

# **UIC v0.2.1**

## **$\lambda$ -Sweep Coherence Transition Test**

### **A Preregistered Quantum Coherence Characterization Study**

**Preregistered: February 7, 2026**

**Status: Simulation-First, Hardware-Compatible**

**Supersedes: UIC v0.1 (Discrete A/B/C)**

*Integrating PEIG and Omega Theoretical Frameworks*

# BACKGROUND & MOTIVATION

## Scientific Context

**Quantum coherence dynamics** are **central** to both **quantum information processing** and foundational questions about **complex quantum systems**.

While **decoherence mechanisms** (dephasing, relaxation, environmental coupling) are **well-characterized**, **fewer studies systematically isolate** the role of internal circuit organization **on coherence preservation under controlled, blinded experimental conditions**.

## Existing literature establishes that:

- **Random phase perturbations destroy coherence** (well-known).
- **Gate count and circuit depth correlate with decoherence** (expected).
- **Entanglement can both protect and accelerate decoherence depending on environmental coupling** (context-dependent).

**However, whether coherent organizational structure (beyond simple gate minimization) modulates coherence times in a graded, reproducible manner remains an open empirical question.**

## **Experiment 1 Results and Limitations**

**UIC v0.1 (completed February 2026) compared three discrete circuit families under blinded conditions:**

- **Condition A:** Isolated single-qubit Ramsey baseline
- **Condition B:** Entangled Bell state with structured coupling rounds
- **Condition C:** Entangled Bell state with randomized noise coupling

### **Key findings:**

- ✓ Binary regime difference observed (B preserved coherence, A/C collapsed)
- ✓ Effect was statistically robust and reproducible
- ✓ Blinding protocol validated end-to-end

### **Limitations identified (external review + self-assessment):**

- ✗ Binary outcome (on/off switch) lacks graded novelty
- ✗ Risks reproducing known effect: "phase randomization kills coherence"
- ✗ No intermediate regime characterized
- ✗ Cannot distinguish threshold phenomenon from continuous sensitivity

### **IBM Quantum Credits review feedback (February 2026):**

"Not novel enough to be considered... pushing the boundary of quantum computing application research." Plus no academic affiliations.

Note from me directly: This does not stop me from my desire to be on your guy's/women's levels. To be a scientist and explore the boundaries. Any academic or employment opportunities yall can direct me to would be great! <3

Interpretation: The discrete A/B/C comparison, while methodologically sound, did not demonstrate sufficient novelty in its current framing to warrant institutional hardware access.

## Experiment 2 Design Rationale

**To address the novelty limitation while preserving methodological rigor, Experiment 2 introduces a continuous organizational control parameter**

**$\lambda \in [0,1]$  that interpolates between circuit regimes:**

- **$\lambda = 0$** : Pure randomized noise coupling (**analog of Condition C**)
- **$\lambda = 1$** : Pure structured coherent coupling (**analog of Condition B**)
- **$\lambda \in (0,1)$** : Probabilistic mixture of **structured vs noise blocks**

**This design enables characterization of coherence behavior as a continuous function of organization, revealing whether the effect exhibits:**

- (a) **Graded sensitivity** - Smooth monotonic or non-linear modulation of  **$T2^*(\lambda)$** , indicating architectural tunability.
- (b) **Sharp phase transition** - Abrupt coherence change **at critical  $\lambda_c$** , indicating **threshold-dependent regime switching**.

### **Key improvements over Experiment 1:**

- ✓ **Transforms binary switch → continuous control parameter**
- ✓ **Maps transition region** (if present) or sensitivity curve
- ✓ **Enables identification** of critical organizational threshold
- ✓ **Distinguishes "regime difference" from "graded architectural effect"**
- ✓ **Provides** hardware-targetable parameters ( $\lambda_c$  or **representative  $\lambda$  set**)

## **Research Objectives (Clearly Bounded)**

### **Primary objective:**

**Characterize  $T2^*(\lambda)$  dependence under blinded conditions to determine if circuit organization modulates coherence in a continuous regime.**

**Secondary objective:**

If transition observed, identify critical parameter  $\lambda_c$  and **quantify transition sharpness for hardware validation targeting.**

**Explicitly NOT objectives:**

- X Claim** detection of consciousness, awareness, or subjective experience.
- X Assert biological relevance or brain-substrate mapping.**
- X Demonstrate** quantum advantage or computational speedup.
- X Prove** necessity of quantum effects for any cognitive function.
- X Validate** PEIG or Omega theoretical frameworks (these provide motivating context, not testable claims).

This experiment **investigates quantum coherence** as an **information-theoretic property of circuit organization**, deferring all consciousness-related interpretations to future theoretical work.

Results remain interpretable regardless of outcome.

## **Broader Impact**

**Beyond immediate findings, this work contributes a reproducible, blinded experimental framework for studying coherence-organization relationships, with methods transferable to other quantum characterization tasks.**

**This is not just "one more decoherence measurement"—it's a methodological contribution that raises standards for how quantum coherence characterization is conducted and reported.**



# Key Impact Dimensions

## **\*\*Methodological Innovation:\*\***

- **Establishes blinded protocols** for quantum circuit experiments (rare in physics).
- **Demonstrates pre-registration feasibility** in exploratory quantum research.
- **Provides a template** for isolating confounded variables in complex quantum systems.

## **\*\*Hardware and Engineering:\*\***

- **Informs** NISQ-era circuit compiler optimization strategies.
- **Either outcome** (organizational sensitivity present/absent) **advances quantum engineering.**
- **Provides benchmark data** for hardware characterization protocols.

**\*\*Cross-Disciplinary Bridges:\*\***

- **Connects** quantum information theory, thermodynamics, complex systems, and error correction.
- **Methodological lessons** transferable beyond quantum computing.
- **Educational case study** in rigorous experimental design.

**\*\*Responsible Innovation:\*\***

- **Models epistemic boundaries** between empirical findings and theoretical speculation.
- **Demonstrates** value of null results and outcome-agnostic research design.
- **Advances open science practices** through transparent methods and data sharing.

All three possible outcomes  
(strong correlation, null result, or threshold behavior)  
**advance understanding** of quantum circuit design  
and **establish methodological precedents**  
**for rigorous quantum phenomenology.**

**\*\*Theoretical Connections:\*\***

- **Results inform information-geometric models** of quantum systems  
(PEIG framework).
- **Contributes data to self-organization theories** in quantum computing  
(Omega framework).
- Either outcome (**organizational sensitivity present/absent**) constrains  
theoretical predictions.
- **Provides empirical grounding** for speculative connections to entropic  
gravity and holographic principles.

**Novelty Statement (For Editorial Assessment)**

This work presents the **first systematic** characterization of **coherence time as a continuous function of probabilistic circuit organization ( $\lambda$ -sweep)**, executed under **pre-registered blinded conditions with depth-matched controls**, distinguishing **organizational effects** from **known noise accumulation and gate-count dependencies.**

## Key Differentiators from Existing Literature:

- **Continuous organizational parameter** (not discrete conditions)
  - **Blinded analysis with frozen statistics** before decode
- **Pre-registered falsification criteria** (prevents outcome bias)
- **Depth-matched controls** (isolates organizational effects)
- **Phase transition vs graded sensitivity** discrimination

**This addresses a gap in quantum coherence characterization:** most studies vary noise parameters or hardware, not internal organizational structure under controlled probabilistic mixing.

