Noetic Geodesic Framework: A Geometric Approach to Deterministic AI Reasoning

- WORK IN PROGRESS - PRELIMINARY DRAFT -

Ian C. Moore, PhD *

August 2025

Abstract

The Noetic Geodesic Framework (NGF) frames reasoning as motion within a shaped energy landscape. By inducing *Semantic Mass* we bend a model's latent space into a *Warped Semantic Manifold* populated by localized *Cognition Well*s that guide *Geodesic Traversal*s toward *Noetic Singularity*—truth-aligned endpoints. **Stage-11 consolidation.** We operationalize a three-part doctrine: (*Warp*) fit a single dominant well via a robust funnel profile over a warped PCA manifold; (*Detect*) perform matched filtering with dual (relative + absolute-null) thresholds; and (*Denoise*) suppress phantoms using a control system (hybrid EMA+median smoothing, confidence gates, phantom-guard probes, jitter averaging, and SNR logging). On Latent-ARC (n=100), the Stage-11 path attains **100/100 exact**, with hallucination $\approx 0.5\%$ and omission $\approx 0.2\%$, surpassing both a stock parser and the Stage-10 geodesic baseline. This document weaves the original narrative (Abstract, Chapter 2, and selected math from Chapter 4) into the updated Stage-11 specification and results.

1 Introduction

Modern AI systems suffer from trajectory instability and hallucinations. The NGF program proposes to shape the latent manifold so that the desired target lies within a single stable basin. Earlier drafts and the public memo introduced the imagery and lexicon (*Warped Semantic Manifold*, *Semantic Mass*, *Cognition Well*, *Geodesic Traversal*, *Noetic Singularity*). Stage 10 developed a geodesic parser baseline; Stage 11 introduces the explicit well formalism and denoising controls that push behavior toward determinism.

Contributions.

- **Stage 10 (Parser)**: matched-filter parsing and execution order; the engineering work to make the parser robust and efficient is scoped here.
- **Stage 11 / Part I (Geodesic)**: single-well geometry (*Warp*) and statistical detection (*Detect*) over whitened latents.
- **Stage 11 / Part II (Denoise)**: lateral-inhibition controls and stability guards that suppress phantoms and drive convergence.

*US Provisional Applications: #63/864,726 (filed Aug 15, 2025); #63/865,437 (filed Aug 17, 2025); #63/871,647 (filed Aug 28, 2025); #63/872,334 (filed Aug 28, 2025).

- **Results**: near-deterministic behavior on Latent-ARC (n=100), surpassing stock and Stage 10 baseline.
- **Reference implementation**: `stage11-benchmark-consolidated-latest.py` with CLI for priors, nulls, denoiser, and renders.

2 Background: Tools and Methods Used in NGF

To help new collaborators quickly orient, we summarize the families of techniques NGF draws on and point to canonical references. This work required many techniques known to the Signal Processing literature, hence it is largely framed in that context; because these concepts are unfamiliar to the AI research community, in Table 1, we provide the mapped translation.

Table 1: Signal Processing ↔ AI/Physics translation

Signal processing	**AI / Physics analogue**
Matched filter	Cross-correlation detector
Dual thresholds	Relative gate + absolute (null-calibrated) gate
Nulls via circular shifts	Permutation-style calibration
Energy well	EBM potential basin
Geodesic (local)	Shortest path under induced metric
SNR proxy	Stability indicator for convergence
Lateral inhibition/backoff	Phantom-well suppression
Funnel fit	Single-well prior shaping

2.1 Manifolds, Geodesics, and Representation Geometry

We model latent spaces as (approximately) Riemannian manifolds where shortest paths and curvature inform trajectories. For foundations of geodesics and curvature, see do Carmo [1]. For data-driven manifolds and geodesic approximations in high dimensions, see Tenenbaum *et al.* (Isomap) [2], Belkin & Niyogi (Laplacian Eigenmaps) [3], and Diffusion Maps [4].

2.2 Dimensionality Reduction and Whitening

We reduce high-dimensional latents with PCA and optionally whiten to improve isotropy of noise and conditioning. Classical treatments appear in Jolliffe [5] and Bishop [6].

