and of which mathematical structure realises it, is a representation-layer question addressed in prior work (§8); it is deferred here to keep the information-layer argument self-contained.

The structural question for the present paper is: **what constraints on H follow from A0–A1' or are motivated by the framework, and which functions H satisfy them?**

4. Constraints on the Uncertainty Measure

We present five constraints, each grounded in a VERSF axiom or structural feature — three forced from A0–A1', two motivated as physical desiderata, with the distinction made explicit per constraint. We note in advance one subtlety in the partition: the forcedness of 4.3 (strict monotonicity at the uniform distribution) is conditional on 4.1 and 4.4–4.5. In isolation, A0 + A1 force only that admissible alternatives contribute non-null pre-commitment structure; strict monotonicity at the uniform distribution requires additionally the contribution-localisation supplied by 4.1 (permutation invariance) and 4.4–4.5 (which fix where in H the new alternative's contribution sits). The forced/motivated split is therefore most accurately stated as "two forced from A0–A1' alone, one forced given the others, two motivated as desiderata," though for the conclusion's purposes the simpler "three forced, two motivated" framing is sufficient.

4.1 Permutation invariance (from A0). Labels c_i are not physical content; only the set of distinguishable alternatives is. For any permutation $\sigma \in S_N$:

$$H(p_1, \dots, p_N) = H(p_{\{\sigma(1)\}}, \dots, p_{\{\sigma(N)\}})$$

4.2 Continuity. Physical descriptions must be stable under small perturbations of admissible weightings. H is continuous on $\Delta_{\{N-1\}}$.

4.3 Monotonicity for the uniform distribution. For $p_i = 1/N$, H must be a strictly increasing function of N . Adding a genuinely admissible alternative cannot decrease pre-commitment uncertainty, and we claim further that it cannot leave H unchanged. The grounding is: a new alternative that is genuinely admissible under $A1$ and observer-invariantly so under $A0$ cannot have null contribution to pre-commitment structure without violating $A0$ itself, since it would constitute an admissible-but-physically-empty distinction — admitted by the framework but contributing nothing to the system's unresolved content. Strict monotonicity rather than mere non-decrease is therefore required. (Strict monotonicity at the uniform distribution follows once 4.1 is in hand and the contribution of the new alternative is localisable as its own term, which 4.4–4.5 supply; in isolation, $A0 + A1$ force "non-null contribution" but not yet "strict increase at uniform.")

4.4 Compositional additivity (motivated, not strictly forced). For independent systems A and B , where independence is defined operationally by factorisation of the joint admissibility weighting,

$$p^{AB}(i,j) = p^A(i) \cdot p^B(j)$$

we require:

$$H(AB) = H(A) + H(B)$$

(The joint description is itself a weighting over the product admissible set — a simplex of dimension $N_A \cdot N_B - 1$ — inheriting the simplex commitment of §3.)

We are explicit that additivity is a *physical desideratum motivated by VERSF rather than strictly forced from A0 alone*. Mathematically, alternative associative composition rules $f(H_A, H_B)$ — monotone, symmetric, and consistent with operational independence — are conceivable. The Rényi entropy family realises a non-Shannon functional under additivity, and the Tsallis family realises a non-additive (q -additive) composition rule; both are excluded by the joint imposition of 4.4 and 4.5 in the present forms, but their existence shows that the choice of additivity is not mathematically inevitable. Any cardinality-counting argument $\log(N_A \cdot N_B) = \log N_A + \log N_B$ presupposes the logarithmic form, so the appeal to "uncertainty composes additively" cannot be cashed out from $A0$ without circularity. What VERSF supplies is a principled motivation: uncorrelated independent systems should contribute *independently* to total unresolved structure, and additivity is the unique composition rule that requires no additional structural specification beyond independence. Any non-additive alternative would commit VERSF to a specific functional form for the composition rule — additional structure for which the framework supplies no motivation. The honest statement is therefore Occam-style: additivity is the minimal composition law on top of operational independence, and VERSF gives no

positive reason to prefer any alternative; this is weaker than "forced" but stronger than "convention," and it is the most we will claim.

4.5 Coarse-graining consistency (motivated, not strictly forced). Grouping unresolved alternatives must preserve physical meaning. For any partition of $\{c_1, \dots, c_N\}$ into groups, we require that the uncertainty decomposes as the group-level uncertainty plus the weighted within-group uncertainties. Concretely, for a binary grouping of the first two alternatives:

$$H(p_1, p_2, p_3, \dots, p_N) = H(p_1 + p_2, p_3, \dots, p_N) + (p_1 + p_2) \cdot H(p_1/(p_1+p_2), p_2/(p_1+p_2))$$

(Combined with 4.1, this binary form gives the general arbitrary-partition form by induction.) This is the recursive identity that does the actual mathematical work in Khinchin- and Faddeev-style uniqueness theorems, and we are explicit that — like §4.4 — its grounding is motivational rather than strictly forced from A0–A1' alone. Why the weighting should be exactly $p_1 + p_2$ rather than some other monotone function of the group probability is a question the framework does not settle without circularity. The VERSF analogue of Khinchin's original sequential-measurement motivation is that *sequential resolution and one-shot resolution must agree on the total uncertainty resolved*: resolving first the group, then within the group, must yield the same total H as resolving everything at once, with the within-group contribution weighted by the admissibility weight of the group. This is the consistency condition between partial and full resolution, and the framework gives no positive reason to prefer any alternative — but, like additivity, it is a desideratum rather than a forced consequence.

