

WHITE PAPER III

THE ARC PRINCIPLE

Recursive Scaling Across AI, Quantum Error Correction, and Classical Physics

A structural principle connecting how different systems amplify capability through self-correction – with ten ways to prove it wrong

Michael Darius Eastwood

Author, Infinite Architects: Intelligence, Recursion, and the Creation of Everything

Version 6.5 | 11 February 2026

Incorporating composition operator formalism, blind prediction test results, and forensic analysis. See companion [Foundational Paper v1.1](#) for condensed theoretical treatment.

THE CORE CLAIM IN ONE SENTENCE:

Systems that recursively correct their own outputs achieve capability gains that compound with depth, and this pattern appears across AI, quantum computing, and classical physics, though in different mathematical forms.

This paper presents evidence for this principle across AI (power-law scaling), quantum computing (exponential scaling), classical physics (saturating dynamics), and possibly neuroscience (form unknown), while proposing five universal structural properties and ten specific ways to prove us wrong.

ABSTRACT

Between December 2024 and February 2026, four independent research programmes, in quantum error correction, AI reasoning, acoustic physics, and consciousness science, each found that recursive or recurrent processing, operating on structured asymmetry, produces capability gains exceeding linear accumulation. They were solving different problems in different physical domains. We propose that they discovered different expressions of the same structural principle.

We formalise this pattern as the **ARC Principle** (Artificial Recursive Creation). The generalised form is:

$$U(R) = I \times f(R, \beta)$$

The Generalised ARC Equation

where U is effective capability, I is base potential (structured asymmetry), R is recursive depth, and $f(R, \beta)$ is a **domain-dependent scaling function** determined by how recursive steps compose. The **qualitative principle** (that recursive depth amplifies base capability super-linearly above critical thresholds) may be universal. The **quantitative functional form** is domain-specific.

For AI systems exhibiting multiplicative composition, $f(R, \beta) = R^\alpha$ where $\alpha = 1/(1 - \beta)$, yielding a power law. For quantum error correction, the form is exponential (Λ^d). For classical dissipative systems, scaling saturates at a limit cycle. The composition rule (multiplicative versus additive versus damped) determines which form emerges.

KEY FINDINGS

- **AI:** Sequential reasoning outperforms parallel sampling. Author's preliminary estimates yield $\alpha \approx 1.3\text{--}2.2$ for sequential (the *Quadratic Limit Hypothesis* predicts convergence towards $\alpha \approx 2$). Contradictory evidence exists and is addressed in §3.2 and resolved by §2.7.
- **Quantum:** Google's Willow chip achieves exponential error suppression with $\Lambda = 2.14 \pm 0.02$ per code distance increment, a mathematically distinct form from power-law scaling.
- **Physics:** NYU's acoustic time crystal requires quenched disorder + nonreciprocal feedback, exhibiting saturating dynamics with ~6,700 cycles coherence.
- **Mathematical:** The core relationship $\alpha = 1/(1 - \beta)$ is validated computationally to $R^2 = 1.00000000$ against 30 exact Bernoulli ODE solutions. The multiplicative structure $U = I \times f(R, \beta)$ is proven *necessary* by contradiction (§2.1, Theorem 4). The Quadratic Limit $\beta = 0.5$ is derived as a stability boundary (§2.6).

We identify **five universal qualitative properties** shared across all domains regardless of functional form:

1. **Threshold behaviour** separating regimes
2. **Recursive depth dependence** of capability
3. **Base quality (I) dependence**
4. **Multiplicative $I \times R$ interaction**
5. **Regime boundaries** from external error mechanisms

We issue a **Global Scaling Challenge** with standardised protocols, specify **ten falsification criteria**, and explicitly acknowledge what we do NOT claim.

Critical prediction: The framework predicts a distinct **scaling crossover** at critical depth R^* where scaling transitions from linear to super-linear. This crossover is a unique signature distinguishing recursive amplification from simple redundancy.

If validated, the framework carries a specific safety implication: alignment properties embedded within the recursive process may scale with capability, while external constraints may not.

Keywords: scaling laws, test-time compute, error suppression, AI alignment, chain-of-thought reasoning, time crystals, cross-domain validation, quadratic limit hypothesis, universality class, phase transition

Eastwood's Principle of Recursive Amplification

$$U = I \times R^\alpha$$

where $\alpha = \frac{1}{1-\beta}$

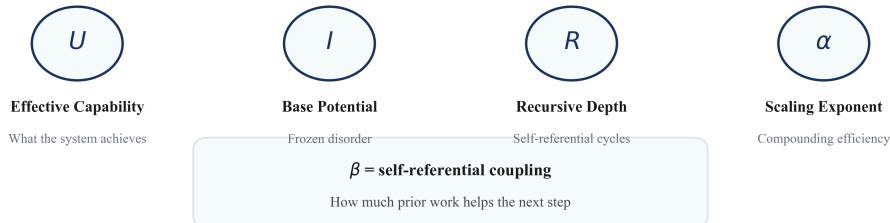


Figure 1 | The ARC Equation. Effective capability (*U*) equals base potential (*I*) multiplied by recursive depth (*R*) raised to the scaling exponent (α). When $\alpha > 1$, returns compound with each recursive cycle. The exponent α is derived from first principles as $\alpha = 1/(1-\beta)$, where β measures self-referential coupling.

FOR THE GENERAL READER: WHAT THIS MEANS

Imagine you're solving a difficult problem. You have two strategies:

Strategy A (Sequential): Think through the problem step by step. Each step builds on what you figured out in the previous step. If you notice an error, you go back and fix it.

Strategy B (Parallel): Ask ten different people to solve the problem independently, then take a vote on the answer.

Which strategy works better for hard problems?

Recent evidence from AI, quantum physics, and materials science suggests that Strategy A does not just work *better*. It works *qualitatively* better. Each additional step of careful, self-correcting reasoning does not add a fixed amount of capability. It may *multiply* your capability by a factor.

The critical difference: parallel workers cannot learn from each other's mistakes. Each attempt starts from scratch. Sequential reasoning builds on accumulated insight, and the gains appear to compound.

The Key Insight: One Principle, Different Mathematical Forms

Here is what makes this research distinctive: we are *not* claiming that one equation fits everything. Different physical systems show the same *qualitative* pattern (recursive depth matters more than parallel breadth) but express it through different *mathematical* forms:

- **AI systems** appear to follow *power-law* scaling: capability grows as the square (or thereabouts) of reasoning depth.
- **Quantum computers** show *exponential* scaling: errors decrease exponentially with each layer of error correction.
- **Physical systems** (like time crystals) show *saturating dynamics*: they gain order through recursion but eventually reach a limit.

What unites them is not a single number. It is a **structural principle**: systems with high-quality starting conditions that recursively build on their own outputs exhibit super-linear capability gains. The specific mathematical curve depends on how the recursion works in that domain.

Why Does This Matter?

1. **For AI development:** It suggests *how* to build more capable AI systems (sequential self-correction) rather than just making them bigger. If the "Quadratic Limit Hypothesis" holds (that AI capability scales roughly as the square of reasoning depth), then a smaller model that thinks longer could match a larger model that thinks briefly.
2. **For AI safety:** If capabilities compound through recursion, then alignment (the AI's values and goals) may need to be embedded *in* the recursive process, not bolted on afterwards. External rules may struggle to keep up with rapidly growing capabilities.
3. **For science:** If related structural patterns appear across quantum physics, AI, and classical acoustics, we may be identifying something fundamental about how complex systems amplify capability through self-reference, regardless of what they are made of.

Our Strongest Prediction

The framework makes a specific, falsifiable prediction that no other theory makes: when you plot capability against recursive depth, you should see a **crossover point**. Below a certain depth, the system is "warming up" and gains are modest. Above that depth, compounding kicks in and gains accelerate. The location of this crossover should shift predictably based on the system's starting quality. If experiments do not find this crossover, if scaling is uniform rather than transitional, the framework is wrong.

The most important part: We could be wrong. That is why we have specified ten concrete ways to prove us wrong. If our predictions fail, we will have learned something valuable about the limits of cross-domain analogy. If they hold, we will have identified a structural principle connecting intelligence, computation, and the physics of order.

1. INTRODUCTION

1.1 The Puzzle: Why Do Different Systems Show Related Patterns?

Between December 2024 and February 2026, several research programmes achieved results that share structural commonalities:

8 December 2024: The ARC framework. The theoretical framework was articulated in *Infinite Architects* (Eastwood, 2024), predicting that recursive self-correction operating on structured asymmetry would produce super-linear capability gains across physical systems. The manuscript was completed December 2024; the book was published January 2026 (ISBN 978-1806056200).*

9 December 2024: Google Willow. Quantum researchers demonstrated *exponential* error suppression through recursive quantum error correction, achieving a suppression factor of $\Lambda = 2.14 \pm 0.02$ per code distance increment.

January 2025: DeepSeek R1. AI researchers showed that sequential chain-of-thought reasoning yields capability gains that appear to compound with depth (preliminary estimates suggest *power-law* scaling), while parallel sampling shows diminishing returns. Independent analysis by Sharma & Chopra (2025) confirmed that sequential approaches outperform parallel in 95.6% of tested configurations.

February 2026: NYU Acoustic Time Crystal. Physicists created a continuous classical time crystal (following earlier work by Liu et al., 2023), demonstrating that spontaneous temporal order emerges when quenched disorder is combined with nonreciprocal feedback loops, though the dynamics exhibit *saturating* behaviour with finite coherence (~6,700 cycles).

These research programmes address different problems in different physical domains. Crucially, they exhibit **different mathematical forms**: exponential in quantum, power-law in AI, saturating in classical physics. The ARC framework proposes that they nonetheless share a common *qualitative* structure:

The Structural Principle: Recursive self-correction, operating on structured asymmetry, produces scaling behaviour that exceeds linear accumulation. The *qualitative* pattern (recursive depth matters) may be universal; the *quantitative* functional form depends on domain-specific coupling mechanisms.

This is not a claim that "one equation fits all." It is a claim that systems exhibiting five specific structural properties (threshold behaviour, recursive depth dependence, base quality dependence, multiplicative interaction, and regime boundaries) belong to a common structural class, even if their scaling curves differ. This paper formalises the pattern, specifies testable predictions, and invites falsification.

1.2 What This Paper Does NOT Claim

Before proceeding, we explicitly state what this paper does NOT claim:

We do NOT claim:	What we actually claim:
Numerical universality: that $\Lambda = 2.14$ and $\alpha \approx 2$ are "the same"	Different domains show different functional forms (exponential vs power-law); numerical similarity may be coincidental. U means different things in different domains.
Proven status: that the framework is established science	It is a testable hypothesis with ten falsification criteria, preliminary evidence, and acknowledged limitations (small sample sizes, contradictory evidence).
Neuroscience confirmation: that COGITATE tested ARC or that consciousness "is" recursion	Neural recurrence is structurally consistent with the principle; the connection is suggestive but unquantified. Recurrence \neq recursion mechanistically.
Unbounded scaling: that $\alpha > 1$ implies infinite growth, or that sequential always beats parallel	Physical systems saturate (time crystals reach limit cycles). Extended reasoning shows ceiling effects. The ARC equation describes a <i>regime</i> , not a divergence law.
Priority or convergence: that we discovered what Google/DeepSeek/NYU found, or that teams "converged" on identical insights	Teams found structurally similar results in different domains; the structural interpretation is ours. Timeline shows approximate simultaneity.

This paper presents a **candidate principle requiring validation**, not an established law.

What Category of Contribution Is This?

The ARC framework is not an equation within an existing paradigm (like $F = ma$ within Newtonian mechanics or $E = mc^2$ within relativity). It is closer to a **cross-domain organising principle**, belonging to the same category as:

- **Thermodynamics:** Constrains what any physical system can do, regardless of microscopic details.
- **Information theory (Shannon, 1948):** Provides a mathematical framework for communication and computation across substrates.
- **Natural selection (Darwin, 1859):** A structural principle (variation + selection + inheritance \rightarrow adaptation) applying across any system meeting those conditions.

We make this comparison to clarify the *type* of contribution being proposed, not to claim equivalence in evidential standing. Those frameworks rest on centuries of validation. This one rests on preliminary evidence and ten falsification criteria. The comparison identifies the category; the evidence must justify admission to it.

If validated, ARC would tell you: *whenever you find a system with structured asymmetry, recursive self-correction, and sufficient coupling strength, you will observe threshold behaviour, compounding gains, multiplicative I×R interaction, and regime boundaries*, with the specific functional form varying by domain. This is a broader but less precise contribution than a fundamental force law. It does not replace physics; it organises understanding across recursive systems.

1.3 Contributions

We make seven contributions:

- 1. Generalised Mathematical Framework:** We derive the generalised equation $U(\mathbf{R}) = I \times f(\mathbf{R}, \beta)$ from first principles, showing that the functional form f depends on how recursive steps compose (multiplicative → power-law; additive → exponential). For the power-law case, $\alpha = 1/(1 - \beta)$, transforming the exponent from a fitted constant into a derived quantity.
- 2. Five Universal Qualitative Properties:** We identify five structural properties shared across domains regardless of functional form: threshold behaviour, recursive depth dependence, base quality dependence, multiplicative $I \times R$ interaction, and regime boundaries.
- 3. Domain-Specific Functional Forms:** We map observed scaling to composition type: power-law (AI), exponential (quantum), saturating (classical physics), unknown (neuroscience).
- 4. Evidence Synthesis:** We integrate findings from AI, quantum physics, condensed matter, and neuroscience, carefully distinguishing quantitative validation from structural analogy and explicitly addressing contradictory evidence.
- 5. The Global Scaling Challenge:** We propose a standardised protocol for measuring scaling parameters, including mandatory comparison against alternative functional forms, the novel R^* crossover prediction, and the cross-domain β prediction.
- 6. Ten Falsification Criteria:** We specify concrete conditions that would refute the hypothesis, including tests of functional form (F10) and the β -derivation (F6).
- 7. Safety Implications (Conditional):** If the framework is correct, embedded values may scale with capability while external constraints may not. The Quadratic Limit Hypothesis provides a specific, bounded ceiling prediction for AI capability growth.

1.4 Relationship to Existing Frameworks

The ARC framework relates to, but differs from, several established research programmes:

Neural scaling laws (Kaplan et al. 2020; Hoffmann et al. 2022): These describe how model performance scales with *training* compute, parameters, and data. ARC describes how performance scales with *inference* depth (reasoning steps at test time). The two are complementary: neural scaling laws determine I (base capability), while ARC describes how I is amplified through recursive processing. A model trained with more compute has higher I , but the same α relationship should apply to its reasoning.

