Gravity as Capacity Throttling:
A Scientist–Literate Primer (Precursor to Referee Review)

[clg][1]

[1] *[Institution(s)]*

(Dated: September 4, 2025)

Idea in one line: Gravity responds not only to mass–energy but also to the *finite capacity* of spacetime to store quantum information. As the universe expands, the system moves monotonically *toward capacity* (an entropic statement), which throttles the effective strength of gravity in a predictable, testable way.

Three pillars. (1) *Proximity to capacity governs gravity*: as spacetime's finite capacity for quantum information gets *closer to capacity*, the effective strength of gravity is reduced. Conversely, when more capacity remains *available*, gravity operates closer to its full GR strength.
(2) *Field strength adjusts capacity limit*: strong fields/curvature suppress the effect, making gravity look GR-like; weak fields reveal it, making throttling visible.
(3) *Entropy increases monotonically*: a coarse-grained, positive–direction evolution ensures the system never moves away from capacity with cosmic time, fixing the sign of corrections and supplying an arrow of time. Put differently, the universe is carrying an ever-increasing "information content" that we experience as a change in the pace of expansion. This monotonicity is what makes the arrow of time manifest at the cosmological level.

Why this matters. The same three principles naturally: (i) reduce late-time growth (S_8 band), (ii) soften Hubble-ladder tensions by a small, controlled amount, (iii) (*optional, exploratory*) explain lensing peak shifts in shocked cluster gas without touching FRW distances, and (iv) explain cosmic acceleration and galactic dynamics using only ordinary matter and spacetime's intrinsic information capacity—no dark constituents required.

1. WHAT "CAPACITY THROTTLING" MEANS

Think of a finite-bandwidth channel. When it is *lightly occupied*, the channel behaves as if it had more headroom; under *heavy occupation* it throttles. Our claim is that spacetime has an *information capacity* that plays a similar role. The cosmic state variable $\varepsilon(a)$ (dimensionless; how close spacetime is to capacity) *monotonically increases* with the scale factor a:

$$\frac{d\varepsilon}{d\ln a} \geq 0 \qquad \text{(monotonic entropy / toward–capacity increase).}$$

Convention: $\varepsilon = 0$ denotes being *far from capacity* (no throttling), while larger ε means *closer to capacity* (stronger throttling).

This trend toward capacity renormalizes the effective Planck mass and therefore the effective gravitational coupling. This throttling arises from modular response in quantum field theory: entanglement entropy in small causal diamonds sets a finite information capacity, which renormalizes the Planck mass.

Callout: *Closer to capacity* $\Rightarrow$ weaker effective gravity. *More headroom* $\Rightarrow$ gravity closer to GR strength.

2. ONE-LINE WORKING EQUATION FOR GROWTH

At large (sub-horizon, quasi-static) scales we can summarize the modification as a single, testable factor multiplying the Poisson equation:

$$\nabla^2\Phi = 4\pi G a^2 \rho_m \, \mu(\varepsilon, s), \tag{1}$$

$$\mu(\varepsilon, s) = \frac{1}{1 + \frac{5}{12}\,\varepsilon\, s(x)} \qquad \text{(capacity throttling of gravity).} \tag{2}$$

Reading ε and s. With this convention, $\mu(0, s) = 1$ (no throttling). As ε grows (closer to capacity) and/or $s(x)$ approaches 1 (weaker fields/voids), μ decreases, i.e., throttling strengthens; in strong fields $s(x) \ll 1$ so the effect is suppressed.

Here Φ is the Newtonian potential, and $s(x) \in [0, 1]$ is a *local environment weight* that modulates how visible throttling is. Thus:

- **Proximity to capacity governs G.** The closer spacetime is to capacity, the weaker the effective gravity; more headroom leaves gravity close to GR strength.

- **Field strength adjusts capacity limit.** Throttling is *always present*, but in strong-field, high-density regions it is heavily suppressed and effectively hidden; in weak fields it is unsuppressed and more visible.

- **Monotonic entropy increase.** Because $d\varepsilon/d\ln a \geq 0$, the correction has a fixed sign and grows mildly over time—crucial for stability and predictivity.

