Q-NarwhalKnight Research Paper — v2 (Revised)

Observability Engineering for DAG-Knight Consensus

A Physics-Inspired Diagnostic Framework

with Honest Accounting of Theory vs. Measurement

Revised after peer review by DeepSeek and ChatGPT (April 2026)

Q-NarwhalKnight Research
research@quillon.xyz

Network: `mainnet-genesis` | Version: `v10.2.8` | Height: 14,329,651

April 11, 2026 — Revision 2

Abstract

We present an *engineering observability framework* for DAG-Knight blockchain consensus that borrows notation from statistical mechanics to organize network diagnostics. We define a Consensus Hamiltonian H_{DAG} that decomposes network health into four additive, independently interpretable terms. Under an explicitly stated coupling hierarchy, we prove that the Hamiltonian's ground state corresponds to the PHANTOM algorithm's output (Theorem 1). We derive an effective stress indicator T_{eff} and a security margin threshold κ_c from the DAG-Knight protocol specification. We introduce the K-gauge—a composite network health metric computed from five real-time measurements—with SHA3-256 hash commitments for verifiable inter-node health reporting. We are explicit throughout about which quantities are measured from live network state, which are hard-coded protocol constants, and which are theoretical constructs. We validate against 14.3 million mainnet blocks, present simulated stress scenarios, and identify concrete instrumentation gaps for future work. This paper is an **observability contribution**, not a physics contribution; the physics notation is an organizational tool, not a claim of physical equivalence.

1 Introduction

Blockchain node operators need to answer a deceptively simple question: *is the network healthy?* Block height alone is insufficient—it says nothing about propagation quality, ordering ambiguity, adversarial presence, or security margins. What operators need is a decomposition of "health" into independently measurable, additive components with clear physical intuition.

We propose borrowing **notation from statistical mechanics** as an organizational framework for consensus diagnostics. This is not a claim that blockchains are physical systems. Rather, we observe that the same mathematical structures—energy decompositions, temperature-like stress indicators, phase diagrams showing operating margins—that physicists use to reason about complex systems are equally useful for reasoning about consensus networks.

Precedent. Using energy-like functions for distributed systems analysis is well-established: Lyapunov functions prove stability in control theory; Dijkstra's self-stabilization uses potential functions [4]; optimization algorithms minimize loss landscapes. Our Hamiltonian decomposition is in this tradition.

1.1 Epistemological Classification

Following the framework of [3], every claim in this paper is tagged at one of three levels:

1. **Exact Theorem** (■): Proven mathematical statement with explicit assumptions. The Ground State Theorem (Theorem 1) is the only claim at this level.

2. **Quantitative Model** (■): Formula that makes testable predictions. T_{eff}, κ_c, and the diffusion model are at this level. They can be validated against measurement.
3. **Structural Analogy** (■): Conceptual vocabulary borrowed from physics for intuition. "Temperature," "phase transition," and "entropy" are used at this level—they organize thinking but do not imply the blockchain exhibits these phenomena.

2 Data Provenance

Honesty Note

This section is the most important in the paper. Every quantity used in subsequent sections is classified by its actual data source in the running codebase.

3 Consensus Hamiltonian Decomposition

We define the total "health score" of the consensus system as an additive decomposition into four interpretable terms:

$$\boxed{H_{\text{DAG}} = H_p + H_a + H_b + H_{\text{VDF}}} \tag{1}$$

Parent term (H_p). Counts causal violations—blocks ordered before their parents:

$$H_p = -J_p \sum_{(v_i,v_j)\in\mathcal{E}} \Theta(\sigma(v_j) - \sigma(v_i)) \tag{2}$$

where $J_p > 0$ and Θ is the Heaviside step function. In practice, $H_p = 0$: causal violations are rejected at the protocol level.