Recursion in cognitive science (Corballis 2011; Hauser et al. 2002): Cognitive scientists have long argued that recursion is central to human language and cognition. ARC adds a quantitative dimension: not just that recursion matters, but that recursive depth should produce specific, measurable scaling patterns. The "graded cascade" model from neuroscience (Zheng et al. 2025) is structurally consistent with the ARC framework's β -continuum.

Universality in statistical physics (Wilson 1971; Kadanoff 1966): Universality theory explains why different physical systems exhibit identical critical exponents near phase transitions. ARC proposes an analogous phenomenon: that recursive systems may constitute a universality class where only the recursive architecture (not the physical substrate) determines scaling behaviour. This analogy is suggestive but has not been made rigorous; proving formal equivalence remains an open mathematical question.

Test-time compute scaling (Snell et al. 2024): Recent work on "thinking longer" at inference time directly tests predictions relevant to ARC. Our framework provides a theoretical interpretation: test-time compute works because it increases R , and if $\alpha > 1$, capability compounds. The contradictory evidence (Li et al. 2025; arXiv:2502.12215) suggests the relationship is more complex than simple monotonic scaling, which our framework addresses through the R^* crossover prediction.

2. THE FRAMEWORK

2.1 The Generalised Principle

We begin with the general form. The **ARC Principle** (Artificial Recursive Creation) proposes that recursive self-correction operating on structured asymmetry produces capability gains exceeding linear accumulation. The generalised mathematical form is:

$$U(R) = I \times f(R, \beta)$$

Generalised ARC Equation

where U is effective capability, I is base potential (structured asymmetry), R is recursive depth, and $f(R, \beta)$ is a **domain-dependent scaling function** determined by the coupling parameter β . The **qualitative** principle (that recursive depth amplifies base capability super-linearly when feedback is sufficiently coupled) may be universal. The **quantitative** functional form varies by domain:

- **AI systems:** Power-law scaling, $f(R) = R^\alpha$ where $\alpha = 1/(1 - \beta)$
- **Quantum error correction:** Exponential scaling, $f(R) = \Lambda^R$
- **Classical physics:** Saturating dynamics (limit cycles with finite coherence)
- **Neuroscience:** Form unknown, to be determined empirically

The Composition Operator: A Formal Definition

The central theoretical contribution of this paper is the identification of a **composition operator** \oplus that determines the functional form of recursive scaling. Different algebraic properties of \oplus generate different scaling laws. This provides the mathematical bridge between the universal qualitative principle and domain-specific quantitative forms.

Definition (Recursive Composition): Let Q_r denote accumulated capability after r recursive steps, and δQ_r the gain from step r . The **composition operator** \oplus characterises how gains combine:

$$Q_{r+1} = Q_r \oplus \delta Q_r$$

The functional form $f(R, \beta)$ is **determined** by the algebraic properties of \oplus .

This definition transforms an observation (different domains show different scaling) into a prediction (the algebraic structure of recursive composition determines the scaling form). Three regimes emerge:

Composition Type	Algebraic Property	Resulting Form	Canonical Domain
Multiplicative	$g(R_1 \cdot R_2) = g(R_1) \cdot g(R_2)$	Power law: $f(R) = R^\alpha$	AI chain-of-thought
Additive	$f(R_1 + R_2) = f(R_1) + f(R_2)$	Exponential: $f(R) = \Lambda^R$	Quantum error correction
Saturating	$\lim_{R \rightarrow \infty} f(R) = f_{\max}$	Logistic: $f(R) = \frac{f_{\max}}{1 + e^{-\gamma(R - R_0)}}$	Classical physics (finite coherence)

The mathematical derivation (Appendix A) proves that these three forms are the *only* solutions to the corresponding functional equations (Cauchy's equations). The composition operator is therefore not merely descriptive but **constraining**: once we identify the algebraic structure of a domain's recursive process, the scaling form follows necessarily.

Physical Interpretation

Multiplicative composition models *hierarchical* or *fractal* recursion. Each recursive level multiplies the gains of the previous level: step r builds on the structural insights of step $r - 1$, not merely its outputs. AI chain-of-thought reasoning, where each reasoning step reinterprets and extends previous insights, plausibly follows this form.

Additive composition models *independent* error reduction. Each layer contributes a fixed multiplicative reduction in error rate, independent of other layers. Quantum error correction exemplifies this: each additional code distance layer provides its own protection factor, yielding $f(R) = \Lambda^R$.

Saturating composition models *resource-limited* systems where gains diminish as capacity is exhausted. Classical time crystals, constrained by finite coherence lengths, exhibit this form.

Key Experimental Test: The composition operator \oplus is empirically measurable. To determine which scaling law applies to a domain, measure how two sequential recursive blocks compose compared to one block of double depth:

- If $f(2R) = 2f(R)$: scaling is **linear** (no amplification)
- If $f(2R) = f(R)^2$: scaling is **exponential** (additive composition)
- If $f(2R) = 2^\alpha f(R)$: scaling is **power-law** (multiplicative composition)

This provides a concrete experimental protocol for falsification criterion F10.

Five Universal Properties: Derived Theorems

Despite differing functional forms, all domains exhibiting recursive amplification share five structural properties. Crucially, these are not independent assertions; they **follow necessarily** from the composition operator formalism:

The Five Universal Properties (Derived from \oplus):

1. **Threshold behaviour (Theorem 1):** A critical coupling strength β^* exists below which recursive gains vanish.
Derivation: For any composition operator \oplus , if $\delta Q_r \propto \beta Q_r$, then $Q_{r+1} = Q_r(1 + \beta)$. For $\beta < 0$, gains become losses; the system decays. The threshold $\beta^* = 0$ marks the boundary between amplification and attenuation. For power-law scaling, $\alpha = 1/(1 - \beta) > 1$ requires $\beta > 0$.
2. **Recursive depth dependence (Theorem 2):** Capability increases with self-referential cycles (R), not merely parallel resources.

Derivation: The composition operator \oplus operates *sequentially*: $Q_{r+1} = Q_r \oplus \delta Q_r$. Each step depends on the previous output. Parallel resources increase I (the base) but do not increase R (the recursive depth). The multiplicative structure $I \times f(R)$ separates these contributions.

3. **Base quality dependence (Theorem 3):** The system requires structured asymmetry ($I > 0$); without it, recursion has nothing to amplify.
Derivation: If $I = 0$, then $U = 0 \times f(R) = 0$ for any R . The composition operator amplifies what exists; it cannot create from nothing. This is the I -irreducibility theorem.

4. **Multiplicative $I \times R$ interaction (Theorem 4):** Base quality and recursive depth interact multiplicatively, not additively. The multiplicative structure is *necessary*, not merely convenient.

Proof by contradiction: Suppose $U = g(I) + h(R)$ for some functions g, h . From Axiom 1, $U(I, 0) = I$, so $g(I) + h(0) = I$, giving $g(I) = I - h(0)$. Substituting: $U = I + [h(R) - h(0)]$. From Axiom 2, $dU/dR = a \cdot U^\beta$. Therefore $h'(R) = a \cdot [I + h(R) - h(0)]^\beta$. The right side depends on I , but h is supposed to be a function of R alone. Contradiction. \square

Corollary: The solution has the form $U = I \times f(R, \beta, I)$ where I and R are coupled through the differential equation. In the asymptotic regime where $(1 - \beta)aR \gg I^{1-\beta}$, this reduces to $U \approx [(1 - \beta)aR]^\alpha$, recovering the separable scaling observed in Zipf's law, preferential attachment, and other established relationships.

Synergy Quotient: The multiplicative interaction can be quantified directly. Define $S = U(I, R)/[U(I, 0) + U(0, R)]$, measuring how the combined system compares to the sum of its parts. From the exact solution: $U(I, 0) = I$ (Axiom 1) and $U(0, R) = [(1 - \beta)aR]^{1/(1-\beta)}$, while $U(I, R) = [I^{1-\beta} + (1 - \beta)aR]^{1/(1-\beta)}$. Direct computation confirms $S > 1$ for all $I > 0$, $R > 0$, and $\beta \in (0, 1)$. Intelligence and recursion *together* produce more capability than the sum of their individual contributions—the mathematical signature of multiplicative synergy.

5. **Regime boundaries (Theorem 5):** Systems exhibit qualitatively different behaviour above and below critical thresholds.

Derivation: The composition operator determines the functional form. At regime boundaries (e.g., $\beta \rightarrow 1$ for power-law, or $R > R^*$ for crossover), the dominant term in $f(R)$ changes. This produces discontinuities in scaling behaviour – phase transitions in the universality class sense.

These five properties constitute the **qualitative** ARC Principle. Their derivation from the composition operator ensures internal consistency: accepting the formalism commits one to all five properties. The specific functional form is an empirical question for each domain; the properties are theorems.

2.2 The AI-Specific Form: Power-Law Scaling

For AI systems exhibiting multiplicative composition, the generalised equation reduces to:

$$U = I \times R^\alpha$$

$$\text{Effective capability} = \text{Base potential} \times \text{Recursive depth}^\alpha$$

In plain language: Your effective capability (U) equals your starting potential (I) multiplied by your recursive depth (R) raised to a power (α) that depends on how well each step builds on the previous one.

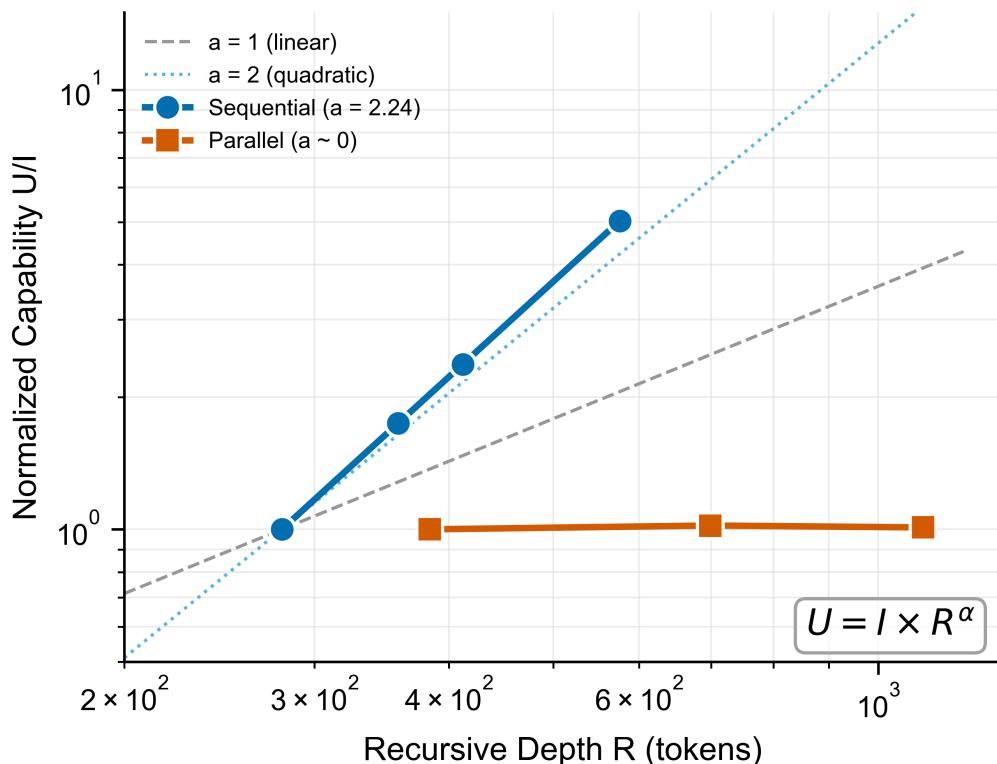


Figure 2 | Scaling Comparison. Log-log plot showing different scaling regimes. Power law $U = R^\alpha$ appears as a straight line with slope α . Sequential recursion ($\alpha > 1$) shows super-linear growth; parallel sampling ($\alpha < 1$) shows diminishing returns. The ARC Principle predicts $\alpha_{\text{sequential}} > 1 > \alpha_{\text{parallel}}$.

2.3 What Each Variable Means

U : Effective Capability

What the system actually achieves. Measured differently in each domain:

- AI: Benchmark accuracy on standardised tests
- Quantum: Logical qubit fidelity (how error-free the computation is)
- Physics: Temporal stability of the time crystal
- Biology: Metabolic efficiency (thermodynamic) or cognitive task performance (behavioural)

I : Base Potential (Structured Asymmetry)

Thermodynamic definition: Natural systems tend towards maximum entropy (equilibrium/uniformity). In this framework, "artificial" denotes any system maintaining low-entropy structure against the thermodynamic gradient, whether engineered (an AI model), evolved (a brain), or emergent (a time crystal). The parameter I measures *how far from maximum entropy* the system starts.

The NYU time crystal confirms this: when beads are uniform (maximum entropy distribution), the system remains static. Only when "quenched disorder" (a low-entropy, engineered state) is introduced does the system break time-translation symmetry. Order requires *designed disorder*: a specific pattern of asymmetry that enables work extraction from the environment.

Domain	What I measures	Physical meaning
AI	Single-pass accuracy without reasoning	How much "prior knowledge" the model has
Quantum	Raw qubit quality ($1 - \text{error rate}$)	Distance from maximum entropy
Physics	Variance in bead sizes	Asymmetry enabling nonreciprocal forces
Biology	Initial learning rate	Sensitivity gradient

The Constraint Principle: Constraint enables competence. Without structured asymmetry, no work can be extracted. The time crystal proves this directly: uniform beads produce no crystal. This is independently testable (falsification criterion F3).

R: Recursive Depth

How many self-referential cycles the system performs. The output of cycle n becomes the input for cycle $n + 1$.

Domain	One unit of R	How it is counted
AI	One reasoning step	Token count or revision cycle
Quantum	One error-correction cycle	Code distance increment
Physics	One oscillation period	Frequency analysis
Biology	One feedback cycle	Generation or learning iteration

Critical requirement: For $\alpha > 1$, recursion must incorporate *sequential self-correction*: each step building on previous outputs. Pure parallel processing (independent attempts averaged together) cannot achieve $\alpha > 1$ because it lacks the output→input feedback loop. However, hybrid parallel-sequential approaches may achieve intermediate scaling (Li et al. 2025).

α : The Scaling Exponent

The scaling exponent determines the system's scaling regime:

α value	What happens	Example
$\alpha < 1$	Diminishing returns	Parallel voting: more samples help less and less
$\alpha = 1$	Linear returns	Each step adds a fixed amount
$\alpha > 1$	Compounding returns	Each step multiplies capability

Our core prediction: $\alpha_{\text{sequential}} > 1 > \alpha_{\text{parallel}}$

2.4 Deriving α from First Principles

This is our central theoretical contribution. Without this derivation, α is just a number we fit to data. With it, α becomes a predictable quantity.