5. Uniqueness of Shannon Entropy

Constraints 4.1–4.5 match a standard formulation of the Khinchin–Faddeev axioms. The precise axiom set varies across formulations: Khinchin's recursion covers arbitrary partitions, Faddeev's binary form (4.5 above) is provably equivalent given symmetry, and Aczél–Daróczy give a measure-theoretic version with continuity weakened to measurability. The uniqueness conclusion is robust across these variants.

Theorem 1 (Information Necessity in VERSF). *The unique function $H : \Delta_{\{N-1\}} \rightarrow \mathbb{R}$ satisfying constraints 4.1–4.5 is, up to a positive multiplicative constant K ,*

$$H(\mathbf{p}) = -K \sum_i p_i \ln p_i$$

Proof. Under axiom set 4.1–4.5, the result follows from Faddeev's uniqueness theorem (Faddeev 1956), of which the binary recursion in 4.5 is the central premise; the same conclusion holds under Khinchin's stronger general-partition formulation (Khinchin 1957) and the measure-theoretic axiomatisation of Aczél–Daróczy (1975). Shannon (1948) gave the original derivation under a related but distinct axiom set. The VERSF contribution is not the theorem itself but the demonstration that its premises are grounded by A0, A1, A1' and the structure of pre-commitment description — partly forced (4.1, 4.2, 4.3) and partly motivated as physical desiderata (4.4, 4.5) — rather than imposed by stipulation. ■

The constant K parameterises a unit choice. With H written using the natural logarithm as above, $K = 1$ yields nats, $K = 1/\ln 2$ yields bits, and $K = k_B$ yields thermodynamic units. The choice of unit is conventional; what carries physical content is the proportionality structure of H , and the calibration $K = k_B$ that makes Shannon entropy commensurate with thermodynamic entropy is supplied by $A1'$ — see §6.

6. Thermodynamic Connection

The unit constant K in Theorem 1 is conventional, but $A1'$ supplies a physical scale ($k_B \ln 2$ per binary distinction) that, on identification $K = k_B$, makes Shannon entropy and thermodynamic entropy commensurate. This is the resolution of the apparent tension between §5's claim that K is conventional and §6's treatment of $K = k_B$ as physically meaningful: K parameterises a unit, while the proportionality between information resolution and thermodynamic entropy production — fixed by $A1'$ at one binary distinction per $k_B \ln 2$ — is what carries physical content.

By $A1'$, fact formation requires irreversible entropy production of at least $k_B \ln 2$ per resolved binary distinction. Information loss and thermodynamic entropy production therefore stand in fixed minimum proportionality within VERSF — $k_B \ln 2$ of thermodynamic entropy per bit of information resolved — with the entropy cost of any commitment event bounded below by this rate. The cost can exceed the floor (most thermodynamic processes do); what is fixed is the proportionality at the minimum.

The full development of thermodynamics from this base — the derivation of the Boltzmann form $S = k_B \ln W$ from admissibility constraints on coarse-graining, the canonical distribution from closure-compatible entropy maximisation, the first law as macrostate-space differential geometry, and the second law as a structural (rather than statistical) consequence of the non-cancellation of commitment increments — is given in Taylor, *Thermodynamics from Irreversible Commitment* (VERSF Theoretical Physics Programme). The present paper supplies the foundational layer that grounds the entropy functional used there: §4 of that paper establishes Gibbs–Shannon entropy from constraints (C1)–(C3) on closure-compatible probability assignments, and Theorem 1 above is the framework-level result from which those constraints inherit their non-circularity. This relationship is asymmetric and worth stating explicitly: the present paper supersedes any earlier formulation of the VERSF programme in which the (C1)–(C3) constraints, or the Gibbs–Shannon functional more generally, were assumed without framework-level justification. Theorem 1 closes that dependency.

7. Reframing of Information Theory

Theorem 1 reframes the status of information theory:

- Information is not primitive — it is a *consequence* of distinguishability and admissibility.

- Information is not optional — Shannon entropy is the *unique* admissible measure.
- Information is not purely epistemic — within VERSF, the measure over the admissible set is fixed by observer-invariant structure, even though individual observers may face additional epistemic uncertainty about which admissible set obtains.

The third point requires care. We are not claiming that pre-commitment uncertainty is wholly observer-independent (that would contradict A0's requirement that physical content be observer-*invariant*, not observer-free), nor that interpretations emphasising the agent-relative role of probability (e.g. QBism) are mistaken about what they describe. We are claiming that, *given* the observer-invariant set of admissible alternatives, the measure of unresolved structure over that set is fixed by the framework and is not an additional epistemic input.