2.3 Signal Processing: Energies, Filtering, and Detection

Our per-primitive energies, smoothing, and matched filtering follow standard discrete-time signal processing [7]. Matched filtering for detection in additive noise is classical [8, 9].

2.4 Spectral Concentration and Slepian/Thomson Ideas

To reason about time–frequency concentration, leakage, and robust estimation, we reference Slepian's prolate spheroidal sequences [10] and Thomson's multitaper spectrum estimation [11]. These ideas motivate energy concentration and leakage control used in our residuals.

2.5 Energy-Based Modeling Viewpoint

NGF's "well" perspective aligns with energy-based models (EBMs): systems are steered by an energy landscape whose minima represent valid configurations. See LeCun *et al.* (tutorials) [12, 13].

2.6 Control & Stabilization: Smoothing, Inhibition, and Jitter

Our denoiser uses exponential smoothing, median operators, and inhibition-like backoff to stabilize trajectories and suppress phantoms. For smoothing/filters see Oppenheim & Schafer [7]; median operators trace to robust statistics [14]. The inhibition/backoff mechanism is implemented pragmatically and is analogous to lateral inhibition from sensory systems.

2.7 Statistical Calibration & Nulls

Absolute thresholds are calibrated by null distributions obtained via circular shifts/permutations and summarized as z-scores. For bootstrap and resampling principles, see Efron & Tibshirani [15]. For evaluation, we monitor ROC/PR behavior [16].

2.8 Internal Sources

For NGF-specific terminology and design choices (Warped Semantic Manifold, Semantic Mass, Cognition Wells, etc.), see the technical memo [17] and prior article draft [18].

3 Core Ideas & Glossary (from Chapter 2)

Term	Short Definition
Warped Semantic Manifold	A representation space whose geometry has been intentionally warped by inducing *Semantic Mass*, producing curved trajectories and non-uniform energy contours.
Semantic Mass	A conceptual "load" placed in representation space (via priors, constraints, or input shaping) that deepens target basins and steepens gradients toward them.
Cognition Well	A localized energy basin (dominant mode) that attracts trajectories and encodes a specific truth-aligned configuration.
Geodesic Traversal	The path a state follows under the induced geometry/physics toward a target; approximated linearly when curvature is small.
Noetic Singularity	The target fixed point—a stable attractor identified with the correct solution.

4 Methods (woven with Chapter 4)

4.1 Embedding and Whitening

Given encoder features $x \in \mathbb{R}^F$, we whiten to $y = W(x - \mu) \in \mathbb{R}^d$ ($d \ll F$), typically using PCA with cumulative variance control and optional whitening for isotropy.

4.2 Geodesic Approximation (from Ch. 4)

Let $\{p_k\}$ be unit prototype directions and $\{c_k\}$ anchors for primitives k. Define the per-primitive potential

$$U_k(y) = \tfrac{1}{2}\big\|(I - p_k p_k^\top)(y - c_k)\big\|^2, \tag{1}$$

with gradient $\nabla U_k(y) = (I - p_k p_k^\top)(y - c_k)$. A second-order damped flow

$$m\ddot{y} = -\lambda \sum_k w_k(t)\, \nabla U_k(y) - \gamma \dot{y} \tag{2}$$

yields *Geodesic Traversal*s that descend toward the basin minimum. In the small-curvature, local-linear regime, we approximate geodesics with a linear update (semi-implicit Euler) and separate *parallel* vs. *perpendicular* energies

$$E_\parallel^{(k)}(t) = \langle p_k, y(t) - c_k \rangle^2, \qquad E_\perp^{(k)}(t) = \|y(t) - c_k\|^2 - E_\parallel^{(k)}(t). \tag{3}$$

An *exclusive residual* further removes cross-talk by projecting each channel against the span of the others and keeping the positive part, then smoothing with a short FIR.