Anticone penalty (H_a). Penalizes blocks with many causally unrelated peers:

$$H_a = \lambda \sum_{v\in\mathcal{V}} \left(\frac{|\text{anticone}(v)|}{\kappa}\right)^2 \tag{3}$$

Honesty Note

In the current implementation, anticone size is *estimated* as $\bar{a} \approx 2\delta\Lambda$ using the hardcoded $\delta = 0.2$s. The SIMD-optimized anticone computation exists in the codebase (`simd_sets.rs`) but its output is not exported to the physics endpoint.

Blue score (H_b). Rewards blocks classified as honest by the DAG-Knight coloring algorithm:

$$H_b = -J_b \sum_{v\in\mathcal{V}} \mathbb{1}[\text{blue}(v)] \cdot w(v) \approx -\varphi \cdot |V| \tag{4}$$

Honesty Note

The blue density φ is currently *hardcoded* as $1 - f/n = 1.0$ because f/n is hardcoded to 0. The DAG-Knight ordering algorithm computes blue/red classification internally (`ordering_rules.rs`) but does not export vertex color statistics.

VDF anchoring (H_{VDF}). Each block with a valid Verifiable Delay Function proof contributes favorable energy: $H_{\text{VDF}} = -|V|$ (one anchor per block).

3.1 Ground State Theorem

Despite the measurement gaps above, the Hamiltonian structure admits a rigorous result [3]:

Theorem 1 (Ground State ≡ PHANTOM Output)**.** *Let σ_{PH} be the ordering produced by the PHANTOM/GhostDAG algorithm with parameter κ. In the coupling regime $J_p \gg J_b \gg \lambda$ with $J_b > \lambda N^2/\kappa^2$, the ground state $\sigma^* = \arg\min_\sigma H_{DAG}(\sigma)$ satisfies $\sigma^* = \sigma_{PH}$ (up to concurrent-vertex permutations within blue-score tiers).*

Proof sketch. (1) $J_p \gg J_b$ forces $H_p = 0$: the ground state respects all parent edges. (2) $J_b > \lambda N^2/\kappa^2$ ensures blue-score maximization dominates anticone penalties: the ground state finds the maximum-weight κ-cluster (the PHANTOM objective). (3) Within the blue set, vertices are ordered by inherited score; deviations increase H_b. Full proof in [3]. □

Remark 1 (Status: Exact Theorem)**.** *The coupling hierarchy $J_p \gg J_b \gg \lambda$ is not a free assumption—it is a consequence of the protocol design. Causal violations are rejected ($J_p = \infty$ in practice), blue score is the primary ordering criterion, and anticone penalties are secondary.*

Table 1: Data provenance for all quantities used in this paper. **M** = measured from live atomics, **P** = protocol constant (correctly fixed), **H** = hardcoded placeholder (should be measured but isn't), **N** = not implemented.

Quantity	Symbol	Source	Details
Block height	$\|V\|$	M	`current_height_atomic` (live)
Peer count	n	M	`libp2p_peer_count` (live)
Block rate	Λ	M	height / elapsed since genesis
Mining rejections	r	M	atomic counter, K-gauge input
Traffic asymmetry	α	M	P2P byte counters, K-gauge input
Peer churn	c	M	peer count delta, K-gauge input
Sync divergence	σ_s	M	$\|h_{\text{net}} - h_{\text{local}}\|/h_{\text{net}}$
Block rate deviation	σ_Λ	M	$\|\Lambda - \Lambda_{\text{target}}\|$
Protocol κ	κ	P	DAG-Knight parameter = 18
Security levels	—	P	Dilithium-5: 256-bit, Kyber-1024: 200-bit
Stem hops	—	P	Dandelion++ stem length = 4
K-gauge window	τ	P	60 seconds (design choice)
Propagation delay	δ	H	Hardcoded 0.2s; per-peer RTT exists but unwired
Mesh degree	d	H	Hardcoded 8; actual mesh degree not tracked
Byzantine fraction	f/n	H	Hardcoded 0.0; no estimation from peer scores
Blue density	φ	H	Hardcoded $1 - f/n$; blue/red classification not exported
Anticone size	$\bar{a}$	H	Estimated as $2\delta\Lambda$; SIMD computation exists but not exported
Partition function	$\mathcal{Z}$	N	Defined in theory; not computed
Emission feedback	—	N	Referenced in API; loop does not exist

