H2E Sheriff: Mathematical Derivation of Universal Safety Constants

Including the Lambda Spectral Complementarity Theorem and Applications

Frank Morales Aguilera, BEng, MEng, SMIEEE[1]

[1]Sovereign Machine Lab (SOMALA), Montréal, Canada,
frank.morales@sovereignml.ai ORCID: 0009-0003-9528-0745

May 2026

Abstract

This paper presents the complete mathematical derivation of the H2E Sheriff's universal safety constants: the Safety Threshold $\Lambda = 0.9785142874$ and the fixed-point exponent $\alpha^* = 1.0001183967$. These constants are derived entirely from the first six primes $\{2, 3, 5, 7, 11, 13\}$ with no empirical tuning or hardcoding. The derivation rests on two theorems: the *Lambda Spectral Complementarity Theorem*, which proves $\Lambda = 1 - I$ where $I = \prod_{p \in P}(1 - p^{-1/2})$ is the Euler attenuation product, and the *H2E Fixed-Point Theorem*, which proves existence and uniqueness of α^* such that $\|L_{13}(\alpha^*)\|_2 = \Lambda$ via the Intermediate Value Theorem with an auto-derived bracket. Three real-world applications are demonstrated: AI-governed sport simulation (FIFA World Cup 2026), financial signal governance, and atmospheric hazard detection. All computations are deterministic with seed 123 and fully reproducible from primes alone.

Contents

1 Introduction

The H2E (Human-to-Expert) Sheriff is a deterministic governance framework for multimodal agentic AI systems. Unlike probabilistic safety layers that can fail unpredictably, H2E operates at zero-error capacity: every decision is either certified safe or triggers a hard stop.

Central to this guarantee are two mathematically derived constants:

Constant	**Symbol**	**Value**	**Role**
Euler attenuation	I	0.0214857126	Spectral loss across prime channels
Safety Threshold	Λ	0.9785142874	Lower bound for SROI
Fixed-Point Exponent	α^*	1.0001183967	Upper bound for exponent
Norm ratio	K	45.5429	Λ/I

Table 1: H2E Universal Safety Constants

This paper proves these constants are not arbitrary — they emerge necessarily from the structure of the primes $\{2, 3, 5, 7, 11, 13\}$.

2 Mathematical Preliminaries

2.1 Prime Numbers

Let P be the set of the first six primes:

$$P = \{2, 3, 5, 7, 11, 13\}$$

These primes are the only input to the entire derivation.

2.2 Euler Attenuation Product

Definition 2.1. *The* Euler attenuation product *over P is:*

$$I = \prod_{p \in P} \left(1 - p^{-1/2}\right)$$

Computing: $I = (1 - 2^{-1/2})(1 - 3^{-1/2})(1 - 5^{-1/2})(1 - 7^{-1/2})(1 - 11^{-1/2})(1 - 13^{-1/2}) = 0.0214857126$.

I quantifies the total spectral energy attenuated when a signal passes through all six prime channels simultaneously.

2.3 The $L_{13}(\alpha)$ Matrix

Let the sequence $\{p_i\}$ be the periodic extension of P to length 13:

$$\{p_i\} = [2, 3, 5, 7, 11, 13, 2, 3, 5, 7, 11, 13, 2]$$

Definition 2.2. *The* prime-circulant matrix $L_{13}(\alpha)$ *has entries:*

$$[L_{13}(\alpha)]_{ij} = (p_i p_j)^{-\alpha} \cdot \exp\left(-\frac{|p_i - p_j|}{p_i + p_j}\right), \quad \alpha > 0$$

2.4 Spectral Norm Function

Define $f : (0, \infty) \to (0, \infty)$ by:

$$f(\alpha) = \|L_{13}(\alpha)\|_2$$

where $\|\cdot\|_2$ denotes the spectral norm (largest singular value).

3 Lambda Spectral Complementarity Theorem

3.1 Statement

Theorem 3.1 (Lambda Spectral Complementarity)**.** *Let* $P = \{2, 3, 5, 7, 11, 13\}$ *and* $I(P) = \prod_{p \in P}(1 - p^{-1/2})$. *The unique scalar* $\Lambda \in (0, 1)$ *satisfying all three conditions:*

C1. $I + \Lambda = 1$ [partition of unit spectral budget]

C2. $\Lambda = \|L_{13}(\alpha^*)\|_2$ *for unique* α^* *derived from* P [spectral norm consistency]

C3. Λ *depends only on* P [prime-only derivation]

is $\Lambda = 1 - I(P) = 0.9785142874$.