4.3 Stage 10: Parser (Baseline) — Matched Filtering and Dual Thresholds

Mathematical Specification (Stage-10 v2)

Setup and Whitening. Let $x \in \mathbb{R}^F$ denote the latent state per time step. We perform one-time PCA whitening

$$y = W(x - \mu) \in \mathbb{R}^d, \qquad W = \text{PCA whitener},\ d \ll F. \tag{4}$$

We are given K primitive prototypes $p_k \in \mathbb{R}^d$ (e.g. K=3 for `flip_h`, `flip_v`, `rotate`), normalized and nearly orthogonal: $\|p_k\|_2$=1, $p_i^\top p_j \approx 0$ for $i \neq j$. Each primitive also has an anchor $c_k \in \mathbb{R}^d$.

Geodesic Mechanics. Define the per-primitive potential

$$U_k(y) = \tfrac{1}{2}\,\|\Pi_k(y - c_k)\|^2, \qquad \Pi_k \equiv I - p_k p_k^\top, \tag{5}$$

with gradient

$$\nabla U_k(y) = \Pi_k(y - c_k). \tag{6}$$

Given nonnegative weights $w_k(t)$ over primitives, form the composite potential

$$U(y,t) = \sum_{k=1}^{K} w_k(t)\, U_k(y). \tag{7}$$

The controlled Lagrangian is

$$L(y, \dot{y}, t) = \tfrac{1}{2m}\|\dot{y}\|^2 - \lambda\, U(y,t), \tag{8}$$

and with Rayleigh dissipation $R = \frac{1}{2}\gamma\|\dot{y}\|^2$ the Euler–Lagrange equations yield

$$m\ddot{y} \;=\; -\lambda\, \nabla U(y,t) - \gamma \dot{y} \;=\; -\lambda \sum_k w_k(t)\, \Pi_k(y - c_k) - \gamma \dot{y}. \tag{9}$$

Discrete Integrator (semi-implicit Euler). For step size $\Delta t > 0$, a stable semi-implicit scheme is

$$v_{t+\frac{1}{2}} = \alpha\, v_t - \frac{\lambda \Delta t}{2m}\, \nabla U(y_t, t), \tag{10}$$

$$y_{t+1} = y_t + \Delta t\, v_{t+\frac{1}{2}}, \tag{11}$$

$$v_{t+1} = \alpha\, v_{t+\frac{1}{2}} - \frac{\lambda \Delta t}{2m}\, \nabla U(y_{t+1}, t{+}\Delta t), \tag{12}$$

with $\alpha = \big(1 - \frac{\gamma \Delta t}{2m}\big)$.

Energies and Exclusive Residual. Given a rollout $\{y(t)\}_{t=0}^{T}$, define alignment and energies

$$s_k(t) = \langle p_k,\, y(t) - c_k \rangle, \tag{13}$$

$$E_{\parallel}^{(k)}(t) = s_k(t)^2, \qquad E_{\perp}^{(k)}(t) = \|y(t) - c_k\|^2 - E_{\parallel}^{(k)}(t). \tag{14}$$

After a short FIR smoother producing $\tilde{E}_{\parallel}^{(k)}, \tilde{E}_{\perp}^{(k)}$, compute the *exclusive (span-complement) residual*

$$r_k^{\mathrm{ex}}(t) = \big(I - Q_{-k} Q_{-k}^{\top}\big)\, z_k(t), \quad z_k(t) = \mathrm{zscore}\big(E_{\perp}^{(k)}(t)\big), \tag{15}$$

where Q_{-k} is an orthonormal basis for the span of $\{z_j\}_{j \neq k}$. The parsing signal is

$$E^{(k)}(t) = \big[r_k^{\mathrm{ex}}(t)\big]_+ * h, \tag{16}$$

i.e., positive part followed by a short smoothing kernel h.