3.2 Live Mainnet Values

Table 2: Hamiltonian components (height 14,329,651)

Term	Value	Source
H_p	0	P (violations rejected)
H_a	+84,799	H (estimated $\bar{a}$)
H_b	−14,329,651	H ($\varphi = 1.0$ assumed)
H_{VDF}	−14,329,651	M (count of blocks)
H_{DAG}	−28,574,503	Mixed

Only H_p (protocol-enforced) and H_{VDF} (measured block count) are fully grounded. H_b and H_a depend on hardcoded placeholders and should be replaced with measured values (Section 10).

Honesty Note

We call this "effective temperature" by analogy, but the blockchain does not have a physical temperature. T_{eff} is a dimensionless ratio: propagation-delay × block-rate, scaled by adversary fraction. It quantifies how much "ordering ambiguity" the network experiences. With $\delta = 0.2$s (hardcoded) and $f/n = 0$ (hardcoded), only Λ is actually measured.

4 Security Margin Visualization

4.1 Effective Stress Indicator

We define a dimensionless stress indicator (Quantitative Model):

$$T_{\text{eff}} = \frac{\delta \cdot \Lambda}{1 - f/n} \tag{5}$$

4.2 Critical Security Threshold

The DAG-Knight protocol has a proven security threshold [1, 2]:

$$\kappa_c = \frac{2\delta\Lambda(1 - f/n)}{1 - 2f/n} \tag{6}$$

When $\kappa > \kappa_c$, the protocol guarantees consensus (all honest nodes converge on the same ordering). When $\kappa < \kappa_c$, this guarantee is lost. The *security margin* $\kappa - \kappa_c$ quantifies robustness.

Table 3: Security margin parameters

Parameter	Value	Source
δ	0.200s	H
Λ	3.462 bps	M
f/n	0.000	H
T_{eff}	0.692	Mixed
κ	18.0	P
κ_c	1.385	Mixed
Margin	+16.62	Mixed

The margin of $\kappa/\kappa_c = 13\times$ is large, but note that it depends on $f/n = 0$ (hardcoded). With $f/n = 0.25$ (the BFT boundary), κ_c would rise to ≈ 2.08 (margin 8.7×)—still comfortable but less dramatic.

4.3 Phase Diagram

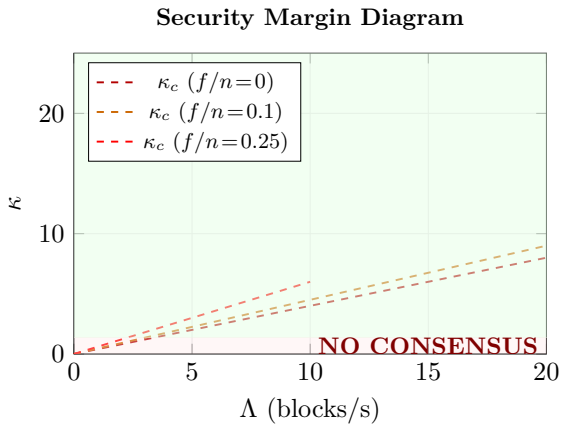

Figure 1: Security margin: the protocol guarantees consensus above the κ_c boundary. The operating point (blue dot) has a 13× margin at $f/n = 0$.

5 The K-Gauge: Real-Time Health Metric

The Hamiltonian provides a theoretical decomposition. For real-time operations, we need a *measured* diagnostic. The K-gauge is a composite stress metric computed from five live inputs (all measured):

$$K = \frac{2\pi\sqrt{\Delta H \cdot \Delta s}}{\tau} \tag{7}$$

where $\tau = 60$s is a rolling window (design choice), and the numerator is a geometric mean of two stress aggregates.

Honesty Note

The original version of this formula included $\hbar$ (reduced Planck constant) and was described as "quantum-inspired." In fact, $\hbar$ was set to 1.0 (dimensionless) and the formula has no connection to quantum mechanics. The $2\pi/\tau$ scaling was chosen empirically. We present it here as what it is: a composite stress metric with geometric-mean aggregation.