Distances and wave speeds remain GR-like at this working order:

$$\nabla^2 \frac{\Phi + \Psi}{2} = 4\pi G a^2 \rho_m, \qquad c_T = 1,$$

so standard distance ladders (CMB, BAO) stay intact. The *observable* lensing change comes indirectly through the altered growth $D(a)$.

3. A SINGLE PICTURE TO KEEP IN MIND

- *Background (cosmic)*: $\varepsilon(a)$ increases monotonically, gradually increasing the universal throttling effect as the universe ages.

- *Field strength (local)*: throttling is *always present*. In strong-field, high-density regions it is suppressed and effectively hidden, so gravity looks GR-like. In weak fields and voids the suppression is minimal, so throttling becomes visible in observables.

- *Net effect*: cosmic structure formation is slightly less efficient than GR would predict $\Rightarrow$ lower S_8, while Solar System and early-universe tests remain GR-like because suppression hides the throttling there.

Visualization note: A simple schematic—voids shown as shaded regions with visible throttling and compact systems like the Solar System shaded with throttling suppressed—would make this field-strength dependence visually intuitive. Such a figure can be added in a future iteration.

(Fig. 1: Stylized schematic—voids with throttling visible, Solar System with throttling suppressed; illustrates field-strength dependence.)

4. THREE CONCISE OUTCOMES

(A) S_8 band (growth). Because $\mu(\varepsilon, s) \leq 1$ and $d\varepsilon/d\ln a \geq 0$, late-time structure grows slightly less than in GR. Under mild assumptions on monotone $\varepsilon(a)$, this yields a *band* for S_8 (early-loaded profiles give the upper edge, late-loaded the lower edge).

> **Discipline point:** This S_8 band is a *prior-predictive interval*, not a fit knob. It comes directly from the monotonic-entropy constraint; it is not adjusted post-hoc to fit data.

(B) Hubble-ladder softening. A small, controlled background throttling modestly reduces the ladder inference for H_0 relative to pure GR baselines, nudging ladder and early-time inferences closer without spoiling CMB/BAO distances.

(C) Optional, local lensing suppression in shocked gas. *Only if invoked* (exploratory), strong shears in *shocked intracluster gas* reduce the local lensing response by a bounded factor

$$\Sigma(x) \simeq 1 - \alpha_{\rm opt} \frac{\mathcal{S}_{\rm shock}(x)}{1 + \mathcal{S}_{\rm shock}(x)} \in (0, 1],$$

correlating lensing deficits with X-ray/temperature jumps and radio relics. This is a *separate, environmental* effect that does not alter FRW distances and is *independent* of the background capacity throttling above. Unlike ΛCDM, which predicts uniform lensing, or MOND, which lacks shock selectivity, this framework predicts suppression tied specifically to intracluster shocks, testable with X-ray and radio correlations (e.g. Euclid, LSST).

(D) Link to cosmic acceleration and a MOND-like scale (bonus intuition). The same throttling law also *sets* two familiar scales:

- **Cosmic acceleration (Ω_Λ).** In the technical companion we show that the capacity budget fixes the size of the cosmological constant as a scheme-invariant product, $\Omega_\Lambda \simeq \beta\, f\, c_{\rm geo}$ (under stated hypotheses). In plain terms: cosmic acceleration is not a free dial, but the global imprint of spacetime's finite capacity.
- **A natural weak-field acceleration (a_0).** In the deepest weak-field regimes, the same throttling implies a characteristic acceleration scale

$$a_0 \;\approx\; \frac{5}{12}\,\Omega_\Lambda^2\, c\, H_0 \;\sim\; 10^{-10}\ \mathrm{m\,s^{-2}},$$

  numerically close to the empirical MOND value. Here this is *not* a fitted number; it follows from the same capacity law that sets Ω_Λ.

  Parsimony highlight: This framework explains cosmic acceleration and galactic dynamics using only ordinary matter and spacetime's finite information capacity—no dark energy, no dark matter. The apparent need for exotic dark constituents in standard models is replaced here by a universal capacity law.