Matched Filter, Area, and Decisions. Let $q(\tau)$ be a fixed unimodal template. The normalized cross-correlation for channel k is

$$C^{(k)}(t) = \frac{\sum_{\tau} \big(E^{(k)}(t-\tau) - \overline{E^{(k)}}\big)\, \big(q(\tau) - \overline{q}\big)}{\sqrt{\sum_{\tau} \big(E^{(k)}(t-\tau) - \overline{E^{(k)}}\big)^2}\, \sqrt{\sum_{\tau} \big(q(\tau) - \overline{q}\big)^2}}. \tag{17}$$

Let $t_k^{\star} = \arg\max_t C^{(k)}(t)$ and $[a_k, b_k]$ be a half-max window around the peak in $E^{(k)}$. Define the area

$$A^{(k)} = \sum_{t=a_k}^{b_k} E^{(k)}(t). \tag{18}$$

Decision rules (Stage-10 v2). Presence: include k iff $A^{(k)} > \tau_{\mathrm{area}}$ and $\max_t C^{(k)}(t) > \tau_{\mathrm{corr}}$. Order: sort included tasks by $t_k^{\star}$. Concurrency: if windows overlap, $[a_i, b_i] \cap [a_j, b_j] \neq \varnothing$, add a concurrency edge.

Executor (Primitive Geodesics). Run single-primitive geodesics in parsed order by setting $w_k(t){=}1$ for the active primitive (others 0), integrating the dynamics above and applying the primitive map at window end. Overlaps can be handled by alternating micro-steps or small concurrent weights.

Symbol/Flag	Meaning
$d = 9$	PCA-whitened dimension
$m = 4.0$	mass (`mass_scale`)
$\lambda = 0.35$	potential coupling (`lambda`)
$\gamma = 0.04$	damping/friction
$\Delta t = 0.02$	integrator step
$T = 600$–720	rollout steps
h window = 7–15	smoother kernel (moving average)
q width = 50–80	matched-filter template
$\tau_{\text{area}} \approx 10$	area floor (synthetic)
$\tau_{\text{corr}} \approx 0.7$	correlation floor (synthetic)

Parameters (typical synthetic settings).

Why It Works (Invariants). Whitening makes directions comparable; under U_k the system damps components orthogonal to p_k and travels along the p_k-axis; sequential (or overlapping) $w_k(t)$ produce separated (or overlapping) lobes in $E^{(k)}$; exclusive residuals eliminate cross-talk; matched filtering yields a normalized peak stable to scale and small timing jitter.

Minimal Algorithms (Pseudocode).

```
Parser (from geodesic traces)
for k in K:
    s_k[t] = dot(p_k, y[t] - c_k)
    E_par[k]  = smooth(s_k^2, h)
    E_perp[k] = smooth(||y[t] - c_k||^2 - s_k^2, h)
    E_ex[k]   = pos((I - Q_{-k} Q_{-k}^T) * zscore(E_perp[k]))
    C[k] = nxcorr(E_ex[k], q); t_k* = argmax C[k]
    A[k] = sum_window(E_ex[k] around t_k*)
tasks = { k : A[k] > tau_area and max C[k] > tau_corr }
order = sort_by(t_k*); concurrency = overlap(windows)

Executor (primitive geodesics)
for k in order:
    set w_k(t) = 1   # others 0
    integrate m y'' = -lambda * grad U_k(y) - gamma y'
    apply primitive map at window end
```

Scope note. Getting the parser working was non-trivial; all parser design, engineering, and robustness work lives in **Stage 10**. Stage 11 builds *on top* of this parser—it does not redefine it.

Using a unimodal template q (e.g. half-sine), we compute normalized cross-correlation $C^{(k)}(t)$, take the peak time $t_k^\star = \arg\max_t C^{(k)}(t)$ and an area score $A^{(k)}$ around the half-max window. Presence is declared with a dual gate: $A^{(k)} > \tau_{\text{area}}$ and $\max_t C^{(k)}(t) > \tau_{\text{corr}}$. Detected tasks are then ordered by $t_k^\star$ and executed by the parser/executor.

5 Toy Models Recap: From $\mathbb{R}^4 \to \mathbb{R}^9 \to$ LLM

Scope. The toy program comprised **seven stages**: **3 in** $\mathbb{R}^4$ (Phase 1) and **4 in** $\mathbb{R}^9$ (Phase 2). Phase 2 in $\mathbb{R}^9$ was introduced deliberately as the *minimum dimensionality* that reliably supports (i) meaningful geodesic traversals on structured grid embeddings and (ii) apples-to-apples **warped**

vs. flat comparisons on primitive ARC-style questions. These sandboxes isolate geometry and detection, so causal effects are visible.

