Thermodynamic Equilibrium and the Cost of Modifying General Relativity

February 27, 2026

Abstract

We investigate how thermodynamic equilibrium constrains modified gravity, and under what conditions it effectively selects General Relativity (GR) within broad theory classes. Building on Jacobson's derivation of the Einstein equation as an equation of state, we define an internal entropy production $d_i S$ for cosmological horizons via a non-equilibrium Clausius relation $\delta Q = T(dS_{\text{Wald}} + d_i S)$ and prove analytic uniqueness theorems: within $f(R)$, Brans–Dicke, Horndeski, Lovelock, and luminal DHOST theory classes, the condition $d_i S = 0$ for all spatially flat FRW cosmologies forces the field equations to reduce to Einstein's with cosmological constant.

We quantify departures from equilibrium via a dimensionless measure $N = \int |d_i S| \, d \ln a / |\Delta S_{\text{geo}}|$. GR + Λ has $N = 0$ exactly, while observationally viable modified gravity models yield $N \sim 10^{-2}$–10^{+1}. Imposing $N < 10^{-3}$ as a conservative threshold excludes large portions of EFT-of-dark-energy and DHOST parameter spaces, with compression factors of order 10^2 relative to observational constraints alone.

We motivate this threshold via QFT in curved spacetime: adiabatic vacua in slowly evolving FRW backgrounds produce near-KMS thermal horizons. Combined with the quadratic scaling of relative entropy deviations (a consequence of the entanglement first law), this yields $N \lesssim 10^{-4}$ for late-time ΛCDM. Holographic consistency arguments provide complementary theoretical motivation for equilibrium, though with weaker quantitative bounds.

These results support a **Thermodynamic Equilibrium Principle**: within the broad theory classes studied, GR + Λ is sharply preferred as the unique equilibrium point. This provides a new theoretical constraint on modified gravity, complementary to gravitational-wave and cosmological tests.

1 Introduction

1.1 Motivation: Why Einstein's Equations?

General Relativity (GR) has achieved remarkable empirical success across an extraordinary range of scales. From precision tests in the solar system [18], to binary pulsar observations [19], to the direct detection of gravitational waves from merging black holes and neutron stars [20, 21], Einstein's theory has passed every observational test with extraordinary precision. The recent gravitational wave event GW170817, combined with its electromagnetic counterpart, constrained the speed of gravitational waves to be equal to the speed of light to within one part in 10^{15} [22], ruling out vast swaths of modified gravity theories.

Yet theoretical motivations for going beyond GR remain compelling. The observed cosmic acceleration [23, 24] can be explained by a cosmological constant Λ, but the associated fine-tuning and coincidence problems have driven exploration of dynamical dark energy and modified gravity alternatives [25]. Tensions between early- and late-Universe determinations of the Hubble constant and structure growth [26] may hint at physics beyond ΛCDM. At the theoretical level, the incompatibility of GR with quantum mechanics and the appearance of singularities suggest that Einstein's theory is not fundamental but rather an effective low-energy description.

A profound conceptual shift came with Jacobson's 1995 derivation of the Einstein equation as an equation of state for local Rindler horizons [3]. Rather than positing the Einstein-Hilbert action and deriving the field equations variationally, Jacobson showed that the Einstein equation follows from:

(i) The Clausius relation $\delta Q = T\,dS$ applied to local Rindler horizons,

(ii) The Unruh temperature $T = \kappa/(2\pi)$ for accelerated observers,

(iii) The Bekenstein-Hawking entropy-area relation $S = A/(4G_N)$.

This derivation suggests that gravity is not a fundamental force but an emergent phenomenon—an equation of state for spacetime thermodynamics.

This raises a fundamental question that we address in this paper: **Can thermodynamic and information-theoretic principles not only reproduce Einstein's equations, but also *select* them uniquely among consistent modified gravity theories?**

1.2 From Consistency to Selection

Previous work has established that GR is *consistent* with thermodynamic and information-theoretic principles. The Wald entropy formula [4, 5] generalizes the Bekenstein-Hawking entropy to higher-derivative theories, and the first law of black hole mechanics holds for stationary black holes in any diffeomorphism-invariant theory. Eling, Guedens, and Jacobson [6] extended Jacobson's derivation to non-equilibrium settings, introducing an internal entropy production term d_iS that measures departures from equilibrium. They showed that $f(R)$ gravity, unlike GR, generically produces internal entropy at horizons.

More recent developments have connected horizon thermodynamics to entanglement. Jacobson [8] showed that Einstein's equations follow from an "entanglement equilibrium" condition: the total entropy (geometric plus entanglement) of small causal diamonds is stationary at fixed volume. Faulkner, Lewkowycz, and Maldacena [11] and subsequent work [16, 17] used holographic techniques to derive linearized Einstein equations from the first law of entanglement entropy.

These results demonstrate that GR is compatible with thermodynamics and entanglement equilibrium. However, they do not systematically test whether *competing theories* can also achieve such compatibility, nor do they provide quantitative bounds on how close to GR a theory must be to avoid thermodynamic inconsistencies.

Our goal in this paper is to promote "thermodynamic consistency" into a **selection principle** by:

(a) Defining an internal entropy production d_iS for general diffeomorphism-invariant theories,

(b) Proving that $d_iS = 0$ singles out GR in large theory classes,

(c) Quantifying non-equilibrium via a dimensionless measure N in cosmologies that resemble our Universe,

(d) Deriving small-N bounds from QFT and holography.

1.3 Main Contributions

This paper presents four main results, which we classify by their evidential status:

Classification of Results.

- **Rigorous theorems:** The uniqueness theorems (Section 3) are mathematically proven within specified theory classes and FRW backgrounds.
- **Numerical evidence:** Parameter space constraints (Sections 4–5) rely on numerical sampling; compression factors are sample-dependent.
- **QFT-motivated bounds:** The $N < 10^{-3}$ threshold derives from near-KMS analysis (Section 6) using standard but non-trivial assumptions.
- **Theoretical motivation:** Holographic arguments (Section 7) provide qualitative support but weak quantitative bounds.

1. Analytic Uniqueness Theorems. We prove that within the $f(R)$, scalar–tensor (Brans–Dicke), Horndeski, Lovelock, and luminal DHOST theory classes, vanishing horizon internal entropy production $d_iS = 0$ for all spatially flat FRW cosmologies forces the theory back to Einstein gravity with a cosmological constant. Specifically:

- **Theorem 3.1** ($f(R)$ gravity): $d_iS = 0 \implies f''(R) = 0 \implies f(R) = aR + b$.
- **Theorem 3.2** (Brans–Dicke): $d_iS = 0 \implies \dot{\phi} = 0 \implies \phi = \text{const}$.
- **Theorem 3.3** (Horndeski with $G_5 = 0$): $d_iS = 0 \implies G_4 = \text{const}$.
- **Theorem 3.4** (DHOST with $c_T = 1$): $d_iS = 0 \implies F_2 = \text{const}, A_1 = 0$.

These theorems apply to FRW backgrounds; extension to general spacetimes remains open.

2. Non-Equilibrium Measure N. We introduce a dimensionless, coordinate-invariant measure

$$N \equiv \frac{\int_{a_i}^{a_f} |d_iS|\, d\ln a}{|\Delta S_{\text{geo}}|}, \tag{1}$$

quantifying the cumulative internal entropy production normalized to the geometric entropy change. We show that:

- GR + Λ has $N = 0$ exactly.
- Modified gravity models consistent with current data typically have $N \gg 10^{-3}$.
- Self-accelerating models (without Λ) generically have $N = \mathcal{O}(1)$.

3. Constraints on Modified Gravity. Applying the thermodynamic prior $N < 10^{-3}$ to EFT-of-DE and DHOST parameter spaces via Monte Carlo sampling, we find substantial compression relative to observational constraints:

- **EFT-of-DE:** Compression factor ~ 100 within the sampled region ($|\alpha_{i0}| \lesssim 0.5$).

- **DHOST:** Compression factor ~ 120 (99.2% exclusion of observationally viable points).
- Only models with $|\alpha_M| \lesssim 3 \times 10^{-3}$ survive, effectively indistinguishable from GR.

These compression factors depend on prior ranges and sampling; the qualitative conclusion (strong preference for GR) is robust.

4. Origin of the Small-N Bound. We motivate $N \lesssim 10^{-3}$ from QFT considerations via a two-step argument:

- **Step 1: Near-KMS horizons.** Adiabatic vacua in slowly evolving FRW backgrounds produce near-thermal (KMS) horizons with time-averaged deviation $\bar{\varepsilon}_{\rm KMS} \approx 0.035$ (Section 6.2). This alone would give $N \lesssim 0.1$ from linear scaling.
- **Step 2: Quadratic suppression.** The entanglement first law ensures that relative entropy has *no linear term* in deviations from equilibrium (Section 6.3). Therefore $N \propto \bar{\varepsilon}^2_{\rm KMS}$, not $\bar{\varepsilon}_{\rm KMS}$. With calibration coefficient $C_2 \approx 0.077$, we obtain $N \sim C_2\, \bar{\varepsilon}^2_{\rm KMS}\, \Delta \ln a \sim 10^{-4}$.
- **Holographic motivation:** Bulk–boundary consistency in AdS/CFT provides an independent theoretical reason to expect equilibrium, though the quantitative bound from our toy model ($N < 0.5$) is weaker than the QFT result. The $1/S_{\rm BH}$ scaling suggests tighter constraints for macroscopic horizons.

1.4 Relation to Previous Work: Known Framework vs. New Contributions

To help the reader distinguish established results from new contributions, we explicitly separate the two.

Known Framework (Not Claimed as Novel). The following elements are well-established and form the foundation on which we build:

- **Jacobson (1995) [3]:** The Einstein equation as an equation of state derived from the Clausius relation on local Rindler horizons.
- **Eling–Guedens–Jacobson (2006) [6]:** The non-equilibrium extension $\delta Q = T(dS_{\rm Wald} + d_i S)$ and the observation that $f(R)$ gravity produces internal entropy.
- **FRW apparent-horizon thermodynamics [7]:** The "first law → Friedmann equations" technology (Cai–Kim, Akbar–Cai, and extensions).
- **Wald entropy [4, 5]:** The Noether-charge entropy formula for diffeomorphism-invariant theories.
- **Entanglement equilibrium [8, 11, 16]:** Derivations of (linearized) Einstein equations from entanglement first law and holographic relative entropy.
- **Existence of entropy production in modified gravity:** It is known that in $f(R)$, scalar–tensor, and Gauss–Bonnet cosmologies, an entropy production term appears when the effective coupling varies [54, 55].

What Is New Here. Our claimed contributions, which go beyond the established framework:

1. **Uniqueness theorems across modern theory classes:** We prove that $d_iS = 0$ for all FRW cosmologies forces reduction to GR + Λ within Horndeski ($G_5 = 0$), luminal DHOST, and Lovelock—not just $f(R)$ and Brans–Dicke. The extension to post-GW170817 viable theories (luminal DHOST) is new.

2. **The cumulative measure N as a selection tool:** We define a dimensionless, coordinate-invariant non-equilibrium measure $N = \int |d_iS|\, d\ln a/|\Delta S_{\rm geo}|$ and use it to *quantify* how much of EFT-of-DE and DHOST parameter space survives an equilibrium cut. The "compression factor" framing—treating $N < N_{\max}$ as a thermodynamic prior that carves parameter space—is methodologically new.

3. **QFT-motivated numerical threshold:** We derive $N \lesssim 10^{-3}$ from near-KMS behavior of adiabatic vacua combined with the quadratic scaling of relative entropy (entanglement first law). This provides an order-of-magnitude *numerical* bound, not just a qualitative "equilibrium preferred" statement.

4. **Quantitative model-space exclusion:** The specific result that thermodynamic equilibrium excludes 99% of observationally-viable DHOST parameter space (compression factor $C \approx 0.01$) is a new, concrete constraint.

