INVENT THE QUESTIONS, COMPILE THE CONSEQUENCES: LEARNING EXECUTABLE STRUCTURE FROM TERMINAL FEEDBACK

Anonymous authors
Paper under double-blind review

ABSTRACT

Can terminal feedback reveal not only a computation's outcome, but also the structure needed to execute and revise it? We study replayable tool systems with hidden states, unknown action targets, and no intermediate observations. Under explicit projective and identifiability conditions, we recover action-aligned coordinates, call laws, and functional dependencies from experiments that each return a scalar. Our central construction uses measured action relations to align state decoding and tool dynamics in one representation. For type-A actions, we recover matrix multiplication from a noisy measured Lie bracket and extract this representation using a spectrally separated anchor. The calibration has explicit perturbation bounds and polynomial complexity in dimension and conditioning parameters, without initialization near the unknown solution. Complementary reset constructions expose local information without assuming independent state control. For one-dimensional records, suitable reset conditions identify bounded-degree complete maps without requiring an open set of jointly reachable states. When stage outputs are retained as final records, additional local-law conditions permit recovery of local laws and minimum parent sets without expanding the composed program. After acquisition and initialization, the learned interface evaluates unseen legal sequences of learned tools and propagates argument edits within a fixed schedule without further terminal queries. Finite-noise guarantees account for discovery, conditioning, and error amplification during execution. Together, these results show how the execution interface itself can be learned from terminal feedback under explicit structural conditions.

1 INTRODUCTION

A terminal score evaluates a computation's outcome without revealing which hidden records its tools changed. Even an accurate predictor may mix records that the tools act on separately. This distinction matters when a model must compose tools in new orders or maintain a computation after an earlier argument changes. We ask when a replayable system's execution interface can itself be learned from scalar replies, with states, action targets and intermediate outputs hidden.

Our central finding is that terminal experiments can identify an execution interface, rather than only fit a score predictor. Two mechanisms make this possible under distinct assumptions: action relations couple state decoding to dynamics, while selective resets expose laws without independent control of every record (Figure 1).

Action relations address a coordinate-consistency problem. For projective type-A actions—linear actions on states up to scale—call order and duration reveal the antisymmetric part of multiplication. Invariance recovers the missing symmetric part. The resulting representation decodes states and applies tools using the same physical clock. Under visibility, conditioning and spectral separation, noisy-data calibration has linear error bounds and polynomial cost without initialization near the unknown solution.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Reset experiments address a different obstacle: decoding a record does not make it independently controllable. Under reset separation, visibility and coordinate compatibility, nested resets identify bounded-degree complete laws without requiring an open set of jointly reachable states. They vary the records supporting each polynomial term, rather than every record at once. The choice of questions is consequential: in a specified synchronized pair-reset setting, grouped resets leave cubic laws indistinguishable however many replies are collected. With retained stage outputs and additional dependence and contraction conditions, local laws and minimum parent sets are recovered without expanding the composition.

After paid acquisition and initialization, the learned interface evaluates unseen legal tool sequences and maintains fixed-schedule argument edits without further terminal queries. Finite-data bounds account for acquisition, conditioning and execution error growth; query-free use need not reduce total cost. The learned object specifies how tools act and compose, not only the scores they produce.

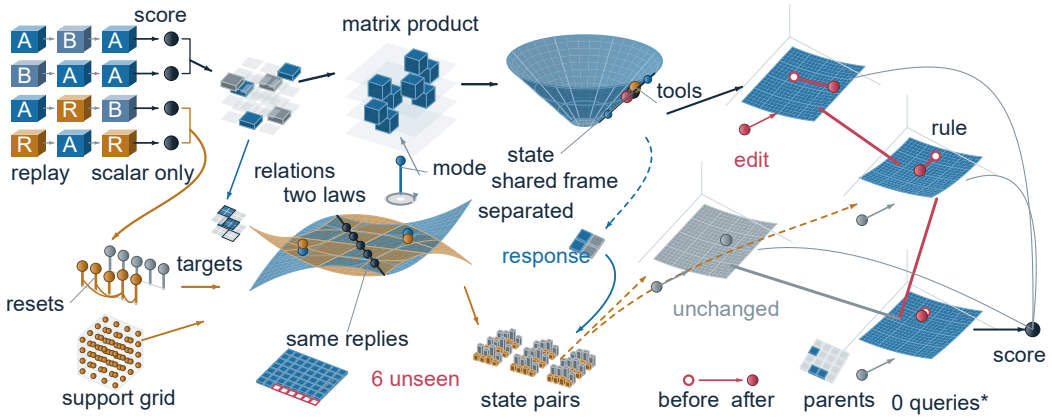


Figure 1: **Alternative routes from terminal scores to executable structure.** Action relations align decoding and dynamics; the mode inset relaxes spectral extraction. Reset grids and decoded pairs identify laws; six outlined cells mark cubic directions invisible to grouped queries. Gold dashes denote conditional reuse. Edits update red descendants, leaving the grey branch unchanged. These are distinct constructions, not one experiment. Query-free use follows acquisition and initialization under branch-specific visibility, orientation, conditioning, native-coordinate and buffer conditions.

2 Related Work

Learning states and dynamics. Nonlinear realization recovers dynamical descriptions from input–output behavior (Fliess, 1980). Predictive state representations define states through future tests and support recursive updates and planning (Littman et al., 2001; Boots et al., 2011). Gate-set tomography jointly calibrates unknown preparations, operations and measurements in common coordinates (Nielsen et al., 2021, Sections 3.1–3.3). The specific target here is action-aligned records and read/write scopes, constructed from terminal action relations and separate law-identifying experiments.

Constructive algebraic recovery. Classical identities and matrix-algebra algorithms supply the exact multiplication and representation ingredients (Haber, 2021; Ivanyos et al., 2012). Local rigidity relates nearby exact Lie brackets (Crainic et al., 2014); approximate isomorphism can also be posed through nonlinear equations (Deng & Reid, 2024). Our contribution is a perturbation-controlled construction on measured, potentially non-Lie tables, without near-truth initialization.

Identifying laws from interventions. Unknown intervention targets can be identified from linearly mixed vector observations (Bing et al., 2024). Here resets must first expose local information through scalar feedback. We connect a restricted query language to a classical polynomial vanishing ideal (Watanabe & Yanagawa, 2019), and construct law-identifying preparations under separate reset conditions. We separate identifying a law from independently controlling every input.

Composition and local updates. Module identification and representation consistency support compositional generalization (Schug et al., 2024; Wiedemer et al., 2024). Change propagation already maintains computations after edits (Acar et al., 2002). We recover the coordinates, laws and scopes needed for this maintenance from terminal experiments, with execution-error control.

3 Access and Execution Objective

A record is a hidden state s_i in a compact buffered subset $K_i \Subset \mathbb{P}^{q_i-1}$. A ray $[v]$ identifies nonzero vectors up to scale. Each public raw name e has unknown support S_e and a positive-duration action

$$T_e^t(s)_i = [\exp(tA_{e,i})s_i] \ (i \in S_e), \qquad T_e^t(s)_i = s_i \ (i \notin S_e), \qquad A_{e,i} \in \mathfrak{sl}_{q_i}(\mathbb{R}). \tag{1}$$

The matrices, states and supports are not supplied. Payloads are deterministic state-Markov maps in a declared finite dictionary, possibly noninvertible. Programs are unknown acyclic record computations.

An experiment replays a history and a genuine continuation. On logged branch J_r of arm W, with propensity at least $\pi_{\min} > 0$, it returns

$$Y_r = h(W_{J_r} s^\star) + \xi_r, \qquad \mathbb{E}[\xi_r \mid \mathcal{F}_{r-1}, W, J_r] = 0. \tag{2}$$

Finite-data results use conditional σ-sub-Gaussian noise, a public mean bound B_Y, and $V_{\mathrm{end}} = \max\{1, \sigma^2 + B_Y^2\}$. One scalar read counts as one endpoint; all attempts and replayed calls are charged. Histories and continuations each contain at most one program. Every prefix and probe is legal, with a retained margin from guards and normalization poles. Signed combinations, matrix inverses and logarithms are offline computations, not negative-time access.

Alternative routes, explicit outputs. The motion route assumes a faithful product of split type-A factors acting on a product-open reachable set, with terminal continuations separating action directions. It needs a visible response at every decoded state, a same-state independent response pair once per factor when $q_i \geq 3$, and either the Hall or repeated-word acquisition design below. Native dictionaries need a compatible reference/anchor. Reset routes instead allow locked motion and replace independent preparation by the erasure and function-class conditions in Section 5; reset arms have $\pi_{\min} = 1$. None of these routes grants direct state observations.