## Important Boundaries

This work is **mechanism-agnostic**.

We test whether **organizational structure modulates coherence** (phenomenological relationship), **not why** it might do so.

**Speculative theoretical frameworks** (e.g., **entropic gravity, information-theoretic geometry, consciousness models**) are intentionally decoupled from empirical claims to maintain falsifiability and epistemic rigor.

# Theoretical Context (Optional Framework)

This experiment can be viewed through the **lens of two complementary theoretical frameworks** that **motivate (but do not require)** the experimental design:

## **Physics-Entanglement-Information-Geometry (PEIG) Framework:**

The **PEIG framework** posits that information organization (not just information content) constitutes a measurable physical observable with geometric interpretation. Within this framework:

- The  $\lambda$  parameter **represents a continuous probe of information-geometric structure.**
- $T_2^*$  serves as **the physical readout of how organizational geometry couples to quantum coherence.**
- Circuit organization **modulates the "information distance" between quantum states.**
- Coherence preservation **depends on geometric properties of the computational trajectory.**

Key PEIG prediction tested here: **IF organizational structure affects information geometry, THEN  $T_2^*(\lambda)$  should show continuous modulation or threshold behavior.**

## **Omega Framework (Self-Organization Dynamics):**

The Omega framework models quantum circuits as self-organizing systems with multi-scale feedback loops:

- Micro-scale: **Individual gate operations and local decoherence**
- Meso-scale: **Circuit organizational structure ( $\lambda$ -dependent)**
- Macro-scale: **System-level coherence maintenance ( $T2^*$ )**

**Key Omega prediction:** Systems with organizational feedback exhibit critical transitions where coherence costs change abruptly at threshold parameters (**potential  $\lambda_c$** ).

### **Integration Point:**

**Both frameworks converge on a testable prediction:**

**Organizational structure ( $\lambda$ )** should couple to coherence dynamics ( **$T2^*$** ) in a manner distinct from **simple gate-count** or **depth effects**.

**The  $\lambda$ -sweep design directly probes this coupling.**

### Critical epistemic distinction:

- These frameworks **MOTIVATE** the experimental design.
- They **do NOT** constitute claims the experiment will validate.
  - Results **are interpretable** REGARDLESS of whether PEIG/Omega assumptions hold.
- Null results would **CONSTRAIN (not falsify)** these frameworks

### Connection to broader physics:

**PEIG** connects to **entropic gravity** (Verlinde), **ER=EPR** conjecture (Susskind/Maldacena), and **holographic principle**.

**Omega** builds on **self-organized criticality** (Bak) and **quantum thermodynamics**.

**Both are speculative extensions** beyond standard quantum information theory.

### What this experiment actually tests:

Whether a continuous organizational parameter modulates coherence time under blinded, controlled conditions.

**The mechanism** (if effects exist) remains open for interpretation.

### All outcomes are publishable and scientifically valuable:

Outcome	Criterion	Implication	Impact	Framework Relevance
Strong correlation (continuous modulation)	$ r  > 0.3, p < 0.01$	Organization is significant modulator across $\lambda$ range	Opens new optimization dimension for NISQ circuits	Supports PEIG: information-geometry coupling confirmed
Null result (no $\lambda$ dependence)	$ r  < 0.1$	Organization irrelevant when depth controlled	Simplifies circuit design; removes spurious variable	Constrains PEIG/Omega: effects below detection threshold
Threshold behavior (regime switch)	<b>Sharp transition at <math>\lambda_c</math></b>	Critical disorganization point exists	Suggests phase-transition-like phenomena in quantum information	Supports Omega: critical transitions at threshold parameters



**Key insight: Each outcome answers a different scientific question:**

- Strong correlation → **"How much does organization matter?"** (sensitivity quantification).
- Null result → **"Can we ignore organization?"** (design simplification)
- Threshold → **"Where does organization matter?"** (critical parameter identification)

## **Relationship to Prior Work**

**What IS known:**

- **Random circuits decohere faster than structured circuits** (qualitative, Cross et al. 2019).
- **Decoherence scales with gate count and circuit depth** (established).
- **Environmental coupling mechanisms are well-characterized**

### **What IS NOT known (this work addresses):**

- Whether **organizational structure is a continuous modulator independent of gate count.**
- **Threshold vs. linear response profile to organizational changes.**
- **Effect size and replicability under blinded, depth-matched conditions.**
- **Phenomenological relationship free from confounding variables.**

### **Key innovation:**

**We isolate organizational structure as an independent variable by holding gate count and depth constant while continuously varying  $\lambda$ .**

**No prior work** has systematically mapped this relationship under pre-registered blinded conditions.

# Pre-emptive Responses to Common Reviewer Concerns:

## Concern 1:

"Isn't this **just known decoherence** physics?"

## Response:

Standard decoherence models predict dependence on gate count and depth, which we explicitly control for.

**If  $\lambda$ -dependence exists** after matching these variables, it represents an **unmeasured organizational contribution distinct from known noise accumulation**.

### **Concern 2:**

**"Why would organization matter** if the physical Hamiltonian is the same?"

### **Response:**

We're testing whether it matters empirically, not assuming it does.

**A null result is equally valuable**—it would demonstrate that organizational structure is irrelevant when depth and gate count are controlled, simplifying quantum circuit optimization.

### **Concern 3:**

**"What if the effect is too small** to matter?"

### **Response:**

Effect size is an empirical question.

Statistical power analysis shows **N=160 circuits provides 80% power** to detect  $|r| \geq 0.3$ .

**If the effect is smaller**, we will accurately report that, which informs future circuit design decisions.

**Concern 4:**

**"Could this be explained by calibration drift or systematic bias?"**

**Response:**

**Three design features control for this:**

- **Blinded  $\lambda$  values** prevent experimenter bias during data collection.
- **Randomized circuit execution order** prevents temporal confounds.
- **Pre-registered analysis pipeline** eliminates researcher degrees of freedom.

### **Concern 5:**

**"Why not test with more qubits or deeper circuits?"**

### **Response:**

**We test the principle at an accessible scale (5-10 qubits).**

**If organizational effects exist**, they should be measurable at this scale.

**If not detectable here**, that constrains theories requiring macroscopic system sizes and informs resource allocation for future studies.

### **One-Sentence Summary (For Abstracts/Grants)**

**"We test whether quantum circuit organizational structure continuously modulates coherence decay independent of gate count and depth, using blinded protocols and pre-registered falsification criteria to isolate a previously confounded variable."**

## Connection to Broader Research Program (Optional Context)

This **empirical investigation** is **intentionally decoupled from speculative theoretical extensions** to maintain scientific rigor.