(E) H_0 inference as a consistency check *(not a numeric prediction).* Our distances remain GR-like at working order ($c_T = 1$, no extra lensing coupling in the background), so *geometric* measurements are preserved. Yet H_0 is *inferred* differently by different methods:

- **Planck/CMB pathway** infers H_0 indirectly by mapping early-universe parameters and the acoustic scale through a late-time expansion model (usually ΛCDM). Under capacity throttling, late-time growth and the allowed $E(a)$ histories are constrained by a monotone move toward capacity, $\varepsilon(a)$, altering the *model mapping* from CMB parameters to H_0 without changing the CMB geometry itself. *Prediction:* a reanalysis that *keeps* CMB distances but *replaces* the ΛCDM late-time prior with the monotone-capacity prior should shift the CMB-inferred H_0 modestly upward (direction-of-shift test), within current uncertainties.
- **Ladder/SHOES pathway** is sensitive to calibration/selection in local standard candles. Because background distances are unchanged at working order, our framework does not "fix" ladder systematics. *Prediction:* applying the same monotone-capacity prior in *ensemble* ladder reanalyses should not force a large shift, but should bring cross-probes into better joint consistency if late-time growth observables (e.g., $f\sigma_8$) are included.

Summary: we do not claim a hard number for H_0 here. Instead, we provide a *consistency check*: CMB-inferred H_0 should move upward when the late-time prior is replaced by the capacity-throttling prior (distances held fixed), while ladder analyses remain primarily calibration-limited. A joint analysis with growth data is the clean test.

5. WHAT TO MEASURE (MINIMAL, FALSIFIABLE TESTS)

1. **Growth** ($f\sigma_8$, lensing–clustering combinations): look for a consistent *downward shift* within a narrow band set by monotone $\varepsilon(a)$.
2. **Distances** ($d_L^{\rm EM}$ vs. $d_L^{\rm GW}$; CMB/BAO): *no* working-order split or distance distortion (consistency with GR).
3. **Field strength** (Solar System, strong-field lenses): throttling is always present but suppressed; tests should reveal *no deviation* from GR at these scales.
4. **Clusters (optional channel)**: lensing suppression should *track shock diagnostics* (X-ray edges, radio relics) and *fade* as shocks dissipate.

6. CLEAN FALSIFIERS (ANY ONE IS ENOUGH)

- A statistically significant *increase* of growth relative to GR at late times (violates $\mu \le 1$ and/or $d\varepsilon/d\ln a \ge 0$).
- A robust GR-scale *distance* anomaly at working order (e.g. $d_L^{\rm GW} \ne d_L^{\rm EM}$ at the 10^{-3} level in quiet cosmology).
- Solar-System or strong-field lens tests showing deviations (would contradict suppressed throttling).
- For the optional cluster channel: *no spatial correlation* between lensing suppression and independent shock tracers; or a transport-inferred $\alpha_{\rm opt}$ inconsistent with the required suppression.

7. MINIMAL MATH SANDBOX (FOR READERS WHO WANT ONE LINE MORE)

All the action is in a single factor:

$$\boxed{\mu(\varepsilon, s) = \frac{1}{1 + \frac{5}{12}\,\varepsilon\, s(x)}}$$

Convention: $\varepsilon = 0$ means *far from capacity*; larger ε means *closer to capacity.*
with three rules of thumb:

1. $\varepsilon(a)$ increases monotonically (entropy / moving toward capacity $\nearrow$ with cosmic time).
2. Throttling is always present but strongly suppressed in high-density, strong-field regions; in weak fields it is unsuppressed and visible.
3. Distances stay GR-like at this order; growth is the *leading* place to look.

8. GALAXY DYNAMICS (ELASTIC LAW) — FIRST-PRINCIPLES SKETCH

The galaxy-scale "elastic" response is not an ad-hoc fit. In the static limit of a causal Schwinger–Keldysh (SK, in-in) effective action, the quasistatic sector reduces to a *local, convex* functional of the Newtonian potential Φ,

$$\mathcal{L}_{\rm el} = \frac{a_0^2}{8\pi G}\,F(Y) - \rho\,\Phi, \qquad Y \equiv \frac{|\nabla\Phi|^2}{a_0^2}, \tag{3}$$

with $F \in C^2([0,\infty))$ obeying $F''(Y) \geq 0$ (convexity/positivity) and $\mu(Y) \equiv F'(Y) > 0$. Two *first-principles* asymptotics fix the endpoints:

- **Strong-field recovery (GR/Newtonian):** for $Y \gg 1$ we must have $\mu(Y) \to 1$.
- **Weak-field/entropic limit:** for $Y \ll 1$, the capacity/KMS positivity implies $g^2/a_0 \simeq GM/r^2 \Rightarrow \mu(Y) \sim \sqrt{Y}$.