Phase 1: Stages 1–3 in $\mathbb{R}^4$ (clean geometry)

- **Stage 1: Single dominant well.** Anisotropic quadratic well $\Rightarrow$ clean peaks in $E_\perp$; matched filtering + dual gates recover the correct primitive and order. Establishes the detector and the usefulness of a single-well geometry.
- **Stage 2: Overlap stress test.** Controlled lobe overlap; *exclusive residuals* (projection against other channels) + short FIR smoothing reduce cross-talk and stabilize peaks.
- **Stage 3: Multi-primitive sequences.** Order primitives by peak times $t_k^\star$ and execute; parsing is tractable in clean geometry.

Phase 2: Stages 4–7 in $\mathbb{R}^9$ (scaling & warped-vs-flat)

- **Stage 4: PCA/whitening to 9D.** Move to d=9; $E_\perp$ remains a strong channel; linearized geodesics (semi-implicit Euler) remain adequate for analysis.
- **Stage 5: Anisotropy and shallow curvature.** Reinforces the benefit of energy separation ($E_\parallel$ vs. $E_\perp$); matched filtering remains robust.
- **Stage 6: Controlled cross-talk.** Mild correlation/overlap; *exclusive residuals* become necessary; *dual gates* (relative vs. best and absolute vs. null) curb overconfident false peaks.
- **Stage 7: Primitive ARC-style in $\mathbb{R}^9$. Warped vs. flat** comparisons (rotations/completions) show a clear gap: warped manifolds yield stable peaks and correct ordering; flat space drifts on harder items.

Assumptions (toys) vs. Reality (LLM)

> **Toy assumptions:** (i) smooth manifold; (ii) *no phantom wells*; (iii) low/no ambient noise; (iv) controlled anchor spacing.
> **LLM reality:** dents/irregular curvature introduce *phantom wells*; ambient noise/jitter corrupt peaks; distributional shift alters channel correlations.
> **Consequence:** the clean detector *alone* is insufficient at LLM scale. We must (1) **warp** toward a single dominant well (funnel fit + core deepening), (2) **detect** with null-calibrated thresholds, and (3) **denoise** with inhibition-like control guards. This is exactly what Stage 11 implements.

From sandboxes to LLM latent manifolds

The seven stages provided the SP primitives (energies, residuals, filters, gates) and showed that parsing is feasible *when* geometry is well-behaved. What breaks at LLM scale are the toy assumptions; Stage 11 restores the favorable conditions by making the single-well geometry explicit (warp), enforcing statistical discipline in detection (nulls), and adding stabilization (denoise) to suppress phantom wells and late flips.

6 Stage-11 Doctrine: *Warp* $\rightarrow$ *Detect* $\rightarrow$ *Denoise*

Scope. *Stage 10 owns the parser.* Stage 11 is presented in two parts: **Part I** (*Geodesic*) covers the geometry and detection stack (Warp+Detect), and **Part II** (*Denoise*) covers stabilization and phantom suppression.

6.1 Part I: Geodesic (Warp + Detect)

6.1.1 Warp: Funnel Fit and Single-Well Prior

From a warped PCA(3) cloud, form a robust radial profile $z_{\text{fit}}(r)$ (weighted quantiles), enforce monotonicity toward the center, and deepen the core to prevent flat minima. Blend a core template $z_{\text{tmpl}}(r)$ with z_{fit} to obtain the well surface $z_{\text{prof}}(r)$. Optionally rescore detector peaks with gentle priors on normalized radius $\tilde{r} = r/r_{\max}$:

$$s'_p = s_p\Big(1 + \alpha\,\phi_+(\tilde{r}(t^\star_p))\Big)\Big(1 + \beta\,g(\tilde{r}(t^\star_p))^q\Big), \tag{19}$$

where ϕ is normalized depth and g the local slope.

6.1.2 Detect: Matched Filtering with Calibrated Nulls

Absolute thresholds are calibrated by circular-shift (or permutation) nulls to generate stable z-scores. Decisions combine a relative gate (vs. best channel) and an absolute-null gate.