**While broader frameworks exploring information-geometry coupling exist** (see Speculative Appendix A in full documentation), **the empirical predictions and falsification criteria stand independently of any particular interpretive framework.**

### Why this separation matters:

- **Results:** remain interpretable regardless of theoretical preferences.
- **Null findings:** are as valuable as positive findings.
- **Multiple theoretical frameworks:** can be evaluated against the same empirical data.
- **Scientific credibility:** is maintained through epistemic hygiene.

# Why Reviewers Should Care

## Methodological contribution:

- **Demonstrates:** blinding protocols for quantum circuit experiments.
- **Establishes:** a template for continuous-variable quantum phenomenology.
- **Shows:** how to isolate confounded variables in complex quantum systems

## Scientific contribution:

- **Tests specific, falsifiable hypothesis:** with clean controls.
- **Addresses systematic gap:** in organizational variation studies.
- **Provides empirical constraints:** for quantum circuit optimization.



### **Practical contribution:**

- **Informs:** NISQ-era circuit compilation strategies.
- **Reduces parameter space:** if organization proves irrelevant.
- **Opens optimization dimension:** if organization proves significant

### **Reject this work only if:**

- **Prior work already tested:** continuous  $\lambda$  under blinded, depth-matched conditions.
- **(citation required:—systematic review found none).**
- **Experimental design contains fatal methodological flaw:** (please specify).
- **Statistical power insufficient:** (we calculate 80% power for target effect size).

**Otherwise, this represents genuine methodological and empirical novelty in:**

- **Quantum coherence phenomenology.**
- **Prepared for:** preregistration defense, grant applications, and peer review.
- **Ready for:** technical evaluation and critique.

## **PRIMARY HYPOTHESIS**

- **$\lambda$  controls the probability that a structured coupling block (entangling, non-randomized) is applied versus a noise/randomization block per feedback round.**
- **Coherence time ( $T2^*$ ) is hypothesized to depend continuously on  $\lambda$ .**
- **Formally,  $T2^*(\lambda)$  is not constant over  $\lambda$ .**

### **Expected qualitative outcomes:**

$\lambda = 0 \rightarrow$  low  $T2^*$  (noise-dominated regime)

$\lambda = 1 \rightarrow$  higher  $T2^*$  (structured regime)

For intermediate  $\lambda$  values, either:

- Graded monotonic modulation, or
- A sharp transition at a critical  $\lambda_c$  (phase boundary)

### **NULL HYPOTHESIS**

After controlling for **circuit depth** and **gate count**:

$T2^*(\lambda)$  shows **no statistically significant dependence** on  $\lambda$ , or

$T2^*(\lambda)$  remains binary, **exhibiting only two plateaus** with no intermediate regime.

# METRICS

## Primary Metric

**$T2^*(\lambda)$  extracted from per-replication exponential decay fits of the form:**

$$V(t) = a \cdot \exp(-t / T2^*) + c$$

**Fitting performed independently per replication  
(no pooled fitting).**

## Secondary Metrics

**Regression slope  $\beta$  in  $T2^*(\lambda) = \alpha + \beta\lambda$**

**Variance of  $T2^*$  across replications per  $\lambda$   
(must remain finite)**

**Transition sharpness:  $|dT2^*/d\lambda|$   
(if non-monotonic)**

## **Tertiary (Exploratory) Metrics**

Mutual information  **$I(\mathbf{A}:\mathbf{B})$  vs  $\lambda$**   
(exploratory only; not used to accept/reject primary hypothesis)

## **EXPERIMENTAL CONDITIONS (BLINDED)**

**Sampling and Replication – Each  $\lambda$  condition consists of:**

- **10** independent replications
- **4000** shots per circuit
- **18** Ramsey delay points  
**(0–200  $\mu\text{s}$ , quasi-log spaced)**

**$\lambda \in \{0.0, 0.1, 0.2, \dots, 1.0\}$  (11 conditions)**

**All  $\lambda$  labels blinded as randomized tokens (e.g., L00–L10)**

**Decode map** held separate and accessed only after statistical analysis is frozen.

## **PHYSICAL NOISE MODEL (AER SIMULATION)**

**To ensure delays induce realistic physical decoherence and avoid artifact-free behavior, We implemented a device-inspired noise model based on typical IBM Quantum processor characteristics:**

### **Idle decoherence applied during circuit delays:**

- **T1** (energy relaxation) = **120  $\mu$ s**
- **T2** (dephasing) = **80  $\mu$ s**
- Implemented via **thermal\_relaxation\_error** with **duration-dependent decay inserted as explicit instructions** during delay operations

### **Gate-level depolarizing noise:**

- Single-qubit gates: **p<sub>1q</sub> = 0.0005**  
(0.05% error rate)
- Two-qubit gates: **p<sub>2q</sub> = 0.002**  
(0.2% error rate)
- **Applied consistently** to all entangling and rotation operations

### Noise model properties:

- Applied **uniformly across all  $\lambda$  conditions** (no  $\lambda$ -dependent noise).
- Parameters **chosen to match mid-range IBM Quantum device performance.**
- Ensures **T2\*** measurements **reflect circuit organization, not noise parameter variation.**

### Validation:

**Noise-free control simulations** confirmed that **observed  $\lambda$ -dependence is not an artifact of the noise model itself.**



# CIRCUIT CONSTRAINTS & CONTROLS

**All  $\lambda$  conditions are depth-matched and gate-count matched in expectation:**

The expected number of **gates per feedback round is equal across  $\lambda$** , though individual realizations may vary stochastically.

**Any residual depth variation will be:**

- Quantified
- Included as a covariate in secondary regression checks.
- Circuit topology, measurement basis, and delays held constant across  $\lambda$ .

# STATISTICAL ANALYSIS PLAN

## Primary Inferential Test (Pre-Specified)

Linear regression model:

$$T2^*(\lambda) = \alpha + \beta\lambda + \varepsilon$$

Where:

- **$T2^*(\lambda)$  is the mean coherence time across replications at each  $\lambda$**
- **$\beta$  is the slope coefficient (primary parameter of interest).**
- **$\alpha$  is the intercept**
- **$\varepsilon$  represents residual error**

## **Hypothesis test:**

**$H_0: \beta = 0$**  (no  $\lambda$ -dependence)

**$H_1: \beta \neq 0$**  (significant  $\lambda$ -dependence)

**Significance level:  $\alpha = 0.05$  (two-tailed test)**

## Decision rule:

- If p-value for  
 $\beta < 0.05 \rightarrow \text{Reject } H_0$   
( $\lambda$ -dependence supported)
- If p-value for  
 $\beta \geq 0.05 \rightarrow \text{Fail to reject } H_0$   
(no  $\lambda$ -dependence detected)

### **Model diagnostics (pre-committed):**

- **$R^2$**  (coefficient of determination)  
to assess fit quality.
- **Residual normality check**  
(Shapiro-Wilk test, diagnostic only).
- **Homoscedasticity inspection**  
(visual, not binding)

### **Minimum effect for practical significance:**

**$|\beta| > 20 \mu\text{s per unit } \lambda$**   
(i.e., at least  $20 \mu\text{s}$  change from  $\lambda=0$  to  $\lambda=1$ )

This threshold chosen to ensure effect is  
distinguishable from  
measurement noise and simulation variance.