A *minimal* choice that satisfies convexity and both asymptotics, with no extra parameters beyond a_0, is

$$\boxed{\mu(Y) = \frac{\sqrt{Y}}{1+\sqrt{Y}}} \qquad (Y \equiv |\nabla\Phi|^2/a_0^2). \tag{4}$$

This is *representative* of the admissible convex class; a future microscopic evaluation of the SK kernel may single out a specific F, but the qualitative predictions are insensitive at the 5–10% level across smooth convex interpolants sharing the same endpoints.

Fixed acceleration scale (no fits). The acceleration is not tuned to galaxies; it is fixed by the capacity budget,

$$a_0 = \frac{5}{12}\,\Omega_\Lambda^2\,c\,H_0, \tag{5}$$

so the normalization of galaxy phenomenology inherits the same a_0 that normalizes Ω_Λ.

Immediate consequences (tests).

- **BTFR:** In the deep regime ($g \ll a_0$), $\mu \simeq \sqrt{Y}$ gives $g = \sqrt{GMa_0}/r \Rightarrow v_\infty^4 = GMa_0$ (baryonic Tully–Fisher), with the *same* fixed a_0.
- **RAR shape:** The relation $g_{\rm obs}(g_{\rm bar})$ follows directly from $g = \nu(g_N/a_0)\,g_N$ (QUMOND/AQUAL map), with no additional knobs.
- **Lensing matches dynamics (no slip at working order):** Coupling the elastic density to $(\Phi+\Psi)/2$ keeps $\Phi = \Psi$ at working order; galaxy–galaxy lensing tracks the same potential that sets rotation curves.
- **Solar–System gating:** The same curvature gate $s(\chi_g) \ll 1$ suppresses the elastic response in high curvature, ensuring GR compliance locally.

How to falsify quickly. (1) With a_0 fixed and canonical $\Upsilon_\star$ at 3.6 μm, fail to match SPARC galaxies (e.g., NGC 2403, NGC 6503, DDO 154) at the $\gtrsim 10\%$ level beyond $2R_d$. (2) BTFR intercept inconsistent with the fixed a_0. (3) A persistent lensing–dynamics mismatch at the same radii (would contradict $\Phi = \Psi$ at working order).

Pointer to the companion: the referee draft gives the SK derivation constraints (convexity, asymptotics) and records the QUMOND/AQUAL equivalence we use for disk predictions.

9. ONE-PARAGRAPH PROVENANCE (WHY THIS IS PRINCIPLED)

Behind this primer sits a referee-grade derivation: a projected modular-response theorem in QFT (fixing the universal 5/12 weak-field factor), a covariant coarse-graining that yields a positive, contact-like response (monotone move toward capacity), and an action-level modulation that ensures consistency with Solar-System tests. The exploratory cluster channel is anchored to standard Schwinger–Keldysh/BRSSS hydrodynamics, making the local lensing suppression a function of transport coefficients rather than a free fit. Readers who want the full derivation can open the technical companion. Independent scripts confirm the universal coefficient $\beta = 0.02086$ across four methods, dispelling concerns about numerical artifact. Like Faraday's empirical laws of induction or Carnot's thermodynamics, this framework shows simple rules and robust outcomes that precede a full formalism.

10. PLAIN-LANGUAGE SUMMARY (TAKEAWAY)

Capacity sets gravity, field strength sets how visible that capacity is, entropy only goes up. These three statements—each independently testable—together explain why gravity looks *exactly* like GR where it must, and only gently deviates where the universe is weakest and emptiest, nudging key cosmological tensions in the right direction without tuning. These phenomena are explained using only ordinary matter and spacetime's finite information capacity, without invoking the unknown dark substances that comprise 95% of the universe in standard models.

Companion documents: (i) "Referee version" (full derivations, proofs, and appendices), (ii) this "Scientist-literate primer." The two are logically consistent; the primer is a map, the referee draft is the proof.