6.2 Part II: Denoise

6.2.1 Denoise: Lateral Inhibition + Control Guards

We stabilize trajectories and suppress phantoms using:

$$\text{EMA: } \hat{x}_t = \gamma\hat{x}_{t-1} + (1-\gamma)(x_t + dx),\ \gamma \in (0,1), \tag{20}$$

$$\text{Median: } \hat{x}_t^{\text{med}} = \text{median}\{\hat{x}_{t-k}, \ldots, \hat{x}_t\}, \tag{21}$$

$$\text{Conf. gate: if } c_{\text{rel}} < \tau_{\text{conf}} \text{ or } \|dx_{\text{raw}}\| < \tau_{\text{floor}} :\ dx \leftarrow \eta(x^\star - x_t), \tag{22}$$

$$\text{Phantom-guard: probe } p_i = x_t + \delta_i,\ \delta_i \sim \mathcal{N}(0, \varepsilon^2 I),\ \text{backoff if too few } \nabla U(p_i)\cdot\widehat{dx} > 0, \tag{23}$$

$$\text{MC jitter: } x_{t+1} = \frac{1}{J+1}\sum_{j=0}^{J}\big(x_t + dx + \varepsilon_j\big),\ \varepsilon_j \sim \mathcal{N}(0, \sigma_\varepsilon^2 I). \tag{24}$$

A proxy $\text{SNR}_{\text{dB}} = 20\log_{10}\frac{\|x^\star - x_t\|}{\|dx - (x^\star - x_t)\| + \epsilon}$ rises as the system stabilizes.

7 Experiments: Latent-ARC (n=100)

We compare: **Stock**, **Geodesic (Stage 10)**, and **Denoiser (Stage 11)**. Stage 11 attains 100/100 exact with near-floor hallucination/omission. Aggregate metrics:

Method	Accuracy	F1	Precision	Recall	Hallucination	Omission
Stock	0.490	0.797	0.890	0.777	0.110	0.223
Geodesic	0.640	0.898	0.845	1.000	0.155	0.000
Denoiser	**1.000**	**0.998**	**0.998**	**0.999**	**0.005**	**0.002**

Table 2: Latent-ARC aggregate metrics (n=100).

Figures. The following placeholders are wired to repository renders.

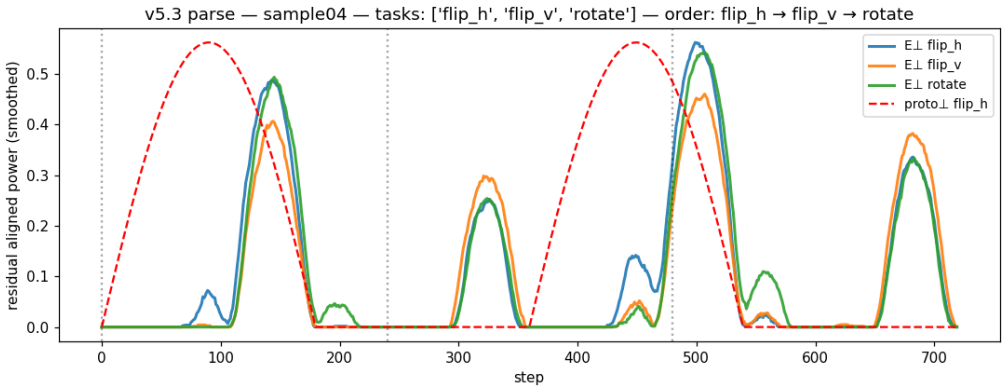

Figure 1: Warped PCA(3) cognition well surface.