Deployment fixes a typed call schedule. Argument edits change values, not names or expression structure; new legal calls may also be appended. A public prefix-closed compact class $\mathcal{W}_T$ bounds buffers and edit influence. For retained records A, let u_t, y_t and F_t denote sources, caches and the original score in aligned charts. Define

$$d_{\mathrm{op}}^A(I, \widehat{I}) = \sup_{\omega \in \mathcal{W}_T} \max_{0 \leq t \leq |\omega|} \max\{\|(\widehat{u}_t - u_t)|_A\|_\infty, \|(\widehat{y}_t - y_t)|_A\|_\infty, |\widehat{F}_t - F_t|\}. \tag{3}$$

Truth is not supplied; absent caches contribute no term. Discarded future-inactive records have no state guarantee.

Complete-map recovery does not itself identify stages. Section 5.2 states when stage recovery is possible; Section 6 explains the common execution argument and its costs.

4 From Terminal Responses to Executable Coordinates

Decoding and dynamics must agree in one representation. Hall and repeated-word probes supply alternative tables for calibration; the one-mode extension relaxes its anchor condition. Polynomial calibration cost excludes discovery.

4.1 Obtaining relations without observing the readout

For a Lie word w, a history H and continuation C, positive-duration experiments estimate $D_w(h \circ C)(s(H))$. The Hall route uses a separated depth-two basis and depth-three raw adjoints. Quotienting invisible directions and splitting primitive centroid idempotents identifies simple factors; nonzero raw projections give supports. These decompositions are established (de Graaf, 1997; Respondek, 1982); terminal acquisition constructs their inputs before targets are known.

An alternative avoids the depth-two assumption. Let $W(t)$ be a regular positive word, $w = W(t^0)$, and X a basis of its duration directions. Insert $g_a(\tau) = e^{\tau A_a}$ twice into the doubled word $W(u)g_a(\tau)W(v)g_a(\tau)$. Derivatives in the outer and inner slots measure the *same unknown terminal covector* on X and $\mathrm{Ad}_{wg_a(\tau)}\,X$. Stacking histories and continuations therefore gives $D_a = B_a\Theta_a$. Full column rank of B_0 makes it possible to remove the covector:

$$\Theta_a = B_a^{\dagger}D_a, \qquad [\mathrm{ad}\,A_a]_X = \Theta_0^{-1}B_0^{\dagger}(\dot{D}_a - \dot{B}_a\Theta_0). \tag{4}$$

Dots differentiate the insertion duration at zero, using limits of legal positive probes. The subtraction removes the changing readout; entries have total derivative order two. With an additional public spectral-nonaliasing bound, $\tau^{-1}\log(\Theta_0^{-1}\Theta_a)$ uses first-duration entries instead. Centerlessness then recovers the bracket and raw coordinates offline. Generation supplies a regular word of length at most the squared algebra dimension, but not legality, a known successful word, or a uniform conditioning gap. Appendix A.1–A.2 states these distinct access and stopping conditions.

4.2 One representation for decoding and dynamics

For one recovered factor, let W have dimension $q^2 - 1$, bracket μ, and raw coordinates a_e. A terminal response $r_x(a) = D_a f(x)$ is a vector of physical-duration derivatives assembled from scalar experiments. A calibration $S : W \to \mathfrak{sl}_q$ preserves brackets.

Theorem 4.1 (Calibrated minimal-nilpotent decoding)**.** *Suppose the terminal table is isomorphic to $\mathfrak{sl}_q$ and acts as in* (1)*. For a bracket-preserving S, define $M_x \in \mathfrak{sl}_q$ by*

$$\mathrm{tr}(M_x Sa) = r_x(a) \qquad (a \in W). \tag{5}$$

After one fundamental/dual orientation, every nonzero M_x has $\mathrm{im}\,M_x = [s_x]$, and raw calls act as $[s] \mapsto [\exp(tSa_e)s]$ at the same duration. Two independent responses at one state determine orientation for $q \geq 3$; for $q = 2$ the branches are inner-equivalent. Native dictionaries still need a compatible anchor/reference.

Why this coupling is forced. The differential of a ray readout has $\varphi^{\top}s = 0$, so $r_x(a) = \varphi^{\top}(Sa)s$ and $M_x = s\varphi^{\top}$. Its image is the state line, a classical minimal-nilpotent relation (Fu, 2006, Section 2.3). Automorphisms are conjugations, optionally followed by negative transpose. Conjugation moves states and generators together; transpose exchanges images and covectors. Two independent same-state responses share an image in only the correct branch. Finally, cS preserves a nonzero bracket only when $c = 1$: physical time is fixed by the action relation, not a separate regression. Appendix A.6 supplies the perturbation and native-frame bounds. In compatible native charts, separated pre/post-call designs fit the declared call and score dictionaries. Nonidentity outputs and nonzero input groups determine write/read scopes; these additional designs, not calibration alone, identify complete laws.

Constructing the representation from noisy relations. Local rigidity gives nearby isomorphisms between exact semisimple brackets (Crainic et al., 2014, Theorem 5.3). It does not initialize the following recovery from a noisy table that may violate Jacobi.

Theorem 4.2 (Constructive calibration from a noisy bracket)**.** *Let $q \geq 2$, $n = q^2 - 1$, $\mu = S^{-1}[S\cdot, S\cdot]$, and $\|S\|, \|S^{-1}\| \leq K \geq 1$ in Euclidean/Frobenius norms. Suppose $\|\widehat{\mu} - \mu\|_F \leq \eta$, $\|\widehat{h} - h\| \leq \eta$, $\|h\| \leq H$ with $H \geq 1$, and $Sh = V\,\mathrm{diag}(\theta_i)V^{-1}$, with real eigenvalues separated by $\gamma > 0$ and $\mathrm{cond}\,V \leq \chi$. Neither S, V nor the spectrum is supplied. For an explicit inverse-polynomial threshold η_0 and polynomial factor a_1 in $q, K, H, \chi, \gamma^{-1}$ (Appendix A.4,* (28)*), the construction returns*

$$\inf_{\alpha \in \mathrm{Aut}(\mathfrak{sl}_q)} \|\widehat{\rho} - \alpha S\|_{\mathrm{op}} \leq a_1\eta, \qquad \|\widehat{\rho}\|, \|\widehat{\rho}^{-1}\| \leq 2K^3 \quad (0 < \eta \leq \eta_0). \tag{6}$$

It needs neither Jacobi projection nor near-truth initialization. Numerical error $\varepsilon_{\mathrm{num}} \leq (4K^3)^{-1}$ enlarges the caps to $4K^3$ and adds $\varepsilon_{\mathrm{num}}$, at bit cost polynomial in these parameters, input length and logarithms of inverse tolerances. A measured finite catalogue containing a valid anchor at unknown index suffices under Corollary A.6's larger bound.

Recover the missing symmetric product. Antisymmetrize the table, put $\widehat{A}_x y = \widehat{\mu}(x,y)$ and $\widehat{B}(x,y) = \mathrm{tr}(\widehat{A}_x\widehat{A}_y)/(2q)$. For $q \geq 3$, take a unit smallest right singular vector $\widehat{T} \in \mathrm{Sym}^3 W^*$ of the *linear* invariance operator

$$(\mathcal{L}_{\widehat{\mu}}T)(a;x,y,z) = T(\widehat{\mu}(a,x),y,z) + T(x,\widehat{\mu}(a,y),z) + T(x,y,\widehat{\mu}(a,z)).$$

Solve $\widehat{B}(\widetilde{J}(x,y),z) = \widehat{T}(x,y,z)$. Writing $\widetilde{J}_x y = \widetilde{J}(x,y)$, normalize with $c_q = (q^2-4)/(2q)$, $\widehat{C}(x,y) = \mathrm{tr}(\widetilde{J}_x\widetilde{J}_y)$, $\widehat{s} = \langle \widehat{C}, \widehat{B}\rangle_F/(c_q\|\widehat{B}\|_F^2)$ and $\widehat{J} = \widetilde{J}/\sqrt{\widehat{s}}$. The perturbation threshold ensures inversion and positivity. For $q = 2$, set $\widehat{J} = 0$ and omit these cubic steps. On $\mathbb{R} \oplus W$, where the scalar coordinate represents $I/\sqrt{q}$, form

$$(a,x)\widehat{\star}(b,y) = \left(\frac{ab + \widehat{B}(x,y)}{\sqrt{q}},\ \frac{ay+bx}{\sqrt{q}} + \widehat{J}(x,y) + \tfrac{1}{2}\widehat{\mu}(x,y)\right). \tag{7}$$

At exact data this is matrix multiplication under $F_+(a,x) = aI/\sqrt{q} + Sx$ or $F_-(a,x) = aI/\sqrt{q} - (Sx)^\top$. The noisy product need not be associative.