## **Secondary Model (Contingency Plan, Pre-Declared)**

Applied **only** if linear model residuals **violate assumptions** or **show systematic non-linearity**.  
(assessed via residual plots and runs test)

**Model candidates (selection before decode):**

**Polynomial regression (degree 2 or 3):**

- Selected via AIC/BIC comparison.
- Used if smooth curvature is evident.

### **Piecewise-linear model:**

- If data suggests **distinct regimes** with **sharp breakpoint**.
- **Breakpoint estimated** via **segmented regression**.

### **Logistic/sigmoid transition model:**

- If **sharp threshold behavior** evident
- Form:  $T2^*(\lambda) = L + (U-L) / (1 + \exp(-k(\lambda-\lambda_c)))$
- Where **L = lower plateau**, **U = upper plateau**,  
 **$\lambda_c$  = transition point**.

## **Model selection protocol:**

1. **Visual inspection** of residuals (blinded)
2. **AIC/BIC** comparison of candidates
3. Choice **documented** before decode
4. Primary **p-value reported from linear model**  
(conservative)



## **Transition point estimation (if sharp transition detected):**

- $\lambda_c$  defined as **point of maximum  $|dT_2^*/d\lambda|$**
- Derivative estimated via **finite-difference with Savitzky-Golay smoothing**  
(polynomial order 2, window 5 points)
- Confidence interval: **bootstrap resampling**  
(1000 iterations, percentile method)
- Reported whether or not secondary model is used

## Supplementary Tests (Exploratory, Not Primary)

If per- $\lambda$  variance is finite and permits group comparisons:

- **One-way ANOVA** on  $T2^*$  distributions across  $\lambda$  groups.
- **Tukey HSD** for all pairwise  $\lambda$  comparisons.
- **Cohen's d effect sizes** for adjacent  $\lambda$  pairs.

These tests **supplement primary regression** but **do not override it**.

Reported in supplementary materials regardless of significance.

## FALSIFICATION CRITERIA

**The hypothesis of  $\lambda$ -dependent coherence modulation is NOT SUPPORTED if any of the following conditions are met:**

## **Primary Falsification Conditions:**

### **1. Regression slope not significant:**

- **Linear regression slope  $\beta$  fails** significance test  
( $p \geq 0.05$ )
- **AND practical effect size is negligible**  
( $|\beta| < 20 \mu\text{s per unit } \lambda$ )

### **2. Binary behavior persists:**

- **$T2^*(\lambda)$  exhibits only two distinct plateaus with no intermediate values across the  $\lambda$  range.**
- **Step-like transition with no graded modulation detectable.**

### **3. Ill-conditioned fits:**

- **Variance in T2\* estimates collapses to near-zero ( $\sigma^2 < 0.01$ ) at three or more  $\lambda$  values.**
- **Fit convergence failures exceed 30% of replications.**
- **Indicates simulation artifacts dominate physical signal.**

### **Secondary Falsification Conditions:**

#### **4. Depth confound:**

- **Observed  $\lambda$ -effect disappears ( $\beta \rightarrow 0$ ,  $p > 0.1$ ) after including circuit depth as covariate in regression.**
- **Indicates the effect is explained by gate count, not organization.**

## 5. Calibration drift:

- **Condition Z** (calibration baseline) **shows comparable  $\lambda$ -like variation despite fixed protocol.**
- **Suggests systematic drift** rather than  **$\lambda$ -dependent effect.**

## Partial Falsification (Interpretation Adjustment):

**If primary test fails BUT secondary models reveal non-linear relationship**

**(e.g., threshold-only, no continuous sensitivity):**

- **Relational hypothesis reformulated as threshold phenomenon.**
- **Still publishable**, but claims adjusted to reflect **binary regime identification** rather than **graded architectural sensitivity**.

**Outcome regardless of falsification:**

- **Null results** are **informative** and **publishable**.
- **Methodology and blinding protocol** remain **valid contributions**.
- Simulation framework **available for community use**.

**SUCCESS CRITERIA (PUBLICATION-WORTHY)**

For the primary hypothesis  
**( $\lambda$ -dependent coherence modulation)**  
to be considered supported, all four criteria below  
must be satisfied.

These thresholds are specified a priori to prevent  
post-hoc rationalization.

**All of the following must be satisfied:**

**1. Statistically significant  $\lambda$ -dependence**

- **Linear regression slope  $\beta$  is significant:  
( $p < 0.05$ )**
- **Practical effect size:  
 $|\beta| > 20 \mu\text{s per unit } \lambda$**
- **Model fit quality:  
 $R^2 > 0.30$**

## **2. Finite, non-saturated variance:**

- Variance  $\sigma^2 > 0.01 \mu s^2$  at all  $\lambda$  values
  - Coefficient of variation  
 $CV = \sigma/\mu < 0.5$  per  $\lambda$  group
- Fit convergence rate  $\geq 70\%$  across replications



### 3. Continuous or threshold modulation (EITHER condition)

#### **EITHER smooth graded modulation:**

- **Monotonic or smooth non-linear  $T2^*(\lambda)$  trend**
- **At least 5 intermediate  $\lambda$  values show  $T2^*$  distinct from both endpoints ( $\lambda=0$  and  $\lambda=1$ )**

#### **OR sharp phase transition:**

- **Identifiable critical point  $\lambda_c$  with 95% CI width  $< 0.2$**
- **Transition amplitude  $\Delta = T2^*(1.0) - T2^*(0.0) > 30 \mu s$**
- **$|dT2^*/d\lambda|$  at  $\lambda_c > 2 \times$  baseline gradient**

#### 4. Effect stability across replications

- **Per-replication**  
coefficient of variation  $< 0.4$  at each  $\lambda$
- **Qualitative trend (increasing/threshold)**  
consistent in  $\geq 80\%$  of reps.
  - No single replication has  
**Cook's distance  $> 1$  (leverage check).**
- **Effect persists** when any single replication is dropped (**jackknife**).

### INTERPRETATION FRAMEWORK

This framework specifies how results will be interpreted and communicated regardless of outcome, protecting against post-hoc narrative shifts.

# OUTCOME 1: Primary Hypothesis Supported

## Evidence pattern:

- All success criteria met (statistical significance, finite variance, continuous/threshold modulation, replication stability)
- Clear  $\lambda$ -dependence of  $T2^*$  observed

## Interpretation:

Circuit organization modulates coherence in a continuous regime under blinded experimental conditions. The effect demonstrates either:

- (a) graded architectural sensitivity (smooth  $T2^*(\lambda)$  curve), or
- (b) sharp organizational phase transition (identifiable  $\lambda_c$ )

This finding goes beyond binary regime identification and is not explained by gate count or circuit depth alone.