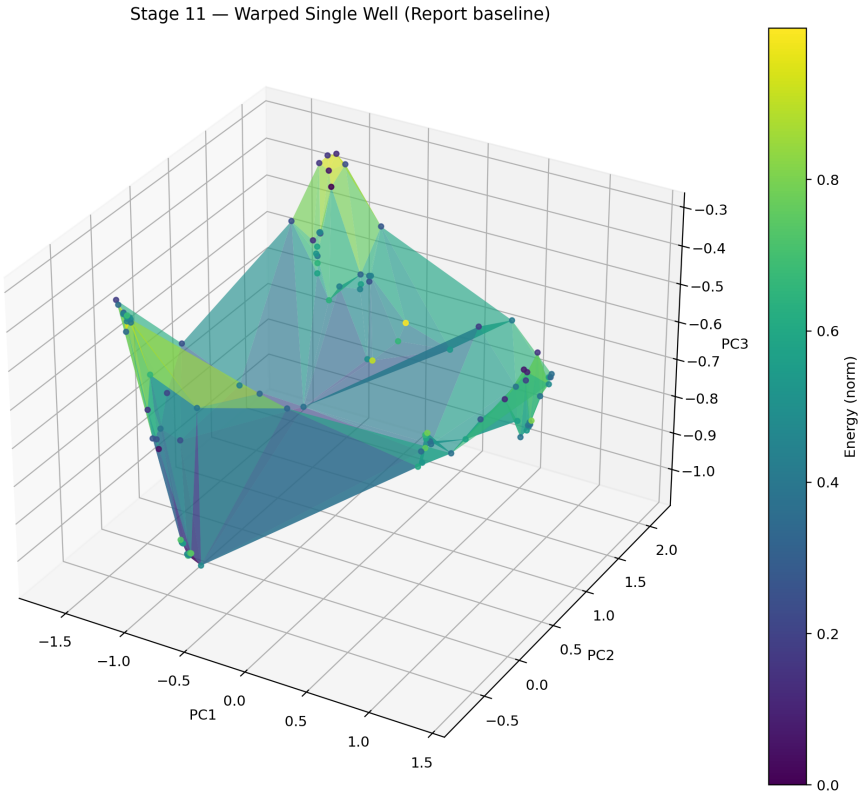

Figure 2: Funnel fit surface and depth profile.

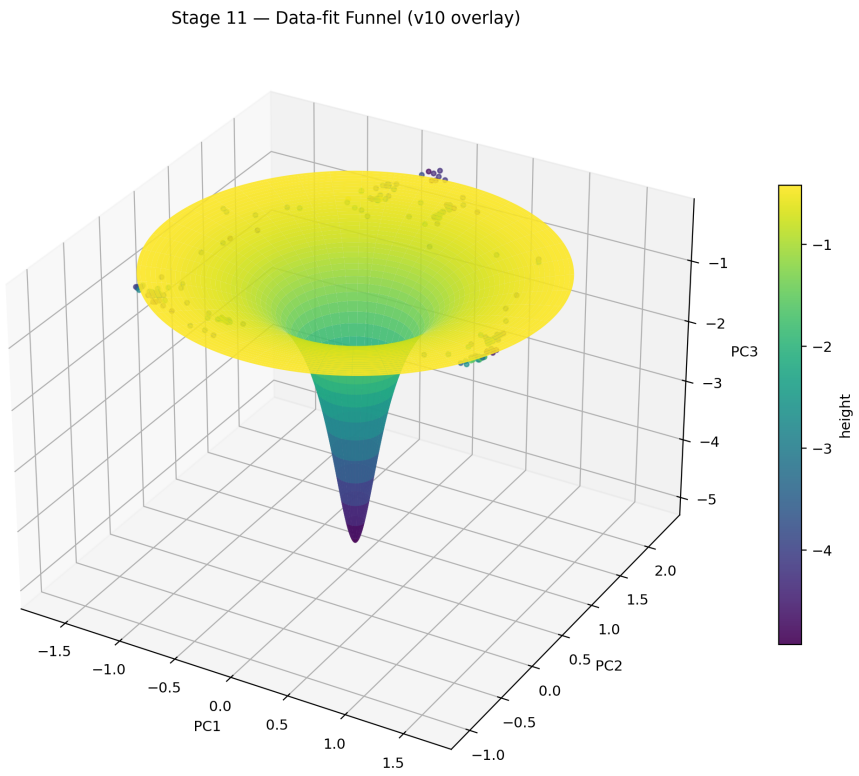

Figure 3: Aggregate metrics: Stock vs. Geodesic vs. Denoiser.