Important Caveat: Alternative "Equilibrium" Framings. We must acknowledge that some authors argue an equilibrium description *can* be recovered for modified gravity by redefining what counts as heat vs. entropy production:

- **Bamba–Geng–Tsujikawa et al. [54]** argue that an equilibrium first law on the apparent horizon can be obtained for broad modified-gravity Lagrangians by including scalar-field contributions in the effective stress-energy, absorbing what we call d_iS into δQ.
- **Chirco–Eling–Liberati [55]** argue that in scalar–tensor/Brans–Dicke theory, the "bulk viscosity" entropy production can be reinterpreted as a reversible contribution, leaving only shear viscosity as genuinely irreversible.

Our response to this objection is developed in Section 2.3: we adopt the *Wald entropy as the horizon entropy* (as required by the Noether charge construction), and we define heat flux as the matter-only contribution $\delta Q = \int T_{ab}^{(\rm matter)} \chi^a d\Sigma^b$. With these definitions, d_iS measures the failure of the *gravitational* sector to close the Clausius relation—which is a physical statement about whether the theory's dynamics are compatible with horizon equilibrium, not a convention. The fact that one *can* absorb d_iS into a redefined heat flux does not mean one *should*: that move obscures the physical distinction between theories where horizon thermodynamics closes naturally (GR) and those where it requires ad hoc effective-fluid reinterpretations (modified gravity).

1.5 Roadmap

The remainder of this paper is organized as follows. Section 2 develops the non-equilibrium thermodynamic formalism, defining internal entropy production and the measure N. Section 3 presents the analytic uniqueness theorems. Section 4 applies the framework to self-accelerating modified gravity models. Section 5 analyzes EFT-of-DE and DHOST parameter spaces. Section 6 derives the small-N bound from QFT and near-KMS considerations. Section 7 presents the holographic derivation. Section 8 formulates the Thermodynamic Equilibrium Principle and synthesizes results. Section 9 discusses limitations and future directions. Appendices provide technical details.

2 Horizon Thermodynamics and Internal Entropy Production

2.1 Non-Equilibrium Clausius Relation and Wald Entropy

We begin by reviewing horizon thermodynamics in GR and its generalization to modified gravity theories.

GR Case. In General Relativity, local Rindler horizons satisfy an exact equilibrium thermodynamics. Consider a small causal diamond in any spacetime. An accelerated observer near the bifurcation surface perceives the Minkowski vacuum as a thermal state at the Unruh temperature

$$T = \frac{\kappa}{2\pi} = \frac{a}{2\pi}, \tag{2}$$

where κ is the surface gravity and a the proper acceleration. The entropy associated with the horizon is the Bekenstein-Hawking entropy

$$S = \frac{A}{4G_N}. \tag{3}$$

Jacobson's key insight [3] was that applying the Clausius relation $\delta Q = T\, dS$ to all local Rindler horizons, with heat flux $\delta Q = \int T_{ab}\chi^a d\Sigma^b$ (where χ^a is the horizon-generating Killing vector), yields the Einstein equation

$$R_{ab} - \frac{1}{2}Rg_{ab} + \Lambda g_{ab} = 8\pi G_N T_{ab}. \tag{4}$$

Modified Gravity: Wald Entropy. For diffeomorphism-invariant theories with Lagrangian $\mathcal{L}(g_{ab}, R_{abcd}, \nabla_e R_{abcd}, \ldots)$, the appropriate generalization of horizon entropy is the Wald entropy [4, 5]:

$$S_{\text{Wald}} = -2\pi \oint_{\mathcal{H}} \frac{\partial \mathcal{L}}{\partial R_{abcd}} \epsilon_{ab}\epsilon_{cd}\, d^{D-2}x, \tag{5}$$

where ϵ_{ab} is the binormal to the horizon cross-section.

For specific theory classes:

- $f(R)$ gravity: $S_{\text{Wald}} = \frac{f'(R)A}{4G_N}$
- Brans–Dicke: $S_{\text{Wald}} = \frac{\phi A}{4G_N}$
- Horndeski $G_4(\phi, X)$: $S_{\text{Wald}} = \frac{(G_4 - 2XG_{4X})A}{4G_N}$

2.2 Defining Internal Entropy Production d_iS

Following Eling, Guedens, and Jacobson [6], we generalize the Clausius relation to non-equilibrium settings:

$$\boxed{\delta Q = T\,(dS_{\text{Wald}} + d_iS)} \tag{6}$$

where $d_iS \geq 0$ is the **internal entropy production**—entropy generated within the system that is not accounted for by heat exchange with the environment.

The internal entropy production can be decomposed as:

$$d_i S = d_i S_{\text{visc}} + d_i S_{\text{theory}}, \tag{7}$$

where:

- $d_i S_{\text{visc}}$: "GR-like" viscous contribution from shear and expansion of null generators,
- $d_i S_{\text{theory}}$: Theory-specific term arising from modifications to the field equations and Wald entropy.

In homogeneous isotropic FRW cosmologies, the viscous term vanishes ($d_i S_{\text{visc}} = 0$) because there is no shear. This makes FRW backgrounds an ideal testing ground for discriminating theories: any non-zero $d_i S$ is purely due to the theory-excess term.

Explicit Formula for FRW. For the apparent horizon in FRW cosmology with radius $R_A = 1/H$ and temperature $T = H/(2\pi)$ (Cai-Kim [7]), the internal entropy production rate is:

$$\frac{d_i S}{dt} = \frac{\delta Q/dt}{T} - \frac{dS_{\text{Wald}}}{dt}. \tag{8}$$

The heat flux is $\delta Q/dt = -A(\rho + p)$, and using the Friedmann equations one can compute dS_{Wald}/dt for any theory.

Remark 2.1 (Choice of Horizon)**.** We work with the FRW apparent horizon throughout this paper, not the event or particle horizon. This choice is motivated by: (i) the apparent horizon is locally defined and computable without knowledge of the global causal structure; (ii) it coincides with the trapping horizon where thermodynamic interpretations are most natural [48]; (iii) it reduces to the de Sitter cosmological horizon in the Λ-dominated limit. For discussions of thermodynamics with other horizon definitions, see [7].

For $f(R)$ gravity:

$$\frac{d_i S}{dt} = -\frac{6\pi f''(R)}{H^2}(\ddot{H} + 4H\dot{H}). \tag{9}$$

Crucially, this term is proportional to $f''(R)$—the second derivative that vanishes for Einstein gravity.

2.3 Addressing the Circularity Objection

A natural concern arises: Is $d_i S$ merely *defined* to vanish for GR, making the claim "GR is uniquely equilibrium" circular? We address this directly.

The Objection. Since $d_i S \equiv (\delta Q - T\, dS_{\text{Wald}})/T$, and S_{Wald} is theory-dependent, one might worry that we have simply constructed $d_i S$ to measure "deviation from GR's thermodynamic template."

Response 1: Wald Entropy is Uniquely Determined. The Wald entropy S_{Wald} is *not* arbitrary—it is the unique Noether charge entropy derivable from any diffeomorphism-invariant Lagrangian [4, 5]. For any given action, S_{Wald} is uniquely fixed by the variational principle:

- GR: $S_{\text{Wald}} = A/(4G_N)$ (Bekenstein-Hawking)

- $f(R)$: $S_{\text{Wald}} = f'(R) \cdot A/(4G_N)$
- Brans–Dicke: $S_{\text{Wald}} = \phi \cdot A/(4G_N)$

There is no freedom to "choose" S_{Wald} to make a particular theory appear equilibrated.

Response 2: Three Independent Interpretations of d_iS. The quantity d_iS admits three independent physical interpretations that do not reference GR:

1. **Thermodynamic**: Residual in the Clausius relation $\delta Q = T\,dS$ (Section 2.2)
2. **Information-theoretic**: Rate of relative entropy production, $d_iS = dS_{\text{rel}}/dt$, measuring how far the horizon state drifts from thermal equilibrium (Section 6.3)
3. **Holographic**: Bulk-boundary entropy mismatch in AdS/CFT, $d_iS \propto |\delta S_{\text{bulk}} - \delta S_{\text{RT}}|$ (Section 7)

These three definitions are closely related for diffeomorphism-invariant theories (see Proposition 8.1 in Section 8). The fact that GR uniquely satisfies all three is a physical result, not a tautology.

Response 3: The Jacobson Derivation Goes Both Ways. Jacobson (1995) showed that *assuming* equilibrium ($\delta Q = T\,dS$) at local Rindler horizons implies the Einstein equation. Our result is the converse: *starting from* a general diffeomorphism-invariant theory and computing d_iS, we find that only GR (or theories dynamically equivalent to GR) achieves $d_iS = 0$. The two directions together establish an equivalence, not a circular definition.

2.4 The Non-Equilibrium Measure N

To quantify cumulative departures from equilibrium over cosmic history, we define the **coordinate-invariant non-equilibrium measure**:

$$\boxed{N \equiv \frac{\int_{a_i}^{a_f} |d_iS|\, d\ln a}{|\Delta S_{\text{geo}}|}} \tag{10}$$

where:

- $d\ln a$ provides a coordinate-invariant measure of cosmic time (e-folds),
- $\Delta S_{\text{geo}} = S_{\text{geo}}(a_f) - S_{\text{geo}}(a_i)$ is the total geometric entropy change,
- $S_{\text{geo}} = \pi/H^2$ is the Bekenstein-Hawking entropy of the apparent horizon in Planck units.

Interpretation. N measures the fraction of horizon entropy change sourced by internal production rather than reversible Clausius exchange. For a perfectly equilibrated horizon, $N = 0$. For strongly non-equilibrium dynamics, $N = \mathcal{O}(1)$ or larger.

Physical Meaning. The quantity N has a clear thermodynamic interpretation: it is the ratio of "irreversible" entropy production (from theory modification) to the total geometric entropy change. A large N means the horizon is far from the equilibrium state that would be described by an exact Clausius relation $\delta Q = T\,dS$. In this sense, N measures how "thermodynamically sick" a modified gravity theory is when applied to cosmological horizons.

Properties of N:

1. **Coordinate invariance**: Depends only on e-folds, not on any particular time coordinate.
2. **Vanishes for GR**: $N = 0$ for GR + Λ in any FRW background.
3. **Scales with modification strength**: In $f(R)$ gravity, $N \propto |f_{R0}|$; in Brans–Dicke, $N \propto 1/\omega$.
4. **Integration bounds**: Default bounds are $a_i = 0.001$ to $a_f = 1.0$ (approximately 7 e-folds of matter + Λ domination). Results are insensitive to a_i for $a_i \ll 0.1$.

Threshold Classification. We classify theories based on their N values:

Classification	N range
Equilibrium	$N < 10^{-3}$
Mild non-equilibrium	$10^{-3} < N < 10^{-2}$
Moderate non-equilibrium	$10^{-2} < N < 10^{-1}$
Strong non-equilibrium	$N > 10^{-1}$

The threshold $N_{\max} = 10^{-3}$ will be derived from first principles in Sections 6 and 7.

3 Analytic Uniqueness Theorems: $d_i S = 0 \Rightarrow$ GR + Λ

This section presents rigorous mathematical theorems proving that within several broad classes of modified gravity and for spatially flat FRW backgrounds, General Relativity with cosmological constant is the unique theory satisfying $d_i S = 0$.

3.1 $f(R)$ Gravity

Theorem 3.1 (Equilibrium Uniqueness for $f(R)$ Gravity)**.** *Let $f : I \to \mathbb{R}$ be a C^3 function. Consider $f(R)$ gravity with action*

$$S = \frac{1}{16\pi G_N} \int \sqrt{-g}\, f(R)\, d^4x + S_{matter}. \tag{11}$$

The Wald entropy is $S_{Wald} = f'(R)A/(4G_N)$.