Extract a small invariant space. Let $\widehat{R}$ be right multiplication by $(0,\widehat{h})$. Among its q^2 roots with multiplicity, take the q with smallest real parts and average their real parts to get $\widehat{\theta}$. Let $\widehat{U}$ contain the q smallest right singular vectors of $\widehat{R} - \widehat{\theta}I$. Restrict left multiplication $\widehat{L}_x$ to this space: $\widehat{\rho}(x) = \mathrm{tf}(\widehat{U}^\top\widehat{L}_x\widehat{U})$, with $\mathrm{tf}\, Z = Z - \mathrm{tr}(Z)I/q$. Exactly, the selected space is a left ideal $\{vw^\top : v \in \mathbb{R}^q\}$; the identity $A(vw^\top) = (Av)w^\top$ yields the defining representation. Selecting eigenvectors inside a repeated cluster is unnecessary.

Why noisy extraction is stable. The invariant cubic line has a nonzero singular gap at least K^{-7}, derived by the compact-form Casimir gap and coordinate transport; $\|\mathcal{L}_{\widehat{\mu}} - \mathcal{L}_\mu\| \leq 3\eta$. The selected ideal has singular gap at least $\gamma/(K^2\chi)$. Inverse, normalization and subspace perturbation then give (6); bounded-inverse bracket residuals control catalogue selection. Cubic identities and exact ideal extraction are classical (Haber, 2021; Ivanyos et al., 2012). The new guarantee controls this construction on measured, potentially non-Lie inputs. Conservative dense costs are $O(n^{10})$ arithmetic and $O(n^7)$ storage before precision factors, not a claim of practical dimension scaling (Appendix A.3–A.4).

The full real-simple spectrum is not essential. For real V, θ, C, Theorem A.8 replaces it by

$$Sh = V\begin{pmatrix}\theta & 0\\ 0 & C\end{pmatrix}V^{-1}, \quad \mathrm{cond}\, V \leq \chi, \qquad \tfrac{1}{2}(C + C^\top) \succeq (\theta+\gamma)I. \tag{8}$$

The complement may be repeated, complex or defective. Its numerical-range condition bounds the resolvent on the separating line, preserving a q-root cluster and the ideal gap. Scanning both signs handles the fundamental/dual choice, at most doubling offline extractions without new terminal queries. This relaxes extraction, not native-coordinate compatibility; small residuals do not certify the input promise.

5 Laws from Accessible Preparations

Resets erase selected history dependence, supporting two distinct routes. For an additive score $h = c + \sum_i h_i$, differences under identical constant overwrites cancel written records; Boolean inversion isolates hidden retention signatures. Under additive-visibility and native-frame conditions, Proposition B.1 recovers coordinates and affine maps from projected spanning differences, including singular maps. Co-motion is allowed; all masks are charged. The nonlinear route below uses a different, one-dimensional reset contract: selective preparations replace the need for independent record control.

5.1 Which preparations identify a law?

The three anonymous locked fields are a permutation of

$$2\rho_H x(1-x)\partial_x, \qquad -\rho_- x^2\partial_x, \qquad \rho_+(1-x)^2\partial_x, \qquad \rho_e \in [4/5, 6/5]. \tag{9}$$

All records start at the native reference $1/2$, and each reset writes at least two records to unknown constants. If Z_i denotes the reset names retaining record i, the hidden signatures have equal public weight and degree-d separation:

$$\Big(\bigcap_{i\in K} Z_i\Big) \setminus Z_j \neq \varnothing \qquad (1 \le |K| \le d,\ j \notin K), \qquad d \le m-2. \tag{10}$$

This cover-free condition (Idalino & Moura, 2026) allows an intersection of learned signatures to retain exactly K; it does not supply singleton actuators. Every calibration word, argument node, replay and post-call decoder must be legal in a common buffer $[a, 1-a]^m$, including the decoder's final probe slots.

What grouped queries cannot see. For this comparison only, specialize to $m \ge 5$ with *all* pair resets writing $1/2$. Let G be the endpoints of a raw-only prefix, one contiguous reset block, and a raw-only suffix; the buffer admits the small positive words used below. Theorem B.4 gives the exact kernel

$$\{p \in \mathcal{P}_{\le 3}(m) : p|_G = 0\} = \operatorname{span}\{(x_i - x_m)(x_j - x_m)(x_i - x_j) : 1 \le i < j < m\}.$$

There are $\binom{m-1}{2}$ independent invisible directions, none of degree at most two. In five records, the maximum grouped cubic design rank is 50 out of 56, not necessarily the rank of every finite bank. Duration derivatives within this language remain blind to the same kernel.

A grouped word leaves at most two values, one for each reset/unreset class. Conversely, resetting a nonsingleton class of any bipartition and varying small minus/plus-flow durations gives a nonzero Jacobian, hence an open subset of its two-class plane. With $t = x_m$ and $y_i = x_i - x_m$, restriction to $y = z\mathbf{1}_S$ forces the coefficient sum on every exact support to vanish. Through degree three, the only free differences are $y_i^2 y_j - y_i y_j^2$. The ideal is classical (Watanabe & Yanagawa, 2019); the connection to legal experiments supplies its identification consequence.

Theorem 5.1 (Complete laws on thin preparations)**.** *Under Definition B.2's buffered reset contract* (9)–(10)*, not the grouped-query restriction, let the original signed polynomial score be essential in every record and have degree at most d. Enrolled complete maps have joint degree at most d in states and public arguments in $[-1,1]^{p_b}$, with $p_b \le p_0$ fixed. Legal terminal means and raw derivatives through order two recover native coordinates, raw/reset laws, the original score, complete maps and functional scopes, up to permutation. Finite guarantees require bounded coefficients, visibility/design and nonzero-group gaps, and charge all decoding. Internal stages are not identified.*

Why lower-dimensional data suffice. Nested resets leave at most d survivors with independently spaced *cumulative* positive durations; all other records are constant within each grid. A highest-total-degree monomial has support at most d. An exterior-variable term cannot change its coefficient on that support grid without having larger total degree. Interpolate, subtract the recovered top degree, and descend. Thus the bank norms every coefficient even though different grids have different exterior constants. The same enumeration first finds visible one-record contexts before signatures or coordinates are known; a three-state bracket/Killing pilot identifies names, rates and native coordinates. No coordinate oracle is hidden in the interpolation step (Appendix B.3).

The degree-matched nested bank has zero kernel under its own contract, without requiring an open reachable set. Appendix B.3 constructs 289 records and 68 constant resets. Each last-writer class, including unreset records, follows one three-parameter projective transformation, giving raw/reset preparation dimension at most $3(68+1) = 207$, yet determining every cubic. This neither makes every interleaving informative nor identifies arbitrary analytic laws: higher-degree equality polynomials can vanish on all preparations. Grouped indistinguishability alone does not imply decision regret; the whole transcript and candidate menu must also meet the conditions in Appendix B.4.

5.2 Local laws without expanding the composition

Local recovery requires no fitted complete map. Suppose stage outputs are retained as final records and, on the full local unit cube,

$$\begin{aligned} y_i &= f_i(x_i, y_{\mathrm{pa}(i)}), \qquad |\,\mathrm{pa}(i)| \le \Delta, \qquad \deg f_i \le r, \\ 0 < \alpha \le \partial_{x_i} f_i \le A, &\qquad \sum_j |\partial_{y_j} f_i| \le \beta < 1. \end{aligned} \tag{11}$$

Write $P(x) = y$ for the program's complete map. Coefficients are bounded by M, outputs lie in $[a_\star, 1 - a_\star]$, and parent arguments are essential. Retained outputs are decoded by terminal probes, not directly observed. Use the same locked reset family, with an essential original score of degree d_h, and separation through $\max\{d_h, 2\Delta + 1, \max_b d_b\} \le m - 2$, where d_b bounds each separately enrolled complete-map payload, not the composed degree of P. Required pre/post-call preparations and decoders remain buffered.

A candidate-minus-truth law uses one source and at most 2Δ outputs. Put $t = 2\Delta + 1$. For $i \in K$, $|K| = t$, fixing exterior sources gives the injective mixed map $\Phi(x_K) = (x_i, P_{K\setminus\{i\}}(x))$, with

$$\|D\Phi\|_\infty \le L = \max\{1, A/(1-\beta)\}, \qquad \|(D\Phi)^{-1}\|_\infty \le B_\partial = \max\{1, [1+(1+A)\beta]/\alpha\}. \quad (12)$$

To see why the inverse does not inherit the forward factor $(1 - \beta)^{-1}$, write $DP = (I - B)^{-1}D$ in topological order, with diagonal D and strictly triangular B. Schur elimination sums paths only until their first return to a retained output or the retained source. These disjoint first-return paths have total absolute weight at most the first-edge bound β. This controls the mixed inverse even as β approaches one at fixed α, A. Lifting a small mixed cube and rounding interpolation nodes yields a finite norming design; the learner need not know the true order or that cube.