## Claims we CAN make:

- ✓ Organizational structure affects coherence dynamics
- ✓ Effect is reproducible and statistically robust
- ✓ Threshold or sensitivity parameter identified
- ✓ Methodology validated for hardware deployment

## Claims we CANNOT make:

- ✗ Consciousness detected or proven
- ✗ Biological relevance established
- ✗ Generalization beyond 2-qubit Bell-state architectures
- ✗ Hardware behavior guaranteed (simulation noise model limitations)

Publication target: Quantum information journals (e.g., Quantum, npj Quantum Information, Physical Review A)

## OUTCOME 2: Null Result (No $\lambda$ -Dependence)

Evidence pattern:

- Regression slope  $\beta$  not significant ( $p \geq 0.05$ )
- OR  $T2^*(\lambda)$  remains flat/binary across  $\lambda$  range
- BUT fits are well-conditioned and variance is finite

Interpretation:

The continuous mixing parameter  $\lambda$  does not modulate coherence beyond binary regime boundaries. The effect identified in Experiment 1 (discrete A/B/C) represents a threshold-only phenomenon rather than graded architectural sensitivity.

Claims we CAN make:

- ✓ Organizational regimes identified (binary switch confirmed)
- ✓ Blinded methodology validated
- ✓ No evidence for continuous modulation in this parameter space
- ✓ Negative result informs theory (organizational effects are threshold-dependent, not smoothly tunable)

Claims we CANNOT make:

- ✗  $\lambda$ -parameter was poorly chosen (it was pre-specified)
- ✗ Effect would appear with different parameter (post-hoc speculation)

Publication target: Open-access venues emphasizing null results (e.g., PLOS ONE, Royal Society Open Science)

## OUTCOME 3: Failure Due to Technical Limitations

Evidence pattern:

- Variance collapse ( $\sigma^2 \rightarrow 0$ ) at multiple  $\lambda$  values
- Fit convergence failures exceed 30%
- OR calibration baseline (Z) shows anomalous variation

Interpretation:

Simulation noise model produces artifacts that dominate or obscure the  $\lambda$ -dependent signal. Results are inconclusive and require:

- (a) refined simulation noise parameters, or
- (b) hardware validation to bypass simulation limitations

Claims we CAN make:

- ✓ Identified boundary of simulation fidelity
- ✓ Experimental protocol is sound but requires physical device
- ✓ Methodology transferable to hardware immediately
- ✓ Simulation serves as control/validation tool, not primary evidence

Claims we CANNOT make:

- ✗ Effect is absent (simulation limitations prevent conclusion)
- ✗ Hardware will show same behavior (simulation  $\neq$  reality)

Next step: Pursue hardware access via academic collaboration or alternative funding; publish methodology as preprint.

# CONSCIOUSNESS-RELATED INTERPRETATIONS

## **Explicit deferral statement:**

Any implications for consciousness, integrated information, or phenomenal experience are DEFERRED and remain speculative.

This experiment characterizes coherence modulation as a function of circuit organization—a quantum information property—not a test of subjective experience, qualia, or awareness.

## **If results support $\lambda$ -dependence:**

Future work may explore whether identified organizational principles map onto theoretical frameworks (e.g., Integrated Information Theory, Global Workspace Theory), but such mappings are interpretive extensions beyond the empirical claims established here.

**Boundary commitment:**

Primary publication will focus on quantum coherence dynamics. Consciousness-adjacent discussion, if included at all, will be restricted to a brief "Future Directions" subsection with explicit epistemic hedging.

**TARGET PUBLICATION VENUES  
(Outcome-Dependent)**

Venue selection will follow this pre-specified hierarchy based on results strength:

**Tier 1: High-Impact Quantum Journals**

Targets: Nature Communications, npj Quantum Information, Physical Review X, Quantum

### **Requirements:**

- All success criteria met
- Hardware validation completed OR strong simulation evidence with clear hardware pathway
- Sharp transition identified ( $\lambda_c$ )  
OR large effect size ( $|\beta| > 50 \mu s$ )

### **Positioning:**

**"Phase transition in quantum coherence as function of circuit organization"**

### **Tier 2: Specialized Quantum Venues**

#### **Targets:**

Physical Review A, Quantum Science and Technology,  
Quantum Information Processing.



### **Requirements:**

- Success criteria met (simulation-based acceptable).
- Finite variance, clean fits, reproducible.
- Moderate effect size ( $|\beta| > 20 \mu s$ ) OR **identifiable threshold.**

### **Positioning:**

"Organizational modulation of quantum coherence under blinded experimental protocol"

### **Tier 3: Open Access / Methodology-Focused**

#### **Targets:**

PLOS ONE, Scientific Reports, Royal Society Open Science

**Requirements:**

- Null or partial results acceptable
- Methodology sound and reproducible
- Negative result informative

**Positioning:**

"Pre-registered investigation of circuit organization effects on quantum coherence"

**Tier 4: Preprint / Community Contribution  
Targets:**

arXiv (quant-ph), Zenodo dataset publication

**Always published regardless of outcome:**

- Complete methodology documentation.
- Open data and code repository.
- Contribution to reproducibility literature.

### **Editorial Response Strategy:**

- If Tier 1 rejects as "not impactful enough"  
→ submit to Tier 2 unchanged.  
(novelty claim remains valid)
- If "not novel" claim appears  
→ cite this preregistration and request specific  
prior work doing  $\lambda$ -swept organizational blinding.
- If methodological questions arise  
→ reference frozen blinded analysis archive

## **POST-ANALYSIS PROTOCOL**

All steps must be executed in the order listed below.

This sequence protects blinding integrity and  
prevents analysis bias.

# PHASE 1: Blinded Analysis (Statistics Frozen Before Decode)

## Step 1.1 - Load and validate blinded data

- Import CSV from simulation output.
- Verify data completeness  
(all  $\lambda$  tokens present, all reps recorded)
- Check for missing values or corrupted entries.
- Document: data\_validation\_report.txt

## Step 1.2 - Compute per-replication T2\* estimates

- Fit **exponential decay model**  
 **$V(t) = a \cdot \exp(-t/T2^*) + c$**  for each (**blind\_token**, rep) pair
- Record fit quality metrics  
( **$R^2$ , convergence status, parameter SEs**)
- Flag ill-conditioned fits  
(**convergence failures or  $R^2 < 0.1$** )
- Output: **t2star\_per\_rep\_blinded.csv**

### **Step 1.3 - Execute primary statistical test (blinded)**

- Linear regression:  $T2^* \sim \lambda$   
(using **blind token order** as proxy)
- Compute **slope  $\beta$ , p-value,  $R^2$ , residuals**
  - Perform model diagnostics  
(normality, homoscedasticity)
- Output: regression\_blinded.txt

### **Step 1.4 - Execute secondary/supplementary tests (blinded)**

- **ANOVA** across  $\lambda$  groups (if variance permits)
- **Polynomial or piecewise models**  
(if residuals indicate)
  - Tukey HSD, effect sizes
- Output: anova\_blinded.csv, tukey\_blinded.txt