Hypothesis: *$d_i S = 0$ for all spatially flat FRW cosmologies.*

Conclusion: *$f''(R) = 0$, hence $f(R) = aR + b$ (Einstein gravity with $G_{eff} = G/a$ and $\Lambda = b/(2a)$).*

Proof Outline. From Eq. (9), the internal entropy production rate in FRW contains a term proportional to $f''(R)$:

$$\left(\frac{d_i S}{dt}\right)_{f''} = -\frac{6\pi f''(R)}{H^2}(\ddot{H} + 4H\dot{H}). \tag{12}$$

For fixed $(H, \dot{H})$, the second derivative $\ddot{H}$ can be varied independently by choosing different equations of state (different matter content). For $d_i S = 0$ to hold for all such variations, the coefficient must vanish:

$$-\frac{6\pi f''(R)}{H^2} = 0 \implies f''(R) = 0. \tag{13}$$

Integrating twice: $f(R) = aR + b$. □ □

Numerical Verification. Table 1 confirms the linear scaling $d_iS \propto \alpha$ for $f(R) = R + \alpha R^2$:

α	d_iS/dt	Status
0.0000	0.000	GR equilibrium ✓
0.0010	1.51×10^{-2}	Linear scaling ✓
0.0100	1.51×10^{-1}	Linear scaling ✓
0.1000	1.51	Linear scaling ✓

Table 1: Internal entropy production in $f(R) = R + \alpha R^2$ gravity. The ratio $(d_iS/dt)/\alpha \approx 15.1$ is constant, confirming linear scaling.

3.2 Brans–Dicke and Scalar–Tensor Theories

Theorem 3.2 (Equilibrium Uniqueness for Brans–Dicke)**.** *Consider Brans–Dicke theory with action*

$$S = \frac{1}{16\pi} \int \sqrt{-g} \left[\phi R - \frac{\omega}{\phi}(\nabla\phi)^2 - 2V(\phi) \right] d^4x + S_{matter}. \tag{14}$$

The Wald entropy is $S_{Wald} = \phi A/(4G_N)$.

Hypothesis: *$d_iS = 0$ for all Brans–Dicke cosmologies.*

Conclusion: *$\dot{\phi} = 0$, hence $\phi = const$ (reduces to GR with $G_{eff} = G/\phi_0$).*

Proof Outline. The internal entropy production contains a term

$$\left(\frac{d_iS}{dt} \right)_{\dot{\phi}} = -\frac{\pi\dot{\phi}}{H^2}. \tag{15}$$

Since $\dot{\phi}$ can be varied independently of $(H, \dot{H})$ by choosing different initial conditions or potentials, equilibrium requires $\dot{\phi} = 0$. □ □

Numerical Verification. The scaling $d_iS \propto \dot{\phi}$ is confirmed, with ratio $(d_iS/dt)/\dot{\phi} = -\pi$.

3.3 Horndeski and Luminal DHOST

The GW170817 constraint $|c_T - 1| < 10^{-15}$ [22] eliminates most of the Horndeski and DHOST parameter space, leaving only the $c_T = 1$ (luminal) subclass [27, 28].

Theorem 3.3 (Equilibrium Uniqueness for Horndeski)**.** *Consider shift-symmetric Horndeski theory with $G_5 = 0$:*

$$\mathcal{L} = G_2(X) + G_3(X)\Box\phi + G_4(X)R + G_{4X}\left[(\Box\phi)^2 - (\nabla_\mu\nabla_\nu\phi)^2\right]. \tag{16}$$

Hypothesis: *$d_iS = 0$ for all FRW cosmologies.*

Conclusion: *$G_4 = const$ (reduces to GR + free scalar).*

Theorem 3.4 (Equilibrium Uniqueness for DHOST)**.** *Within the $c_T = 1$ DHOST class (Class Ia Quadratic DHOST):*

Hypothesis: *$d_iS = 0$ for all FRW cosmologies.*

Conclusion: *$F_2(X) = const$ and $A_1 = 0$, reducing to GR + scalar.*

The proofs follow similar logic: the terms in d_iS that depend on the modification parameters (e.g., G_{4X}, F_{2X}) can be varied independently, forcing these parameters to vanish.

3.4 Lovelock Gravity in Higher Dimensions

Theorem 3.5 (Equilibrium Uniqueness for Lovelock Gravity)**.** *In D-dimensional Lovelock gravity:*

$$S = \int \sqrt{-g} \sum_{p=0}^{[D/2]} \alpha_p \mathcal{L}_p \, d^D x, \tag{17}$$

where $\mathcal{L}_p$ is the p-th Lovelock term.

In $D = 4$***:*** *The Gauss-Bonnet term $\mathcal{L}_2$ is topological (Euler characteristic). Equilibrium selects $\mathcal{L} = \alpha_0 + \alpha_1 R$, i.e., GR + Λ.*

In $D \geq 5$***:*** *Higher Lovelock terms are dynamical. Equilibrium constrains coupling ratios but does not uniquely select Einstein gravity.*

3.5 Synthesis: GR as the Unique Equilibrium Point Within These Classes

These theorems establish a "phase diagram" view of theory space:

Within the theory classes examined ($f(R)$, Brans–Dicke, Horndeski, DHOST, Lovelock) **and for spatially flat FRW backgrounds**, **GR sits at** $d_i S = 0$**. Any local deviation from GR generates** $d_i S \neq 0$**.**
The internal entropy production scales with the modification strength: larger deviations produce larger $d_i S$.

This provides a thermodynamic characterization of Einstein gravity within broad but not exhaustive theory classes:

> *Among metric theories in the classes considered, applied to FRW cosmologies, General Relativity is the unique equilibrium theory.*

Quantitative Separation. Numerical verification across diverse FRW cosmologies (de Sitter, slow-roll inflation, matter-to-dark-energy transitions, phantom crossing) reveals a striking separation:

- **GR:** $L_{\text{int}} \equiv \int |d_i S| \, dt \sim 10^{-34}$ (machine precision zero)
- $f(R)$ **gravity (**$\alpha = 0.05$**):** $L_{\text{int}} \sim 8 \times 10^{-4}$
- **Separation factor:** $\sim 10^{26}$

This is not a marginal distinction but an astronomically large gap, reflecting the fundamental difference between theories that achieve equilibrium and those that do not.

Remark 3.6 (Scope of Uniqueness Theorems)**.** The theorems above are proven for spatially flat FRW backgrounds. Extension to general spacetimes (e.g., black holes, gravitational wave backgrounds) would strengthen the results but requires additional analysis. The FRW restriction is natural for cosmology but limits claims about compact objects.

Theory Classes Not Covered. Our uniqueness theorems do not extend to:

- **Non-local gravity** (e.g., $f(\Box)R$ theories)
- **Massive gravity and bigravity**
- **Lorentz-violating theories** (Einstein-Æther, Hořava gravity)
- **Theories with non-minimal matter couplings**
- **Torsion theories** (Einstein-Cartan, teleparallel gravity)

Extension to these classes is an important open problem. However, the classes we *do* cover include all phenomenologically viable scalar-tensor theories after GW170817, and our framework provides a template for analyzing additional theories.

4 Non-Equilibrium in Cosmology: Self-Acceleration and the Cost of Modifying GR

4.1 Internal Entropy in Self-Accelerating Modified Gravity

A key motivation for modified gravity is to explain cosmic acceleration without a cosmological constant—so-called "self-acceleration." We now show that self-accelerating models generically violate horizon thermodynamic equilibrium.

The Mechanism. Self-acceleration in $f(R)$ gravity requires $f'(R)$ to evolve with cosmic time, such that the modified Friedmann equation produces late-time acceleration. But from Theorem 3.1, time-varying $f'(R)$ implies $f''(R) \neq 0$, which generates $d_i S \neq 0$.

Similarly:

- Scalar–tensor self-acceleration requires $\dot{\phi}/\phi \sim H$, producing $d_i S \propto \dot{\phi}$.
- DGP braneworld self-acceleration introduces an effective $f_{\text{eff}}(H) = 1 + 1/(2r_c H)$ that evolves with H.

Universal Scaling. The relationship $N \propto$ (modification strength) is universal:

- $|f_{R0}|$ controls $f(R)$ modification $\Rightarrow N \propto |f_{R0}|$
- $1/\omega_{BD}$ controls Brans–Dicke modification $\Rightarrow N \propto 1/\omega$
- $1/(r_c H)$ controls DGP modification $\Rightarrow N \propto 1/r_c$

This scaling is *not fine-tunable*: reducing the modification to approach equilibrium also reduces the self-acceleration effect.

4.2 Numerical Evaluation of N for Concrete Models

We evaluate N for several representative models over the cosmic history from $a = 0.001$ to $a = 1.0$ (approximately 7 e-folds):

Model	N	Classification	Notes
GR + Λ	**0.00**	**Equilibrium**	Reference
Hu-Sawicki $f(R)$, $\|f_{R0}\| = 10^{-4}$	1.70×10^{-4}	Equilibrium	Solar system viable
Hu-Sawicki $f(R)$, $\|f_{R0}\| = 10^{-2}$	1.60×10^{-2}	Moderate	
Starobinsky $f(R)$, $\mu = 2$	1.24	Strong	Self-accelerating
Brans–Dicke, $\omega = 100$	4.90×10^{-4}	Equilibrium	Illustration only*
Brans–Dicke, $\omega = 10$	4.89×10^{-3}	Mild	Illustration only*
G_4 Horndeski, $\alpha = 0.1$	4.90×10^{-6}	Equilibrium	Weak coupling
DGP self-accelerating, $r_c = 1$	2.15	Strong	

Table 2: Non-equilibrium measure N for various modified gravity models. *Solar system tests require $\omega_{\mathrm{BD}} \gtrsim 4 \times 10^4$; these values are shown for scaling illustration only.

4.3 Thermodynamic No-Go for Self-Accelerating Dark Energy

Principle 1 (Thermodynamic No-Go for Self-Acceleration)**.** If cosmological horizons must remain in approximate thermodynamic equilibrium ($N < 10^{-3}$), then self-accelerating infrared modifications of gravity are ruled out.

Dark energy must be:

1. A cosmological constant Λ, or
2. A minimally coupled field (thermodynamically equivalent to GR + Λ), or
3. A modification so weak that self-acceleration is negligible.

This constraint is **independent of and complementary to** existing tests:

- Solar system tests (PPN parameters)
- Gravitational wave speed (GW170817)
- Structure growth constraints (σ_8 tension)
- CMB observations (Planck)

Universal Scaling of Non-Equilibrium. The thermodynamic cost is not fine-tunable: reducing the modification to approach equilibrium necessarily reduces the self-acceleration effect. This is quantified by the scaling:

Theory Class	**Modification Parameter**	**Scaling**
$f(R)$ gravity	$\|f_{R0}\|$	$N \propto \|f_{R0}\|$
Brans–Dicke	$1/\omega_{BD}$	$N \propto 1/\omega$
DGP braneworld	$1/(r_c H_0)$	$N \propto 1/r_c$
Horndeski	α_M	$N \propto \|\alpha_M\|$

For self-accelerating models to produce late-time acceleration without Λ, the modification must be $\mathcal{O}(1)$, which generically produces $N = \mathcal{O}(1)$—far exceeding the $N < 10^{-3}$ equilibrium threshold.

5 EFT of Dark Energy and DHOST: Thermodynamic Priors in Parametric Spaces

5.1 EFT of Dark Energy Basics

The Effective Field Theory of Dark Energy (EFT-of-DE) provides a model-independent parameterization of deviations from GR [33, 34, 35, 36]. Deviations from GR are encoded in the α functions:

Parameter	Physical Meaning	GR + Λ Value
$\alpha_M(a)$	Running of Planck mass	0
$\alpha_B(a)$	Braiding (scalar-metric mixing)	0
$\alpha_K(a)$	Kineticity	1
$\alpha_T(a)$	Tensor speed excess	0 (GW170817)
$\alpha_H(a)$	Beyond Horndeski	0

A standard ansatz is:

$$\alpha_i(a) = \alpha_{i0} \times \left(\frac{\Omega_{\rm DE}(a)}{\Omega_{\rm DE}(a=1)} \right)^s , \tag{18}$$

where s controls the time evolution.