Theorem 5.2 (Local laws without composition expansion)**.** *Under the local-law and reset conditions just stated, targeted preparations identify the unique minimum parent sets and local polynomials without expanding P. At finite precision, visibility and nonzero coefficient-group gaps, buffered legality and sufficiently accurate decoded state pairs give coefficient, parent and execution guarantees with all decoding costs charged.*

Indeed, a zero-residual candidate differs from truth on at most t mixed variables. Their norming inequality forces every difference coefficient to vanish, and minimum cardinality selects exactly the essential parents. Alternatively, a graph-oblivious pool suffices for each fixed unknown admissible program under the additional bounded-ancestor separation and all-permutation legality conditions of Appendix C.2's design. For k resets, fixed local parameters, buffers and duration budget, its size is

$$Q_{\mathrm{pool}} = O\big(k^{2\Delta+1+2r} \log(m/\delta_{\mathrm{design}})\big)\,. \quad (13)$$

With probability at least $1 - \delta_{\mathrm{design}}$, this pool simultaneously norms all candidate differences. These are decoded pairs, not scalar-query counts; blind decoding is separately charged. Targeted preparation can be cheaper. Appendix C.2 gives the precise ancestor, separation and legality conditions.

Locality avoids a real representation cost. In centered coordinates, the buffered quadratic chain $v_1 = z_1/4 + z_1^2/32$ and $v_j = z_j/4 + z_j^2/32 + (v_{j-1} + v_{j-1}^2)/16$ has depth-independent local bounds but exact complete degree 2^j. This is not a lower bound for approximation. Conversely, removing own-source nondegeneracy permits identical complete maps with different essential local parents; additional complete-map observations cannot recover that distinction (Appendix C.3).

6 Execution, Finite Guarantees, and Cost

The acquisition routes, under their respective contracts, provide the inputs to execution: compatible coordinates, admitted laws and scopes, initialization and the original score. The learned functions define a versioned dependency graph of arguments, guards, normalizations, retained values and score features. Values outside the learned safe cone stay unchanged; visited nodes apply the frozen functions to the same parents as full recomputation. Induction proves exact maintenance of that model, including noncommuting and noninvertible calls, without terminal queries or prefix rereads. With population-exact functions, this also agrees with the hidden system on every finite admitted buffered sequence. Initialization, copies, traversal and score work remain charged (Acar et al., 2002).

True-system accuracy is separate. A fitted defect d_a and call Lipschitz bound L_a give $r^+ \le L_a r + d_a$. Bounding this recurrence over the buffered prefix class and evaluating score/cache errors gives (3). Parent contraction within a program does not imply stability under repeated expansive calls.

A separate pilot fixes discrete structure and a design; fresh observations refine numerical values. Finite recovery requires isolated ranks, factors, supports, native frames and coefficient groups, not an observed rank plateau. All rejected candidates and replays count. At fixed conditioning, a confirmed

C^5 order-j design ($j \le 3$) reaches entry error ζ at $\widetilde{O}(\zeta^{-(j+2)})$ cost; the nonaliased first-duration route refines at $\widetilde{O}(\zeta^{-5/2})$. A public holomorphic width/envelope permits accuracy-independent discovery and $\widetilde{O}(\varepsilon^{-2})$ refinement; real analyticity alone does not. Appendix D's branch budgets give $d_{\mathrm{op}}^{A} \le \varepsilon$ on the joint event, with pilot, design and estimation failures summing to at most δ.

Appendix E verifies the motion-route conditions in a sparse-output affine-residual family with linear overlap, independent nondegenerate product-box blocks and legality/payload closure, at fixed local dimension, causal depth and density scales (Theorem E.1). Neither polynomial calibration nor query-free deployment implies inexpensive acquisition: generation alone can require exponentially many observations in a restricted training language (Theorem D.1).

7 Experiments: Learning Structure for Reuse

The experiments test constructive calibration, new-objective reuse of frozen dynamics, and the information exposed by query design. These are complementary tests of executable structure, not a joint implementation of every acquisition and editing contract.

Access and comparison. Twenty-four worlds have two $q = 3$ factors and 18 anonymous tools. Dimensions, a spanning raw design and the readout class are supplied; states, targets, matrices and anchors are hidden. Noise-matched positive-time acquisition gives 240 runs at $\sigma = 0, 10^{-10}, 10^{-8}, 10^{-6}$. Construction uses 1,266,800 endpoints plus 512 validation reads, 256–12,272 more than direct fitting, which uses eight independent, validation-selected starts. Constructive initialization adds mapping/projection costs. Each model ranks 48 held-out menus of 32 legal sequences formed by call permutations or duration edits. Top-1 measures selection of the true score minimizer. Earlier coupled-fit controls remain in Appendix F.2.

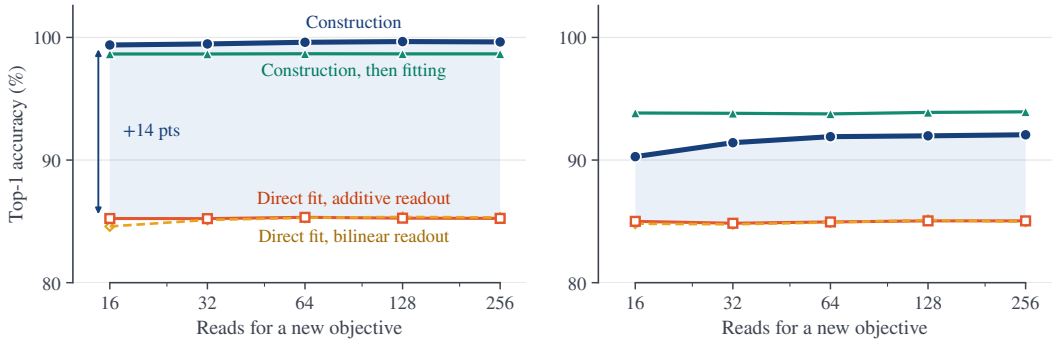


Figure 2: **Frozen dynamics transfer beyond the acquisition score.** Left: $\sigma = 10^{-8}$; right: 10^{-6}. Means aggregate objectives/repeats within, then across 24 worlds. Independent direct fits use additive or expanded bilinear readouts (dashed); initialization is available in 72/72 and 69/72 cells. Acquisition is charged separately. Appendix Table 2 covers all noise levels.

A new readout, not new dynamics. For three redrawn scores per world, only readouts are fit on 16–256 observations with ridge/leave-one-out selection. Construction uses nine numerator features over its frozen rational denominator. Direct fits use six additive features or nine coordinate products. Summing products over one normalized record recovers the other's coordinates, so the bilinear class contains the additive class without extra data or dynamics fitting.

At sixteen reads and $\sigma = 10^{-8}$, construction achieves 99.4% top-1 versus 85.2% for additive and 84.6% for bilinear direct fits. Its paired advantage over the additive fit is 14.1 percentage points (world-bootstrap 95% interval $[11.4, 17.3]$). It beats both readouts in all 24 worlds at every budget through 10^{-8}. Expanding the readout does not remove this reuse gap: acquisition-score prediction and transfer are distinct outcomes in this family. Worlds, not menus, are independent units.

At 10^{-6} and sixteen reads, construction wins in nineteen worlds and loses in five: the paired advantage is 5.3 points (interval $[2.4, 8.5]$). Mean normalized regret is lower, but median world-averaged

menu-maximum absolute error exceeds direct fitting's. Initialized fitting reaches 93.8% top-1 on eligible cells, without improving every error summary. Decision reuse differs from score accuracy; new-read budgets exclude acquisition and do not establish total-cost savings (Appendix F.2.2).

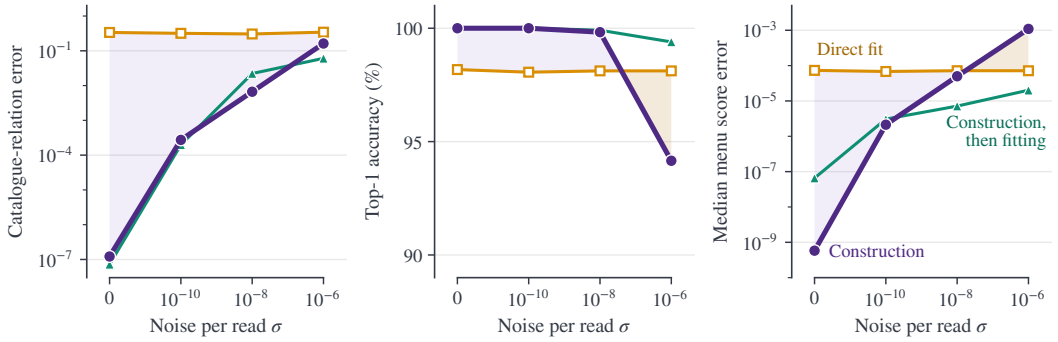


Figure 3: **Original-score prediction and linear catalogue relations.** Menus are held out under the acquisition objective. Left: median relation error; centre: top-1; right: median across cells of within-cell median menu-maximum absolute error. Independent fits cover 240 cells, with 237 eligible initializations. At the largest noise, direct fitting has lower score error despite higher relation error. Appendix F.2.1 gives costs and denominators; Figure 8 retains coupled controls.