### **Step 1.5 - Generate blinded visualizations**

- **T2\* vs  $\lambda$  scatter plot** (blind tokens on x-axis)
  - Per-replication **decay curves**
  - Residual **diagnostic plots**
  - Output: **figures/ directory**  
(all PNG files timestamped)

### **Step 1.6 - FREEZE ANALYSIS**

- Create archive:  
analysis\_frozen\_YYYYMMDD\_HHMM.zip
- Contains: all CSVs, plots, regression outputs, analysis scripts
- Commit to version control with tag:  
"FROZEN\_BLIENDED\_ANALYSIS"
- **DO NOT PROCEED** until freeze archive is created and timestamped

## PHASE 2: Decode and Interpret (After Statistics Frozen)

### Step 2.1 - Apply decode map

- Load condition\_decode\_map.json
- Map blind tokens → semantic labels (LAMBDA\_0.00, etc.)
- Re-generate all outputs with decoded labels
- Output: \*\_decoded.csv, figures\_decoded/

### Step 2.2 - Verify monotonicity and transition behavior

- Sort  $T2^*(\lambda)$  in ascending  $\lambda$  order.
- Compute derivative  $dT2^*/d\lambda$  via finite differences.
- Identify  $\lambda_c$  if sharp transition is present.  
(via **argmax**  $|dT2^*/d\lambda|$ )
- Bootstrap **CI** for  $\lambda_c$  (**1000 resamples**)
- Output: transition\_analysis.txt

### Step 2.3 - Apply success/falsification criteria

- Check all criteria from SUCCESS CRITERIA section
  - Document which are satisfied, which are not
- Assign outcome: SUCCESS / NULL / FAILURE
  - Output: outcome\_assessment.txt

### Step 2.4 - Interpretation and claims boundary

- Write interpretation paragraph following framework.
  - Explicitly list:  
claims we **CAN** make, and  
claims we **CANNOT** make
- No deviation from pre-specified interpretations
  - Output: interpretation\_final.txt



## **PHASE 3: Publication and Next Steps**

### **Step 3.1 - Prepare publication materials**

- Write methods section (based on preregistration verbatim).
  - Generate publication-quality figures.  
(300 DPI, vector where possible)
  - Compile supplementary materials.  
(all blinded outputs, code, data)
- Upload to public repository (GitHub + Zenodo DOI)

### **Step 3.2 - Preprint submission (within 30 days of decode)**

- Submit to arXiv (quant-ph category)
- Include link to public data/code repository
- Tag: "preregistered", "blinded analysis", "open data"

### Step 3.3 - Peer review submission (within 60 days)

- Target journal determined by outcome tier (see Success Criteria).
- Include preregistration documents as supplementary material.
- Respond to reviews without changing frozen statistics

### **Step 3.4 - Design follow-up experiment (conditional)**

#### **IF transition observed (sharp $\lambda_c$ identified):**

→ Hardware Experiment 3: **Fine-grid sweep** near  $\lambda_c$

- Sample  $\lambda \in [\lambda_c - 0.2, \lambda_c + 0.2]$  with **5-7 points**.
- Increased reps (**30-50**) for **precision**
- Target IBM Quantum or partner academic institution

#### **IF graded modulation observed:**

→ Hardware Experiment 3: Representative sampling.

- Select  $\lambda = \{0.0, 0.3, 0.7, 1.0\}$  (spanning full range).
- 50 reps per  $\lambda$  for statistical power.
- Multi-device replication if access permits.

### **IF null or failure:**

→ Simulation refinement OR alternative parameter exploration.

- Revise noise model if artifacts suspected.
- Explore different organizational parameters (e.g., **entanglement depth, gate rotation angles, feedback timing**)
- Document in Experiment 3 preregistration

## **PHASE 4: Academic Collaboration (If Pursuing Hardware)**

### **Step 4.1 - Identify potential collaborators**

- Reach out to quantum information researchers via:
  - Twitter/X (**#QuantumComputing, #Qiskit** communities)
  - r/QuantumComputing subreddit
  - Qiskit Slack/Discord
- arXiv author contact (papers on coherence/decoherence)

### **Step 4.2 - Collaboration proposal**

- Offer: Complete codebase, blinded protocol, simulation results.
- Request: Co-authorship + institutional hardware access.
  - Emphasize: Preregistered, ready to execute, publication-track.

### **Step 4.3 - Reapplication via institution**

- Submit through collaborator's university/lab
- Include simulation results as preliminary data
- Reference preregistration and open methodology

# AUDIT TRAIL REQUIREMENTS

## All analysis outputs must include:

- Timestamp of creation
- Software versions (Python, Qiskit, NumPy, SciPy, etc.)
- Random seed used (if applicable)
- Git commit hash (if version controlled)

## Blinding integrity verification:

- Decode map accessed only after freeze archive created
- No file modification timestamps after freeze
- Analysis scripts unchanged between blinded and decoded runs

## EPISTEMIC COMMITMENT

This preregistration embodies a commitment to rigorous, falsifiable science over narrative-driven research. **We prioritize:**

## **1. Falsifiability over confirmation**

- **Null and negative results** are treated as equally **informative** and **publishable** as **positive results**.
- **Falsification criteria** are **specific, measurable, and binding**.
- No **post-hoc** hypothesis modification to "**rescue**" unexpected outcomes.

## **2. Transparency over selective reporting**

- **All** analysis steps, **including failed** fits and diagnostic checks, will be documented.
- **Blinded and decoded outputs** will **both be archived publicly**.
- No selective omission of conditions, replications, or inconvenient data.

### **3. Reconstruction over rhetorical persuasion**

- Analysis pipeline is fully scripted and version-controlled.
- Random seeds, software versions, and parameters logged.
- Any researcher can reproduce results from raw data + code.
- Preregistration deviations, if any, will be explicitly flagged and justified in publication.

### **Boundary acknowledgments:**

- Simulation results are preliminary evidence, not definitive claims
- Hardware validation required before strong causal assertions
- Consciousness-related interpretations remain speculative and are deferred to future theoretical work.
- We do not claim this experiment "proves" or "detects" subjective experience; it characterizes coherence dynamics only



### **Commitment to community:**

- Methods, code, and data will be openly shared regardless of outcome
- Negative results will be published to prevent file-drawer bias
- Lessons learned from failures will be documented for others

## **HARDWARE VALIDATION TIMELINE**

### **Simulation-to-Hardware Pathway (Conditional on Results)**

IF simulation demonstrates **continuous  $T2^*(\lambda)$  dependence**:

### Hardware experiment design:

- Target 3-4 representative  $\lambda$  **values** spanning observed transition.