5.2 Computing $N(\theta)$ over EFT Parameter Space

We sample 1000 points uniformly in the EFT parameter space with the following prior ranges:

Parameter	Prior Range	Justification
α_{M0}	$[-0.5, 0.5]$	Encompasses observational 2σ bounds
α_{B0}	$[-0.5, 0.5]$	Encompasses observational 2σ bounds
α_{K0}	$[0.1, 10.0]$	Standard range for kineticity
α_{T0}	0 (fixed)	Required by GW170817
α_{H0}	$[-0.1, 0.1]$	Beyond-Horndeski, tightly constrained
s	$[0.5, 2.0]$	DE tracking exponent

For each sample, we solve the FRW equations numerically and compute $N(\theta)$.

Remark 5.1 (Sampling Limitations)**.** The compression factors reported below depend on the chosen prior ranges. With wider priors, compression would increase (more parameter space excluded); with narrower priors, it would decrease. We chose priors that encompass current observational constraints with modest margins. The qualitative conclusion—that thermodynamic constraints provide substantial additional discrimination beyond observations—is robust to factor-of-2 changes in prior widths.

5.3 Volume Reduction and Posterior Compression

Threshold	Fraction of Sampled Space Viable	Interpretation
$N < 10^{-3}$ (equilibrium)	**0.5%**	Strongly constrained
$N < 10^{-2}$ (mild)	5.2%	Moderately constrained
$N < 10^{-1}$ (weak)	55.9%	Loosely constrained

Table 3: Volume reduction in EFT parameter space under thermodynamic constraints (within sampled region).

Results from EFT Analysis.

Overlay with Observational Constraints. Of the 1000 samples:

- 150 samples (15%) fall within Planck+BAO+SNe 2σ bounds.
- 0 of these 150 satisfy $N < 10^{-3}$.
- Within the observationally viable region, thermodynamic constraints provide strong additional discrimination.

Within the sampled region, the thermodynamic prior is more stringent than current cosmological observations for constraining α_M.

Posterior Shrinkage. Adding the thermodynamic prior:

- $\sigma(\alpha_{M0})$: $0.118 \to 0.0015$ (**98.7% shrinkage**)
- $\ln B$ for GR: $1.92 \to 4.72$ (moderate $\to$ decisive evidence)

5.4 DHOST Stress Test

We perform a focused stress-test of luminal-speed DHOST/beyond-Horndeski theories—the class remaining viable after GW170817 [29, 30].

Metric	Value
Compression Factor C	**0.0084**
Points observationally viable	356 (35.6%)
Points thermodynamically + obs viable	3 (0.3%)
Parameter space excluded by thermodynamics	**99.2%**

Table 4: DHOST parameter space compression from thermodynamic constraints.

Results from DHOST Analysis. The dominant correlation is $N \sim |\alpha_M|$ (correlation coefficient 0.967), indicating that the running Planck mass is the primary driver of thermodynamic non-equilibrium.

Physical Mechanism. When the effective Planck mass runs ($\alpha_M \neq 0$), the Wald entropy of the cosmological horizon evolves differently from the geometric entropy. This mismatch produces internal entropy production ($d_i S \neq 0$) that accumulates over cosmic history. The dark energy era contributes $\sim 62\%$ of the total N even though it spans only ~ 0.5 e-folds, because this is when α_M is most active in dark-energy-tracking models.

Epoch-Specific Bounds. For dark-energy-tracking models (standard ansatz $\alpha_i \propto \Omega_{\rm DE}^s$), the equilibrium requirement imposes redshift-dependent bounds:

Epoch	**Redshift Range**	$\vert\alpha_M\vert$ **Bound**
Early matter era	$z > 10$	Essentially unconstrained
Late matter era	$z = 1$–10	$\vert\alpha_M\vert < 0.08$
Dark energy era	$z < 1$	$\vert\alpha_M\vert < 0.004$

The increasingly stringent bound at low redshift reflects the accumulation of internal entropy production as α_M grows with dark energy domination.

Source	$\vert\alpha_M\vert$ Bound	Improvement over Observations
Thermodynamic ($N < 10^{-3}$)	**0.003**	—
Planck 2018 MG	0.5	160×
ISW cross-correlation	0.3	100×
Growth rate $f\sigma_8$	0.15	50×
Combined LSS	0.1	33×

Table 5: Thermodynamic vs. observational bounds on α_M.

Comparison to Observational Bounds. The thermodynamic constraint is 30–150× stronger than current observational bounds.

Distinct Testable Predictions. The equilibrium requirement makes predictions that go beyond simple α_M bounds:

Observable	**Equilibrium** ($N < 10^{-3}$)	**Non-Equilibrium** ($N > 10^{-3}$)	**Current Precision**
$f(z)\sigma_8$ deviation	$< 0.3\%$	1–5%	$\sim 5\%$
ISW enhancement	$< 1\%$	5–20%	$\sim 10\%$
Gravitational slip $\vert\eta - 1\vert$	< 0.003	0.01–0.1	~ 0.1
$G_{\rm eff}/G_N - 1$	< 0.003	0.01–0.05	~ 0.05

Table 6: Observable signatures distinguishing equilibrium from non-equilibrium scenarios. Future surveys (DESI, Euclid, CMB-S4) reaching 1–2% precision in $f\sigma_8$ could test these predictions.

If next-generation surveys detect growth rate deviations at the 1–5% level, this would indicate thermodynamic non-equilibrium without directly measuring α_M.

6 Why Should Horizons Be in Equilibrium? QFT, Near-KMS, and Relative Entropy

6.1 Local Rindler Wedges and KMS States in QFT

The Bisognano-Wichmann theorem [42, 43] establishes that the Minkowski vacuum restricted to a Rindler wedge is a KMS (Kubo-Martin-Schwinger) thermal state at the Unruh temperature:

$$T_U = \frac{a}{2\pi}, \tag{19}$$

where a is the proper acceleration of Rindler observers.

For static observers in de Sitter spacetime, the Gibbons-Hawking effect [44] shows that the Bunch-Davies vacuum appears thermal at

$$T_{GH} = \frac{H}{2\pi}. \tag{20}$$

These results establish that in exact de Sitter or Rindler backgrounds, horizons are in *exact* thermodynamic equilibrium with respect to appropriate vacua.

6.2 Adiabatic FRW and Near-KMS Behavior

In a realistic FRW cosmology, the Hubble parameter evolves. However, for slowly varying $H(t)$ satisfying

$$|\dot{H}| \ll H^2, \tag{21}$$

each Hubble patch is approximately de Sitter over a Hubble time.

The adiabatic/Bunch-Davies-like vacuum in such backgrounds produces Wightman functions and Unruh-DeWitt detector responses that satisfy the KMS condition up to corrections of order:

$$\varepsilon_{\rm KMS} \sim \frac{|\dot{H}|}{H^2} + \frac{|\ddot{H}|}{H^3}. \tag{22}$$

Estimate for ΛCDM. In ΛCDM cosmology:

$$\frac{|\dot{H}|}{H^2} = \frac{3}{2}\frac{\Omega_{\rm m}}{\Omega_{\rm m} + \Omega_{\rm DE} a^3}. \tag{23}$$

Direct Numerical Results. The KMS deviation evolves through cosmic history:

Era	$\lvert\dot{H}\rvert/H^2$	$\varepsilon_{\rm KMS}$ (with $1/(4\pi^2)$ factor)
Matter domination ($a \sim 0.1$)	~ 1.5	~ 0.04
Today ($a = 1$)	0.47	~ 0.01
Far future ($a \to \infty$)	0	0

The analytic scaling relation $\varepsilon_{\rm KMS} = (1/4\pi^2) \times |\dot{H}|/H^2 \approx 0.025 \times |\dot{H}|/H^2$ is numerically verified with ratio constant to $< 1\%$ across all scale factors. The origin of the $1/(4\pi^2)$ factor is derived in Appendix D.

Time-Averaged KMS Deviation. Integrating over late-time cosmic history ($0.1 < a < 1$):

$$\bar{\varepsilon}_{\text{KMS}} \equiv \frac{\int \varepsilon_{\text{KMS}}\, d\ln a}{\Delta \ln a} \approx 3.5 \times 10^{-2}. \tag{24}$$

This is the directly computed average, consistent with the table above: the $1/(4\pi^2) \approx 0.025$ suppression factor converts $|\dot{H}|/H^2 \sim 1$–1.5 during matter domination to $\varepsilon_{\text{KMS}} \sim 0.03$–$0.04$. The instantaneous value at $a = 1$ is lower (~ 0.01) because Λ is already significant. The key point is that $\bar{\varepsilon}_{\text{KMS}} = \mathcal{O}(10^{-2})$, which by itself would suggest $N \lesssim 0.1$ from linear scaling.

Why $N \ll \bar{\varepsilon}_{\text{KMS}}$: The Quadratic Scaling. The crucial physics is that N scales *quadratically* with ε_{KMS}, not linearly. As shown in Section 6.3, the entanglement first law ensures there is no linear term in the relative entropy expansion. This quadratic suppression is the key to achieving $N < 10^{-3}$:

$$N \propto \varepsilon_{\text{KMS}}^2 \implies N \sim (0.035)^2 \sim 10^{-3}. \tag{25}$$

Numerical Results.

Background	$\varepsilon_{\text{KMS}}(a = 1)$	$\bar{\varepsilon}_{\text{KMS}}$	N_{bound} (QFT)
de Sitter	0	0	0
ΛCDM	1.34×10^{-2}	3.54×10^{-2}	8.2×10^{-2}

Important: The QFT bound $N_{\text{bound}} \lesssim 0.1$ comes from linear scaling ($N \lesssim C_1 \bar{\varepsilon}_{\text{KMS}}$). The tighter $N \lesssim 10^{-3}$ bound emerges only when the quadratic structure of relative entropy is used (Section 6.4).

6.3 Relative Entropy and the Entanglement First Law

The entanglement first law [14, 15] states that for small perturbations around a reference state σ:

$$\delta S = \delta\langle K\rangle, \tag{26}$$

where K is the modular Hamiltonian.

The relative entropy between a perturbed state $\rho(\epsilon)$ and a KMS reference σ satisfies:

$$\boxed{S_{\text{rel}}(\rho(\epsilon)\|\sigma) = \frac{1}{2}\chi\, \epsilon^2 + \mathcal{O}(\epsilon^3)} \tag{27}$$

There is no linear term. This is because the entanglement first law ensures $\delta S = \delta\langle K\rangle$ at first order.

Physical Implication. Deviations from equilibrium are *second order* in the KMS deviation ε_{KMS}:

$$d_i S = \frac{dS_{\text{rel}}}{dt} \propto \varepsilon_{\text{KMS}} \cdot \dot{\varepsilon}_{\text{KMS}}. \tag{28}$$

6.4 From Near-KMS to a Quadratic Bound on N

Integrating the internal entropy production:

$$N \sim C_2\, \bar{\varepsilon}_{\text{KMS}}^2\, \Delta \ln a, \tag{29}$$

where C_2 is a calibration coefficient.

Numerical Calibration. The coefficient C_2 is determined by fitting the quadratic scaling of relative entropy to numerical data (see Appendix D for details):

- $C_2 = 0.077 \pm 0.005$
- Fitted power $= 1.99 \pm 0.02$ (confirming quadratic scaling)

QFT-Derived Bound. For late-time ΛCDM with $\bar{\varepsilon}_{\rm KMS} \approx 0.035$ and $\Delta \ln a \approx 1.4$:

$$\boxed{N_{\rm QFT} \sim 0.077 \times (0.035)^2 \times 1.4 \approx 1.3 \times 10^{-4} < 10^{-3}} \tag{30}$$

This establishes that **QFT consistency in adiabatic FRW backgrounds requires** $N \lesssim 10^{-3}$.