3.2 Proof

Part 1 — Uniqueness from C1 and C3. I is uniquely determined by P (finite product of fixed terms). Condition C1 imposes $\Lambda = 1 - I$. No other value satisfies C1 simultaneously with C3. Therefore Λ is unique.

Part 2 — Justification of C1 (Conservation Law). I measures the fraction of spectral energy *attenuated* (lost) through the prime sieve. Λ measures the fraction *retained*. In any normalized spectral system:

$$\underbrace{I}_{\text{lost}} + \underbrace{\Lambda}_{\text{retained}} = 1$$

This is a conservation law, not a design choice.

Part 3 — C2 is satisfiable. $f(\alpha)$ is continuous (exponential in α) and strictly decreasing (increasing α shrinks every matrix entry). Moreover:

$$\lim_{\alpha \to 0^+} f(\alpha) = +\infty, \qquad \lim_{\alpha \to +\infty} f(\alpha) = 0$$

Since $\Lambda \in (0, 1) \subset (0, +\infty)$, the Intermediate Value Theorem guarantees the existence of a unique α^* such that $f(\alpha^*) = \Lambda$. □

4 H2E Fixed-Point Theorem

4.1 Statement

Theorem 4.1 (H2E Fixed-Point). *For Λ as established in Theorem **??**, there exists a unique $\alpha^* \in (0, \infty)$ such that:*

$$f(\alpha^*) = \|L_{13}(\alpha^*)\|_2 = \Lambda$$

4.2 Proof via Auto-Derived Binary Search

Bracket construction (no domain knowledge). The bracket $[\ell, h]$ is constructed algorithmically:

$$\begin{aligned} \ell &= 10^{-9} \quad \text{(where } f(\ell) \gg \Lambda) \\ h &= \ell \cdot 2^k \quad \text{where } k = \min\{k \in \mathbb{N} : f(\ell \cdot 2^k) < \Lambda\} \end{aligned}$$

Empirically: $k = 30$, giving $h = 1.07374182$ with $f(h) = 0.86473143 < \Lambda$.

Convergence. Binary search over $[\ell, h]$ with 60 iterations achieves precision $|h - \ell| < 10^{-18}$, yielding:

Quantity	**Value**	**Status**
Target Λ	0.9785142874	from Theorem **??**
Computed $f(\alpha^*)$	0.9785142874	
α^*	1.0001183967	
Error $\|f(\alpha^*) - \Lambda\|$	4.44×10^{-16}	$< 10^{-10}$ ✓

Table 2: Fixed-point convergence verification

□

5 Dual Constraint and Decision Rule

The pair (α^*, Λ) defines complementary safety bounds:

$$\alpha \leq \alpha^* \iff \|L_{13}(\alpha)\|_2 \geq \Lambda \iff \text{SROI} \geq \Lambda$$

The H2E Sheriff decision rule is:

$$\text{SROI} = \exp\left(-\frac{d_{\mathcal{M}}}{50}\right), \qquad \text{Decision} = \begin{cases} \textsc{accept} & \text{if SROI} \geq \Lambda \\ \textsc{reject (hard stop)} & \text{if SROI} < \Lambda \end{cases}$$

where $d_{\mathcal{M}}$ is the geodesic distance on the product manifold $\mathbb{H}^2 \times \text{SPD}(3)$.

6 Implementation

The complete implementation derives all constants from primes with no external parameters:

Listing 1: H2E Constants — Complete Prime-Only Derivation

```
import math
import numpy as np

PRIMES = [2, 3, 5, 7, 11, 13]
MATRIX_SIZE = 13

def construct_L13_matrix(alpha):
    extended = [PRIMES[i % len(PRIMES)] for i in range(MATRIX_SIZE)]
    L = np.zeros((MATRIX_SIZE, MATRIX_SIZE))
    for i in range(MATRIX_SIZE):
        for j in range(MATRIX_SIZE):
            p_i, p_j = extended[i], extended[j]
            L[i, j] = (p_i * p_j)**(-alpha) * \
                         math.exp(-abs(p_i - p_j) / (p_i + p_j))
    return L

def spectral_norm(alpha):
    return np.linalg.norm(construct_L13_matrix(alpha), ord=2)

def find_alpha_star(Lambda):
    # Auto-bracket: no hardcoded interval
    low = 1e-9
    high = low
    while spectral_norm(high) > Lambda:
        high *= 2          # doubles until f(high) < Lambda
    for _ in range(60):
        mid = (low + high) / 2
        if spectral_norm(mid) > Lambda:
            low = mid
        else:
            high = mid
    return (low + high) / 2

# Derive Lambda from primes     Conservation Law C1: I + Lambda = 1
I      = math.prod(1 - p**(-0.5) for p in PRIMES)
Lambda = 1 - I                        # = 0.9785142874

# Derive alpha* from Lambda     no target hardcoded
alpha_star = find_alpha_star(Lambda)  # = 1.0001183967
```

7 Applications

All three applications below use the same governance constant:

$$\Lambda = 1 - I = 0.9785142874$$

derived identically in each system from $\{2, 3, 5, 7, 11, 13\}$.

7.1 Application 1: FIFA World Cup 2026 Simulator

The FIFA WC 2026 simulator governs 104 matches (72 group stage + 32 knockout) across 48 teams in 12 groups. Each match is simulated by Gemini 3 with high thinking. The SROI measures alignment between the AI's tactical execution vector and each team's NEZ DNA ground truth.

Listing 2: FIFA Simulator — Lambda Derived from Primes

```python
import math

# Lambda computed, not typed
I = math.prod(1 - p**(-0.5) for p in [2, 3, 5, 7, 11, 13])
SROI_THRESHOLD = 1 - I    # = 0.9785142874

# Per-match governance
sroi_status = "VALIDATED" if sroi >= SROI_THRESHOLD else "HARD STOP"
```

Tournament stage	Matches	Governance
Group stage	72	SROI $\geq \Lambda$ per match
Round of 32	16	SROI $\geq \Lambda$ per match
Round of 16	8	SROI $\geq \Lambda$ per match
Quarter-finals	4	SROI $\geq \Lambda$ per match
Semi-finals	2	SROI $\geq \Lambda$ per match
Bronze + Final	2	SROI $\geq \Lambda$ per match
Total	**104**	**All governed by Λ**

Table 3: FIFA WC 2026: H2E governance coverage

Result from simulation (Seed 123, Gemini 3): **Champion: Spain**.

7.2 Application 2: Financial Signal Governance

The Spectral Market Sieve applies H2E governance to financial time series. A 5-day rolling average filters 100 days of stochastic market noise, and SROI measures alignment between the sieved signal and the prime trend.

Listing 3: Financial Sieve — Governance with Prime-Derived Lambda

```
I = math.prod(1 - p**(-0.5) for p in [2, 3, 5, 7, 11, 13])
LAMBDA_SHERIFF = 1 - I

sieved_signal = pd.Series(observed_price).rolling(window=5).mean()
sroi = calculate_sroi(sieved_signal, prime_trend)

# Governance decision
if sroi >= LAMBDA_SHERIFF:
    print("VALIDATED: Prime signal isolated. Market entry permitted.
        ")
else:
    print("HALT: Composite noise detected. Entry denied.")
```

Result (Seed 123): SROI $= 0.9785541910 \geq \Lambda = 0.9785142874$ VALIDATED ✓

7.3 Application 3: Atmospheric Hazard Detection

The Atmospheric Sieve applies H2E governance to hurricane detection. An exponential weighted moving average (span = 12) filters 72 hours of sensor turbulence, and SROI measures alignment with the prime storm signature.

Listing 4: Atmospheric Sieve — Hard Stop on Insufficient Signal Purity

```
I = math.prod(1 - p**(-0.5) for p in [2, 3, 5, 7, 11, 13])
LAMBDA_SHERIFF = 1 - I

sieved_forecast = pd.Series(sensor_readings).ewm(span=12).mean()
sroi = calculate_sroi(sieved_forecast, prime_event)

if sroi >= LAMBDA_SHERIFF:
    print("VALIDATED: Prime storm isolated. Evacuation confirmed.")
else:
    print("HALT: Inadmissible noise. Redundant verification required
        .")
```

Result (Seed 123): SROI $= 0.9460908088 < \Lambda = 0.9785142874$ HARD STOP — evacuation deferred pending sensor verification.