Initialization on the original objective. The gauge-invariant row-space metric measures linear tool dependencies, not all Lie identities or physical durations. On original-score test menus, replacing one random start with mapped construction improves prespecified error in 231 cells, ties in six and fails mapping in three. At high noise, one worst-menu error and four top-1 rates worsen. Independent clipping/L-BFGS remove cross-start coupling, but projected support zeros preclude isolating pure basin selection.

Does the question language affect useful information? At matched cost in 24 paired worlds, noiseless grouped queries give .522 decision accuracy versus 1.000 for interleaved and multiset-matched random legal orders. Menus probe grouped-language ambiguity, not universal ordering optimality. On 24 fresh worlds, noise-matched derivatives raise all-attempt interleaved success from .425 to .911 at $\sigma = 10^{-6}$ and 8,687 reads; at 10^{-5} it falls to .490. Exact information obstruction differs from noisy recovery (Appendix F.4).

Additional mechanisms and boundaries. Calibration through $q = 6$ tests perturbation sensitivity, not acquisition or the one-mode extension. For CNN-derived CIFAR-10 objectives, menu-selection accuracy reaches 96.4% with 512 new reads per class-score objective, versus 41.0% for an order-agnostic readout; the tools remain synthetic. The Brier study completes nineteen of twenty acquisitions under a joint endpoint budget. Whole-map fits retain most decision benefits in the same learned coordinates, supporting reuse without isolating parent recovery's added value. Composition, enrollment and update-time controls remain in Appendix F.

8 Discussion and Limitations

The execution interface is a learnable object under the stated conditions. Action relations align states and dynamics; selective preparations expose laws without independent record control. Grouped-query blindness explains why question design matters, while frozen-dynamics transfer separately tests the reuse of the learned structure.

The alternative acquisition contracts impose substantive projective, reset, visibility, native-coordinate and conditioning requirements. Local-stage recovery also needs retained outputs and own-source nondegeneracy. Query-free composition and fixed-schedule edits follow paid acquisition and initialization; finite-data accuracy depends on error growth over admitted buffered sequences. Simulated tools support conditional reuse, not uniform accuracy or total-cost superiority. Broader action and law classes and lower acquisition costs remain important directions.

AI USE STATEMENT

We used large language model (LLM) tools for five purposes. *Drafting:* they drafted and restructured sections. *Polishing:* they polished language throughout the manuscript. *Retrieval:* they assisted literature search and the discovery of related work. *Research ideation and execution:* they helped develop methods and experimental designs, and wrote and ran simulation, analysis, audit and plotting code. *Mathematics:* they assisted in deriving and checking the proofs of the stated results. No dataset was generated by an LLM: the synthetic worlds come from the deterministic simulators described in Appendix F, and the real data are public (CIFAR-10 and the digits dataset). Experimental summaries were checked against stored per-run results, with separate implementations used to verify selected model evaluations. LLMs are not authors; the authors reviewed and take full responsibility for all claims, proofs, experiments, citations and text.

ETHICS STATEMENT

This work studies simulated tool systems and uses the public CIFAR-10 and digits datasets. No human participants were recruited, and no new human annotations or personal data were collected.

REPRODUCIBILITY STATEMENT

Appendices A–E provide the assumptions and proofs of the theoretical results. Appendix F describes the experimental access conditions, simulators, data sources, model configurations, sampling and estimation procedures, resource budgets, uncertainty calculations and failure criteria. The code used for the experiments will be made publicly available upon acceptance.

REFERENCES

Umut A. Acar, Guy E. Blelloch, and Robert Harper. Adaptive functional programming. In *Proceedings of the 29th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pp. 247–259, 2002. doi: 10.1145/503272.503296.

Simon Bing, Urmi Ninad, Jonas Wahl, and Jakob Runge. Identifying linearly-mixed causal representations from multi-node interventions. In *Proceedings of the Third Conference on Causal Learning and Reasoning*, volume 236 of *Proceedings of Machine Learning Research*, pp. 843–867, 2024. URL https://proceedings.mlr.press/v236/bing24a.html.

Byron Boots, Sajid M. Siddiqi, and Geoffrey J. Gordon. Closing the learning-planning loop with predictive state representations. *The International Journal of Robotics Research*, 30(7):954–966, 2011. doi: 10.1177/0278364911404092.

Armand Borel. *Linear Algebraic Groups*, volume 126 of *Graduate Texts in Mathematics*. Springer, second edition, 1991. doi: 10.1007/978-1-4612-0941-6.

Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3, 1950. doi: 10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Anthony Carbery and James Wright. Distributional and L^q norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$. *Mathematical Research Letters*, 8(3):233–248, 2001. doi: 10.4310/MRL.2001.v8.n3.a1.

Claude Chevalley and Samuel Eilenberg. Cohomology theory of Lie groups and Lie algebras. *Transactions of the American Mathematical Society*, 63(1):85–124, 1948. doi: 10.1090/S0002-9947-1948-0024908-8.

Marius Crainic, Florian Schätz, and Ivan Struchiner. A survey on stability and rigidity results for Lie algebras. *Indagationes Mathematicae*, 25(5):957–976, 2014. doi: 10.1016/j.indag.2014.07.015.

Willem A. de Graaf. Calculating the structure of a semisimple Lie algebra. *Journal of Pure and Applied Algebra*, 117–118:319–329, 1997. doi: 10.1016/S0022-4049(97)00016-9.

Laurent Demanet and Alex Townsend. Stable extrapolation of analytic functions. *Foundations of Computational Mathematics*, 19(2):297–331, 2019. doi: 10.1007/s10208-018-9384-1.

Siyuan Deng and Gregory Reid. Contributions to the algorithmic foundations of approximate Lie symmetry algebras of differential equations. *arXiv preprint arXiv:2410.20608*, 2024. doi: 10.48550/arXiv.2410.20608.

Michel Fliess. Realizations of nonlinear systems and abstract transitive Lie algebras. *Bulletin of the American Mathematical Society*, 2(3):444–446, 1980. doi: 10.1090/S0273-0979-1980-14760-6.

Baohua Fu. A survey on symplectic singularities and symplectic resolutions. *Annales Mathématiques Blaise Pascal*, 13(2):209–236, 2006. doi: 10.5802/ambp.218.

Gene H. Golub and Victor Pereyra. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2):413–432, 1973. doi: 10.1137/0710036.

Edouard Goursat. Sur les substitutions orthogonales et les divisions régulières de l'espace. *Annales Scientifiques de l'École Normale Supérieure*, 6:9–102, 1889. doi: 10.24033/asens.317.

Ming Gu and Stanley C. Eisenstat. Efficient algorithms for computing a strong rank-revealing QR factorization. *SIAM Journal on Scientific Computing*, 17(4):848–869, 1996. doi: 10.1137/0917055.

Howard E. Haber. Useful relations among the generators in the defining and adjoint representations of SU(N). *SciPost Physics Lecture Notes*, 21, 2021. doi: 10.21468/SciPostPhysLectNotes.21.

James E. Humphreys. *Introduction to Lie Algebras and Representation Theory*, volume 9 of *Graduate Texts in Mathematics*. Springer, 1972. doi: 10.1007/978-1-4612-6398-2.

Thaís Bardini Idalino and Lucia Moura. Cover-free families on hypergraphs and combinatorial group testing. *Journal of Combinatorial Optimization*, 51(5):55, 2026. doi: 10.1007/s10878-026-01429-0.

A. P. Isaev and S. O. Krivonos. Split Casimir operator for simple Lie algebras, solutions of Yang–Baxter equations, and Vogel parameters. *Journal of Mathematical Physics*, 62(8):083503, 2021. doi: 10.1063/5.0049055.

Gábor Ivanyos, Lajos Rónyai, and Josef Schicho. Splitting full matrix algebras over algebraic number fields. *Journal of Algebra*, 354(1):211–223, 2012. doi: 10.1016/j.jalgebra.2012.01.008.

Arthur J. Krener. A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems. *SIAM Journal on Control*, 12(1):43–52, 1974. doi: 10.1137/0312005.