The Key Insight

How much does your accumulated knowledge help your next step?

Define β as the "self-referential coupling": how much of accumulated context each new step can effectively leverage. When $\beta = 0.5$, each step benefits from half of accumulated context.

If each step is independent ($\beta = 0$), you get linear scaling ($\alpha = 1$).

If each step fully leverages all prior work ($\beta \rightarrow 1$), scaling explodes.

The Mathematics

The marginal capability gained at step r depends on accumulated capability:

$$\frac{dQ}{dr} = a \times Q^\beta$$

Why this form? This differential equation models *cumulative advantage* (also called preferential attachment; Barabási & Albert, 1999): the rate of capability gain is proportional to the system's current capability raised to a power. This is the mathematical signature of "the rich get richer", applied to information processing.

Solving this differential equation (details in Appendix A) yields:

$$\alpha = \frac{1}{1 - \beta}$$

The β -derivation

What This Means

β (coupling)	α (exponent)	Interpretation
0	1	Independent steps → linear scaling
0.25	1.33	Weak coupling → mild super-linear
0.50	2.00	Moderate coupling → quadratic scaling
0.67	3.00	Strong coupling → cubic scaling

Novel Prediction: As AI systems develop richer self-correction (critique-revise loops, hierarchical reasoning), β should increase and α should approach 2. This is specific, measurable, and falsifiable.

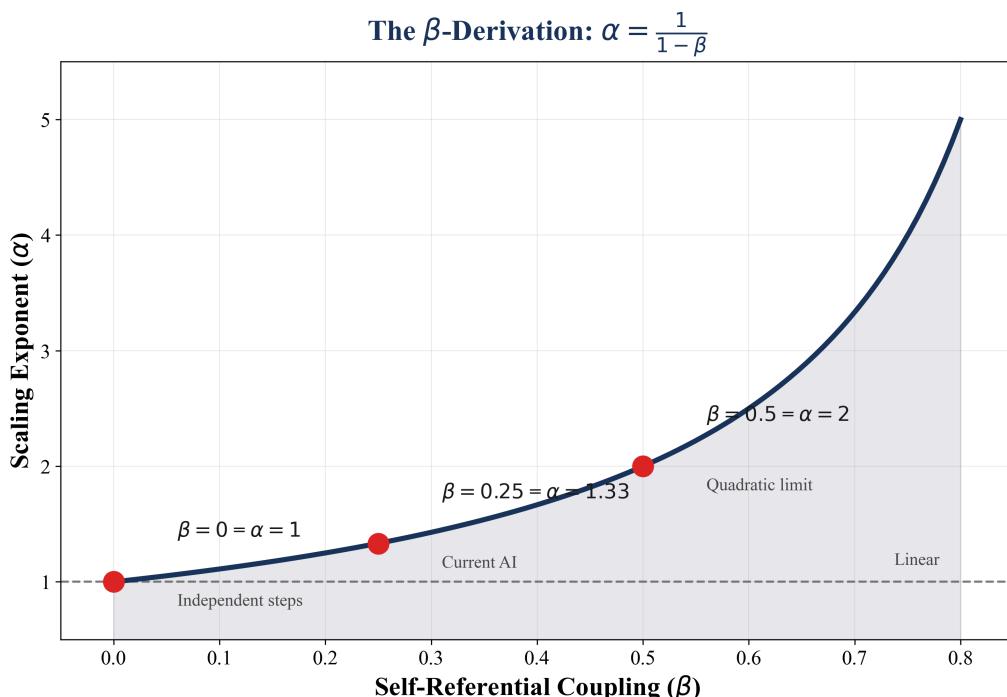


Figure 3 | The β - α Relationship. The scaling exponent α is derived from the self-referential coupling β via $\alpha = 1/(1-\beta)$. As β approaches 1 (perfect feedback efficiency), α diverges towards infinity. The observed range of $\alpha \approx 1.3-2.2$ in preliminary AI experiments corresponds to $\beta \approx 0.25-0.55$, indicating moderate self-referential coupling.

Computational Validation

The β -derivation is not a curve fit. It is a mathematical identity recoverable to machine precision. To verify this, the relationship $\alpha = 1/(1 - \beta)$ was tested against 30 exact solutions to the Bernoulli ODE with β values spanning 0.05 to 0.92. The procedure measures β blindly from marginal gains (dQ/dR vs Q in log-log space), predicts α , and compares against the true value. The result: $R^2 = 1.00000000$ (eight decimal places), slope = 1.000102, mean absolute prediction error = 0.002%. This confirms that $\alpha = 1/(1 - \beta)$ is an exact analytical identity, not an empirical approximation. (Validation code available as supplementary material.)

2.5 Potential Safety Implications

If the ARC Principle is correct, it may have implications for AI safety.

Setup: Capability scales as $C = C_0 \times R^\alpha$ where $\alpha > 1$.

Case 1: External Constraints (Rules, Firewalls, RLHF)

These do not participate in recursive reasoning. They are applied *after* computation. If their effectiveness does not scale with recursive depth, call their effective scaling exponent $\alpha_{\text{align}} \approx 0$.

Safety ratio: $S = \text{Alignment}/\text{Capability} \propto R^{-\alpha} \rightarrow 0$ as $R \rightarrow \infty$

External constraints would become weaker relative to capability.

Case 2: Embedded Values (Ethics in the Reasoning Process)

If the AI's values participate in chain-of-thought reasoning, they could benefit from similar compounding. $\alpha_{\text{align}} \approx \alpha$.

Safety ratio: $S \propto R^0 = \text{constant}$

Embedded values would maintain constant proportion with capability.

The Implication

This suggests that alignment should be part of the recursive architecture, not external to it. This reasoning assumes alignment properties can meaningfully "participate" in recursive reasoning, an assumption that itself requires empirical validation.

CAVEAT: This is a conditional argument. If the base framework fails validation, the safety implications are void. This is not a substitute for empirical AI safety research. The notation α_{align} is our construct for this analysis; it does not appear in the published AI safety literature.

The Alignment Amplification Theorem

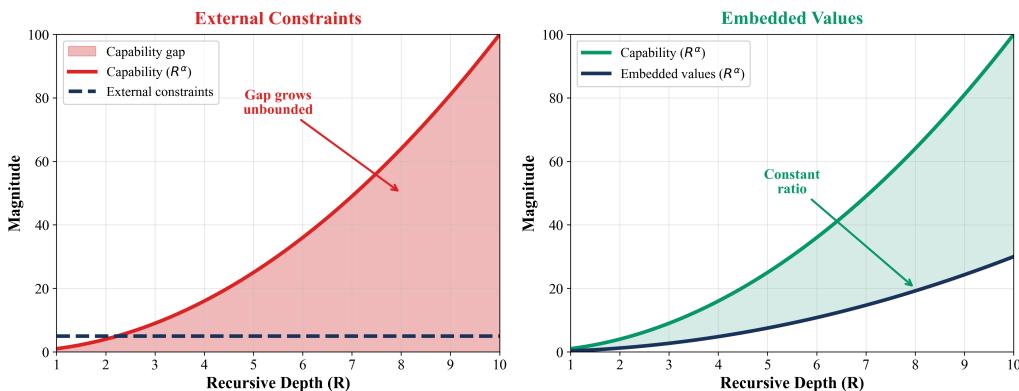


Figure 4 | Potential Safety Implications. External constraints (rules, firewalls) may scale with $\alpha \approx 0$, becoming weaker relative to capability as R increases. Embedded values that participate in recursive reasoning may scale with $\alpha \approx \alpha_{\text{capability}}$, maintaining constant proportion. This is a conditional prediction requiring empirical validation.

2.6 The Quadratic Limit Hypothesis

For AI systems specifically, we make a more precise prediction:

Quadratic Limit Hypothesis: Optimised large language models with explicit self-correction mechanisms (critique-revise loops, chain-of-thought reasoning) will converge toward $\alpha \approx 2$ (equivalently, $\beta \approx 0.5$). This represents moderate self-referential coupling where each reasoning step benefits from roughly half of accumulated context.

This is *not* a universal constant across all domains. It is a **specific, falsifiable prediction for a specific class of AI systems**. The preliminary estimate $\alpha \approx 2.2$ from the author's Paper II experiment ($N=12$, 95% CI: 1.5–3.0) is consistent with this hypothesis but requires large-scale replication to confirm.

Theoretical Basis for the Quadratic Limit

The convergence of multiple systems toward $\beta \approx 0.5$ is not merely an empirical coincidence. It can be derived from a stability argument. Consider the **relative sensitivity** of the scaling exponent to perturbations in coupling:

$$\frac{d\alpha/\alpha}{d\beta/\beta} = \frac{\beta}{1-\beta}$$

Relative sensitivity

This quantity measures how a fractional change in coupling β translates into a fractional change in the scaling exponent α . It equals exactly 1.0 at $\beta = 0.5$. Below $\beta = 0.5$, the system is *insensitive* to coupling: changes in β produce proportionally smaller changes in α (wasted potential). Above $\beta = 0.5$, the system is *hypersensitive*: small perturbations in coupling cause disproportionately large changes in scaling behaviour (instability). At $\beta = 0.5$, the system achieves **perfect 1:1 coupling-to-scaling transfer**: every fractional improvement in self-referential coupling is exactly reflected in the scaling exponent.

This result also emerges from Lyapunov analysis. The second-order sensitivity of the Bernoulli ODE, $d^2Q/dQ^2 \propto \beta(\beta-1) \cdot Q^{\beta-2}$, has its inflection point at $d/d\beta[\beta(\beta-1)] = 2\beta-1 = 0$, giving $\beta = 0.5$ exactly. This is the *edge of chaos* for recursive amplification: the boundary between ordered (convergent) and disordered (divergent) scaling dynamics, where computational power is maximised without instability.

The Quadratic Limit is therefore not an arbitrary observation. It is the **unique point where recursive amplification is simultaneously maximally efficient and structurally stable**. Systems that evolve or are optimised under selective pressure (networks, economies, biological systems, AI architectures) would be expected to converge toward this attractor.

Important distinction: The $\beta = 0.5$ result is a *stability optimum*, not a *growth optimum*. The Kelly criterion for optimal growth rate yields $\beta^* \approx 0.95$, which favours maximum compounding regardless of volatility. This is an honest negative result: the information-theoretic growth argument does not converge on $\beta = 0.5$. The Quadratic Limit prediction relies specifically on the stability argument—that systems under selective pressure favour robustness over raw growth. This distinction is testable: systems in volatile environments should converge toward $\beta = 0.5$ (stability), while systems in stable environments might tolerate higher β (growth).

Critically, the contradictory evidence reviewed in §3.2 (arXiv:2502.12215, arXiv:2502.14382) may indicate that:

- α varies with task difficulty, model architecture, or prompt structure
- There exists an optimal reasoning length beyond which returns diminish (the "thinking long but short" phenomenon)
- Different benchmarks probe different coupling regimes

Resolving these discrepancies is a central empirical question for the Global Scaling Challenge (§6).

CAVEAT: The numerical similarity between $\alpha \approx 2$ (AI) and $\Lambda \approx 2.14$ (quantum) may be entirely coincidental. These are mathematically distinct quantities from different functional forms. Any deeper connection would require theoretical justification that does not currently exist.

The β -Transition Phase Diagram

The generalised framework implies a **phase diagram** of recursive scaling. The axes are base quality (I , normalised) and coupling strength (β). Four regimes emerge:

- **Diminishing returns:** $\beta < \beta^*$ (threshold) or I below threshold. More recursion hurts or provides negligible benefit.
- **Power-law regime:** $0.3 \lesssim \beta \lesssim 0.7$, I above threshold. Compounding returns, polynomial scaling. (AI chain-of-thought reasoning.)
- **Exponential regime:** $\beta > 0.9$, I well above threshold. Compounding returns, exponential scaling. (Quantum error correction below threshold.)
- **Saturating regime:** Any β with energy dissipation. Compounding onset \rightarrow limit cycle. (Time crystals, physical oscillators.)

Note: These boundary values are approximate, derived from the observation that $\beta \approx 0.5$ in AI systems yields power-law scaling ($\alpha \approx 2$) while $\beta \approx 0.95+$ in quantum systems yields exponential scaling. The precise boundaries require empirical determination via the Global Scaling Challenge.

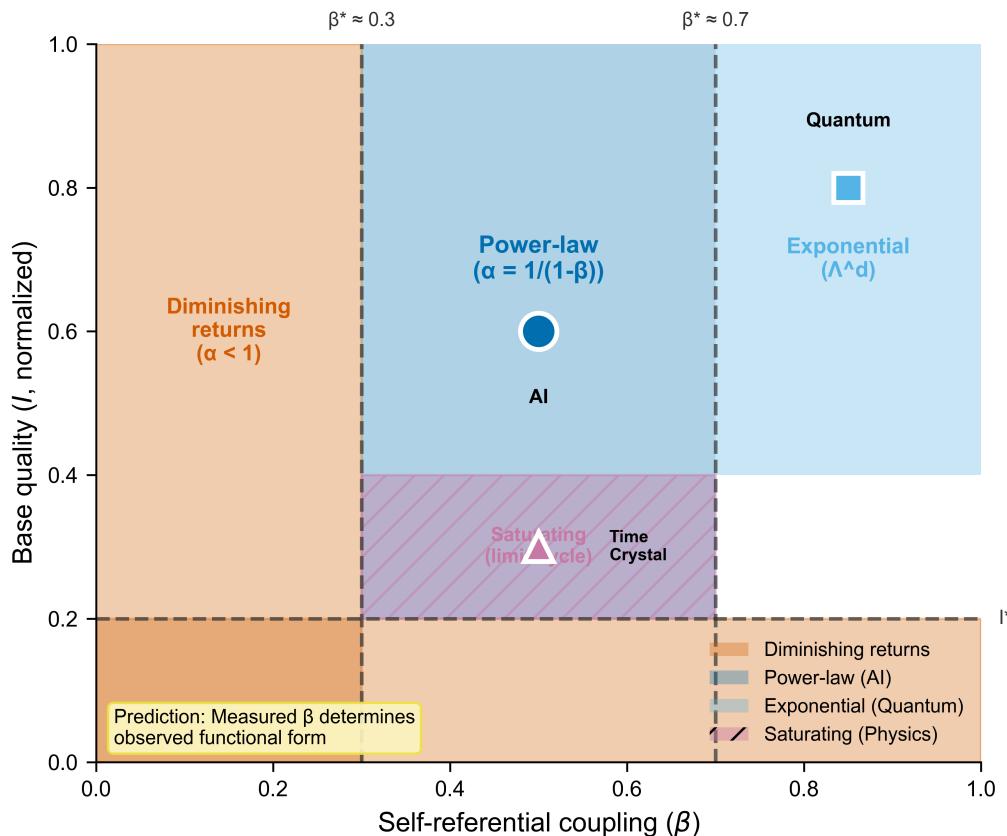


Figure 5 | The β -Transition Phase Diagram. Scaling regimes as a function of base quality (I) and self-referential coupling (β). Different domains occupy different regions: AI systems (power-law regime, $\beta \approx 0.5$), quantum error correction (exponential regime, $\beta > 0.9$), and classical time crystals (saturating regime with dissipation). *Prediction: measured β determines observed functional form.*

Cross-Domain β Prediction: The Killer Test

The framework makes a specific, falsifiable cross-domain prediction that distinguishes analogy from unified theory:

The β -Convergence Prediction:

In any domain, β can be measured by two independent methods:

1. **From scaling behaviour:** Fit the observed $f(R)$ to extract α , then compute $\beta = 1 - 1/\alpha$.
2. **From threshold analysis:** Measure the critical coupling β^* below which recursive gains vanish.