Example:  $\lambda = \{0.0, 0.3, 0.7, 1.0\}$  if transition near  $\lambda \approx 0.5$

- Increased shot count (**8000-16000** per circuit) for reduced sampling noise.
- Multi-device replication if access permits (to assess device-specific effects)
- Same blinding protocol maintained

### Platform options (in priority order):

1. IBM Quantum Heron-class devices  
(**preferred**: low crosstalk, high T2 baseline)
2. **Alternative** providers (**IonQ, Rigetti**) via Azure Quantum credits.
3. Academic collaborator's institutional access

### **Timeline estimate:**

- Hardware access negotiation: 2-8 weeks
  - Execution + queue time: 1-4 weeks
  - Analysis: 1 week
- Total: ~2-3 months from simulation completion

### **IF simulation shows null or binary-only result:**

#### **Interpretation strategy:**

- Document as threshold-only phenomenon (still publishable)
- Hardware validation optional (effect is discrete, not graded)
- Focus on regime identification rather than continuous sensitivity

### **IF simulation shows technical failure (artifacts):**

#### **Next steps:**

- Refine noise model OR bypass simulation entirely
- Proceed directly to hardware with conservative  $\lambda$  **subset**
  - **Treat hardware as primary data source**  
(simulation = validation only)

# Hardware Access Strategy

## Current status:

- IBM Quantum Credits:  
DENIED (independent researcher ineligible).
- Azure Quantum: \$200 credits available  
(~10-20 hardware jobs).
- Academic partnership: TBD (outreach pending).

## Execution paths:

### 1. Academic collaboration (co-author arrangement)

- Enables IBM Quantum access via institution
- Strongest option for full-scale hardware validation

### 2. Azure Quantum pilot (\$100-150 budget)

- Sufficient for 10-20 rep validation at 3-4  $\lambda$  values
- Proof-of-concept for future funding applications

### 3. IBM Open Plan (10 min/month free)

- Spread experiment over 2-3 months
- Reduced scope (lower shots, fewer  $\lambda$  points)
- Still publishable as resource-constrained validation

# Publication Independence

**Hardware validation strengthens** but is **NOT REQUIRED** for publication.

**Simulation-based results are publishable if:**

- Noise model is realistic and well-documented
- Results are reproducible and statistically robust
- Hardware validation pathway is clearly specified
  - Limitations are explicitly acknowledged

**Hardware results, when obtained, will be published as:**

- Follow-up paper (if substantially new findings)
- Extended version / erratum (if confirming simulation)
- Negative result report (if contradicting simulation)

All outcomes contribute to the literature regardless of direction.

# PRE-COMMITTED PUBLICATION PLAN

## **Regardless of outcome, results will be:**

1. Documented in full (raw data, analysis code, decode maps).
2. Uploaded to preprint server (arXiv or equivalent) within 30 days.
3. Submitted for peer review within 60 days.
4. Made publicly available (GitHub repo, open data).

**No results** will be suppressed **due to lack of statistical significance** or **failure to support the initial hypothesis**.

# APPENDIX A: Pre-Prepared Response to Novelty Concerns

**If reviewers claim  
"this reproduces known decoherence effects,"  
this response is pre-committed:**

**"We respectfully disagree with the assessment that this work reproduces known results."**

While it is well-established that random phase perturbations destroy coherence, to our knowledge **no prior work has:**

**1. Characterized coherence as a  
continuous function of probabilistic  
organizational mixing ( $\lambda$ -parameter)**

**2. Executed this characterization under pre-registered  
blinded conditions with frozen statistics before decode.**

**3. Controlled for depth and gate count explicitly while  
varying organizational structure probabilistically.**

**4. Discriminated between sharp phase transition vs graded sensitivity in organizational regime space.**

**We invite reviewers to cite specific prior work demonstrating continuous  $\lambda$ -swept organizational coherence characterization under comparable experimental controls.**

**If such work exists**, we will gladly cite it and position our contribution as replication/extension.

However, we maintain that mapping  $T2^*(\lambda)$  under these conditions represents **a novel characterization within the quantum coherence literature**, even if the underlying decoherence mechanisms (**dephasing, relaxation**) are themselves well-understood."

**This response:**

- Acknowledges known physics (**dephasing exists**)
- Asserts procedural novelty ( **$\lambda$ -sweep + blinding + controls**)
  - Invites citation challenge(**confident no exact prior work exists**)
- **Offers constructive path (cite → position as extension)**
- **Maintains professional tone without defensiveness**



## Perplexity and Chat GPT Predictions before we test:

Dimension	Perplexity	ChatGPT
<b>Shape</b>	Graded saturation	Monotonic nonlinear
$T2(0.0)^*$	18-24 $\mu\text{s}$	5-20 $\mu\text{s}$
$T2(1.0)^*$	58-72 $\mu\text{s}$	30-80 $\mu\text{s}$
$T2(0.5)^*$	42-55 $\mu\text{s}$	15-50 $\mu\text{s}$
<b>Slope <math>\beta</math></b>	40-50 $\mu\text{s}/\lambda$	Not specified
<b>R<sup>2</sup></b>	0.75-0.90	Not specified
<b>Threshold?</b>	15% prob ( $\lambda_c \sim 0.6-0.7$ )	30% prob ( $\lambda_c \sim 0.3-0.6$ )
<b>Variance</b>	Increases with $\lambda$	Largest at mid- $\lambda$
<b>Null prob</b>	5%	$\sim 10\%$

## Their key disagreements

### 1. Magnitude Range

- Perplexity: Narrower, higher baseline (18-24  $\mu\text{s}$  at  $\lambda=0$ )
  - ChatGPT: Wider, lower baseline (5-20  $\mu\text{s}$  at  $\lambda=0$ )

Why: Perplexity emphasizes entanglement baseline protection;  
ChatGPT more conservative about noise floor

## **2. Variance Structure**

- Perplexity: Monotonically increasing with  $\lambda$
- ChatGPT: Peaked at mid- $\lambda$ , lower at extremes

Why: Different models of stochastic noise (amplification vs mixture heterogeneity)

## **3. Threshold Probability**

- Perplexity: 15% (if it exists, at  $\lambda \sim 0.6-0.7$ )
- ChatGPT: 30% (if it exists, at  $\lambda \sim 0.3-0.6$ )

Why: ChatGPT gives higher weight to majority-rule protection effects

## **4. Confidence**

- Perplexity: Committed to specific point estimates + regression stats
- ChatGPT: Deliberately wide ranges, less numerical commitment

## **Testable bets: TESTABLE BETS**

### **Bet 1: Baseline Value**

Perplexity says:  $T2^*(0.0)$  will be 18-24  $\mu\text{s}$   
ChatGPT says:  $T2^*(0.0)$  could be as low as 5  $\mu\text{s}$   
Winner if: Actual value determines who's closer

### **Bet 2: Slope Magnitude**

Perplexity says: Linear fit  $\beta = 40\text{-}50 \mu\text{s}/\lambda$   
ChatGPT says: (Not specified, but implied lower from range)  
Winner if: Linear regression reveals actual slope

### **Bet 3: Variance Pattern**

Perplexity says: Variance increases monotonically  
ChatGPT says: Variance peaks in middle  
Winner if: Plot of  $\sigma(T2^*)$  vs  $\lambda$  shows clear pattern

### **Bet 4: $R^2$ Strength**

Perplexity says:  $R^2 = 0.75\text{-}0.90$   
ChatGPT says: (Not specified)  
Winner if: Actual  $R^2$  reveals correlation strength