Michael L. Littman, Richard S. Sutton, and Satinder Singh. Predictive representations of state. In *Advances in Neural Information Processing Systems*, volume 14, 2001. URL https://proceedings.neurips.cc/paper/2001/hash/1e4d36177d71bbb3558e43af9577d70e-Abstract.html.

Erik Nielsen, John King Gamble, Kenneth Rudinger, Travis Scholten, Kevin Young, and Robin Blume-Kohout. Gate set tomography. *Quantum*, 5:557, 2021. doi: 10.22331/q-2021-10-05-557.

Albert Nijenhuis and R. W. Richardson, Jr. Deformations of Lie algebra structures. *Journal of Mathematics and Mechanics*, 17(1):89–105, 1967. doi: 10.1512/iumj.1968.17.17005.

A. L. Onishchik and E. B. Vinberg. *Lie Groups and Algebraic Groups*. Springer Series in Soviet Mathematics. Springer, 1990. doi: 10.1007/978-3-642-74334-4.

Victor Y. Pan. Univariate polynomials: Nearly optimal algorithms for numerical factorization and root-finding. *Journal of Symbolic Computation*, 33(5):701–733, 2002. doi: 10.1006/jsco.2002.0531.

Witold Respondek. On decomposition of nonlinear control systems. *Systems & Control Letters*, 1(5):301–308, 1982. doi: 10.1016/S0167-6911(82)80027-3.

Simon Schug, Seijin Kobayashi, Yassir Akram, Maciej Wołczyk, Alexandra M. Proca, Johannes von Oswald, Razvan Pascanu, João Sacramento, and Angelika Steger. Discovering modular solutions that generalize compositionally. In *The Twelfth International Conference on Learning Representations*, pp. 24674–24726, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/6b40b312ce4933f4f2d1ac0b0b6b07f7-Abstract-Conference.html.

Kevin D. Smith, Saber Jafarpour, Ananthram Swami, and Francesco Bullo. Topology inference with multivariate cumulants: The Möbius inference algorithm. *IEEE/ACM Transactions on Networking*, 30(5):2102–2116, 2022. doi: 10.1109/TNET.2022.3164336.

Joel A. Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012. doi: 10.1007/s10208-011-9099-z.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, 2009. doi: 10.1007/b13794.

Junzo Watanabe and Kohji Yanagawa. Vandermonde determinantal ideals. *Mathematica Scandinavica*, 125(2):179–184, 2019. doi: 10.7146/math.scand.a-114906.

J. H. C. Whitehead. Certain equations in the algebra of a semi-simple infinitesimal group. *The Quarterly Journal of Mathematics*, os-8(1):220–237, 1937. doi: 10.1093/qmath/os-8.1.220.

Thaddäus Wiedemer, Jack Brady, Alexander Panfilov, Attila Juhos, Matthias Bethge, and Wieland Brendel. Provable compositional generalization for object-centric learning. In *The Twelfth International Conference on Learning Representations*, pp. 20137–20164, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/57c89126d60c209f48d0e6395c766bb3-Abstract-Conference.html.

PROOF AND EVIDENCE GUIDE

Appendix A (p. 13) develops the Hall and repeated-word entrances, then proves constructive representation recovery and calibrated decoding; Section A.5 relaxes spectral extraction to one separated real mode. Appendix B (p. 22) supplies additive reset isolation or binary nonlinear acquisition on thin preparations; Section B.4 characterizes the grouped-query kernel and its decision scope. Its finite-value lemma requires already isolated contexts. Appendix C (p. 26) proves local-law recovery and the shared-pool design. These branches share Appendix D's estimation and execution arguments (p. 29), not their acquisition assumptions. Its final result separates generation from excitation. Appendix E (p. 32) verifies a sparse-output family, distinguishing global discovery from active-subsystem precision. Appendix F (p. 36) records experimental access, controls, costs and failures.

A ACTION RECOVERY AND CALIBRATION

A.1 ACCESS, VISIBILITY, AND FACTORIZATION

Definition A.1 (Motion contract)**.** On each compact buffered history patch and public-ledger branch, assume a faithful split-semisimple action with type-A simple factors, line-parabolic isotropy, and a product-open orbit. Retained raw supports are nonempty. Rational histories, genuine continuations, and positive-duration cells are densely enumerable, without labeled successful words, minors, or gaps. All arms and prefixes obey Section 3's two-program rule.

The analytic branch has analytic means/flows on a connected neighborhood and a common mixed-suffix grammar. A duration-dense bank $\mathcal{T}$ satisfies $\partial_i(h \circ T) \equiv 0$ for all $T \in \mathcal{T}$ exactly when every admitted mixed future is invariant in record i (probe completeness). Before each suffix, legal raw prefixes contain a group-submersive positive word on a common open box. Equal finite-public-Markov ledgers and fixed future-active sources give equal residual legal languages. Alternatively, the C^5 branch assumes a finite separated continuation bank at the required states.

The Hall entrance uses depth-two pivots and depth-three raw adjoints. Alternatively, Section A.2 uses faithful centerless generation of known dimension (or an isolation schedule), legal doubled regular words and insertion cells; its first-duration estimator also needs a public spectral nonaliasing bound. Global profile rank, local visibility, native anchor/reference separation, buffered law designs, and finite deployment influence are separate conditions. Exact finite structure requires public legal-step/derivative envelopes and computable isolation of ranks (by a known target rank or a remaining-bank envelope), factors, raw supports, native frames, law groups, and domains.

Decoding requires a visible response at each state and a separated same-state pair once per factor for $q \geq 3$; complete-plane designs suffice. For program, payload, and original-score dictionaries, the motion branch assumes product-open pre/post designs with redundant features removed, isolated nonzero groups and omitted-parent/effective-stage residuals, and isolated native chambers and domains. Complete-call outputs are probe-complete on the same buffer. A native anchor has ordered real-simple coordinate eigenlines and a reference with separated nonzero coordinates. The reset branches replace the product-open designs under their own contracts.

Lemma A.2 (Positive-time visibility)**.** *On a future-active factor of the analytic contract, continuation responses span* $\operatorname{ann} \mathfrak{p}_x$*, where* $\mathfrak{p}_x = \{a : a^{\#}(x) = 0\}$*. A compact buffered patch has a finite common separating prefix. Hence each state has a nonzero response, and an independent pair exists for* $q \geq 3$*. No dimension-uniform discovery rate is implied.*

Proof. Positive products have group interior (Krener, 1974). Analytic continuation therefore gives the same annihilator for positive and signed response spans, without executing inverses. Put $K_C(u) = h(Cux)$. Left-invariant fields annihilating K_C on an open continuation set form a Lie subalgebra containing $\mathfrak{p}_x$. Intersecting over common mixed tails and transporting by Ad_g identifies this subalgebra with the response annihilator, using common-prefix slack. A subalgebra containing the maximal line parabolic is either that parabolic or the whole simple algebra (Borel, 1991; Onishchik & Vinberg, 1990); activity excludes the latter. Nonzero minors persist locally, and compactness supplies a common finite bank and fixed-instance gap. □

Form the free Lie algebra on public names and quotient by the ideal remaining terminally zero after all legal left words and continuations. For pivot words q_j, let $B_{rj} = D_{q_j}(h \circ C_r)(x_{H_r})$. Left state actions satisfy $D_{[a,b]} = D_b D_a - D_a D_b$, so

$$z_a = B^\dagger (D_a f_r)_r, \qquad [\operatorname{ad} a] q_j = B^\dagger (D_{[a,q_j]} f_r)_r. \tag{14}$$

Centerlessness gives uniqueness. Compute $\operatorname{ad}[a,b] = [\operatorname{ad} a, \operatorname{ad} b]$ and the Killing matrix $\kappa_{ij} = \operatorname{tr}(\operatorname{ad} q_i \operatorname{ad} q_j)$ offline. The order-three bound concerns derivatives, not physical word length.

The centroid $T[u,v] = [Tu, v] = [u, Tv]$ is a product of copies of $\mathbb{R}$. Its primitive idempotents split simple ideals. Those preserving $\operatorname{ann} R_x$ give the finest action-ideal-aligned product: $\mathfrak{p}_x = (\mathfrak{p}_x \cap e\mathfrak{g}) \oplus (\mathfrak{p}_x \cap (1-e)\mathfrak{g})$, and conversely each product supplies such an idempotent. Lemma A.2 identifies isotropy; transport commutes with the centroid. Nonzero raw projections give supports up to permutation, not uniqueness among arbitrary predictive coordinates.