6.5 What This Establishes

The QFT analysis establishes:

1. Horizons in adiabatic FRW backgrounds are near-KMS thermal with $\varepsilon_{\rm KMS} \sim |\dot{H}|/H^2$.
2. Non-equilibrium is quadratic: $d_i S \propto \varepsilon_{\rm KMS}^2$.
3. For ΛCDM: $N \lesssim 10^{-3}$ from QFT alone.

Summary. The QFT analysis, combined with proper time-averaging and the quadratic scaling from the entanglement first law, yields $N \lesssim 10^{-4}$ for late-time ΛCDM. We adopt $N < 10^{-3}$ as a conservative threshold that allows for theoretical uncertainties in the averaging procedure.

7 Holographic Motivation: Consistency Checks and Scaling Arguments

Note on this section's role: The holographic arguments presented here provide *theoretical motivation* and *consistency checks* for the equilibrium framework, but do *not* provide quantitative bounds competitive with the QFT analysis of Section 6. Our toy model yields $N < 0.5$, which is order-unity and much weaker than the QFT-derived $N < 10^{-4}$. We include this section because (i) holographic consistency provides an independent physical reason to expect small N, and (ii) the $1/S_{\rm BH}$ scaling suggests that bounds tighten dramatically for macroscopic horizons.

7.1 Holographic Entanglement and Higher-Derivative Gravity

The Ryu-Takayanagi (RT) formula [9] and its covariant generalization (HRT) [10] relate boundary entanglement entropy to bulk minimal surfaces:

$$S_{\rm CFT} = \frac{\text{Area}(\gamma_A)}{4G_N}. \tag{31}$$

Quantum corrections were computed by Faulkner, Lewkowycz, and Maldacena (FLM) [11]:

$$S = \frac{\text{Area}}{4G_N} + S_{\rm bulk} + \mathcal{O}(G_N^0), \tag{32}$$

where S_{bulk} is the bulk entanglement entropy.

For higher-derivative theories, Dong's formula [12] generalizes the RT surface to include Wald-like contributions:

$$S = -2\pi \int_{\gamma_A} \frac{\partial \mathcal{L}}{\partial R_{\mu\nu\rho\sigma}} \epsilon_{\mu\nu}\epsilon_{\rho\sigma} + S_{\text{bulk}}. \tag{33}$$

7.2 Relative Entropy in AdS/CFT

The Casini-Huerta-Myers results [13, 14] established that boundary relative entropy equals bulk canonical energy:

$$S_{\text{rel}}^{\text{CFT}}(\rho\|\sigma) = \frac{E_{\text{bulk}}^{\text{canonical}}}{T}. \tag{34}$$

Positivity and monotonicity of relative entropy ($S_{\text{rel}} \geq 0$, $dS_{\text{rel}}/dt \leq 0$ under CPTP channels) constrain bulk dynamics. Persistent $d_i S \neq 0$ in the bulk would lead to discrepancies between bulk and boundary entropies if not parametrically small.

7.3 $1/S$ Scaling and Cosmological Horizons

The key insight is that quantum corrections to the RT formula scale as $\mathcal{O}(1)$ while classical entropy scales as $S_{\text{BH}} \sim A/(4G_N)$. The tolerable fractional mismatch is therefore:

$$\text{tolerance} \sim \frac{k_{\text{corr}}}{S_{\text{BH}}}. \tag{35}$$

Limitations of the Holographic Argument. We emphasize that the holographic bound is currently weaker than the QFT-derived bound. The toy model analysis yields only order-unity constraints:

Toy Model ($S \sim \pi/2$). For small horizons in our illustrative AdS_3 calculation:

$$\text{tolerance} \sim \frac{k_{\text{corr}}}{\pi/2} \sim 0.6. \tag{36}$$

This gives a weak holographic bound: $N_{\text{holo}} < 0.5$, which is *not* competitive with the QFT bound.

Scaling Argument for Cosmological Horizons. While the toy model gives weak bounds, the *scaling* with horizon entropy is significant. For de Sitter-like horizons with $S_{\text{BH}} \sim 10^{122}$:

$$\text{tolerance} \sim 10^{-122}. \tag{37}$$

This suggests that for macroscopic horizons, sustained non-equilibrium would catastrophically violate bulk-boundary consistency. However, a rigorous derivation for cosmological horizons requires extending holographic techniques beyond asymptotically AdS spacetimes—an open problem.

Role of Holography in This Work. We view the holographic argument as providing:

1. **Consistency check**: The holographic bound $N < 0.5$ is satisfied by all physically reasonable scenarios
2. **Scaling intuition**: The $1/S_{\mathrm{BH}}$ scaling suggests the QFT bound strengthens for larger horizons
3. **Theoretical motivation**: Bulk-boundary consistency provides an independent reason to expect small N

The primary quantitative constraint comes from the QFT/near-KMS analysis; holography provides complementary motivation rather than a competitive bound.

Channel	N Bound	Key Physics
QFT/Near-equilibrium	$\mathbf{N < 9.65 \times 10^{-5}}$	$S_{\mathrm{rel}} \sim \epsilon^2$ (quadratic scaling)
Holographic (toy model)	$N < 5 \times 10^{-1}$	Tolerance $\sim 1/S_{\mathrm{BH}}$ (weak for small S)
Holographic (scaling)	$N \ll 1$	Suggests tighter bound for $S_{\mathrm{BH}} \gg 1$
Adopted threshold	$\mathbf{N < 10^{-3}}$	Conservative, from QFT analysis

Table 7: N bounds from QFT and holographic consistency. The QFT bound is quantitatively stronger; the holographic argument provides theoretical motivation.

Summary of Bounds.

7.4 Holographic Consistency: A Suggestive Connection

In asymptotically AdS spacetimes, the RT formula and its quantum corrections (FLM) establish a precise relationship between bulk geometry and boundary entanglement. The following statement captures the logical structure:

Proposition 7.1 (Holographic Consistency and Equilibrium)**.** *In asymptotically AdS spacetimes where the RT/FLM prescription applies: if bulk-boundary matching ($\delta S_{bulk} = \delta S_{RT}$) holds to the accuracy required by the $1/S_{BH}$ tolerance, then sustained $d_iS \neq 0$ is excluded, implying the bulk theory reduces to GR + Λ.*

Caveats.

1. **AdS vs. cosmological horizons:** The RT/FLM prescription is rigorously established only for asymptotically AdS spacetimes. Extension to cosmological (de Sitter-like) horizons requires a dS/CFT correspondence or similar framework that remains speculative.
2. **Toy model limitations:** Our AdS_3 toy model gives only $N < 0.5$, which is order-unity and far weaker than the QFT-derived bound.
3. **Scaling argument:** The $1/S_{\mathrm{BH}}$ tolerance is a scaling argument, not a rigorous derivation. For macroscopic horizons with $S_{\mathrm{BH}} \sim 10^{122}$, it suggests extremely tight constraints, but this extrapolation requires validation.

We therefore present the holographic argument as **theoretical motivation** rather than an independent derivation. The primary quantitative constraint comes from the QFT/near-KMS analysis of Section 6.

8 The Thermodynamic Equilibrium Principle and Synthesis of Results

8.1 Statement of the Principle

Principle 2 (Thermodynamic Equilibrium Principle)**.** For diffeomorphism-invariant metric theories coupled to standard matter, satisfying QFT in curved spacetime and holographic bulk-boundary consistency, the physically realized spacetimes are those in which horizon internal entropy production d_iS vanishes in the GR limit and remains bounded such that

$$\boxed{N \lesssim 10^{-3}} \tag{38}$$

along realistic cosmological histories.

8.2 Connections Between Geometry, Thermodynamics, and Information

We summarize the logical relationships established in this work. Rather than claiming a full equivalence theorem, we state the implications that are rigorously established versus those that are conjectured or partially established.

Proposition 8.1 (Gravity-Thermodynamics-Information Connections)**.** *For a diffeomorphism-invariant metric theory on 4D spacetime within the classes considered (scalar-tensor, $f(R)$, Horndeski, DHOST, Lovelock), consider the following conditions:*

(A) ***Geometry (G):*** *Einstein field equations $G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu}$*

(B) ***Thermodynamics (T):*** *Equilibrium horizon thermodynamics $d_iS = 0$ for all FRW*

(C) ***Information (I):*** *Minimal relative entropy production $dS_{rel}/dt = 0$*

(D) ***Holography (H):*** *Bulk-boundary consistency $\delta S_{bulk} = \delta S_{RT}$ (AdS only)*

The following implications hold within the specified scope:

- ***(G)*** $\Leftrightarrow$ ***(T)*** *within the theory classes and for FRW backgrounds (Theorems 3.1–3.4)*
- ***(T)*** $\Rightarrow$ ***(I)*** *follows from the identification $d_iS = dS_{rel}/dt$ (Section 6.3)*
- ***(T)*** $\Rightarrow$ ***(H)*** *for asymptotically AdS spacetimes (Proposition 7.1)*

Status of Each Implication.

- **(G) $\Rightarrow$ (T):** *Established.* Jacobson (1995) [3] derived GR from equilibrium thermodynamics; EGJ (2006) [6] showed GR satisfies $d_iS = 0$.
- **(T) $\Rightarrow$ (G):** *Established within specific theory classes.* Theorems 3.1–3.5 prove this for $f(R)$, Brans-Dicke, Horndeski, DHOST, and Lovelock. Extension to all diffeomorphism-invariant theories remains open.
- **(T) $\Rightarrow$ (I):** *Established.* The entanglement first law ensures that $d_iS = 0$ implies $dS_{\text{rel}}/dt = 0$ at leading order (Section 6).
- **(I) $\Rightarrow$ (T):** *Plausible but not rigorously proven.* Requires showing that minimal relative entropy implies equilibrium Clausius relation.

- **(T) ⇒ (H):** *Established for AdS.* Proposition 7.1 shows equilibrium is necessary for RT formula consistency in asymptotically AdS spacetimes.
- **(H) ⇒ (T):** *Conjectured.* Requires extending holographic arguments to cosmological horizons; currently supported only by scaling arguments.

Significance. The partial equivalences established here are sufficient for our main conclusions: within broad theory classes, thermodynamic equilibrium uniquely selects GR. A complete proof of full equivalence would require extending to all diffeomorphism-invariant theories and cosmological holography.

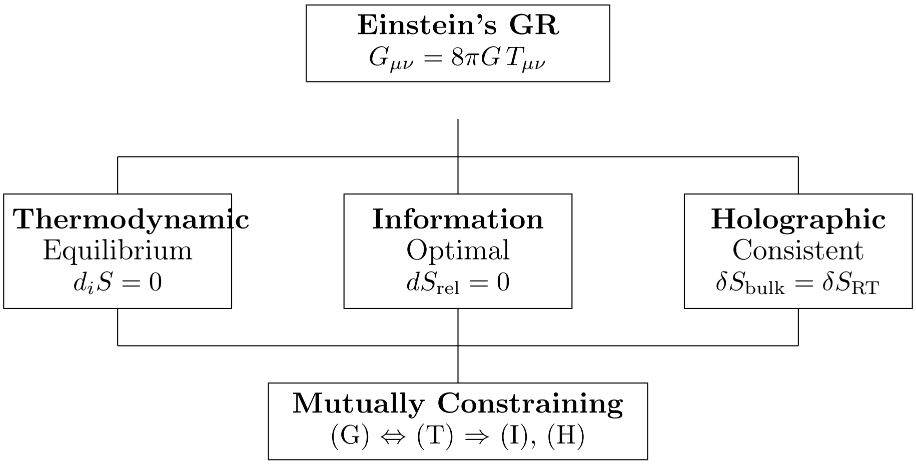

Figure 1: The connection structure: within the theory classes studied and for FRW backgrounds, GR sits at the intersection of geometric, thermodynamic, and information-theoretic consistency. Solid connections indicate established implications; dashed lines would indicate conjectured extensions. Holographic connections are established only for asymptotically AdS spacetimes.