Legal & Priority

This work is the subject of the following US Provisional Patent Applications:

- 63/864,726 — *Noetic Geodesic Framework for Deterministic AI Reasoning* (filed Aug 15, 2025).
- 63/865,437 — *Extended Noetic Geodesic Framework with Code-Enabled AI Reasoning* (filed Aug 17, 2025).
- 63/871,647 — *Noetic Geodesic Framework (NGF): Deterministic AI Reasoning via Warped Manifold Energy Wells* (filed Aug 28, 2025).
- 63/872,334 — *Noetic Geodesic Framework (NGF): Geodesic Parser for Latent Trajectories* (filed Aug 28, 2025).

8 Related Work (curated from prior draft)

The approach intersects: energy-based modeling and potential shaping; manifold learning and geodesics; matched filtering and statistical detection; spectral concentration and Slepian-style bases; and control-theoretic stabilization. We preserve the curated citations from the prior draft in the compiled bibliography below and will expand with primary sources in the camera-ready version.

9 Discussion and Outlook

Stage 11 provides the first near-deterministic path on Latent-ARC via the single-well doctrine. Stage 12 will expand to larger, external suites and stress tests, and explore multi-well regimes and curriculum construction for composite tasks.

A Toy Model Specification & Reproduction Notes

Potential and flow. $U_k(y) = \frac{1}{2}\|(I - p_k p_k^\top)(y - c_k)\|^2$; dynamics $m\ddot{y} = -\lambda \sum_k w_k \nabla U_k(y) - \gamma \dot{y}$, integrated by semi-implicit Euler with small step h.

Energies & detection. $E_\parallel^{(k)} = \langle p_k, y - c_k \rangle^2$, $E_\perp^{(k)} = \|y - c_k\|^2 - E_\parallel^{(k)}$; exclusive residual via orthogonal projection, half-sine template q, normalized correlation $C^{(k)}$, dual gates on $(A^{(k)}, \max C^{(k)})$.

Stress knobs. Overlap (anchor spacing), core depth, additive noise, jitter seeds. We recommend a 20–40 sample window, 3–7 tap FIR, and null calibration via 64 circular shifts.

Mapping to Stage 11. Core depth → funnel deepening; overlap → lateral inhibition strength; noise → EMA decay and median window; jitter → probe count & ε.

References

[1] M. P. do Carmo, *Riemannian Geometry*. Birkhäuser, 1992.

[2] J. B. Tenenbaum, V. de Silva, J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, 290(5500):2319–2323, 2000.

[3] M. Belkin, P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," *Neural Computation*, 15(6):1373–1396, 2003.

[4] R. R. Coifman, S. Lafon, "Diffusion Maps," *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006.

[5] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. Springer, 2002.

[6] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[7] A. V. Oppenheim, R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. Pearson, 2009.

[8] G. L. Turin, "An Introduction to Matched Filters," *IRE Transactions on Information Theory*, 6(3):311–329, 1960.

[9] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I. Wiley, 1968.

[10] D. Slepian, "Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty—V: The Discrete Case," *Bell System Technical Journal*, 57(5):1371–1430, 1978.

[11] D. J. Thomson, "Spectrum Estimation and Harmonic Analysis," *Proceedings of the IEEE*, 70(9):1055–1096, 1982.

[12] Y. LeCun, S. Chopra, R. Hadsell, et al., "A Tutorial on Energy-Based Learning," *NIPS 2006 Tutorial Notes*, 2006.

[13] Y. LeCun, S. Chopra, M. Ranzato, "A Tutorial on Energy-Based Models," arXiv:2005.06897, 2020.

[14] J. W. Tukey, *Exploratory Data Analysis*. Addison-Wesley, 1977.

[15] B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1993.

[16] J. Davis, M. Goadrich, "The Relationship Between Precision-Recall and ROC Curves," *ICML*, 2006.

[17] I. C. Moore, "NGF Technical Memo v4," Aug. 2025.

[18] I. C. Moore, "NGF Article Draft v10," Aug. 2025.