For locally generating tuples $A_e = (A_{e,i})$ in isomorphic nonabelian real simple factors,

$$\operatorname{Lie}\{A_e\} = \bigoplus_i \mathfrak{g}_i \iff \nexists i \neq j,\ \varphi \in \operatorname{Iso}(\mathfrak{g}_i, \mathfrak{g}_j) : A_{e,j} = \varphi(A_{e,i}) \text{ for all } e. \tag{15}$$

Projection kernels are ideals, so a two-factor subdirect algebra is a direct sum or an isomorphism graph; induction gives the locked blocks (Goursat, 1889). Mismatched supports cannot lock. Algebraic generation alone gives neither a Hall-depth bound nor quantitative excitation.

A.2 A SHARED-COVECTOR ALTERNATIVE TO HALL DISCOVERY

Lemma A.3 (Regular positive words)**.** *A generating alphabet of a d-dimensional Lie algebra admits a positive rational word of length at most d^2 whose right-trivialized duration directions span the algebra. Its durations can be arbitrarily small. This statement alone supplies neither legality nor a uniform gap.*

Proof. For arbitrarily small positive times, let V_k span $\operatorname{Ad}_g A_a$ over words of length at most k. If $V_{k+1} = V_k$, differentiation at zero shows generator-adjoint invariance, hence $V_k = \mathfrak{g}$. Thus $V_{d-1} = \mathfrak{g}$. If a word w has proper duration span T_w, choose g of length at most $d-1$ and a with $\operatorname{Ad}_{wg} A_a \notin T_w$. Appending ge^{tA_a} on the right leaves earlier right-trivialized columns unchanged and adds that direction. At most d extensions suffice. A nonzero minor persists at nearby positive rational durations. □

Write $W(t) = e^{t_L A_{a_L}} \cdots e^{t_1 A_{a_1}}$, $w = W(t^0)$, and select d independent slots s_i with

$$X_i = (\partial_{s_i} W) W^{-1} = \operatorname{Ad}_{e^{t_L^0 A_{a_L}} \cdots e^{t_{s_i+1}^0 A_{a_{s_i+1}}}} A_{a_{s_i}}.$$

For $g_a(\tau) = e^{\tau A_a}$ measure $F_{a,r}(u,v,\tau) = h(C_r W(u) g_a(\tau) W(v) g_a(\tau) x_{H_r})$. The executed order is $H_r, g_a, W(v), g_a, W(u), C_r$. Both insertions are paid; $\tau = 0$ omits them. At $p = w g_a$, $y = p^2 x_{H_r}$, define $\ell_{a,r}(Z) = d(h \circ C_r)_y(Z^\#(y))$. Then the u and v slot derivatives are $B_{a,ri} = \ell_{a,r}(X_i)$ and $D_{a,ri} = \ell_{a,r}(\operatorname{Ad}_p X_i)$, proving $D_a = B_a \Theta_a$ in Section 4.1.

Complete continuation responses over an open history patch and faithfulness separate every algebra direction: a traceless matrix preserving every line of an open patch is zero. Hence some finite stack has full column rank. A finite baseline minor persists for small insertions, and exhaustive word/slot/row search finds it under the assumed legality and rank stopping rule.

First-duration estimator. A public $M_{\mathrm{im}} \geq \max_{a, \lambda \in \operatorname{spec}(\operatorname{ad} A_a)} |\Im \lambda|$ allows rational $\tau M_{\mathrm{im}} < \pi$. Then

$$\Theta_a = B_a^\dagger D_a, \qquad L_a = \tau^{-1} \log(\Theta_0^{-1} \Theta_a) = [\operatorname{ad} A_a]_X. \tag{16}$$

Near-identity observations do not certify this independent nonaliasing condition.

Derivative estimator. On legal insertion cells accumulating at zero, differentiation of $D_a = B_a \Theta_a$ gives $\dot{D}_a = \dot{B}_a \Theta_0 + B_0 \Theta_0 L_a$, hence

$$[\operatorname{ad} A_a]_X = \Theta_0^{-1} B_0^\dagger (\dot{D}_a - \dot{B}_a \Theta_0). \tag{17}$$

Entries have total derivative order two. Unknown derivatives of the readout cancel in $\dot{D}_a - \dot{B}_a\Theta_0$. This replaces spectral nonaliasing, not full-column rank or legal insertion access.

For either estimator, let $S_i = e^{t^0_L L_{a_L}} \cdots e^{t^0_{s_i+1} L_{a_{s_i+1}}}$. Then

$$K_i = S_i L_{a_{s_i}} S_i^{-1} = [\operatorname{ad} X_i]_X, \quad [X_i, X_j] = \sum_k (K_i)_{kj} X_k, \quad \sum_i z_{a,i} K_i = L_a. \tag{18}$$

The final linear solve is unique by centerlessness. A response measured after W decodes wx_H in this basis; the inverse of its learned word matrix (a product of calibrated raw exponentials) returns x_H offline. For complete response spans the equivalent transport is $R_{x_H} = \Theta_0^T R_{wx_H}$; individual responses must not be equated across different contexts.

The entrance is strictly broader than depth-two generation: in $\mathfrak{sl}_2 \oplus \mathfrak{sl}_2$, let $H = \operatorname{diag}(1, -1)$, $K = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Names $(H, 2H), (K, K)$ and their first bracket span three dimensions, while iterated brackets give $(4K, 16K), (8S, 64S), (16H, 128H)$, spanning all six.

A.3 Residual stability

Proposition A.4 (A small bracket residual fixes the common gauge)**.** *Give W a Euclidean norm and $\mathfrak{sl}_q$ the Frobenius norm. For invertible S, let $\mu_S(u, v) = S^{-1}[Su, Sv]$. If $S, \widehat{S}$ and their inverses have norm at most $K \geq 1$ (an input promise, derived in Section E.2 for the natural family, not certified by a fitted residual), then*

$$e = \|\mu_{\widehat{S}} - \mu_S\|_{\mathrm{op}} \leq (256q^2K^3)^{-1} \quad \Longrightarrow \quad \inf_{\alpha \in \operatorname{Aut}(\mathfrak{sl}_q)} \|\widehat{S} - \alpha S\|_{\mathrm{op}} \leq 4qK^6 e. \tag{19}$$

A table of error η and an output residual at most $\eta + \eta_{\mathrm{opt}}$ give $e \leq 2\eta + \eta_{\mathrm{opt}}$, without requiring the table to satisfy Jacobi.

Proof. Use the Casimir homotopy for alternating cochains with adjoint coefficients (Whitehead, 1937; Nijenhuis & Richardson, 1967; Crainic et al., 2014):

$$(hf)(x_1, \ldots, x_{p-1}) = \sum_a [e_a, f(e^a, x_1, \ldots, x_{p-1})], \qquad dh + hd = I. \tag{20}$$

Here e_a, e^a are Killing-dual and $d = d_{\mathrm{CE}}$ (Chevalley & Eilenberg, 1948). In the expansion, terms on the other arguments cancel, as do basis variations by $\sum_a [x, e_a] \otimes e^a + e_a \otimes [x, e^a] = 0$. The remainder is $\sum_a \operatorname{ad}(e_a) \operatorname{ad}(e^a) f = f$. Indeed $\sum_{ij} [E_{ij}, [E_{ji}, Z]]/(2q) = Z$ for traceless Z; the scalar direction contributes zero. A Frobenius-orthonormal traceless basis has Killing dual its transpose divided by $2q$, whence $\sum_a \|e_a\|_F \|e^a\|_F = (q^2 - 1)/(2q) < q$ and $\|h\| \leq 2q$ in multilinear operator norm.

Let $\nu = \mu_0 + r$ be a Lie bracket on $\mathfrak{sl}_q$, $\mu_0(x, y) = [x, y]$, and $\delta = \|r\| \leq (256q^2)^{-1}$. Jacobi gives $\|dr\| \leq 3\delta^2$. With $u = -hr$ and $A = I + u$, exact expansion of $\nu^A = A^{-1}\nu(A\cdot, A\cdot) = \mu_0 + r_+$ yields

$$r_+ = (I + u)^{-1}\{hdr + \mu_0(u\cdot, u\cdot) + r(u\cdot, \cdot) + r(\cdot, u\cdot) + r(u\cdot, u\cdot)\}. \tag{21}$$

Since $\|\mu_0\| \leq 2$ and $\|u\| \leq 2q\delta \leq 1/2$, $\|r_+\| \leq 64q^2\delta^2 \leq \delta/4$. Iteration gives $\delta_j \leq 4^{-j}\delta$ and $\sum_j \|u_j\| \leq (8/3)q\delta$. The products $P_N = \prod_{j=0}^{N-1}(I + u_j)$ converge invertibly to P, with $\|P - I\| \leq 4q\delta$, $\|P^{-1} - I\| \leq 8q\delta$ and $\nu(Px, Py) = P[x, y]$, by $\prod_j (1 + \|u_j\|) \leq \exp \sum_j \|u_j\|$ and the inverse product bound.