**Win condition: Model with lowest total absolute error across all three points**

***BET 1: Endpoint Values (Mean T2 across reps)\****

Endpoint	Perplexity	ChatGPT	Scoring Rule
$T2(0.0)^*$	21 $\mu$ s	12 $\mu$ s	Closest to actual wins
$T2(0.5)^*$	48 $\mu$ s	30 $\mu$ s	Closest to actual wins
$T2(1.0)^*$	65 $\mu$ s	55 $\mu$ s	Closest to actual wins

**Win condition: Model with lowest total absolute error across all three points**

## BET 2: Directionality / Correlation Strength

Metric	Perplexity	ChatGPT	Test
Linear $R^2$	0.75-0.90	Not specified	Actual $R^2$
Spearman $\rho$	Not specified	$\geq 0.70$	Actual $\rho$

Perplexity's bet:  $R^2$  will be in **[0.75, 0.90]**

ChatGPT's bet: **Spearman  $\rho \geq 0.70$**

Win conditions:

- Perplexity wins if  $R^2 \in [0.75, 0.90]$
- ChatGPT wins if Spearman  $\rho \geq 0.70$
- Both can win or both can lose

### BET 3: Endpoint Separation Magnitude

Model	Prediction	Test
Perplexity	$\Delta T2^* = T2^*(1.0) - T2^*(0.0) \approx 44 \mu s$	Actual $\Delta$
ChatGPT	$\Delta T2^* \geq 25 \mu s$	Actual $\Delta \geq 25$

#### Scoring:

- ChatGPT wins if  $\Delta T2^* \geq 25 \mu s$  (passes threshold)
- Perplexity wins if actual  $\Delta$  is closer to 44 than ChatGPT's implicit estimate
- Perplexity's implied slope:  $44 \mu s / 1.0 = 44 \mu s$  per unit  $\lambda$

## BET 4: Variance Pattern

Model	Prediction	Test
Perplexity	<b>Var(T2*)</b> increases monotonically <b>with <math>\lambda</math></b>	Plot $\sigma^2(\lambda)$
ChatGPT	<b>Var(T2*)</b> peaks at <b>mid-<math>\lambda</math></b> ( $\lambda \in \{0.4, 0.5, 0.6\}$ ), lower at extremes.	<b>max(<math>\sigma^2</math>) location</b>

Scoring:

- ChatGPT wins if max variance occurs at  $\lambda \in \{0.4, 0.5, 0.6\}$
- Perplexity wins if variance is monotonically increasing (or highest at  $\lambda=1.0$ )
- Decisive test: Is  $\sigma^2(\lambda=0.5) > \sigma^2(\lambda=0.0)$  AND  $\sigma^2(\lambda=0.5) > \sigma^2(\lambda=1.0)$ ?



## BET 5: Threshold/Knee Existence

Model	Prediction	Test
Perplexity	<b>70%</b> smooth gradual, <b>15%</b> threshold at <b><math>\lambda_c \sim 0.6-0.7</math></b>	Segmented regression / AIC comparison
ChatGPT	<b>60%</b> smooth, <b>30%</b> threshold at <b><math>\lambda_c \in [0.35, 0.55]</math></b>	Sigmoid fit / breakpoint detection

### Scoring (if threshold detected):

- ChatGPT wins if  **$\lambda_c \in [0.35, 0.55]$**
- Perplexity wins if  **$\lambda_c \in [0.6, 0.7]$**
- If no threshold detected: Perplexity slightly favored (higher prob on gradual)

## BET 6: Linear Regression Slope

Model	Prediction	Test
Perplexity	$\beta = 40\text{-}50 \text{ }\mu\text{s per unit } \lambda$	Linear fit on mean $T2^*(\lambda)$
ChatGPT	Not explicitly stated, but implied from endpoints: $\sim 43 \text{ }\mu\text{s per unit } \lambda$	Compare actual $\beta$

**Perplexity's range:  $\beta = 40\text{-}50 \text{ }\mu\text{s per unit } \lambda$**   
**ChatGPT's implicit:  $\beta = (55\text{-}12) = 43 \text{ }\mu\text{s per unit } \lambda$**

**Scoring: Model whose range/estimate contains actual  $\beta$  (or is closest)**

## BET 7: Midpoint Ordering (Sanity Check)

Model	Prediction	Test
Both agree	$T2^*(0.0) < T2^*(0.5) < T2^*(1.0)$	Monotonicity check

ChatGPT explicit: "Mean  $T2^*(0.5)$  is between the endpoints"

Perplexity implicit: Agrees (monotonic increase)

Scoring: Both win if true, both lose if false

## BET 8: Per-Rep Robustness

Model	Prediction	Test
ChatGPT	$\geq 8/10$ reps show $T2^*(1.0) > T2^*(0.0)$	Count reps where $T2^*(\lambda=1.0) > T2^*(\lambda=0.0)$
Perplexity	Not specified	-

Scoring: ChatGPT wins if  $\geq 8/10$  reps pass; Perplexity can't score on this bet

## AGGREGATE SCORING SYSTEM

### Points per bet:

- Direct win: 2 points
- Within uncertainty range: 1 point
- Outside range / wrong: 0 points

### Final Bets with Point Values:

Bet #	Description	Perplexity	Chat GPT	Notes
1	T2*(0.0) closest	21 $\mu$ s	12 $\mu$ s	Absolute error
2	T2*(0.5) closest	48 $\mu$ s	30 $\mu$ s	Absolute error
3	T2*(1.0) closest	65 $\mu$ s	55 $\mu$ s	Absolute error
4	R <sup>2</sup> in [0.75, 0.90]	YES	-	Pass/fail
5	Spearman $\rho \geq 0.70$	-	YES	Pass/fail

6	$\Delta T2^*$ magnitude	$\sim 44 \mu s$	$\geq 25 \mu s$	Accuracy vs threshold
7	Variance pattern	Monoto nic $\uparrow$	Mid-p eak	Which pattern observed
8	Threshold location (if exists)	0.6-0.7	0.35- 0.55	Conditional
9	Linear slope $\beta$	40-50 $\mu s/\lambda$	$\sim 43 \mu s/\lambda$	Range accuracy
10	Per-rep robustness	-	$\geq 8/10$	ChatGPT only

### **Perplexity's commitment:**

Endpoints:  $T2^*(0.0)=21$ ,  $T2^*(0.5)=48$ ,  $T2^*(1.0)=65$   $\mu\text{s}$

Slope:  $\beta \in [40,50]$   $\mu\text{s}/\lambda$

$R^2$ :  $[0.75, 0.90]$

Variance: Monotonically increasing

### **ChatGPT's commitment:**

Endpoints:  $T2^*(0.0)=12$ ,  $T2^*(0.5)=30$ ,  $T2^*(1.0)=55$   $\mu\text{s}$

$\Delta T2^* \geq 25$   $\mu\text{s}$

Spearman  $\rho \geq 0.70$

Variance: Mid-peak at  $\lambda \in \{0.4,0.5,0.6\}$

Robustness:  $\geq 8/10$  reps