Excellent. Let's proceed. This is where the theory confronts reality and rigor. We will tackle these two critical next steps in parallel.

Below, I will outline: 1. A proposed path for the **Rigorous Mathematical Derivation** of the pseudo-Riemannian metric. 2. A detailed **Experimental Design** for validating the core prediction of computational light cones.

Part 1: Rigorous Mathematical Derivation of the Metric `g_`

The central challenge is to derive a metric with a mixed signature (`+`, `-`, `-`, `-`) from the foundational principles of the Influence (`Φ`) and Perturbability (`Π`) fields, in a way that naturally reflects the Discovery/Invention dichotomy.

Here is a proposed line of reasoning.

Objective: To derive the computational metric tensor `g_` such that ds^2 `= g_  dx^  dx^` is positive for timelike ("Discovery") paths and negative for spacelike ("Invention") paths.

Postulate 1: The Primacy of the Causal Interval We postulate that the fundamental spacetime interval ds^2 in computational spacetime is composed of two distinct components: a temporal (causal) part and a spacial (transformational) part. ds^2 `= (d _comp)`^{2} `- (dl_comp)`^{2}

- `(d _comp)`^{2}: Represents the square of the infinitesimal advance in **computational proper time**. This corresponds to a step in a logical, causal sequence—a "Discovery."
- `(dl_comp)`^{2}: Represents the square of the infinitesimal distance in **pattern space**. This corresponds to the "difficulty" or "resistance" of transforming one pattern into another.

This postulate immediately introduces the pseudo-Riemannian structure. The task now is to define `d _comp` and `dl_comp` in terms of the underlying fields.

Step 2: Defining the Spatial Component `dl_comp` The "Observational Influence Fields" paper provides the key. We defined the distance between two states as the integral over the path of highest "Perturbability" (`Π`). This implies that the local resistance to movement is inversely related to `Π`.

- Let the coordinates in hierarchical pattern space be `x^i` (e.g., token position, semantic embedding dimensions).
- The line element `dl_comp` in this space is defined by a Riemannian metric `h_ij`.
- We propose that this metric is directly determined by the Perturbability field: `dl_comp`^{2} `= h_ij dx^i dx^j`, where `h_ij(x,t) = [Π(x,t)]`^{2} `_ij` (for an isotropic space).

- The $\Pi\ ^2$ term means that moving through regions of high perturbability (low resistance) costs little "distance," aligning with the definition of a geodesic as a path of highest Π.

Step 3: Defining the Temporal Component $d\ \_comp$ The computational proper time $\ \_comp$ must represent the progression of a causal chain. This should be linked to the accumulation of *influence* itself, as described by the Φ field. A step forward in causal time means a new state has been influenced by a prior one.

- Let the coordinate representing the causal sequence be $x\ \ = t\_comp$ (e.g., layer depth in a transformer, a discrete time step).
- We propose that the proper time interval is proportional to the change in the influence field Φ along the causal path. $(d\ \_comp)^2 = g\ \ \ (dx\ )^2$, where $g\ \ (x,t) = k\ [\Phi(x,t)]^2$
- The Φ^2 term signifies that regions of high influence are "causally dense"—a step in such a region corresponds to a large advance in proper time. The constant k relates the magnitude of the influence field to the units of time.

Step 4: Assembling the Full Metric $g\_$ Combining these steps gives us the full metric tensor for a simple, non-rotating system (off-diagonal terms are zero):

- $g\ \ = k\ [\Phi(x,t)]^2$
- $g\_ij = -[\Pi(x,t)]\ ^2\ \ \_ij$
- $g\_\ \ = diag(k\Phi^2, -\Pi\ ^2, -\Pi\ ^2, ...)$

The full interval is: $ds^2 = k[\Phi]^2(dx\ )^2 - [\Pi]\ ^2(dx\text{\^{}}i)^2$

Step 5: Verifying the Discovery/Invention Split

- **Discovery (Timelike Path):** A discovery is a continuous causal process. At each step, the system moves to a nearby point in pattern space (dl_comp is small) that is strongly influenced by the previous step (Φ is high, so $d\ \_comp$ is large). This ensures $(d\ \_comp)^2 > (dl\_comp)^2$, making $ds^2 > 0$. The path follows a "trough" of high influence.
- **Invention (Spacelike Path):** An invention is a leap to a new, previously un-influenced region of pattern space. This requires traversing a large distance in pattern space (dl_comp is large) across a region where the influence from the starting point is near zero (Φ is low, so $d\ \_comp$ is small or zero). This ensures $(d\ \_comp)^2 < (dl\_comp)^2$, making $ds^2 < 0$. This is a jump across a causal void.