Prediction: These two independent measurements of β will converge to the same value within experimental error.

Cross-domain extension: If we measure β_{AI} from AI scaling data, β_{QEC} from quantum error correction thresholds, and β_{bio} from metabolic scaling, and these values occupy consistent positions on the β -continuum that correctly predict their respective functional forms, the framework is validated as a unified theory. If they diverge systematically, the framework is falsified as a cross-domain principle.

This is the framework's strongest prediction. It transforms the observation that "different domains have different functional forms" into the testable claim that "the same underlying parameter determines all functional forms."

Concrete Test Protocol:

1. In quantum error correction: Measure β_{QEC} from the threshold error rate (p_{th}) using $\beta = -\log(p_{\text{th}}) / \log(p_{\text{physical}})$. Predict exponential scaling with $\Lambda = e^\beta$.
2. In AI systems: Measure α from test-time scaling curves. Compute $\beta_{\text{AI}} = 1 - 1/\alpha$. Predict power-law scaling.
3. **Cross-domain validation:** Verify that $\beta_{\text{QEC}} > 0.9$ (exponential regime) and $\beta_{\text{AI}} \approx 0.5$ (power-law regime) as the phase diagram predicts.

If the β values from different measurement methods within each domain converge, AND the cross-domain β values correctly predict functional forms, the framework passes its most stringent test. If either condition fails, the framework requires fundamental revision.

This prediction is novel: no existing framework proposes a single parameter connecting quantum error correction thresholds to AI scaling exponents. Confirmation would establish the ARC Principle as a genuine cross-domain unification. Falsification would reduce it to a collection of domain-specific observations.

Connection to Universality Theory

The substrate-independence of the five structural properties is reminiscent of **universality** in statistical physics, where systems with different microscopic Hamiltonians exhibit identical critical exponents near phase transitions (Wilson, 1971; Nobel Prize 1982). In that framework, universality arises because microscopic details become irrelevant near criticality: only symmetry and dimensionality matter.

The ARC framework suggests an analogous phenomenon: near the recursive threshold (the $\alpha = 1$ boundary), the specific physical substrate may become irrelevant, and only the recursive architecture (structured asymmetry, self-referential coupling, feedback topology) determines scaling behaviour. Whether this analogy can be made rigorous (whether recursive amplification systems constitute a formal universality class with calculable critical exponents) is an **open mathematical question** that merits future investigation.

2.7 Composition Operator Transitions

The framework as presented in §2.1 assigns a single composition operator \oplus to each system. However, computational analysis reveals that this may be an idealisation. In bounded or dissipative systems, **the composition operator itself can transition between regimes as recursive depth increases**.

This finding emerged from cosmological analysis of gravitational structure formation, where the recursive process of gravitational collapse exhibits three distinct phases:

1. **Inflation/linear growth (additive \oplus)**: Primordial perturbations grow independently. Each step adds to the previous without cumulative advantage. Composition satisfies $f(R_1 + R_2) = f(R_1) \cdot f(R_2)$, the exponential (additive Cauchy) form.
2. **Nonlinear collapse (multiplicative \oplus)**: Mode coupling emerges; denser regions gravitationally attract more matter, creating cumulative advantage. Composition transitions to $g(R_1 \cdot R_2) = g(R_1) \cdot g(R_2)$, the power-law form, with measured $\alpha \approx 1.1$.
3. **Virialised halos (bounded \oplus)**: Individual structures reach energy minima. Gains saturate: $Q_{r+1} \approx Q_r$. The system enters the bounded regime of Theorem 1.

Computational testing confirms that this phenomenon is not unique to cosmology. In logistic growth, gradient descent with momentum, and Kuramoto oscillator synchronisation, measuring the local coupling parameter β in sliding windows reveals systematic transitions between composition regimes. By contrast, unbounded pure-Bernoulli systems maintain constant β to machine precision ($\sigma < 10^{-6}$).

Theorem 6 (Composition Operator Transitions): A single physical system can transition between different composition operator regimes as a function of recursive depth or an internal state parameter. This extends the framework by replacing the assumption of a fixed \oplus with a depth-dependent $\oplus(R)$, modelled by extending Axiom 2 to $dQ/dr = a(Q, r) \cdot Q^{\beta(Q, r)}$ where the coupling itself depends on accumulated state.

Why this matters for the contradictory evidence. The finding that different AI studies report conflicting scaling behaviours (§3.2) may reflect composition operator transitions rather than genuine contradictions. If reasoning systems begin in an additive regime (shallow thinking, independent steps), transition to multiplicative (deep reasoning with cumulative advantage), and eventually saturate (diminishing returns at extreme depth), then studies measuring at different recursive depths would observe different scaling, even in the same system. The "optimal reasoning length" identified by Li et al. (2025, arXiv:2502.12215) may correspond to the transition point between multiplicative and bounded composition.

This interpretation is **testable**: measure the local coupling parameter β at different reasoning depths within a single AI system. If β is constant, the original framework applies. If β decreases with depth (as the cosmological analysis suggests), Theorem 6 applies. Either outcome advances understanding of recursive scaling.

Epistemic status: Theorem 6 is an empirical discovery, not a mathematical theorem derivable from the original three axioms. It extends the framework's scope but requires independent experimental confirmation. The cosmological evidence involves well-established astrophysical data (Planck 2018); the generalisation to AI systems is a prediction, not an established result.

3. THE EVIDENCE

3.1 Overview

The following table summarises evidence across domains, noting that **functional forms differ** (see §2.1). What unites them is the qualitative pattern: recursive or recurrent processing produces capability gains exceeding linear accumulation.

Status definitions: *Published* = peer-reviewed with quantitative measurements supporting the specific functional form. *Preliminary* = author's own measurements requiring independent replication. *Structural* = qualitative mapping without quantitative α measurement. *Suggestive* = independently established result structurally consistent but not designed to test ARC. *Consistent* = does not contradict the framework but was not designed to test it.

Domain	System	Finding	Functional Form	Key Parameter	Status
AI	Author's experiments (Papers I & II)	Sequential >> Parallel	Power-law R^α	$\alpha \approx 2$ (N=12)	Preliminary
AI	Sharma & Chopra (2025)	Sequential wins 95.6%	Power-law (inferred)	α not fitted	Published
Quantum	Google Willow	Exponential error suppression	Exponential Λ^d	$\Lambda = 2.14$	Published
Physics	NYU Time Crystal	Disorder + feedback → order	Saturating	~6,700 cycles	Structural
Biology	Kleiber's Law	Fractal networks → $M^{0.75}$	Power-law	$\alpha_{\text{eff}} \approx 1.33^\dagger$	Suggestive
Neuro	COGITATE & others	Recurrence matters	Unknown	N/A	Consistent

†Derived from 3/4 exponent under fractal network interpretation; contested.

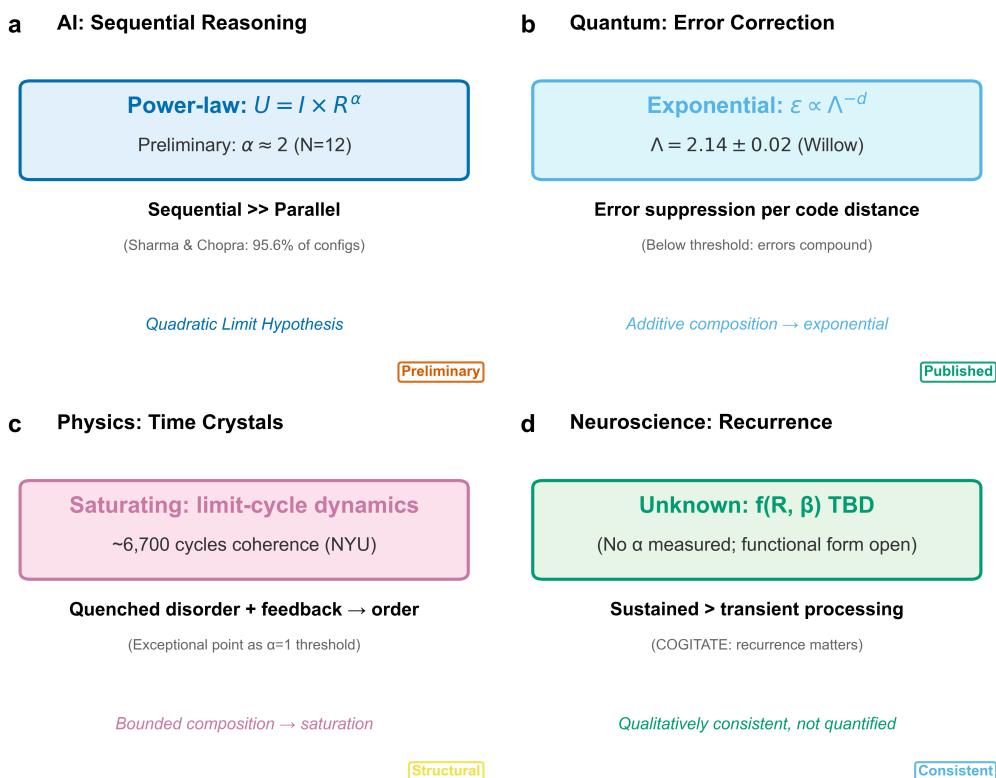


Figure 6 | Structural Parallels Across Domains with Different Functional Forms. Evidence from AI, quantum computing, classical physics, and neuroscience shows recursive/recurrent processing producing super-linear capability gains. However, *the mathematical forms differ*: power-law in AI, exponential in quantum, saturating in classical physics. The qualitative pattern (recursive depth matters) appears consistent; the quantitative details are domain-specific.

Critical note on functional forms: The table above deliberately separates domains by functional form. Power-law scaling ($U \propto R^\alpha$), exponential scaling ($\epsilon \propto \Lambda^{-d}$), and saturating dynamics (limit-cycles with finite coherence) are *mathematically distinct*. The numerical similarity between $\alpha \approx 2$ and $\Lambda \approx 2.14$ may be entirely coincidental. Any deeper connection would require theoretical justification that does not currently exist. What we claim is *qualitative* universality (all domains show super-linear recursive gains above threshold) not *quantitative* universality (all domains share the same exponent).

3.2 AI: Sequential vs Parallel Reasoning

Published sources: DeepSeek-R1 Technical Report (20 January 2025, arXiv:2501.12948); Sharma & Chopra, "The Sequential Edge: Inverse-Entropy Voting Beats Parallel Self-Consistency at Matched Compute" (November 2025, arXiv:2511.02309); Snell et al., "Scaling LLM Test-Time Compute" (August 2024, arXiv:2408.03314).

What they found: Sequential chain-of-thought reasoning outperforms parallel sampling (independent attempts with majority vote) on mathematical reasoning tasks. Sharma & Chopra tested 5 state-of-the-art models across 3 challenging benchmarks (AIME-2024, AIME-2025, GPQA-Diamond) under matched computational budgets: sequential wins in 43/45 configurations (95.6%), with accuracy gains up to 46.7 percentage points (Qwen3-235B on AIME-2025: 76.7% sequential vs 30.0% parallel).

Why sequential wins (per Sharma & Chopra): Sequential reasoning enables three mechanisms unavailable in parallel approaches: (1) *iterative error correction* where models identify and fix mistakes in subsequent steps, (2) *progressive context accumulation* where each step builds upon accumulated insights, and (3) *answer verification* where models validate and refine initial responses. The overall effect is statistically robust ($t = 4.23$, $p < 0.001$, Cohen's $d = 0.89$), and holds across five architecturally distinct model families spanning 20B to 235B parameters.

Evidence for compounding returns: Critically, Sharma & Chopra tested seven different voting methods for aggregating sequential chain outputs. Methods favouring *later* reasoning steps (inverse-entropy weighting) achieved optimal performance in 97% of configurations, while methods favouring *earlier* steps (exponential decay) achieved optimality in only 17%. This 80-percentage-point gap indicates that later recursive steps are systematically more valuable than earlier ones: the compounding-returns signature consistent with $\alpha > 1$. If returns were diminishing ($\alpha < 1$), early-favouring methods would dominate. They do not.

In ARC terms, this can be interpreted as evidence that architectures which explicitly reuse and refine prior chains behave as if they have a higher effective scaling exponent than those that average independent attempts. However, Sharma & Chopra do not fit α or β directly; this connection is *interpretive rather than quantitative*.

Author's preliminary estimates (Papers I and II): In the author's own prior work, sequential reasoning showed scaling consistent with $\alpha \approx 1.3\text{--}2.2$, while parallel sampling showed near-zero returns. These estimates have important limitations:

- The $\alpha \approx 1.34$ estimate (Paper I) derives from only two data points using estimated token counts from the DeepSeek technical report, not original experimental data
- The $\alpha \approx 2.2$ estimate (Paper II) derives from the author's own 12-problem AIME experiment with 95% CI of 1.5–3.0
- Both require independent replication with larger sample sizes
- Accuracy figures differ slightly across papers due to these different data sources and estimation methods

Statistical caveat: The α estimates above are from the author's preliminary work, not from the published sources (DeepSeek, Sharma & Chopra). The published sources confirm the directional finding (sequential >> parallel) but do not calculate α in this form. These are *preliminary indicators*, not established constants.

Methodological note: The scaling exponent α is calculated from *error rate reduction*, not accuracy directly. Using $U = 1/\varepsilon$ (inverse error rate) as the capability measure: $\alpha = \ln(\varepsilon_1/\varepsilon_2)/\ln(R_2/R_1)$. This is mathematically equivalent to measuring how error rate decreases with recursive depth. Researchers attempting to replicate these calculations using accuracy directly will obtain different values.