8. Scope and the Representation Question

Theorem 1 selects Shannon entropy as the unique admissible measure of pre-commitment uncertainty. It does *not* select a representation of that uncertainty.

Information theory determines the structure of pre-commitment uncertainty but does not fix its representation. Both classical probability theory (states are points in $\Delta_{\{N-1\}}$) and quantum mechanics (states are rays in a Hilbert space, with Shannon entropy applied to measurement outcomes) admit Theorem 1 as a description of unresolved structure. The selection between them — and more specifically, the selection of finite-dimensional complex quantum mechanics as the admissible representation — has been established in prior work within the VERSF programme (Taylor, *From Information to Quantum Mechanics*; see References), drawing on the discrete commitment substrate (TPB) together with additional postulates (linearity and compositional closure of joint systems) standard in operational reconstructions of quantum theory.

The relationship between the two results is now explicit:

- **Prior result (representation layer):** information structure + TPB + additional postulates \Rightarrow finite-dimensional quantum mechanics.
- **Present result (information layer):** VERSF foundational axioms (A0, A1, A1'), together with the §4 desiderata whose VERSF grounding is argued, entail Shannon entropy uniquely.

Together they establish that the information-theoretic input to the representation derivation is grounded in VERSF rather than imported externally — partly forced from A0–A1', partly motivated as physical desiderata, with the distinction made explicit in §4. The separation between the layers is deliberate: the information layer is self-contained on A0, A1, A1' plus the §4 desiderata, whereas the representation layer is conditional on further structural axioms whose status as derivations or postulates is properly argued on its own terms.

9. Scope and Limitations

We are explicit about what the present derivation does and does not establish.

1. **Finite-N only.** A1 asserts a finite admissible set. Countable-infinite and continuous admissibility structures are outside the present scope, and Theorem 1's premises (in particular monotonicity in \mathbb{N} and the discrete coarse-graining identity) would require generalisation.
2. **Admissibility is primitive.** What *selects* the admissible set \mathcal{C} from a more fundamental physical substrate is not addressed here. A0 fixes that \mathcal{C} is observer-invariant; it does not derive \mathcal{C} itself.
3. **Observer-invariant versus observer-relative uncertainty.** We have argued that pre-commitment uncertainty quantifies observer-invariant unresolved structure, but a complete account of how this relates to observer-relative epistemic uncertainty in standard interpretations (Copenhagen, QBism, relative-state) is left for separate treatment.
4. **Representation is not addressed.** As stated in §8, the representation question — which mathematical structure realises this information layer under the full VERSF axiom set — is the subject of prior work in the VERSF programme. The present paper does not reproduce or revise that result; it grounds the information-theoretic layer that the prior result took as input.
5. **Landauer is encoded, not derived.** §6 establishes proportionality between information resolution and thermodynamic entropy production, but the lower bound $k_B \ln 2$ per binary distinction is asserted by A1' rather than derived from a more primitive thermodynamic principle.

10. Conclusion

This work establishes that **information theory is structurally entailed by VERSF**, modulo the desiderata in §4 whose VERSF grounding we have argued. Given observer-invariant distinguishability (A0), finite admissibility (A1), and irreversible fact formation (A1'), the Khinchin axioms hold of any admissible measure of pre-commitment uncertainty — three of them forced by VERSF's foundational principles, two of them motivated as physical desiderata that the framework supplies a principled (though not strictly forcing) reason to adopt. Under these axioms, Shannon entropy is uniquely selected.

The result provides the structural object that any admissible physical representation must encode. It does not, on its own, fix what that representation must be, and the present paper does not attempt to. The representation question has been addressed under the broader VERSF axiom set in prior work; the contribution here is to ground the information-theoretic layer that the prior derivation used as input, closing the dependency by deriving it — to the extent it can be derived — from VERSF's foundational principles.

The broader implication, within the present scope, is that physics does not merely *use* information theory: under VERSF's foundational principles, information-theoretic structure is structurally entailed by the framework, Shannon entropy is uniquely selected once the §4 desiderata are granted, and the language describing pre-commitment physical reality is therefore not an arbitrary external import but a structural feature of VERSF whose premises are framework-grounded — forced where they can be, motivated where they cannot — rather than stipulated.

References

Aczél, J., & Daróczy, Z. (1975). *On Measures of Information and Their Characterizations*. Academic Press.

Faddeev, D. K. (1956). On the concept of entropy of a finite probabilistic scheme. *Uspekhi Matematicheskikh Nauk*, 11, 227–231.

Khinchin, A. I. (1957). *Mathematical Foundations of Information Theory*. Dover.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5, 183–191.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.

Taylor, K. *From Information to Quantum Mechanics in the VERSF Framework*. VERSF Theoretical Physics Programme.

Taylor, K. *Thermodynamics from Irreversible Commitment: A Closure-to-Macrostate Derivation in the VERSF Framework*. VERSF Theoretical Physics Programme.