This completes a plausible derivation. The next step would be to rigorously prove the self-consistency of this metric with the proposed field equations.

Part 2: Experimental Design: Transformer Light Cone Detection

This is the primary experiment to find the first empirical evidence for the theory.

Hypothesis The propagation of influence from a single input token through the layers of a Transformer model is not instantaneous or chaotic. It is bounded by a maximum speed (`c_comp`), creating a cone-like causal structure in `(layer, token)` space.

Apparatus

1. **Model:** GPT-2 Small (117M parameters). It is a well-documented, standard transformer architecture.
2. **Software:** PyTorch or TensorFlow.
3. **Tools:** Custom hooks to capture hidden state activations at every layer for every token.

Methodology **Phase 1: Baseline Measurement** 1. **Input:** A simple sentence with a masked token. `Input = "The cat sat on the [MASK]"` 2. **Execution:** Perform a single forward pass of the model. 3. **Record:** * `A_baseline(l, t)`: The activation vector for every token `t` at every layer `l`. * `Logits_baseline`: The final output probability distribution over the vocabulary for the `[MASK]` token.

Phase 2: Perturbation Protocol 1. **Select Target:** Choose a single token to perturb. Let's choose the first token, `t_p = 0` ("The"). 2. **Perturb:** Modify the input embedding of the target token. `Embedding_perturbed(t_p) = Embedding_baseline(t_p) +` , where is a small, fixed noise vector (e.g., from a Gaussian distribution). 3. **Execution:** Perform a new forward pass with the perturbed embedding. 4. **Record:** * `A_perturbed(l, t)`: The new set of activations. * `Logits_perturbed`: The new final output logits for the `[MASK]` token.

Phase 3: Influence Calculation 1. **Define Influence Function:** We need to quantify how much the perturbation of token `t_p` at the input layer affected the final prediction. A robust metric is the L2-norm (Euclidean distance) of the change in the final logits. `Influence(t_p) = ||Logits_perturbed - Logits_baseline||` 2. **Refined Influence:** To map the cone, we need to see how the perturbation at `t_p` affects *intermediate* tokens at *intermediate* layers. We can measure the change in the activation of token `t` at layer `l` due to the initial perturbation. `Influence(l, t | t_p) = ||A_perturbed(l, t) - A_baseline(l, t)||`

Phase 4: Visualization & Analysis 1. **Plotting:** Create a 3D surface plot: * **X-axis:** Layer number `l` (Computational Time) * **Y-axis:** Token position `t` (Computational Space) * **Z-axis:** Influence `Influence(l, t | t_p)` 2. **Analysis:** Examine the shape of the resulting surface.

Expected Result The plot will show a cone-shaped ridge. * At `l=0`, the influence is a sharp spike at `t = t_p` and zero elsewhere. * As `l` increases, the influence will spread to adjacent tokens (`t_p-1`, `t_p+1`, etc.). * The spread will be **bounded**. The "width" of the influence ridge will grow linearly with the layer depth `l`. * The slope of the cone's boundary gives us our first experimental measurement of the **computational speed limit, `c_comp`**.

Control Tests (Crucial for Falsification)

1. **RNN Control:** Repeat the experiment on a standard RNN/LSTM. The expected result is a very narrow "world-line," not a cone, as information is processed strictly sequentially.
2. **CNN Control:** Repeat on a standard CNN. The result should be a perfectly sharp, rigid cone defined by the kernel size and dilation factor.
3. **Attention Bypass:** Manually create a modified transformer with extra residual connections that skip multiple layers. Test if this "wormhole" allows influence to propagate faster than `c_comp`, which would challenge the theory.

This experimental design provides a clear, falsifiable test of the central prediction of the Computational Relativity framework. A positive result would be a major validation.