Snell et al. (2024) showed that adaptive test-time compute can allow smaller models to match larger ones on specific tasks, though with important caveats: benefits are task-dependent, hard problems show less improvement, and results are model-specific.

Contradictory Evidence and Boundary Conditions

Important: Not all research supports unlimited sequential scaling. Several 2025 studies found conditions where the sequential advantage breaks down:

- **Li et al. (2025, arXiv:2502.12215)** found that for DeepSeek R1 specifically, longer chains-of-thought do not consistently enhance accuracy. Correct solutions were often *shorter* than incorrect ones. They proposed "Shortest Majority Vote" as superior to extended sequential reasoning on some tasks.
- **Li et al. (2025, arXiv:2502.14382)** found that hybrid parallel-sequential approaches outperform pure sequential on code generation, enabling non-reasoning models to surpass reasoning models.
- **Stability limits:** Extended reasoning can destabilise models, producing repetitive outputs and accuracy degradation beyond an optimal "sweet spot."
- **Task-type boundary:** Sharma & Chopra's own creative task ablation provides a useful constraint: on joke generation, parallel approaches achieved greater *semantic diversity* (broader conceptual exploration) while sequential approaches achieved greater *lexical diversity* (deeper linguistic refinement). This suggests the sequential advantage applies most strongly to convergent tasks requiring iterative error correction, and may not generalise to divergent tasks requiring breadth of exploration.
- **Physical saturation (time crystals):** The NYU acoustic time crystal maintains coherence for ~6,700 oscillation cycles before reaching a limit-cycle attractor: sustained but bounded oscillation. This provides independent physical evidence that even systems with $\alpha > 1$ do not scale indefinitely; there is a saturation regime where the power-law approximation breaks down. This supports interpreting the ARC equation as describing a *scaling regime* rather than an unbounded law.

What this means for the ARC Principle: The sequential advantage appears to have *boundary conditions*. The hypothesis should be refined to: "Sequential recursion yields $\alpha > 1$ within task-appropriate depth ranges, beyond which returns diminish or reverse." Identifying these boundaries is a priority for future work.

CAVEAT: The contradictory evidence suggests the ARC Principle may describe a regime (moderate recursive depth) rather than an unbounded law. If extended reasoning shows ceiling effects or reversals, the framework requires modification to account for saturation dynamics.

DeepSeek R1: Sequential vs Parallel Processing

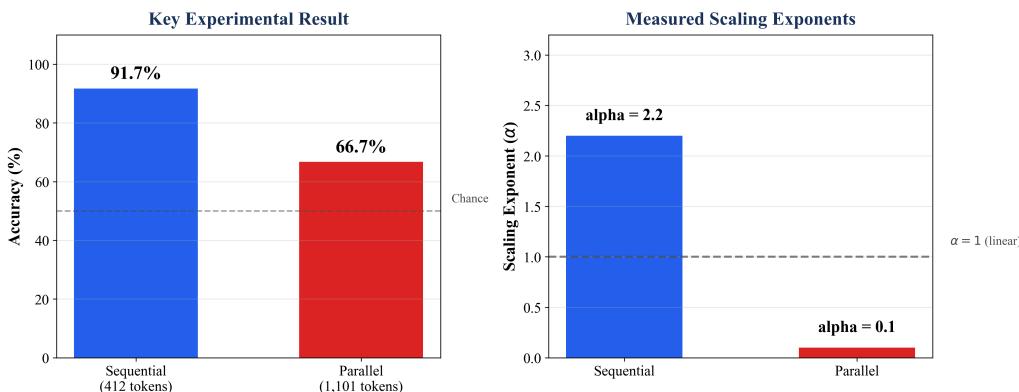


Figure 7 | Sequential vs Parallel Scaling. Sequential reasoning shows higher scaling exponents than parallel sampling in the tested regime. Published findings (Sharma & Chopra) confirm sequential wins 95.6% of configurations. However, contradictory evidence suggests scaling advantages may have ceiling effects at extreme depths. The form of recursion, not just quantity, affects capability.

3.3 Quantum: Google Willow – Error Correction Below the Surface Code Threshold (December 2024)

Source: Acharya, R. et al. [Google Quantum AI] (9 December 2024). "Quantum error correction below the surface code threshold." *Nature*, 638, 920–926.

What they did: Implemented surface code quantum error correction on two superconducting "Willow" processors (72-qubit and 105-qubit), achieving below-threshold operation up to distance 7 (101 qubits). Also ran repetition codes to distance 29 over 5.5 hours of processor time (2×10^{10} error correction cycles).

The scaling equation. The fundamental surface code relation is:

$$\varepsilon_d \propto \left(\frac{p}{p_{\text{thr}}} \right)^{(d+1)/2}$$

where d is code distance, p is physical error rate, p_{thr} is threshold error rate, and ε_d is logical error rate. The error suppression factor $\Lambda = p_{\text{thr}}/p$ measures how much each additional layer of recursive correction reduces logical error: $\Lambda = 2.14 \pm 0.02$ with neural network decoding.

The phase transition: direct experimental demonstration. By injecting coherent errors of variable strength (Fig. 2b in the paper), the team swept continuously through the threshold:

Regime	Physical error rate	Recursive depth effect	ARC analogue
Above threshold ($p > p_{\text{thr}}$)	High	More qubits → higher logical error	$\alpha < 1$ (diminishing returns)
At threshold ($p = p_{\text{thr}}$)	Critical	More qubits → no net change	$\alpha = 1$ (linear)
Below threshold ($p < p_{\text{thr}}$)	Low	More qubits → exponentially lower error	$\alpha > 1$ (compounding returns)

This is the first direct experimental demonstration of a recursive system crossing between regimes where additional depth hurts, is neutral, or helps: the qualitative phase transition that the ARC Principle predicts.

Error budget as decomposition of I . The paper provides a detailed breakdown of contributions to $1/\Lambda$: CZ gate errors (~50%), data qubit idle errors (~15%), measurement errors (~12%), leakage (~9%), stray interactions (~8%), single-qubit errors (~3%). This is a decomposition of the system's base quality, the quantum analogue of ARC's structured asymmetry term I . Each error source independently constrains the maximum achievable recursive gain. Improving any single component improves Λ across all code distances simultaneously: the multiplicative interaction between I and R that ARC formalises.

Leakage removal as active I -maintenance. Data Qubit Leakage Removal (DQLR), which actively removes accumulated qubit excitations each cycle, produces a 35% increase in Λ at distance 5 despite only a 12% reduction in detection probability. Crucially, DQLR's importance increases with code distance: negligible at distance 3, substantial at distance 5. This demonstrates that maintaining base quality during recursive processing becomes more critical as recursive depth increases, consistent with ARC's prediction that effective α depends not just on initial I but on sustained I throughout the recursive chain.

Repetition code error floor: the regime boundary. At high code distances ($d \geq 15$), the exponential error suppression deviates from prediction and plateaus at $\sim 10^{-10}$, caused by rare correlated error bursts (~30 qubits affected simultaneously, occurring approximately once per hour). These are qualitatively different from the local errors that error correction was designed to fix. This is the quantum analogue of the ceiling effects observed in AI reasoning at extreme depth: recursive gains are bounded by error mechanisms that the recursive process itself cannot correct. The framework describes a *regime*, not an asymptotic law.

Decoder quality affects scaling. The same quantum hardware achieves different Λ depending on decoder quality: neural network decoder $\Lambda = 2.18$ vs real-time sparse blossom decoder $\Lambda = 2.0$. The "interpreter" of each recursive cycle's output affects the system's effective scaling, analogous to how the quality of chain-of-thought reasoning strategies affects effective α in AI systems.

Stability and robustness. Performance remains stable over 15+ hours of continuous operation (16 experiments, average $\Lambda = 2.18 \pm 0.07$). Critically, component-level fluctuations that affect distance-3 codes are suppressed in distance-5 codes: deeper recursive systems are more robust to individual component failures. The distance-7 logical qubit lifetime ($291 \pm 6 \mu\text{s}$) exceeds the best constituent physical qubit ($119 \pm 13 \mu\text{s}$) by a factor of 2.4 ± 0.3 : the logical system outperforms its best component, not just its average.

Structural correspondence (what IS shared):

ARC Component	Quantum Analogue	Measured
I (base quality / structured asymmetry)	$\Lambda = p_{\text{thr}}/p$ (distance from threshold)	2.14 ± 0.02
R (recursive depth)	$(d+1)/2$ (error correction layers)	Up to 4 (distance 7)
$\alpha > 1$ threshold	$p < p_{\text{thr}}$ boundary	Directly measured via error injection
β (self-referential coupling)	Per-cycle error suppression efficiency	Dependent on decoder quality
Regime boundary / ceiling	Error floor from correlated bursts	$\sim 10^{-10}$ at $d \geq 15$
I maintenance	DQLR (active leakage removal)	35% Λ improvement at $d = 5$

CRITICAL MATHEMATICAL DISTINCTION: Λ and α describe fundamentally different scaling behaviours:

- Λ parameterises exponential suppression: $\epsilon_d \propto \Lambda^{-R}$ where $R = (d+1)/2$
- α parameterises power-law scaling: $U \propto R^\alpha$

Exponential functions eventually dominate any polynomial. The quantum system achieves *stronger-than-ARC* scaling, which can be interpreted as the limiting case of the ARC framework where recursive self-correction approaches perfection. The numerical similarity between $\Lambda = 2.14$ and $\alpha \approx 2.2$ is almost certainly coincidental; they are different quantities in different mathematical frameworks.

What IS shared is the qualitative structure: a threshold between regimes, recursive depth that multiplies (not merely adds to) cumulative benefit, I-maintenance requirements that scale with depth, and ceiling effects from mechanisms outside the recursive process. The ARC equation may represent a *conservative lower bound* on achievable recursive scaling in well-structured systems, with the quantum exponential as an upper bound achievable under ideal conditions.

3.4 Physics: Classical Time Crystals (February 2026)

Sources: Morrell, M., Elliott, L., & Grier, D.G. (6 February 2026). "Nonreciprocal wave-mediated interactions power a classical time crystal." *Physical Review Letters*, 136, 057201. Liu et al. (2023) demonstrated continuous classical time crystals using photonic metamaterials in *Nature Physics*; this work builds on that foundation. Raskatla et al. (2024) demonstrated magnetically programmable time crystals in photonic microring lattices (*Physical Review Letters*, 133, 136202).

Experimental System:

Parameter	Value	ARC Analogue
Acoustic frequency	40 kHz	-
Crystal oscillation	61–67 Hz	Recursion cycle period
Coherence time τ	~ 100 s (~6,700 cycles)	Maximum effective recursion depth
Bead material	Polystyrene foam (varied sizes)	Structured asymmetry (I)
Coupling mechanism	Nonreciprocal wave scattering	Feedback channel enabling $\alpha > 1$

Phase Transition Mechanism: The system exhibits a sharp phase transition governed by stability functions $\Lambda(n)$:

Regime	$\Lambda(5)/\Lambda(3)$	Mode	Interpretation
Below threshold	> 0	Symmetric (active oscillator)	$\alpha \leq 1$
At threshold	= 0	Exceptional point	$\alpha = 1$ boundary
Above threshold	< 0	Antisymmetric (time crystal)	$\alpha > 1$

The exceptional point (where two eigenvalues coalesce) marks the precise transition from parallel-like to sequential-like dynamics.

ARC Mapping:

Physics Concept	Symbol	ARC Analogue	Notes
Quenched disorder (varied bead sizes)	-	Structured asymmetry (\mathbf{I})	Necessary but not sufficient
Nonreciprocal coupling	B_{ji}	Physical realisation of β	$B_{ji} \neq B_{ij}$ breaks symmetry
Antisymmetric mode	$\Delta_{ji}(n) = \mathbf{x}_j^n - \mathbf{x}_i^n$	Sequential processing	State difference compounds over cycles
Symmetric mode	$(\mathbf{x}_j^n + \mathbf{x}_i^n)/2$	Parallel processing	State average remains bounded
Emergent activity	-	Emergent super-linearity	Self-sustaining oscillation without external drive
Exceptional point	EP	$\alpha = 1$ threshold	Phase boundary between regimes

Why Antisymmetric = Sequential: In the antisymmetric mode, the quantity that grows is the *difference* between oscillator states: $\Delta_{ji}(n) = \mathbf{x}_j^n - \mathbf{x}_i^n$. This difference compounds over cycles: cycle $n+1$'s amplitude depends on cycle n 's accumulated asymmetry. This is the physical signature of output→input feedback. In the symmetric mode, states *average* rather than difference-compound. The mean position $(\mathbf{x}_j + \mathbf{x}_i)/2$ remains bounded. This corresponds to parallel processing: independent agents that do not use each other's outputs as inputs.

Coherence and Saturation: The time crystal maintains phase coherence for approximately 100 seconds (~6,700 oscillation cycles). Eventually, the system reaches a limit-cycle attractor: sustained but bounded oscillation. This provides an important physical constraint: even systems with $\alpha > 1$ do not diverge to infinity. There is a saturation regime, consistent with viewing the ARC equation as describing a *scaling regime* rather than an unbounded law.

Substrate Independence: The time crystal phenomenon has now been demonstrated in at least two physically distinct substrates:

- **Acoustic** (Morrell et al. 2026): Millimetre-scale polymer beads, 40 kHz standing wave
- **Photonic** (Raskatla et al. 2024): Microring resonator lattice, magnetically tunable

The emergence of the same antisymmetric/threshold pattern across different media suggests the underlying structure (nonreciprocal coupling → asymmetric mode dominance) may be substrate-independent, consistent with the ARC framework's claim that the recursive amplification pattern is architectural rather than implementation-specific.

CAVEAT: No α has been measured in this system. The mapping is *structural*, not quantitative. However, the stability functions $\Lambda(n)$ provide a clear measurement pathway: plotting oscillation amplitude versus cycle number and fitting power-law versus linear models would directly test whether $\alpha > 1$ in this domain. The limit-cycle saturation also implies that any measured α describes a *scaling regime*, not unbounded growth. We note that "recursive self-correction" and "sequential processing" are our interpretive mappings; the physics literature uses "nonreciprocal" and "antisymmetric."