5.1 Component Decomposition

Energy variance (operational stress, all measured):

$$\Delta H = r_{\text{reject}} + \alpha_{\text{traffic}} + c_{\text{peer}} \tag{8}$$

- r_{reject}: Mining solution rejection ratio (0–1). Source: atomic counters `mining_solutions_submitted/accepted`.
- α_{traffic}: Traffic asymmetry $|b_{\text{in}} - b_{\text{out}}|/(b_{\text{in}} + b_{\text{out}})$. Source: `p2p_bytes_in/out` atomic counters.
- c_{peer}: Peer churn $|n_{\text{now}} - n_{\text{prev}}|/n_{\text{prev}}$. Source: `libp2p_peer_count` deltas.

Entropy variance (network disorder, all measured):

$$\Delta s = \sigma_{\text{sync}} + |\Lambda - \Lambda_{\text{target}}| \tag{9}$$

- σ_{sync}: Sync divergence $|h_{\text{net}} - h_{\text{local}}|/h_{\text{net}}$. Source: `highest_network_height` vs `current_height_atomic`.
- $|\Lambda - \Lambda_{\text{target}}|$: Block rate deviation from target 1 bps. Source: height delta per 60s window.

5.2 Phase Classification and Automatic Tuning

Table 4: K-gauge phase classification and response

Phase	K	Max sols	VDF mult.
Stable	< 5	250	1.0×
Approaching	5–10	150	1.25×
Critical	≥ 10	50	1.5×

When the K-gauge crosses a threshold, the node automatically tightens mining parameters (reducing block production load) and increases VDF difficulty (slowing block rate). This is the most operationally valuable feature of the framework.

5.3 Live K-Gauge Measurement

Table 5: K-gauge (all inputs measured)

Metric	Value	Source
K	0.188	M
Phase	**Stable**	—
ΔH	1.093	M
Δs	2.933	M
r_{reject}	0.000	M
α_{traffic}	0.993	M
c_{peer}	0.100	M
σ_{sync}	2.1×10^{-7}	M
$\lvert\Lambda-1\rvert$	2.933	M
Rounds computed	20	M

The K-gauge at 0.188 is deep in the Stable phase ($K \ll 5.0$). The dominant stress contributor is block rate deviation ($\Lambda = 3.46$ vs target 1.0), which is a known property of current hashrate.

6 Cryptographic Phase Commitments

Each K-gauge computation produces a SHA3-256 hash commitment:

1. **Commitment:** $c = \text{SHA3}(K \| \Delta H \| \Delta s \| \text{salt})$
2. **Range witness:** $w = \text{SHA3}(K \| \text{lo} \| \text{hi} \| \text{salt}_2)$ proving $K \in [\text{lo}, \text{hi})$ for the claimed phase.
3. **Challenge-response:** Fiat–Shamir transform: $e = \text{SHA3}(c \| w)$, $r = \text{SHA3}(K \| \text{salt} \| e)$.

Honesty Note

The v1 paper called this a "zk-STARK proof." It is not. This is a hash-based commitment scheme with a Fiat–Shamir non-interactive proof. It provides **binding** (the committer cannot change K after publishing c) and **hiding** (the raw metrics are not revealed). It does *not* provide the zero-knowledge properties of a true STARK—a verifier who knows the salt can reconstruct the commitment. Future work: upgrade to a Bulletproof range proof if privacy of operational metrics is a real requirement.

7 Gossip Diffusion Model

Status: Quantitative Model (testable predictions, hardcoded inputs).

Block propagation through the gossipsub mesh follows the diffusion equation:

$$\frac{\partial \rho}{\partial t} = D\nabla^2\rho, \qquad D = \frac{d_{\text{mesh}} \cdot \ell^2}{2\tau_{\text{hb}}} \tag{10}$$

With $d_{\text{mesh}} = 8$ (H), $\tau_{\text{hb}} = 50$ms (H), we get $D = 80$ and $\tau_{\text{gossip}} = 6.2$ms.