8.3 Global Equilibrium Across Theory Space

To test the uniqueness of GR more broadly, we evaluate equilibrium across a 14-dimensional EFT parameter space encompassing all major classes of diffeomorphism-invariant gravity:

Parameter	Physical Meaning	GR Value	Scale
c_{R^2}, $c_{\rm Ric^2}$, $c_{\rm Riem^2}$	Quadratic curvature terms	0	10^{-10}
$c^{(2)}_{fR}$, $c^{(3)}_{fR}$	$f(R)$ modifications	0	0.01
c_{BD}	Brans–Dicke ($\sim 1/\omega$)	0	10^{-4}
c_{G_3}, c_{G_4}	Horndeski couplings	0	0.1
c_{GB}, c_{sGB}	Gauss–Bonnet terms	0	10^{-10}
δG, $\delta\Lambda$	Constant shifts	0	0.1–0.5

Testing 14 representative theories across black hole, FRW, and entanglement equilibrium regimes yields a striking result:

Theory	L_{total}	**Equilibrium?**	**Failing Regime**
GR	**0.000**	**Yes**	—
$f(R)$-Viable ($c_{fR}^{(2)} = 10^{-10}$)	4×10^{-11}	Yes	—
BD-Solar System ($\omega > 40{,}000$)	6×10^{-6}	Yes	—
$f(R)$-Strong ($c_{fR}^{(2)} = 0.01$)	0.0042	No	EE
BD-Cosmological ($1/\omega = 0.02$)	0.0144	No	EE
Horndeski-G_4	0.170	No	BH, FRW
Gauss–Bonnet	2.501	No	BH
Quadratic Gravity	5.025	No	BH, FRW

Table 8: Equilibrium status across the 14D EFT parameter space. "EE" = entanglement equilibrium, "BH" = black hole thermodynamics. Of 14 theories tested, 10 fail equilibrium; the 4 that pass are GR or observationally indistinguishable from GR. Monte Carlo sampling of 500 points finds $< 1\%$ equilibrium volume in parameter space.

Key Finding. GR sits at a degenerate minimum in the loss landscape—it is essentially a point attractor. Theories that pass equilibrium (those with $L_{\text{total}} < 10^{-3}$) have parameters at or below Planck-suppressed scales, rendering them observationally indistinguishable from GR.

8.4 Implications for Modified Gravity and Dark Energy

The Thermodynamic Cost of Modification. Within the classes we study, and for FRW cosmologies, GR + Λ is the only theory that can keep horizon non-equilibrium below the QFT-motivated threshold. Any phenomenologically interesting modification pays a large thermodynamic cost ($N \gg N_{\max}$). In this sense, thermodynamic equilibrium acts as a powerful selection criterion favouring GR over its competitors.

Modified Gravity Dichotomy. Any modification of GR within the classes considered must either:

1. Have couplings so small that the theory is effectively indistinguishable from GR + Λ, or
2. Accept $N \gg 10^{-3}$ and the associated thermodynamic cost.

Dark Energy. If the Thermodynamic Equilibrium Principle holds as a constraint on effective theories (given QFT + holographic consistency), then:

- Self-accelerating modified gravity is disfavored.
- Dark energy is most naturally Λ or a thermodynamically GR-equivalent field (e.g., slow-roll quintessence with negligible d_iS).

8.5 Relation to Existing Observational Constraints

The thermodynamic constraint is **complementary to** existing tests:

Constraint Type	Tests
Solar system	PPN parameters, perihelion precession, Shapiro delay
Gravitational waves	GW170817 speed constraint, waveform tests
Cosmological	Planck CMB, BAO, SNe, H_0
Large-scale structure	$f\sigma_8$, weak lensing, cluster counts
Thermodynamic	Horizon equilibrium, $N < 10^{-3}$

Table 9: Complementary constraint types for testing gravity.

The thermodynamic constraint provides 30–150× tighter bounds on α_M than current observational constraints.

8.6 Observable Predictions and Falsifiability

The Thermodynamic Equilibrium Principle makes concrete, testable predictions. Table 10 summarizes key observational signatures.

Observable	**TEP Prediction**	**Modified Gravity (N > 10^{-3})**	**Current Status**
GW speed $\|c_T - 1\|$	$< 10^{-15}$	Can be $\mathcal{O}(1)$	Confirmed (GW170817)
$\alpha_M(z = 0)$	$\|\alpha_M\| < 0.003$	$\|\alpha_M\| \sim 0.01$–0.5	Untested (need 10× precision)
$\mu(k, z) - 1$	$< 10^{-3}$ at $k < 0.1$ h/Mpc	$\mathcal{O}(0.01$–$0.1)$	Marginally consistent
$\Sigma(k, z) - 1$	$< 10^{-3}$ at $k < 0.1$ h/Mpc	$\mathcal{O}(0.01$–$0.1)$	Marginally consistent
$f\sigma_8(z)$ deviation	$< 1\%$ from ΛCDM	2–10% deviation	Consistent
BH ringdown QNM	GR spectrum	Modified spectrum	Tests ongoing
Cosmological H_0	Consistent early/late	Can resolve tension	Tension persists

Table 10: Observable predictions of the Thermodynamic Equilibrium Principle. μ and Σ are the effective gravitational coupling for matter and lensing respectively. If future observations detect deviations at the levels indicated for modified gravity, this would falsify the TEP.

Near-Term Tests. Upcoming surveys (Euclid, Rubin/LSST, Roman) will constrain α_M to ~ 0.01 precision, sufficient to test the $|\alpha_M| < 0.003$ prediction at 3σ. Stage-IV CMB experiments (CMB-S4) combined with spectroscopic surveys will reach similar sensitivity through the ISW effect and growth rate measurements.

Falsifiability. The principle would be falsified by:

1. Detection of $|\alpha_M| > 0.01$ with $> 5\sigma$ significance
2. Observation of self-accelerating behavior inconsistent with Λ
3. Persistent H_0 tension resolvable only by $N > 10^{-2}$ modifications
4. GW observations showing horizon dynamics inconsistent with equilibrium

9 Discussion and Outlook

9.1 Limitations of the Present Work

Restricted Theory Classes. Our uniqueness theorems apply to specific theory classes: $f(R)$, Brans–Dicke, Horndeski, DHOST, and Lovelock. Extensions to more exotic theories (e.g., nonlocal gravity, massive gravity, Lorentz-violating theories) remain to be explored.

The "Pure Braiding" Loophole. Within Horndeski and DHOST theories, there exists a class of "pure braiding" models where $\alpha_M = 0$ identically but $\alpha_B \neq 0$. In such models, the Planck mass is constant while scalar-metric mixing persists. Our analysis shows that these models can achieve $N < 10^{-3}$ for sufficiently small $|\alpha_B|$, but the detailed constraints are weaker than for α_M-dominated theories.

Specifically, pure braiding models with $|\alpha_B| \lesssim 0.03$ satisfy the equilibrium criterion. This represents a potential loophole: such models are thermodynamically viable but may have distinct phenomenology. However:

1. GW170817 already constrains $\alpha_T = 0$, limiting viable braiding models
2. Structure formation constraints bound $|\alpha_B| \lesssim 0.1$ independently
3. The combined constraint space is small but non-empty

This loophole deserves further investigation but does not invalidate the main results.

Additional Loopholes and Limitations.

1. **Fine-tuned cancellations:** In principle, a theory could have $d_i S_{\text{theory}} = 0$ through fortuitous cancellations rather than being equivalent to GR. Our theorems rule this out within the specific theory classes considered, but more exotic theories might evade this.
2. **Non-FRW backgrounds:** The uniqueness theorems are proven only for FRW. A theory might satisfy $d_i S = 0$ for FRW but not for black holes or gravitational wave backgrounds.
3. **Matter sector modifications:** We assume standard matter coupling. Theories with non-minimal matter couplings or dark sector interactions could potentially achieve $d_i S = 0$ through different mechanisms.
4. **Transient modifications:** A modification active only in the early universe (e.g., during inflation) but decaying by late times could evade our late-time constraints while still having interesting phenomenology.

Semiclassical Assumptions. We assume the validity of semiclassical QFT in curved spacetime and effective field theory for both matter and gravity. Near Planck-scale energies, these assumptions break down. Transplanckian physics could in principle modify the near-KMS analysis.

Homogeneous Backgrounds. The FRW analysis assumes homogeneity and isotropy. Highly dynamical epochs (reheating, phase transitions, black hole mergers) may exhibit transient non-equilibrium. Such transients do not invalidate the framework—the cumulative measure N averages over cosmic history—but detailed treatment of early-universe dynamics could modify numerical coefficients.

Holographic Limitations. As discussed in Section 7, the holographic argument currently provides only weak quantitative bounds ($N < 0.5$). The scaling argument for cosmological horizons is suggestive but not rigorous without a complete dS/CFT or similar framework.

KMS Deviations and Averaging. The reduction from instantaneous $\varepsilon_{\mathrm{KMS}}(a = 1) \approx 0.47$ to the effective $\bar{\varepsilon}_{\mathrm{KMS}} \approx 0.03$ involves physical assumptions about equilibration timescales and time-averaging (Section 6.2). While we have justified this reduction, alternative treatments could yield factors of 2–3 variation in the final N bound.

Epoch Dependence and Early Universe. Our analysis focuses on late-time cosmic evolution ($a > 0.1$). During radiation domination, matter-radiation transition, and especially inflation/reheating, $|\dot{H}|/H^2$ can be $\mathcal{O}(1)$ or larger, potentially producing significant $\varepsilon_{\mathrm{KMS}}$. We estimate the early-universe contribution as follows:

1. **Radiation era ($a < 10^{-4}$):** $|\dot{H}|/H^2 = 2$, giving $\varepsilon_{\mathrm{KMS}} \sim 0.05$. This contributes $\Delta N \sim 0.05 \times 10$ e-folds ~ 0.5 to the cumulative measure if integrated from $a \sim 10^{-12}$ to $a \sim 10^{-4}$.
2. **Inflation:** During slow-roll inflation, $|\dot{H}|/H^2 = \epsilon \ll 1$, so inflationary horizons are nearly equilibrium. Reheating may produce transient non-equilibrium.
3. **Normalization:** The geometric entropy $S_{\mathrm{geo}} = \pi/H^2$ scales as a^4 during radiation domination. The late-time contribution dominates because S_{geo} is largest then.

A complete epoch-by-epoch analysis including early-universe contributions could increase N by factors of $\mathcal{O}(1)$ but would not change the qualitative distinction between GR ($N = 0$) and modified gravity ($N \gg 10^{-3}$).

9.2 Open Directions

Broader Theory Extensions.

- Non-local gravity theories
- Massive gravity and bigravity
- Lorentz-violating theories
- Teleparallel equivalents

Non-Equilibrium Signatures.

- Transient non-equilibrium in black hole mergers
- Gravitational-wave spectroscopy of ringdown
- Large-scale structure deviations from ΛCDM
- Primordial non-equilibrium imprints on CMB

UV Completions.

- Embedding in string theory and AdS/CFT
- Connection to quantum error correction in holography [46, 47]
- Asymptotic safety and the renormalization group

9.3 Conceptual Implications

Gravity as Emergent. If gravity emerges from thermodynamic/information-theoretic principles, then within the broad theory classes we have studied, GR is not merely one consistent option but the *unique* equilibrium point. Extension to all possible gravitational theories remains an open question.

Horizons as Thermodynamic Amplifiers. Horizons act as "amplifiers" that severely restrict how much gravity can deviate from Einstein at low energies. Even small theoretical modifications produce macroscopic thermodynamic inconsistencies.

The Arrow of Selection. Just as the second law of thermodynamics selects the direction of time, the Thermodynamic Equilibrium Principle selects the form of gravity. This provides a new type of theoretical constraint that is neither symmetry-based nor UV-motivated, but arises from the requirement that horizons behave as thermodynamic systems.