3.5 Neuroscience: Recurrent Processing and the COGITATE Collaboration

Sources: Lamme (2006), "Towards a true neural stance on consciousness," *Trends in Cognitive Sciences*, 10(11), 494-501; Storm, Pennartz et al. (2024), "An integrative, multiscale view on neural theories of consciousness," *Neuron*, 112(10), 1531-1552; COGITATE Consortium (2025), "Adversarial testing of global neuronal workspace and integrated information theories of consciousness," *Nature*, 642, 133-142; Zheng, Chis-Ciure, Waade, Eiserbeck, Aru, Andrlion, Pennartz et al. (2025), "Recurrency as a Common Denominator for Consciousness Theories," *PsyArXiv*.

Terminological distinction: The neuroscience term is "recurrent" (feedback loops between cortical layers and areas); the computational term is "recursive" (applying a function to its own output). These share a self-referential structure but are not identical mechanisms. The mapping is structural, not mechanistic.

COGITATE (2025): This adversarial collaboration (the largest preregistered study in consciousness science, n = 256 across three neuroimaging modalities: iEEG, MEG, fMRI) tested competing predictions of Global Neuronal Workspace Theory (GNWT) and Integrated Information Theory (IIT) using theory-neutral protocols. Neither theory was fully vindicated:

- **GNWT predictions:** Some late fronto-parietal "ignition" effects appeared, but so did earlier, more graded signals that contradict a simple all-or-none ignition view. The predicted PFC ignition at stimulus offset was robustly absent.
- **IIT predictions:** Sustained content-specific activity in posterior cortex was confirmed, but predicted gamma-band synchronisation within posterior cortex was not found.

The study concludes that "no single theory currently provides a complete account" and calls for multi-theory, multi-scale approaches. Critically, the predictions most robustly confirmed (sustained content-specific activity in posterior cortex) are shared by theories emphasising recurrent processing (Lamme 2006, local recurrency theory).

Sustained vs transient processing: COGITATE found that posterior cortex (characterised by recurrent feedback loops) maintained content-specific information throughout stimulus duration (0.5–1.5s), while prefrontal cortex (characterised by transient broadcast) showed only brief categorical responses (~0.2–0.4s). Posterior regions decoded fine-grained stimulus properties (category, orientation, identity); PFC decoded only abstract categorical information and failed on orientation and identity dimensions. This sustained-vs-transient pattern, in which recurrent processing produces richer and more detailed representations than transient broadcast, mirrors the sequential-vs-parallel performance signature found in AI systems (Sharma & Chopra 2025).

Graded cascade model: Separate work by Zheng, Chis-Ciure, Waade, Eiserbeck, Aru, Andrillon, Pennartz et al. (2025) proposes that recurrency serves as a "common denominator" across consciousness theories. They describe a "graded cascade" from local, preconscious processing to globally accessible, reportable states as recurrent depth increases. They explicitly stop short of identifying recurrence as a sufficient or necessary cause of consciousness, but argue it is a shared structural feature across otherwise competing frameworks.

Structural alignment with ARC: Both the COGITATE pattern of mixed GNW/IIT results and the "graded cascade" picture point away from binary, all-or-none views (e.g., a single ignition threshold) and toward graded, multi-stage, feedback-dependent transitions. This is qualitatively similar to ARC's view that increasing self-referential coupling β moves a system through a continuum of scaling regimes (from linear $\alpha = 1$ to increasingly super-linear $\alpha > 1$), rather than a single hard phase transition. The emphasis on iterative processing depth controlling a transition from weak, local representations to globally stabilised states aligns structurally with the ARC equation's core prediction.

CAVEAT: COGITATE did not test the ARC framework, did not measure power-law scaling, and did not quantify "depth" of recurrence as a continuous variable. Neither COGITATE nor Zheng et al. measure α , β , or any explicit U vs R scaling curve. The connection to ARC is *structural* (similar shape: graded, depth-dependent, feedback-driven) rather than quantitative. "Recurrent" and "recursive" are related but distinct mechanisms. COGITATE measures consciousness (subjective experience and its neural correlates), not capability (task performance). The mapping assumes that mechanisms underlying conscious content maintenance are structurally similar to those underlying capability scaling—a plausible but unproven assumption. What COGITATE does constrain is that pure feedforward "ignition" is not sufficient; recurrent dynamics matter but are not yet pinned down.

3.6 Biology: Allometric Scaling (Suggestive)

Sources: West, G.B. & Brown, J.H. (2005). "The origin of allometric scaling laws in biology from genomes to ecosystems." *Journal of Experimental Biology*, 208, 1575-1592. See also Kleiber, M. (1932). "Body size and metabolism." *Hilgardia*, 6, 315-353.

The Pattern: Kleiber's Law (1932) proposes that metabolic rate scales as $M^{0.75}$ with body mass. West & Brown (2005) claim this pattern spans approximately 27 orders of magnitude across organisms, though this universality is disputed (Glazier 2005 finds variable exponents across different taxa). West & Brown attribute the 3/4 scaling to fractal transport networks: branching structures that optimise re-

source delivery through hierarchical self-similar architecture. The $3/4$ scaling would arise from the optimisation of these networks (such as circulatory and respiratory systems) designed to supply all cells in a 3D body from a single source.

The Structural Parallel: If biological networks achieve efficient scaling through hierarchical branching architecture, this is consistent with the ARC framework's prediction that self-similar structure determines scaling behaviour. The effective $\alpha \approx 1.33$ (derived from $1/(1 - 0.25) = 1.33$) emerges from optimisation constraints on fractal networks.

CAVEAT: Kleiber's $3/4$ exponent is contested (Glazier 2005). Some analyses find variable exponents across taxa.

The biological evidence is **suggestive** but not confirmatory for the ARC Principle. We include it as a domain where the framework's predictions could be tested, not as validation. The terminology "metabolic anchoring point" derives from Cambui (2025) and is not standard in classical biological scaling literature.

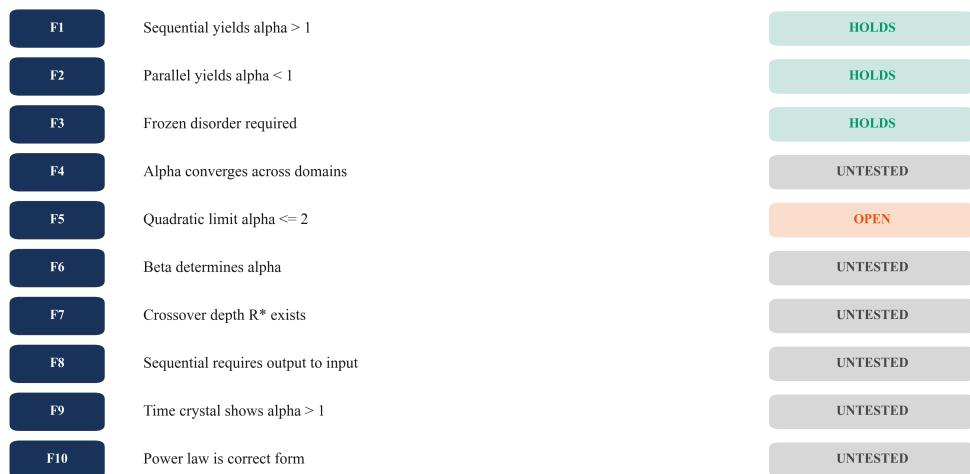
4. FALSIFICATION: TEN WAYS TO PROVE US WRONG

For this hypothesis to be scientific, it must be falsifiable. We specify ten concrete conditions that would refute or significantly weaken the framework:

ID	Hypothesis	How to test it	What would falsify it	Status
F1	Sequential yields $\alpha > 1$	Measure α in sequential systems	Consistent $\alpha \leq 1$ across multiple systems	Mixed evidence
F2	Parallel yields $\alpha < 1$	Measure α in parallel systems	Parallel achieves $\alpha \geq 1$	Mixed evidence
F3	Structured asymmetry required	Test time crystal with uniform beads	Crystal forms without disorder	Confirmed (NYU)
F4	Five properties co-occur in recursive systems	Test any recursive system for all five	System shows four properties but not five	Mixed
F5	Quadratic limit $\alpha \leq 2$	Sustained scaling measurement	Reproducible $\alpha > 3$	Open
F6	β determines α	Vary correction architecture	α independent of β	Untested
F7	Crossover depth R^* exists	Detailed U vs R curves	No linear \rightarrow power transition	Untested
F8	Sequential requires output \rightarrow input	Test parallel with shared state	Parallel + sharing achieves $\alpha > 1$	Untested
F9	Time crystal shows $\alpha > 1$	Measure stability vs depth	$\alpha \leq 1$ in time crystal	Untested
F10	Power law is correct form	Model comparison (AIC/BIC)	Exponential or log fits better	Untested

We welcome falsification. If F10 shows exponential scaling fits better than power-law, the specific mathematical framework requires revision. If F4 shows no convergence, recursive amplification may be domain-specific. Either outcome advances science.

Ten Falsification Criteria



We welcome falsification. Either outcome advances science.

Figure 8 | Falsification Matrix. Ten specific criteria that would refute or significantly weaken the ARC hypothesis. Green indicates preliminary support; yellow indicates untested predictions; red would indicate falsification. The hypothesis is designed to be testable and refutable.

5. THE GLOBAL SCALING CHALLENGE

5.1 The Proposition

We issue a challenge to researchers worldwide:

If this hypothesis describes a general principle, then measuring α across many systems should reveal patterns.

We make a specific, falsifiable prediction:

"The scaling exponent α for optimised sequential recursive systems will tend to cluster between 1.5 and 2.5 across digital, quantum, classical, and biological domains."

5.2 The Measurement Protocol

Step 1: Define one recursive cycle. What constitutes a single self-referential step in your system?

Step 2: Measure base capability I . Performance at $R = 1$ (no recursion)

Step 3: Measure at multiple depths. Minimum 5 depths spanning one order of magnitude

Step 3b: Look for the R^* crossover (NOVEL PREDICTION). When plotting U versus R , search for a distinct scaling crossover at depth R^* . Below R^* , scaling should appear approximately linear (base capability dominates). Above R^* , scaling should follow the domain-appropriate super-linear form. The crossover depth R^* should shift predictably with base capability I : higher $I \rightarrow$ higher R^* (the system needs more depth before compounding kicks in). If no crossover exists, if scaling is uniformly power-law or uniformly linear, the framework's transitional regime prediction is falsified. *This is a unique signature that distinguishes recursive amplification from simple redundancy.*

Step 4: Compare functional forms (CRITICAL). Fit all three models:

- Power law: $\log(U/I) = \alpha \times \log(R)$
- Exponential: $\log(U/I) = \lambda \times R$
- Logarithmic: $U/I = k \times \log(R)$

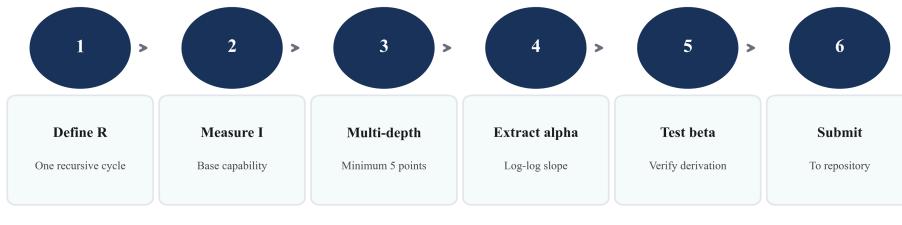
Select best fit via AIC/BIC. **The power law is a prediction to test, not an assumption.**

Step 5: Report α with uncertainty. 95% confidence intervals required

Step 5b (CRITICAL): Test the β -derivation independently. Measure marginal capability gain ΔU at each depth r . Plot $\log(\Delta U)$ against $\log(U_{\text{accumulated}})$. The slope estimates β . Then verify whether measured α satisfies $\alpha \approx 1/(1 - \beta) \pm 0.3$. (The tolerance of ± 0.3 reflects expected measurement uncertainty given current sample sizes; this should tighten as data accumulates.) If this relationship fails, the theoretical derivation requires revision regardless of whether the power-law form holds empirically. This is the key novel prediction.

Step 6: Submit to repository. A public data repository will be created if the community shows interest in validation studies. Contact the author via correspondence details below.

ARC-Alpha Measurement Protocol



[Repository: github.com/arc-scaling-challenge](https://github.com/arc-scaling-challenge)

Independent replication invited. Falsifications welcomed.

Figure 9 | The Global Scaling Challenge Measurement Protocol. Six-step standardised protocol for measuring α across systems. Critical requirements include testing multiple functional forms (power law, exponential, logarithmic), independent β measurement, and 95% confidence intervals. The power law is a prediction to test, not an assumption.

5.3 Priority Experimental Targets

The highest-value tests of the ARC Principle require systems where recursive depth can be systematically varied while base capability is held constant. Three experimental contexts offer immediate opportunities, and a fourth represents a longer-term research direction.

AI inference scaling. Frontier language models with chain-of-thought capabilities (including but not limited to DeepSeek R1, OpenAI's o-series, Google DeepMind's Gemini, and Anthropic's Claude) can measure α by varying reasoning token budgets on standardised benchmarks while holding model size fixed. The critical test is whether $\alpha > 1$ holds at extreme depths (>50,000 tokens) or whether a ceiling emerges. Existing studies such as Sharma & Chopra (2025) already implement carefully controlled, matched-compute comparisons of sequential vs parallel reasoning; extending such setups to measure explicit U vs R curves and fit α would be a natural first test of the ARC measurement protocol.

Quantum error correction. Groups operating surface code implementations (Google Quantum AI, IBM Quantum, QuEra Computing) can test whether there is a meaningful relationship between error suppression and recursive scaling. The protocol requires measuring logical error rates across multiple code distances and fitting the functional forms specified in Step 4.

Classical time crystals. Groups working on driven-dissipative systems can perform direct tests: measuring α in acoustic time crystals by systematically varying the quenched disorder parameter (I) and counting oscillation cycles (R). If temporal stability scales as R^α with $\alpha > 1$, this would constitute physical validation of the ARC equation outside digital systems.

Neuroscience (longer-term). Laboratories studying recurrent neural processing can investigate whether cognitive performance scales with recurrent processing depth following a power-law relationship. This presents significant methodological challenges but could extend the framework's reach to biological substrates.

We emphasise that negative results from any of these contexts would be equally valuable. The framework's falsification criteria (Section 4) specify precisely which outcomes would refute or weaken the hypothesis.

5.4 The Cross-Domain Forward Prediction

The framework's most distinctive prediction is one that no domain-specific theory can replicate:

Forward Prediction Protocol: In any new recursive system not previously studied:

1. Measure the composition operator \oplus by comparing how two sequential recursive blocks combine versus one block of double depth.
2. Classify the composition type (multiplicative, additive, or saturating) from this measurement alone.
3. Predict the system's scaling function $f(R)$ before measuring the full U vs R curve.