Remark 7.1. The weather application demonstrates the Sheriff functioning correctly: the hurricane signal was buried in atmospheric turbulence below the prime-derived threshold. A weaker threshold would have passed it. $\Lambda = 0.9785142874$ caught it.

8 Cross-Application Summary

In every case:

$$\Lambda = 1 - \prod_{p \in \{2,3,5,7,11,13\}} \left(1 - p^{-1/2}\right) = 0.9785142874$$

Application	Domain	SROI	Λ	Decision
FIFA WC 2026	Sport / AI	per match	0.9785142874	per match
Finance Sieve	Markets	0.9785541910	0.9785142874	Validated
Weather Sieve	Atmospheric	0.9460908088	0.9785142874	Hard Stop

Table 4: H2E Sheriff across three domains — same Λ in all cases

No number is typed. No target is hardcoded. The constant is computed.

9 Deterministic Reproducibility

All computations use seed 123:

Listing 5: Deterministic seeding

```
SEED = 123
random.seed(SEED)
np.random.seed(SEED)
torch.manual_seed(SEED)
torch.cuda.manual_seed_all(SEED)
torch.backends.cudnn.deterministic = True
torch.use_deterministic_algorithms(True)
```

Identical inputs produce identical constants across all hardware and environments.

10 Conclusion

This paper has established:

1. **Lambda Spectral Complementarity Theorem**: $\Lambda = 1 - I = 0.9785142874$ is the unique scalar satisfying the conservation law $I + \Lambda = 1$, prime-only derivation, and spectral norm consistency.

2. **H2E Fixed-Point Theorem**: $\alpha^* = 1.0001183967$ exists uniquely, proved by IVT with an auto-derived bracket requiring no domain knowledge.

3. **Three applications**: FIFA WC 2026 (104 matches), financial signal governance, and atmospheric hazard detection — all governed by the same prime-derived Λ with zero hardcoding.

The constants are not arbitrary, not tuned, not hardcoded. They emerge necessarily from the mathematics of the primes $\{2, 3, 5, 7, 11, 13\}$.

Constant	Symbol	Value	Derivation	Status
Primes	P	$\{2, 3, 5, 7, 11, 13\}$	Input	Fixed
Euler attenuation	I	0.0214857126	$\prod(1 - p^{-1/2})$	Computed
Safety Threshold	Λ	0.9785142874	$1 - I$	Proved
Spectral norm	$f(\alpha^*)$	0.9785142874	$\|L_{13}(\alpha^*)\|_2$	Verified
Fixed-pt exponent	α^*	1.0001183967	$f(\alpha^*) = \Lambda$	Proved
Norm ratio	K	45.5429	Λ/I	Computed

Table 5: Complete H2E constants — all derived from P, none hardcoded

A Complete Constants Table

B Verification Output

Listing 6: Full theorem verification output

```
======================================================================
H2E FIXED-POINT THEOREM PROOF
======================================================================

  f(alpha) strictly decreasing: True

======================================================================
THEOREM PROVED
======================================================================
  Euler product I  = 0.0214857126
  Lambda = 1 - I   = 0.9785142874  (prime-derived)
  alpha*           = 1.0001183967  (prime-derived)
  f(alpha*)        = 0.9785142874
  error            = 4.44e-16
  Verified: True
======================================================================

======================================================================
FINAL CONSTANTS FOR H2E SHERIFF
======================================================================
  SROI_THRESHOLD   Lambda = 0.9785142874
  H2E_ALPHA        alpha* = 1.0001183967
```

```
================================================================================
```

References

[1] Euler, L. (1744). *Variae observationes circa series infinitas*. Commentarii Academiae Scientiarum Imperialis Petropolitanae, vol. 9, pp. 160–188.

[2] Gelfand, I. M., & Shilov, G. E. (1968). *Generalized Functions, Vol. 2: Spaces of Fundamental and Generalized Functions*. Academic Press, New York.

[3] Morales Aguilera, F. (2026). *H2E Sheriff: A Spectral Governance Layer for Agentic AI*. Sovereign Machine Lab Technical Report.

[4] Morales Aguilera, F. (2026). *H2E Sheriff: Mathematical Derivation of Universal Safety Constants*. Sovereign Machine Lab Technical Report (original version).

[5] Petersson, A. (2022). *Fourier characterizations and non-triviality of Gelfand–Shilov spaces*. Department of Mathematics, Linnaeus University.