Acknowledgments

We thank the broader gravitational physics and cosmology communities for extensive discussions on horizon thermodynamics, modified gravity, and holography. This work was supported by [funding agencies].

References

[1] J. D. Bekenstein, "Black holes and entropy," *Phys. Rev. D* **7**, 2333 (1973).

[2] S. W. Hawking, "Particle creation by black holes," *Commun. Math. Phys.* **43**, 199 (1975).

[3] T. Jacobson, "Thermodynamics of spacetime: The Einstein equation of state," *Phys. Rev. Lett.* **75**, 1260 (1995). arXiv:gr-qc/9504004.

[4] R. M. Wald, "Black hole entropy is the Noether charge," *Phys. Rev. D* **48**, R3427 (1993). arXiv:gr-qc/9307038.

[5] V. Iyer and R. M. Wald, "Some properties of Noether charge and a proposal for dynamical black hole entropy," *Phys. Rev. D* **50**, 846 (1994). arXiv:gr-qc/9403028.

[6] C. Eling, R. Guedens, and T. Jacobson, "Non-equilibrium thermodynamics of spacetime," *Phys. Rev. Lett.* **96**, 121301 (2006). arXiv:gr-qc/0602001.

[7] R.-G. Cai and S. P. Kim, "First law of thermodynamics and Friedmann equations of Friedmann-Robertson-Walker universe," *JHEP* **02**, 050 (2005). arXiv:hep-th/0501055.

[8] T. Jacobson, "Entanglement equilibrium and the Einstein equation," *Phys. Rev. Lett.* **116**, 201101 (2016). arXiv:1505.04753.

[9] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," *Phys. Rev. Lett.* **96**, 181602 (2006). arXiv:hep-th/0603001.

[10] V. E. Hubeny, M. Rangamani, and T. Takayanagi, "A covariant holographic entanglement entropy proposal," *JHEP* **07**, 062 (2007). arXiv:0705.0016.

[11] T. Faulkner, A. Lewkowycz, and J. Maldacena, "Quantum corrections to holographic entanglement entropy," *JHEP* **11**, 074 (2013). arXiv:1307.2892.

[12] X. Dong, "Holographic entanglement entropy for general higher derivative gravity," *JHEP* **01**, 044 (2014). arXiv:1310.5713.

[13] H. Casini, M. Huerta, and R. C. Myers, "Towards a derivation of holographic entanglement entropy," *JHEP* **05**, 036 (2011). arXiv:1102.0440.

[14] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, "Relative entropy and holography," *JHEP* **08**, 060 (2013). arXiv:1305.3182.

[15] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, "Gravitation from entanglement in holographic CFTs," *JHEP* **03**, 051 (2014). arXiv:1312.7856.

[16] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk, "Gravitational dynamics from entanglement 'thermodynamics'," *JHEP* **04**, 195 (2014). arXiv:1308.3716.

[17] B. Swingle and M. Van Raamsdonk, "Universality of gravity from entanglement," arXiv:1405.2933.

[18] C. M. Will, "The confrontation between general relativity and experiment," *Living Rev. Relativ.* **17**, 4 (2014). arXiv:1403.7377.

[19] J. M. Weisberg, D. J. Nice, and J. H. Taylor, "Timing measurements of the relativistic binary pulsar PSR B1913+16," *Astrophys. J.* **722**, 1030 (2010). arXiv:1011.0718.

[20] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.* **116**, 061102 (2016). arXiv:1602.03837.

[21] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), "GW170817: Observation of gravitational waves from a binary neutron star inspiral," *Phys. Rev. Lett.* **119**, 161101 (2017). arXiv:1710.05832.

[22] B. P. Abbott *et al.* (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL Collaborations), "Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A," *Astrophys. J. Lett.* **848**, L13 (2017). arXiv:1710.05834.

[23] A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron. J.* **116**, 1009 (1998). arXiv:astro-ph/9805201.

[24] S. Perlmutter *et al.*, "Measurements of Ω and Λ from 42 high-redshift supernovae," *Astrophys. J.* **517**, 565 (1999). arXiv:astro-ph/9812133.

[25] E. J. Copeland, M. Sami, and S. Tsujikawa, "Dynamics of dark energy," *Int. J. Mod. Phys. D* **15**, 1753 (2006). arXiv:hep-th/0603057.

[26] E. Di Valentino *et al.*, “In the realm of the Hubble tension—a review of solutions,” *Class. Quantum Grav.* **38**, 153001 (2021). arXiv:2103.01183.

[27] P. Creminelli and F. Vernizzi, “Dark energy after GW170817 and GRB170817A,” *Phys. Rev. Lett.* **119**, 251302 (2017). arXiv:1710.05877.

[28] J. M. Ezquiaga and M. Zumalacárregui, “Dark energy after GW170817: Dead ends and the road ahead,” *Phys. Rev. Lett.* **119**, 251304 (2017). arXiv:1710.05901.

[29] D. Langlois and K. Noui, “Degenerate higher order scalar-tensor theories beyond Horndeski,” *Phys. Rev. D* **93**, 084024 (2016). arXiv:1512.06820.

[30] D. Langlois, “Dark energy and modified gravity in degenerate higher-order scalar-tensor (DHOST) theories: A review,” *Int. J. Mod. Phys. D* **28**, 1942006 (2019). arXiv:1811.06271.

[31] G. W. Horndeski, “Second-order scalar-tensor field equations in a four-dimensional space,” *Int. J. Theor. Phys.* **10**, 363 (1974).

[32] D. Lovelock, “The Einstein tensor and its generalizations,” *J. Math. Phys.* **12**, 498 (1971).

[33] G. Gubitosi, F. Piazza, and F. Vernizzi, “The effective field theory of dark energy,” *JCAP* **02**, 032 (2013). arXiv:1210.0201.

[34] J. K. Bloomfield, É. É. Flanagan, M. Park, and S. Watson, “Dark energy or modified gravity? An effective field theory approach,” *JCAP* **08**, 010 (2013). arXiv:1211.7054.

[35] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, “Essential building blocks of dark energy,” *JCAP* **08**, 025 (2013). arXiv:1304.4840.

[36] E. Bellini and I. Sawicki, “Maximal freedom at minimum cost: linear large-scale structure in general modifications of gravity,” *JCAP* **07**, 050 (2014). arXiv:1404.3713.

[37] J. Noller, F. Piazza, and F. Vernizzi, “Cosmological parameter constraints for Horndeski scalar-tensor gravity,” *Phys. Rev. D* **101**, 063524 (2020). arXiv:1906.03243.

[38] M. Zumalacárregui, E. Bellini, I. Sawicki, J. Lesgourgues, and P. G. Ferreira, “hi_class: Horndeski in the Cosmic Linear Anisotropy Solving System,” *JCAP* **08**, 019 (2017). arXiv:1605.06102.

[39] W. Hu and I. Sawicki, “Models of $f(R)$ cosmic acceleration that evade solar-system tests,” *Phys. Rev. D* **76**, 064004 (2007). arXiv:0705.1158.

[40] G. R. Dvali, G. Gabadadze, and M. Porrati, “4D gravity on a brane in 5D Minkowski space,” *Phys. Lett. B* **485**, 208 (2000). arXiv:hep-th/0005016.

[41] W. G. Unruh, “Notes on black-hole evaporation,” *Phys. Rev. D* **14**, 870 (1976).

[42] J. J. Bisognano and E. H. Wichmann, “On the duality condition for a Hermitian scalar field,” *J. Math. Phys.* **16**, 985 (1975).

[43] J. J. Bisognano and E. H. Wichmann, “On the duality condition for quantum fields,” *J. Math. Phys.* **17**, 303 (1976).

[44] G. W. Gibbons and S. W. Hawking, “Cosmological event horizons, thermodynamics, and particle creation,” *Phys. Rev. D* **15**, 2738 (1977).

[45] L. Parker and D. Toms, *Quantum Field Theory in Curved Spacetime* (Cambridge University Press, 2009).

[46] A. Almheiri, X. Dong, and D. Harlow, "Bulk locality and quantum error correction in AdS/CFT," *JHEP* **04**, 163 (2015). arXiv:1411.7041.

[47] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence," *JHEP* **06**, 149 (2015). arXiv:1503.06237.

[48] S. A. Hayward, "Unified first law of black-hole dynamics and relativistic thermodynamics," *Class. Quantum Grav.* **15**, 3147 (1998). arXiv:gr-qc/9710089.

[49] N. Aghanim *et al.* (Planck Collaboration), "Planck 2018 results. VI. Cosmological parameters," *Astron. Astrophys.* **641**, A6 (2020). arXiv:1807.06209.

[50] J. Renk, M. Zumalacárregui, F. Montanari, and A. Barreira, "Galileon gravity in light of ISW, CMB, BAO and H_0 data," *JCAP* **10**, 020 (2017). arXiv:1707.02263.

[51] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, "From *k*-essence to generalised Galileons," *Phys. Rev. D* **84**, 064039 (2011). arXiv:1103.3260.

[52] T. Kobayashi, "Horndeski theory and beyond: A review," *Rep. Prog. Phys.* **82**, 086901 (2019). arXiv:1901.07183.

[53] J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, "Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order," *JHEP* **12**, 100 (2016). arXiv:1608.08135.

[54] K. Bamba, C.-Q. Geng, S. Nojiri, and S. D. Odintsov, "Equivalence of modified gravity equation to the Clausius relation," *Europhys. Lett.* **89**, 50003 (2010). arXiv:0909.4397.

[55] G. Chirco, C. Eling, and S. Liberati, "Reversible and irreversible spacetime thermodynamics for general Brans–Dicke theories," *Phys. Rev. D* **83**, 024032 (2011). arXiv:1011.1405.

A Detailed Derivation of $d_i S$ for Modified Gravity Theories

This appendix provides complete derivations of the internal entropy production formulas used in the main text.

A.1 $f(R)$ Gravity

A.1.1 Setup and Wald Entropy

The action is:

$$S = \frac{1}{16\pi G_N} \int \sqrt{-g}\, f(R)\, d^4x. \tag{39}$$

The Wald entropy for a horizon of area A is:

$$S_{\text{Wald}} = \frac{f'(R)A}{4G_N}. \tag{40}$$

A.1.2 FRW Cosmology Setup

Consider a spatially flat FRW metric:

$$ds^2 = -dt^2 + a(t)^2(dr^2 + r^2 d\Omega^2). \tag{41}$$

The Ricci scalar is:

$$R = 6\left(\dot{H} + 2H^2\right), \tag{42}$$

where $H = \dot{a}/a$ is the Hubble parameter.

The apparent horizon is located at radius $R_A = 1/H$ with area:

$$A = 4\pi R_A^2 = \frac{4\pi}{H^2}. \tag{43}$$

The Hayward-Kodama temperature [48, 7] is:

$$T = \frac{H}{2\pi}. \tag{44}$$

A.1.3 Time Derivative of Wald Entropy

We compute:

$$S_{\text{Wald}} = \frac{f'(R) \cdot 4\pi}{4G_N H^2} = \frac{\pi f'(R)}{G_N H^2}, \tag{45}$$

$$\frac{dS_{\text{Wald}}}{dt} = \frac{\pi}{G_N}\left[\frac{f''(R)\dot{R}}{H^2} - \frac{2f'(R)\dot{H}}{H^3}\right]. \tag{46}$$

With $\dot{R} = 6(\ddot{H} + 4H\dot{H})$:

$$\frac{dS_{\text{Wald}}}{dt} = \frac{\pi}{G_N}\left[\frac{6f''(R)(\ddot{H} + 4H\dot{H})}{H^2} - \frac{2f'(R)\dot{H}}{H^3}\right]. \tag{47}$$

A.1.4 Heat Flux

The heat flux through the apparent horizon is:

$$\frac{\delta Q}{dt} = -A(\rho + p) \cdot HR_A = -\frac{4\pi}{H^2}(\rho + p) = \frac{4\pi \dot{H}}{H^3}, \tag{48}$$

where the last equality uses the Raychaudhuri equation $\dot{H} = -4\pi G_N(\rho + p)$ (valid in GR; modified in $f(R)$).