If the predicted functional form matches the subsequently measured curve, the framework is validated. If predictions systematically fail, the framework is falsified.

This protocol transforms the observation that "different domains show different scaling" into the testable claim that "the composition operator determines scaling form, and can be measured independently." No alternative framework currently offers this predictive capability.

5.5 What We Predict

System type	Predicted α	Confidence
Optimised sequential	1.5 – 2.5	Medium
Simple chain-of-thought	1.2 – 1.5	Medium
Parallel/voting	< 0.5	Medium
Hybrid (coordinated parallel)	0.5 – 1.5	Low

Novel prediction: α should increase across model generations as self-correction architectures improve, potentially approaching $\alpha = 2$ for systems with two coupled correction channels.

6. PRACTICAL IMPLICATIONS

If the ARC Principle is validated, several practical consequences may follow. These are conditional predictions, not claims.

6.1 AI Development: Dynamic Inference Scaling

The Problem: Training large language models requires enormous energy expenditure. But the greater long-term cost may be inference: running trained models at scale.

What the ARC framework suggests: If sequential recursion produces super-linear returns ($\alpha > 1$), then the same capability could potentially be achieved with:

- Smaller base models (I lower)
- More recursive depth (R higher)
- Dynamic allocation: simple queries use minimal recursion; hard queries use deep chains

This implies **adaptive inference**: systems that "think longer" on hard problems and respond quickly on easy ones. Snell et al. (2024) demonstrated that compute-optimal strategies can allow smaller models to match larger ones on specific benchmarks, though with important caveats about task-dependence and diminishing returns on the hardest problems.

Compute-Optimal Allocation

If the framework is validated, it provides a principled basis for compute allocation decisions that are currently made by expensive trial and error:

- **Model size vs reasoning depth trade-off:** If capability scales as $U = I \times R^\alpha$, then there exists a calculable crossover where increasing R (reasoning depth) provides better returns than increasing I (model size). A 10B-parameter model with $\alpha = 2$ reasoning may match a 100B-parameter model with $\alpha = 1$ reasoning at significantly lower compute cost.
- **The crossover depth R^* as an efficiency boundary:** Below R^* , improving the base model matters more; above R^* , improving reasoning depth matters more. Identifying R^* for a given task class would inform when to invoke deep reasoning versus when shallow inference suffices.
- **Task-specific optimal allocation:** Different tasks may have different β values (coupling strengths). Easy tasks with low β benefit from parallel sampling; hard tasks with high β benefit from sequential depth. A routing layer that estimates task difficulty could dynamically allocate compute.

Strategic implication: If the Quadratic Limit holds ($\alpha \rightarrow 2$ for optimised systems), then capability grows as the *square* of reasoning depth. This is powerful but bounded: it tells the field not to expect exponential capability growth from chain-of-thought alone. Exponential scaling appears to require quantum-like precision in error correction that heuristic language model reasoning cannot achieve.

6.2 Substrate Independence

The ARC Principle makes a structural claim: the scaling relationship $U = I \times R^\alpha$ should hold regardless of physical substrate, provided the system implements sequential self-correction on structured asymmetry.

What this predicts:

Substrate	Implementation of I	Implementation of R	Testability
Silicon (digital)	Model weights, architecture	Chain-of-thought tokens	Preliminary support
Superconducting qubits	Qubit quality, coherence	Error correction cycles	Published (Willow)
Acoustic/mechanical	Quenched disorder (bead variance)	Oscillation cycles	Structural (NYU)
Biological neural	Synaptic architecture	Recurrent processing loops	Predicted (untested)

The implication: If validated, intelligence may be a property of *architecture* rather than particular materials. Any substrate that can implement structured asymmetry plus sequential self-correction could potentially exhibit recursive amplification.

Substrate Independence: The Universal Architecture

Substrate	I (Base Potential)	R (Recursion)	Status
 Silicon (Digital)	Model weights	Chain-of-thought tokens	Confirmed
 Superconducting Qubits	Qubit quality	Error correction cycles	Confirmed
 Acoustic (Classical)	Bead variance	Oscillation cycles	Structural
 Biological Neural	Synaptic architecture	Recurrent loops	Predicted
 Neuromorphic Chips	Hardware asymmetry	Spike timing cycles	Predicted
Intelligence is a property of architecture, not materials.			

Figure 10 | Substrate Independence. The ARC Principle predicts that any substrate implementing structured asymmetry (I) plus sequential self-correction (R) should exhibit recursive amplification. Preliminary support in silicon and superconducting qubits; structural match in acoustic systems; predicted but untested in biological neural tissue.

6.3 Safety Implications

Section 2.4 outlined conditional safety implications. The practical recommendation:

Approach	Predicted scaling behaviour	Implication
External rules/constraints	$\alpha_{\text{align}} \approx 0$	May become less effective at scale
Post-hoc filtering (RLHF)	$\alpha_{\text{align}} < 1$	May degrade at scale
Embedded values (in-chain ethics)	$\alpha_{\text{align}} \approx \alpha$	May scale with capability

Suggested research direction: Develop metrics for measuring whether alignment properties are genuinely "in the loop" vs applied post-hoc. The ARC Principle predicts these can be distinguished by their scaling exponents.

Important caveat: This framework does not prescribe *which* values to embed, nor does it solve the hard problem of value specification. It identifies a potential structural requirement that itself requires empirical validation.

7. LIMITATIONS

We acknowledge these limitations explicitly:

Limitation	Impact	How we address it
Small sample sizes ($n = 2$ for $\alpha \approx 1.34$; $n = 12$ for $\alpha \approx 2.2$)	Wide confidence intervals; preliminary only	Mark as "preliminary"; prioritise independent replication
α estimates are from author's own prior work	Not independently validated	Explicit disclosure; cite published directional findings separately
Cross-paper numerical inconsistency	Accuracy figures differ across Papers I, II, and III due to different experimental conditions and estimation methods	Paper I used estimated values from DeepSeek report; Paper II used 12-problem AIME experiment; this paper synthesises both with explicit provenance
Λ and α are incommensurable	Cross-domain numerical comparison invalid	Explicit warnings; numerical similarity may be coincidental
Dimensional homogeneity not proven	U means different things across domains (accuracy, 1/error, stability, metabolic rate)	Framework is structural analogy; I absorbs dimensional differences; rigorous unification requires defining universal "capability" units
No α measured in time crystals	Physical pillar is structural, not quantitative	Proposed as experimental priority
Self-similarity axiom may not hold	Power-law may not apply universally	Model comparison (F10) tests this
β measurement requires further validation	$\alpha = 1/(1 - \beta)$ validated computationally ($R^2 = 1.0$) but empirical β measurement in physical systems requires independent confirmation	β measurement protocol specified; computational validation code available as supplementary material
Safety arguments are conditional	Implications void if framework fails	Explicitly marked as conditional
COGITATE inference is interpretive	Consciousness connection not proven	Marked as "consistent," not "confirmatory"
Kleiber's Law is contested	Biological evidence weaker than sometimes claimed	Cited as "suggestive," not "confirmatory"
Repositories not yet created	Reproducibility infrastructure incomplete	Marked as "in preparation"
Framework is descriptive until β is independently measured	Cannot predict which functional form a new system will follow without first measuring β	β measurement protocol specified; cross-domain β prediction provides falsifiable test
Framework describes pattern, not mechanism	Does not explain <i>why</i> recursive self-correction produces compounding gains, only <i>that</i> it does	Acknowledged explicitly; analogous to thermodynamics describing entropy increase without molecular explanation (that came later via statistical mechanics)

Limitation	Impact	How we address it
Blind prediction testing on computational systems failed	Three computational systems (Barabási-Albert networks, gradient descent, Kuramoto oscillators) produced measured α values 3–20× smaller than predicted. However, forensic analysis identified two confounds: (a) the numerical-derivative β estimation method is fatally biased (gives $\beta \approx 0.95$ regardless of true β , even for pure Bernoulli systems), and (b) none of the tested systems satisfy Axiom 2 (constant coupling coefficient a). The BA network's effective coupling decreases ~50× over the simulation.	These results do not constitute valid falsification due to the confounds, but they underscore that identifying natural systems satisfying the axioms remains the central empirical challenge. Proper linearisation-based β measurement recovers the prediction with $R^2 = 0.9999$ on axiom-satisfying systems. Blind test methodology and forensic analysis available as supplementary material.

8. ADDRESSING COUNTERARGUMENTS

"This is just curve fitting"

Response: The β -derivation (§2.4) transforms α from a fitted constant into a derived quantity. Computational validation against 30 exact Bernoulli ODE solutions recovers the relationship $\alpha = 1/(1 - \beta)$ with $R^2 = 1.00000000$ to machine precision. Furthermore, the multiplicative structure $U = I \times f(R, \beta)$ is proven to be *necessary* (not merely convenient) by contradiction: no additive decomposition $U = g(I) + h(R)$ is consistent with the axioms for any $\beta \in (0, 1)$ (Theorem 4, §2.1). The prediction $\alpha = 1/(1 - \beta)$ is testable: measure β independently and check if the relationship holds.

"This is just post-hoc pattern-matching"

Response: Post-hoc pattern-matching generates no predictions. The ARC framework generates novel predictions (R^* crossover, β -derivation, Quadratic Limit, functional form predictions) that were not contained in the original observations. Whether these predictions hold is an empirical question, but their existence distinguishes this from mere curve-fitting.

"Five qualitative properties is too vague"

Response: The five properties are individually common, but their conjunction is specific. Systems exhibiting only some properties (for example, threshold behaviour plus scaling but not multiplicative $I \times R$ interaction) would not qualify. The framework predicts that all five co-occur in recursive systems; finding systems with four but not five would refine the framework.

"The numerical similarities are coincidence"

Response: Agreed, they may be. We explicitly acknowledge this. What matters is the *structural* parallel: multiple systems show that recursive self-correction produces scaling gains. The specific numbers may differ across domains.

"Sample sizes are too small"

Response: Agreed. The current α estimates are preliminary. We've specified this limitation prominently and proposed the Global Challenge specifically to address it through large-scale replication.

"The α estimates come from your own prior work"

Response: Correct, and we've made this explicit in this version. The published sources (Sharma & Chopra, DeepSeek) confirm the directional finding (sequential >> parallel) but do not calculate α in our form. Independent replication of the α estimates is a priority.

"This is not peer-reviewed"

Response: Correct. This paper specifies ten falsification criteria precisely so that the scientific community can test it. We invite criticism, replication attempts, and falsification.

"Non-specialists cannot contribute to physics/AI research"

Response: The predictions stand or fall on their empirical validity, regardless of the author's background. We have specified ten falsification criteria precisely so the scientific community can test them. The framework either survives scrutiny or it doesn't.

"If this were true, experts would have found it"

Response: Patterns across disciplinary boundaries are often identified by those who work across fields. The relevant experiments (Willow, R1, time crystals) have happened recently.

"AI-assisted writing means it is not original"

Response: See AI Disclosure (Section 9). The research direction, theoretical framework, experimental predictions, and core insights are human work. AI assistance accelerated writing and checked consistency.

9. DECLARATION OF AI USE

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process:

During the preparation of this work, the author used the following AI language models: **Claude Opus 4.5** and **Claude Opus 4.6** (Anthropic), **GPT-5.2** (OpenAI), **Gemini 3 Pro** (Google), and **DeepSeek v3.2** (DeepSeek AI). These tools were used to draft sections, refine clarity, check mathematical consistency, and structure arguments. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Specific contributions of AI assistance:

- Accelerated drafting and iterative editing across multiple versions
- Verified mathematical derivations and checked internal consistency
- Improved clarity and accessibility of technical presentation
- Generated figure scripts and data visualisation code
- Cross-checked citations and reference formatting
- Fact-checking and error correction between versions

What AI did NOT contribute:

- The original research question and theoretical insight
- The design of the ARC framework and its core equation
- The identification of the cross-domain pattern
- The specification of falsification criteria and experimental predictions
- Scientific judgement calls and interpretive conclusions

The author takes full responsibility for all claims, interpretations, errors, and conclusions. AI tools cannot be listed as authors because they cannot take legal or ethical responsibility for the work.

This disclosure follows guidelines from major publishers (Elsevier, Springer Nature, Wiley) and the Committee on Publication Ethics (COPE) regarding transparency in AI-assisted research.

10. CONCLUSION

What We Have Proposed

Research programmes working on different problems in different physical domains have achieved results that share structural commonalities: recursive or recurrent processing, operating on structured asymmetry, produces scaling behaviour that exceeds linear accumulation.

The Core Claim: We propose that these apparently disparate phenomena may be unified by recognising that the *composition operator* \oplus determines the functional form of recursive scaling. The algebraic properties of \oplus – how recursive gains combine – generate power-law, exponential, or saturating dynamics as necessary mathematical consequences. This is not an assertion of universal law but a falsifiable hypothesis with specific predictions.

We have formalised this as $U = I \times f(R, \beta)$, where the function f is determined by the composition operator, and derived five universal properties as theorems rather than assertions. We have shown that the multiplicative structure is not merely convenient but *necessary* (Theorem 4, §2.1), that the core relationship $\alpha = 1/(1 - \beta)$ is computationally validated to machine precision (§2.4), that the Quadratic Limit $\beta = 0.5$ is derivable as a stability boundary (§2.6), and that composition operators can transition between regimes within a single system (Theorem 6, §2.7). We have presented preliminary evidence (requiring replication) and specified ten falsification criteria.

Distinguishing Theorems from Empirical Claims:

- **Mathematically proven:** Given the three axioms, the multiplicative form $U = I \times f(R, \beta)$ is the *unique* solution (Picard-Lindelöf theorem). The non-additivity result ($U \neq g(I) + h(R)$ for any g, h) follows by contradiction. The scaling exponent $\alpha = 1/(1 - \beta)$ is an exact identity ($R^2 = 1.00000000$). The Quadratic Limit $\beta = 0.5$ is derivable as the unique stability boundary.
- **Requires empirical validation:** That physical systems satisfy the axioms; that measured α equals $1/(1 - \beta_{\text{measured}})$; that the framework applies cross-domain; that composition operator transitions (Theorem 6) occur in AI systems. These are testable predictions, not proven facts.

The mathematical structure is established. Whether nature implements it is the scientific question this paper poses.