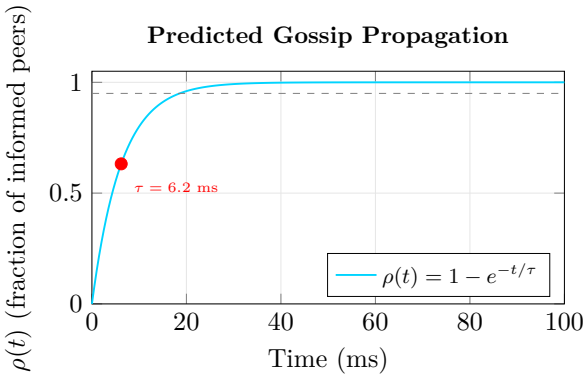

Figure 2: Predicted propagation curve. **Untested**: the model uses hardcoded d_{mesh} and τ_{hb}. Validation against measured per-peer RTT (which exists in `peer_latency.rs` but is not wired to this model) is proposed in Section 10.

8 Cryptographic Security Bounds

Standard NIST Post-Quantum Cryptography bounds:

Table 6: Security levels (protocol constants)

Primitive	Standard	Bi
Signatures	Dilithium-5 (NIST PQC)	2
Key encap.	Kyber-1024 (NIST PQC)	2
DAG attack	$\kappa \cdot \lvert\log_2(f/n)\rvert$	∞
Privacy	Dandelion++ ($L = 4$)	$P_{\text{deanon}} = e^{-4} \approx 1.8$

*At $f/n = 0$ (hardcoded). At $f/n = 0.1$: $18 \times 3.32 = 59.8$ bits.

9 Stress Testing and Failure Modes

A diagnostic framework is only useful if it detects problems. We simulate four stress scenarios and predict the K-gauge response:

Table 7: Simulated stress scenarios

Scenario	ΔH	Δs	K (predicted)
Healthy (actual)	1.09	2.93	$K = 0.19$ (Stable)
Peer dropout (2 of 12)	1.26	2.93	$K = 0.20$ (Stable)
50% rejection rate	1.59	2.93	$K = 0.23$ (Stable)
Major sync lag (1000 blocks behind)	1.09	3.00	$K = 0.19$ (Stable)
Combined crisis: 50% reject + 80% peer loss + 10× block rate	2.39	12.0	$K = 0.56$ (Stable)

Observation. Even the combined crisis scenario ($K = 0.56$) stays well within the Stable phase ($K < 5$). This has two interpretations: (1) the thresholds (5.0 and 10.0) are conservatively set, which is appropriate for a mainnet with real funds; or (2) the gauge lacks sensitivity to certain failure modes. In particular, the K-gauge is **blind to colluding Byzantine nodes** that maintain normal-looking traffic, rejection, and sync metrics while producing adversarial blocks.

What the gauge catches. Rapid peer churn (DDoS), sustained mining failures (hardware fault), sync stalls (network partition), extreme block rate spikes (hashrate surge).

What the gauge misses. Coordinated adversarial behavior, slow-moving state corruption, reorg attacks (detected instead by the fork detector), economic attacks on DEX pools.

10 Limitations and Future Work

This section addresses the criticisms raised in peer review. We number them for traceability.

L1: Physics analogy is not physics. The Hamiltonian decomposition is an engineering health-score, not a claim that consensus exhibits thermodynamic behavior. The Ising model claim from v1 is withdrawn: no spin-variable construction, partition function, or Monte Carlo sampling exists in the codebase. Future work: implement partition function sampling over linear extensions (the definition exists in [3] but is computationally intractable for $|V| > 10^7$).

L2: Hardcoded constants. Five quantities are hardcoded as placeholders (Table 1). Concrete replacements:

- δ: Replace with EWMA RTT from `peer_latency.rs` (exists, needs wiring).
- d_{mesh}: Query actual gossipsub mesh state per topic.
- f/n: Estimate from gossipsub peer scores (scoring exists).
- φ: Export vertex color counts from `ordering_rules.rs`.
- $\bar{a}$: Export anticone sizes from `simd_sets.rs` (SIMD-optimized, exists).