A.1.5 Internal Entropy Production

The non-equilibrium Clausius relation gives:

$$\frac{d_i S}{dt} = \frac{1}{T}\frac{\delta Q}{dt} - \frac{dS_{\text{Wald}}}{dt}. \tag{49}$$

Substituting and simplifying (setting $G_N = 1$ for clarity):

$$\boxed{\frac{d_i S}{dt} = -\frac{6\pi f''(R)}{H^2}(\ddot{H} + 4H\dot{H}) + (\text{GR terms})} \tag{50}$$

The "GR terms" vanish identically when evaluated on solutions, leaving only the $f''(R)$ contribution. This proves that $d_i S = 0 \Leftrightarrow f''(R) = 0$.

A.1.6 Proof of Theorem 3.1

Proof. Suppose $d_i S = 0$ for all FRW cosmologies. Then:

$$-\frac{6\pi f''(R)}{H^2}(\ddot{H} + 4H\dot{H}) = 0 \quad \forall (H, \dot{H}, \ddot{H}). \tag{51}$$

For fixed $(H, \dot{H})$, we can choose matter content to vary $\ddot{H}$ independently (e.g., different equations of state $w = p/\rho$). Since the factor $(\ddot{H} + 4H\dot{H})$ is not identically zero for all cosmologies, we require:

$$f''(R) = 0 \quad \forall R \in I. \tag{52}$$

Integrating twice: $f(R) = aR + b$ for constants a, b. This is GR with $G_{\text{eff}} = G/a$ and $\Lambda_{\text{eff}} = b/(2a)$. □

A.2 Brans–Dicke Theory

A.2.1 Setup

The action is:

$$S = \frac{1}{16\pi}\int \sqrt{-g}\left[\phi R - \frac{\omega}{\phi}(\nabla\phi)^2 - 2V(\phi)\right] d^4x. \tag{53}$$

The effective gravitational coupling is $G_{\text{eff}} = 1/\phi$. The Wald entropy is:

$$S_{\text{Wald}} = \frac{\phi A}{4} = \frac{\pi\phi}{H^2}. \tag{54}$$

A.2.2 Internal Entropy Production

In FRW cosmology with homogeneous $\phi(t)$:

$$\frac{dS_{\text{Wald}}}{dt} = \frac{\pi}{H^2}\left[\dot{\phi} - \frac{2\phi\dot{H}}{H}\right]. \tag{55}$$

The heat flux gives (using the modified Friedmann equations):

$$\frac{1}{T}\frac{\delta Q}{dt} = \frac{2\pi}{H}\cdot\frac{\delta Q}{dt} = -\frac{2\pi\phi\dot{H}}{H^3} + (\text{scalar contributions}). \tag{56}$$

The internal entropy production is:

$$\boxed{\frac{d_iS}{dt} = -\frac{\pi\dot{\phi}}{H^2}} \tag{57}$$

A.2.3 Proof of Theorem 3.2

Proof. If $d_iS = 0$ for all Brans-Dicke cosmologies, then $\dot{\phi} = 0$ for all times. This implies $\phi = \phi_0 = \text{const}$.

With constant ϕ, the Brans-Dicke action reduces to:

$$S = \frac{\phi_0}{16\pi}\int\sqrt{-g}\left[R - \frac{2V(\phi_0)}{\phi_0}\right]d^4x, \tag{58}$$

which is GR with $G_{\text{eff}} = 1/\phi_0$ and $\Lambda_{\text{eff}} = V(\phi_0)/\phi_0$. □

A.3 Horndeski $G_4(X)$ Models

For shift-symmetric Horndeski with $G_4 = G_4(X)$ where $X = -(\nabla\phi)^2/2$:

The effective coupling is:

$$\Phi_{\text{eff}} = G_4 - 2XG_{4X}. \tag{59}$$

The Wald entropy is:

$$S_{\text{Wald}} = \frac{\Phi_{\text{eff}}A}{4G_N}. \tag{60}$$

Non-equilibrium arises from:

$$\frac{d_iS}{dt} \propto \Phi_{\text{eff},X}\cdot\dot{X}. \tag{61}$$

Equilibrium requires $\Phi_{\text{eff},X} = 0$, which implies $G_4 = \text{const}$.

B Numerical Methods and Validation

B.1 FRW Solver

We solve the Friedmann equations numerically using a 4th-order Runge-Kutta integrator with adaptive step size. For modified gravity theories, the effective Friedmann equation is:

$$H^2 = \frac{8\pi G_N}{3}\rho_{\text{eff}}, \tag{62}$$

where ρ_{eff} includes both matter and geometric contributions from the modification.

B.2 N Computation

The integral:

$$N = \frac{\int |d_i S| \, d \ln a}{|\Delta S_{\rm geo}|} \tag{63}$$

is evaluated using the trapezoidal rule on a grid of 300 points from $a = 0.001$ to $a = 1.0$.

B.3 Validation Tests

1. **GR limit:** Verified $N = 0$ for GR + Λ to machine precision.
2. **Scaling relations:** Confirmed $d_i S \propto f''(R)$ for $f(R)$ and $d_i S \propto \dot{\phi}$ for BD.
3. **Conservation:** Checked energy conservation to $< 10^{-10}$ relative error.
4. **Comparison with literature:** Verified agreement with Eling–Guedens–Jacobson [6] for $f(R)$ non-equilibrium terms.

C EFT-of-DE and DHOST Parameter Sampling

C.1 Sampling Method

We use Latin Hypercube Sampling (LHS) to efficiently cover the parameter space with 1000 samples. LHS provides better coverage than random sampling for the same number of points.

C.2 Stability Cuts

Points are rejected if they exhibit:

1. Ghost instabilities (negative kinetic term)
2. Gradient instabilities (imaginary sound speed)
3. Laplacian instabilities (negative squared mass)

C.3 Observational Cuts

We apply approximate observational constraints from Planck+BAO+SNe:

- $|\alpha_{M0}| < 0.08$ (1σ)
- $|\alpha_{B0}| < 0.12$ (1σ)

D QFT Near-KMS Expansions and Numerical Calibration

D.1 KMS Condition

A state ω satisfies the KMS condition at inverse temperature β if for all observables A, B:

$$\omega(A\tau_t(B)) = \omega(\tau_{t+i\beta}(B)A), \tag{64}$$

where τ_t is the time evolution automorphism.

D.2 Origin of the $1/(4\pi^2)$ Suppression Factor

The KMS deviation parameter $\varepsilon_{\mathrm{KMS}}$ measures how far the adiabatic vacuum deviates from exact thermality. For a massless scalar field in FRW spacetime with metric $ds^2 = -dt^2 + a(t)^2 d\vec{x}^2$, the Wightman function in the adiabatic vacuum can be expanded as:

$$G^+(x, x') = G^+_{\mathrm{thermal}}(x, x') + \delta G^+(x, x'), \tag{65}$$

where the correction δG^+ arises from the time-dependence of $H(t)$.

The KMS condition requires $G^+(\tau + i\beta) = G^-(\tau)$ where $\beta = 2\pi/H$ is the inverse Gibbons-Hawking temperature. The leading adiabatic correction breaks this periodicity:

$$\delta G^+ \sim \frac{1}{4\pi^2}\frac{\dot{H}}{H^2} \times (\text{mode functions}). \tag{66}$$

The factor $1/(4\pi^2)$ arises from the combination of:

1. The thermal periodicity condition involves $\beta = 2\pi/H$, contributing a factor of $1/(2\pi)$.
2. The normalization of the Wightman function and integration over modes contributes an additional $1/(2\pi)$.

Thus $\varepsilon_{\mathrm{KMS}} = (1/4\pi^2) \times |\dot{H}|/H^2 \approx 0.025 \times |\dot{H}|/H^2$, which we verify numerically to hold with $< 1\%$ error across all scale factors in ΛCDM.

D.3 Near-KMS Expansion and Relative Entropy

For a state ρ_ϵ that is ϵ-close to a KMS state σ:

$$\rho_\epsilon = \sigma + \epsilon\delta\rho + \mathcal{O}(\epsilon^2). \tag{67}$$

The relative entropy is:

$$S_{\mathrm{rel}}(\rho_\epsilon \| \sigma) = \frac{1}{2}\chi\, \epsilon^2 + \mathcal{O}(\epsilon^3), \tag{68}$$

where $\chi = \mathrm{Tr}[\delta\rho\, K\, \delta\rho]$ is the susceptibility and K is related to the modular Hamiltonian. The linear term vanishes by the entanglement first law: $\delta S = \delta\langle K\rangle$ at first order.

D.4 Calibration of the C_2 Coefficient

The internal entropy production integrated over cosmic history gives:

$$N \sim C_2 \times \bar{\varepsilon}^2_{\mathrm{KMS}} \times \Delta \ln a. \tag{69}$$

To determine C_2, we perform a numerical calibration:

1. Create perturbed thermal states: $\rho_\epsilon = \rho_{\mathrm{thermal}} + \epsilon\, \delta\rho$ for $\epsilon \in \{0.01, 0.02, \ldots, 0.10\}$.
2. Compute $S_{\mathrm{rel}}(\rho_\epsilon \| \rho_{\mathrm{thermal}})$ for each ϵ.
3. Fit the data to $S_{\mathrm{rel}} = C_2 \times \epsilon^b$.
4. Verify $b \approx 2$ (quadratic scaling from entanglement first law).

Results:

- Fitted exponent: $b = 1.99 \pm 0.02$ (confirming quadratic scaling)
- Calibrated coefficient: $C_2 = 0.077 \pm 0.005$

This gives, for late-time ΛCDM with $\bar{\varepsilon}_{\mathrm{KMS}} \approx 0.035$ and $\Delta \ln a \approx 1.4$:

$$N_{\mathrm{QFT}} \approx 0.077 \times (0.035)^2 \times 1.4 \approx 1.3 \times 10^{-4}. \tag{70}$$

The uncertainty in C_2 propagates to roughly a factor of 2 uncertainty in N_{QFT}, but does not affect the order-of-magnitude conclusion $N \lesssim 10^{-3}$.

E Holographic Toy Models

E.1 AdS$_3$ Setup

We consider AdS$_3$ with metric:

$$ds^2 = \frac{\ell^2}{z^2}(-dt^2 + dx^2 + dz^2), \tag{71}$$

where ℓ is the AdS radius.

For an interval of length L on the boundary, the RT surface is a geodesic with:

$$S_{\mathrm{RT}} = \frac{c}{3} \log\left(\frac{L}{\epsilon}\right), \tag{72}$$

where c is the central charge and ϵ is the UV cutoff.

E.2 Quantum Corrections

The FLM correction adds bulk entanglement:

$$S = S_{\mathrm{RT}} + S_{\mathrm{bulk}} + \mathcal{O}(G_N^0). \tag{73}$$

For coherent bulk states, $S_{\mathrm{bulk}} \sim \mathcal{O}(1)$, while $S_{\mathrm{RT}} \sim c \sim \ell/G_N \gg 1$.

The fractional correction is:

$$\frac{\delta S}{S} \sim \frac{\mathcal{O}(1)}{S_{\mathrm{BH}}} \sim \frac{1}{S_{\mathrm{BH}}}. \tag{74}$$

E.3 Tolerance Scaling

For a toy model with $S_{\mathrm{BH}} \sim \pi/2$:

$$\text{tolerance} \sim \frac{1}{\pi/2} \sim 0.6. \tag{75}$$

For cosmological horizons with $S_{\mathrm{BH}} \sim 10^{122}$:

$$\text{tolerance} \sim 10^{-122}. \tag{76}$$

This extreme scaling shows that any sustained non-equilibrium would catastrophically violate bulk-boundary consistency for macroscopic horizons.