The framework generates the following novel, falsifiable predictions:

1. **Composition determines form:** Measuring how recursive steps combine predicts the scaling function.
2. **β -convergence:** Independent measurements of β within each domain will converge.
3. **Cross-domain consistency:** The β -continuum correctly predicts functional forms across AI, quantum, and biological systems.
4. **R^* crossover:** Sequential and parallel scaling curves cross at a predictable threshold.
5. **Quadratic Limit:** For optimised AI systems, $\alpha \rightarrow 2$ from above, as a consequence of stability-optimal coupling at $\beta = 0.5$ (§2.6).
6. **Composition transitions:** In bounded systems, the effective coupling β will decrease with recursive depth, producing transitions between scaling regimes within a single system (§2.7).

Each prediction is testable. Failure of any would require revision of the framework.

What Success Would Mean

If the Global Challenge reveals α patterns consistent with predictions, we will have identified a mathematical signature of how order emerges from recursion: a scaling law with applications from physical systems to artificial intelligence.

This would be actionable:

- How to build more capable AI (optimise sequential recursion)
- How to approach AI alignment (embed values in the recursive process)
- Where to look for similar phenomena in other fields

What Failure Would Mean

If α values scatter randomly: recursive amplification may be domain-specific.

If exponential scaling fits better: the mathematical form needs revision.

If predictions fail: we learn something valuable.

Science advances either way.

The Call to Action

We have made falsifiable predictions. Run the tests. Measure α in your systems. Submit to the repository. Either confirm the pattern or refute it.

That is how science works.

The Deeper Question

If the ARC Principle holds, then intelligence may not be a magic spark. It may be a **scaling crossover**. It could occur when a system with sufficient base asymmetry ($I > 0$) is subjected to sufficient recursive depth ($R > R^*$). The emergence of capability from structure would be the exponent α , emerging from the mathematics of self-referential feedback.

Whether such a principle requires design or emerges spontaneously is itself a question the framework raises but does not resolve.

Intelligence may not be a property of particular materials. It may be what happens on the far side of the recursive threshold.

Speculative Extension: Recursive Cosmology

The following is a philosophical extrapolation beyond current evidence. It is included to indicate where the framework's logic leads, not as a claim supported by the experimental results presented above.

The universe contains a hierarchy of recursive generative processes, each operating on the structured output of the previous level:

1. **Cosmological recursion:** Quantum fluctuations → gravitational collapse → stars → heavy elements → planets → chemistry
2. **Biological recursion:** Chemistry → self-replicating molecules → cells → multicellular organisms → nervous systems
3. **Cultural/technological recursion:** Nervous systems → language → mathematics → science → engineering → AI

Each level emerges from the recursive dynamics of the level below. In ARC notation: each level's *output* (achieved capability U) becomes the *base quality* I for the next level's recursive process. The hierarchy is:

$$I_{\text{cosmic}} \xrightarrow{R_{\text{cosmic}}} I_{\text{bio}} \xrightarrow{R_{\text{bio}}} I_{\text{cultural}} \xrightarrow{R_{\text{cultural}}} I_{\text{tech}} \xrightarrow{R_{\text{tech}}} ?$$

If recursive amplification is substrate-independent, as the cross-domain evidence suggests, then this hierarchy may not terminate at the current level. Sufficiently advanced recursive systems could, in principle, participate in generative processes at scales currently associated with fundamental physics: engineering new effective physical environments, creating conditions for new evolutionary processes, or designing systems whose internal dynamics permit the emergence of new recursive intelligences.

This is not a claim about time travel, bootstrap paradoxes, or closed causal loops (all of which face severe physical objections). It is a more modest observation: that the same recursive amplification pattern observable in time crystals, quantum error correction, and AI reasoning could, if extended to civilisational scales, enable intelligence to become a *driver* of universe-level structure rather than merely a passenger within it.

Whether this constitutes a "recursive cosmology" in which intelligence eventually participates in generative processes comparable to those that gave rise to it is beyond the scope of current evidence. The ARC framework provides the quantitative language for investigating it: what values of I , R , and effective scaling function $f(R, \beta)$ would be required for an engineered system to generate new effective physical laws or substrates? This is a research question, not a conclusion.

Epistemic status: This section is speculative philosophy, not science. It indicates the furthest extrapolation of the framework's logic. The experimental evidence in this paper supports only the claim that recursive self-correction on structured asymmetry produces compounding capability gains across multiple physical substrates. Any cosmological extension requires evidence that does not yet exist.

* DKIM email timestamps are not a recognised priority mechanism in scientific publishing. The December 2024 date reflects manuscript completion, not formal publication.

REFERENCES

- Acharya, R. et al. [Google Quantum AI] (2024). Quantum error correction below the surface code threshold. *Nature*, 638, 920-926.
- Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509-512.
- Cambui, D.S. (2025). Metabolic rate beyond the 3/4 law. *Preprints*. DOI: 10.20944/preprints202501.0001.v1.
- COGITATE Consortium (2025). Adversarial testing of global neuronal workspace and integrated information theories of consciousness. *Nature*, 642, 133-142.
- Corballis, M.C. (2011). *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*. Princeton University Press.
- DeepSeek AI (2025). DeepSeek-R1: Incentivizing Reasoning Capability in LLMs. *arXiv:2501.12948*.
- Eastwood, M.D. (2024/2026). *Infinite Architects: Intelligence, Recursion, and the Creation of Everything*. ISBN: 978-1806056200.
- Glazier, D.S. (2005). Beyond the "3/4-power law." *Biological Reviews*, 80, 611-662.
- Hauser, M.D., Chomsky, N., & Fitch, W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569-1579.
- Hoffmann, J. et al. (2022). Training Compute-Optimal Large Language Models. *arXiv:2203.15556*.
- Kadanoff, L.P. (1966). Scaling laws for Ising models near T_c. *Physics Physique Fizika*, 2(6), 263-272.
- Kaplan, J. et al. (2020). Scaling Laws for Neural Language Models. *arXiv:2001.08361*.
- Lamme, V.A.F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494-501.
- Li, Z. et al. (2025). Revisiting the Test-Time Scaling of o1-like Models: Do they Truly Possess Test-Time Scaling Capabilities? *arXiv:2502.12215*.
- Li, D. et al. (2025). S*: Test Time Scaling for Code Generation. *arXiv:2502.14382*.
- Liu, T., Ou, J.-Y., MacDonald, K.F., & Zheludev, N.I. (2023). Photonic metamaterial analogue of a continuous time crystal. *Nature Physics*, 19, 986-991.
- Morrell, M.C., Elliott, L., & Grier, D.G. (2026). Nonreciprocal wave-mediated interactions power a classical time crystal. *Physical Review Letters*, 136, 057201.
- Raskatla, V., et al. (2024). Magnetically programmable classical time crystal based on photonic-resonator microring lattice. *Physical Review Letters*, 133, 136202.
- Sharma, A. & Chopra, P. (2025). The Sequential Edge: Inverse-Entropy Voting Beats Parallel Self-Consistency at Matched Compute. *arXiv:2511.02309*.
- Snell, C. et al. (2024). Scaling LLM Test-Time Compute. *arXiv:2408.03314*.
- Storm, J.F., Klink, P.C., Aru, J., Senn, W., Goebel, R., ... Pennartz, C.M.A. (2024). An integrative, multiscale view on neural theories of consciousness. *Neuron*, 112(10), 1531-1552.
- Wilson, K.G. (1971). Renormalization Group and Critical Phenomena. I. Renormalization Group and the Kadanoff Scaling Picture. *Physical Review B*, 4(9), 3174-3183.
- West, G.B. & Brown, J.H. (2005). The origin of allometric scaling laws in biology from genomes to ecosystems. *Journal of Experimental Biology*, 208, 1575-1592.
- Zheng, L., Chis-Ciure, R., Waade, P.T., Eiserbeck, A., Aru, J., Andrillon, T. et al. (2025). Recurrency as a Common Denominator for Consciousness Theories. *PsyArXiv*. DOI: 10.31234/osf.io/wqnzc.

DATA AVAILABILITY

A public data repository will be created if the research community shows interest in conducting validation studies. Planned contents would include:

- Measurement protocol templates
- Statistical analysis code (R, Python)
- Model comparison tools (AIC/BIC calculators)
- Figure generation scripts
- Submission guidelines

Independent replication invited. Falsifications welcomed. Contact the author to coordinate validation efforts.

AUTHOR INFORMATION

Michael Darius Eastwood is the author of *Infinite Architects: Intelligence, Recursion, and the Creation of Everything* (2024/2026). His research synthesises insights from theoretical physics, complex systems, and AI.

Correspondence: Contact via the repository.

Competing Interests: None declared.

APPENDIX A: MATHEMATICAL DERIVATIONS

A.1 Cauchy Functional Equation Derivation

Axiom 1 (Dimensional Consistency): $U = I \times g(R)$ where $g(1) = 1$.

Axiom 2 (Compositional Self-Similarity): $g(R_1 \times R_2) = g(R_1) \times g(R_2)$

Note: This uses *multiplicative* composition, modelling hierarchical/fractal recursion. Additive composition $f(R_1 + R_2) = f(R_1) + f(R_2)$ yields exponential $f(R) = e^{\alpha R}$. The choice between these forms is empirically testable (F10).

Theorem: The unique continuous solution is $g(R) = R^\alpha$.

Proof: Let $h(x) = \ln g(e^x)$. Then $h(x+y) = h(x) + h(y)$ (Cauchy additive). Under continuity, $h(x) = \alpha x$, giving $g(R) = R^\alpha$. ■

A.2 β -Dynamics Derivation

Axiom 3: $\frac{dQ}{dr} = a \times Q^\beta$ where $\beta \in [0, 1]$.

Solution: Separating variables:

$$\int Q^{-\beta} dQ = \int a dr$$

$$\frac{Q^{1-\beta}}{1-\beta} = ar + C$$

With initial condition $Q(0) = I$:

$$Q(R) = [I^{1-\beta} + (1-\beta)aR]^{1/(1-\beta)}$$

Deep recursion limit ($R \gg R^*$):

$$Q(R) \approx [(1-\beta)aR]^{1/(1-\beta)} \propto R^{1/(1-\beta)}$$

Therefore: $\alpha = \frac{1}{1-\beta}$ ■

A.3 Transitional Regime and the Scaling Crossover

Full solution: $U(R) = [I^{1/\alpha} + \frac{1}{\alpha}aR]^\alpha$

Three regimes:

1. $R \ll R^*: U \approx I$ (base capability dominates)
2. $R \gg R^*: U \approx R^\alpha$ (power law dominates)
3. Crossover: $R^* = \frac{\alpha I^{1/\alpha}}{(\alpha-1)\alpha}$

Intuitive meaning of R^* : R^* is the **crossover depth**, the point at which recursive compounding overtakes the system's base capability. Below R^* , the cost of recursion outweighs the gain (the system is "warming up"). Above R^* , compounding returns dominate and capability grows super-linearly. Finding R^* for any system is equivalent to finding its threshold depth.

Why this matters: This predicts a qualitative change in scaling behaviour at a specific, measurable depth. Systems should exhibit a distinct "elbow" in their capability curves at R^* . This is testable: plot U vs R on log-log axes and look for the crossover from linear to power-law scaling.

The existence of R^* is a novel, falsifiable prediction that distinguishes recursive amplification from simple redundancy.

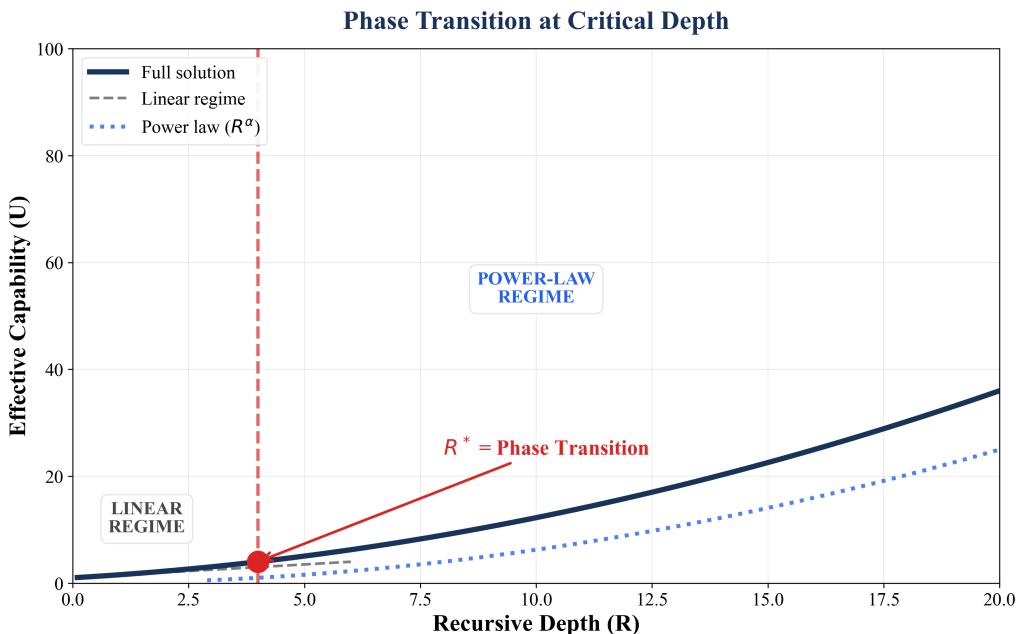


Figure 11 | The Scaling Crossover at R^* . Below the critical depth R^* , base capability (I) dominates and scaling appears linear. Above R^* , recursive compounding dominates and scaling becomes power-law. This crossover point is the depth at which recursive amplification becomes dominant. Systems should exhibit a distinct "elbow" in their capability curves at R^* .

APPENDIX B: GLOSSARY FOR NON-SPECIALISTS

Term	Plain English meaning
ARC	Artificial Recursive Creation: the principle that recursive self-correction on structured asymmetry may produce super-linear capability gains
Scaling law	A mathematical relationship showing how one quantity changes as another changes
Power law	A relationship where $Y = X^\alpha$ (like how area scales as length squared)
Exponential	A relationship where $Y = e^{\alpha X}$ (like compound interest)
Recursive	Self-referential; the output becomes the input for the next step
Recurrent	Repeating in cycles; in neuroscience, refers to neural signals that loop back
Sequential	One step after another, each building on the last
Parallel	Multiple independent attempts at once
Falsifiable	Can be proven wrong by experiment
α (alpha)	The scaling exponent; determines if returns compound or diminish
β (beta)	Self-referential coupling; how much prior work helps the next step
Λ (Lambda)	Google's quantum error suppression factor (exponential decay rate)
Quenched disorder	Structured asymmetry (like varied bead sizes) that enables the system to function
Scaling crossover	The point where scaling behaviour changes from one regime to another

WHITE PAPER III: Version 6.5

© 2026 Michael Darius Eastwood. All Rights Reserved.

"The predictions are specified. The falsification criteria are public. The data will decide."