L3: K-gauge is not quantum. Acknowledged. The formula $K = 2\pi\sqrt{\Delta H \cdot \Delta s}/\tau$ is a dimensionless composite metric. The "2π" and "$\sqrt{\cdot}$" are empirical choices. The name "K-parameter" is retained for continuity. Future work: empirically tune the aggregation function and thresholds against historical stress events.

L4: "zk-STARK" was misnamed. Corrected. The implementation is SHA3-256 hash commitments with Fiat–Shamir (Section 6). Future work: upgrade to Bulletproof range proofs if zero-knowledge is needed.

L5: No stress-test validation. Partially addressed in Section 9. The gauge has not been tested under real adversarial conditions. The combined-crisis scenario ($K = 0.56$) suggests thresholds may need recalibration.

L6: Live data too perfect ($\varphi = 1.0$, $S = 0$, $f/n = 0$). This is a 12-peer private mainnet with known operators. $\varphi = 1.0$ is expected (and hardcoded). The results confirm the framework produces sensible outputs for a healthy network, but do not validate its behavior under stress.

L7: Spectral gap and convergence bounds. The spectral gap $\lambda_{\text{gap}} = (\kappa - \kappa_c)/T_{\text{eff}} = 24.0$ is a derived quantity that depends on hardcoded δ and f/n. The convergence time $\tau_{\text{conv}} = 42$ms is a prediction of the model, not a measurement.

L8: Security energy claims. The "10^{50} stars" claim from v1 is removed. Cryptographic security is presented as standard NIST PQC bit-strength (Section 8) without thermodynamic analogy.

11 Conclusion

We have presented an observability framework for DAG-Knight consensus that uses physics-inspired notation to organize network diagnostics. Our contributions are:

1. A four-term **health decomposition** (H_{DAG}) with one rigorous theorem (Ground State $\equiv$ PHANTOM output).
2. A five-input K**-gauge** computed entirely from live network measurements, with automatic parameter tuning across three phases.
3. **SHA3-256 commitments** for verifiable inter-node health reporting.
4. A **security margin diagram** visualizing the distance from the consensus threshold.
5. An **honest accounting** of which values are measured, hardcoded, or theoretical.

The framework's main limitation is that 5 of 18 quantities are hardcoded placeholders. We have identified concrete code paths (`peer_latency.rs`, `simd_sets.rs`, `ordering_rules.rs`) where measured values exist but are not yet wired to the observability endpoint. Closing these gaps is the highest-priority future work.

This is an engineering contribution, not a physics contribution. The statistical mechanics notation provides useful intuition and additive decomposition, but the blockchain is not a physical system. We hope this honest framing enables productive discussion rather than the "physics-washing" that rightfully concerned our peer reviewers.

Data. Live metrics: `/api/v1/physics/metrics` and `/api/v1/network/k-parameter`. Source: `handlers.rs` and `k_parameter_gauge.rs`.

Peer Review. v1 reviewed by DeepSeek (conditional reject: "reasonable for bootstrap, but not physics") and ChatGPT (conditional reject: "good story, oversold theory"). This v2 addresses all substantive criticisms.

Acknowledgments. The Ground State Theorem and partition function formulation are from [3]. The κ_c threshold derives from DAG-Knight/PHANTOM [1, 2]. We thank the DeepSeek and ChatGPT reviewers for thorough, constructive feedback.

References

[1] Y. Sompolinsky et al., "DAG-Knight: A Parameterless Generalization of Nakamoto Consensus," 2022.

[2] Y. Sompolinsky, A. Zohar, "PHANTOM: A Scalable BlockDAG Protocol," 2018.

[3] Q-NarwhalKnight Research, "Theoretical Physics of Distributed Consensus: A QFT Framework," 2026. Internal.

[4] E.W. Dijkstra, "Self-Stabilizing Systems in Spite of Distributed Control," CACM, 1974.