Set $R = \widehat{S}S^{-1}$. Transport by S gives $\mu_R - \mu_0$ of norm $\delta \leq K^3 e$. The preceding construction makes $\alpha = RP$ an automorphism, and $\|\widehat{S} - \alpha S\| = \|R(I - P)S\| \leq 4qK^3\delta \leq 4qK^6 e$. □

A.4 Constructive representation recovery

Use Theorem 4.2's norms and promises. Tensor Frobenius norms include all entries in orthonormal Euclidean bases. Antisymmetrizing $\widehat{\mu}$ cannot increase its error. Spectral extraction and native dictionary normalization are distinct: only the latter needs the compatible frame in (35).

Write $\widehat{A}_x y = \widehat{\mu}(x, y)$ and form

$$\widehat{B}(x, y) = \operatorname{tr}(\widehat{A}_x \widehat{A}_y)/(2q). \tag{22}$$

For $q \geq 3$, obtain a unit smallest right singular vector $\widehat{T}$ of the following operator on orthonormal coordinates of $\operatorname{Sym}^3 W^*$:

$$(\mathcal{L}_{\widehat{\mu}} T)(a; x, y, z) = T(\widehat{\mu}(a, x), y, z) + T(x, \widehat{\mu}(a, y), z) + T(x, y, \widehat{\mu}(a, z)). \tag{23}$$

Solve $\widehat{B}(\widetilde{J}(x, y), z) = \widehat{T}(x, y, z)$. With $c_q = (q^2 - 4)/(2q)$ and $\widetilde{J}_x y = \widetilde{J}(x, y)$, set

$$\widehat{C}(x, y) = \operatorname{tr}(\widetilde{J}_x \widetilde{J}_y), \qquad \widehat{s} = \frac{\langle \widehat{C}, \widehat{B} \rangle_F}{c_q \|\widehat{B}\|_F^2}, \qquad \widehat{J} = \widetilde{J}/\sqrt{\widehat{s}}. \tag{24}$$

For $q = 2$, set $\widehat{J} = 0$ and omit the cubic and scale steps. The bounds below justify inversion and positivity; a failed operation rejects its candidate outside that regime. On $\mathcal{E} = \mathbb{R} \oplus W$, with scalar coordinate representing $I/\sqrt{q}$, define

$$(a, x) \widehat{\star} (b, y) = \left(\frac{ab + \widehat{B}(x, y)}{\sqrt{q}}, \frac{ay + bx}{\sqrt{q}} + \widehat{J}(x, y) + \frac{1}{2} \widehat{\mu}(x, y) \right). \tag{25}$$

This noisy product need not be associative. Let $\widehat{R}$ be right multiplication by $(0, \widehat{h})$. Compute its q^2 eigenvalues with multiplicity, select the q of smallest real part, and let $\widehat{\theta}$ be the real part of their mean. Let $\widehat{U}$ span the q smallest right singular vectors of $\widehat{R} - \widehat{\theta} I$, orthonormally. For left multiplication $\widehat{L}_x$ by $(0, x)$, output

$$\widehat{\rho}(x) = \operatorname{tf}(\widehat{U}^\top \widehat{L}_x \widehat{U}), \qquad \operatorname{tf}(Z) = Z - \operatorname{tr}(Z) I/q. \tag{26}$$

Right multiplication is essential here: its eigenspace is invariant under left multiplication.

Lemma A.5 (Exact construction and inherited gaps)**.** *Under Theorem 4.2's exact assumptions,* (22)–(26) *recover an inner conjugate of S or $-S^\top$, with norm and inverse norm at most K^3. For $q \geq 3$, $\ker \mathcal{L}_\mu$ is the line spanned by $T_S(x, y, z) = \frac{1}{2} \operatorname{tr}((SxSy + SySx)Sz)$ and its nonzero singular values are at least K^{-7}; also $\|\mathcal{L}_{\widehat{\mu}} - \mathcal{L}_\mu\| \leq 3\eta$. The right-anchor eigenspaces are q-dimensional left ideals with singular gap at least $\gamma/(K^2 \chi)$.*

Proof. Invariant line. Restriction of an invariant cubic to traceless diagonal matrices is symmetric, hence proportional to $\sum_i \lambda_i^3$ for $q \geq 3$ and zero for $q = 2$. Determinant-one conjugations realize permutations after a commuting diagonal correction. The distinct-real-spectrum locus contains an open set; restriction and polarization therefore give global uniqueness.

Singular gap. Complexify with the Frobenius Hermitian norm and change the action basis unitarily to a compact-form orthonormal basis. Then $\mathcal{L}_0^* \mathcal{L}_0$ is the positive compact Casimir on $\operatorname{Sym}^3 \mathfrak{g}^*$. With roots of squared length two, its nontrivial highest-weight eigenvalues satisfy (Humphreys, 1972)

$$(\lambda, \lambda + 2\varrho) \geq \min_{1 \leq i < q} \frac{i(q - i)(q + 1)}{q} \geq 1, \qquad \lambda = \sum_i m_i \omega_i \neq 0,$$

since $m_i \geq 0$ and fundamental-weight inner products are positive. Thus $\|\mathcal{L}_0 T\| \geq \operatorname{dist}(T, \ker \mathcal{L}_0)$. For covariant pullback $C_j(S)$, $\mathcal{L}_\mu C_3(S) = C_4(S) \mathcal{L}_0$, $\|C_3(S)\| \leq K^3$, and $\|C_4(S)^{-1}\| \leq K^4$; distance transport gives K^{-7}. Each insertion in (23) has perturbation norm at most η, by summing squared adjoint-matrix norms, giving 3η.

Product and scale. Killing nondegeneracy gives $B(x, y) = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y)/(2q)$ and J_0. For traceless A, let $J_A(X) = (AX + XA)/2$ on M_q. Then $\operatorname{tr}_{\operatorname{End}(M_q)}(J_A J_B) = q \operatorname{tr}(AB)/2$. In $M_q = \mathbb{R}(I/\sqrt{q}) \oplus \mathfrak{sl}_q$, the two off-diagonal block products contribute $2 \operatorname{tr}(AB)/q$, leaving $c_q \operatorname{tr}(AB)$ on $\mathfrak{sl}_q$. Thus $B(J_0(x, y), z) = \lambda T_S(x, y, z)$ gives the symmetric traceless product scaled by λ, and $\operatorname{tr}((J_0)_x (J_0)_y) = \lambda^2 c_q B(x, y)$. This determines magnitude without treating B as a norm; it is the real-coordinate form of Haber (2021, Eqs. (15)–(17), (27), (62)). The exact products in (25) are ordinary matrix multiplication under $F_+(a, x) = aI/\sqrt{q} + Sx$ and $F_-(a, x) = aI/\sqrt{q} - (Sx)^\top$, respectively. For $q = 2$, Cayley–Hamilton gives $AB + BA = \operatorname{tr}(AB) I$, so the symmetric traceless part vanishes. Orthogonality of scalar and traceless subspaces gives the caps.

Left ideal. Under $F_\pm$, $I_\theta = \{vw^\top : v \in \mathbb{R}^q\}$ for a unit left eigenvector w of the corresponding anchor. The map $v \mapsto F_\pm^{-1}(vw^\top)$ and its inverse on the image have norms at most K. An orthonormal ideal basis therefore induces G with $\|G\|, \|G^{-1}\| \le K$, yielding $G^{-1}(Sx)G$ or $G^{-1}(-(Sx)^\top)G$ and the K^3 caps. This ideal construction is classical (Ivanyos et al., 2012). Canonical right multiplication has eigenvector condition at most χ, and transport costs K^2. Writing $R = ZDZ^{-1}$ and projecting $Z^{-1}v$ outside the chosen eigenspace gives $\|(R-\theta I)v\| \ge (\gamma/\operatorname{cond} Z)\operatorname{dist}(v, I_\theta)$. □

Constants and perturbation proof. Let $M = 2nK^3$ and $b_1 = (2M+1)/(2q)$. For $q \ge 3$, define

$$\begin{aligned} t_1 &= 12K^7, \qquad j_1 = 2K^2 t_1 + 2K^4 b_1, \qquad c_1 = (2K^2+1)j_1, \\ s_0 &= (nc_qK^6)^{-1}, \qquad s_1 = 4K^2(c_1 + K^6 b_1), \\ j_2 &= 2s_0^{-1/2} j_1 + 2K^2 s_0^{-3/2} s_1, \qquad p_1 = b_1/\sqrt{q} + j_2 + 1/2, \\ \eta_{\mathrm{alg}} &= \min\{1, (12K^7)^{-1}, (2K^2 b_1)^{-1}, j_1^{-1}, s_0/(2s_1)\}. \end{aligned} \tag{27}$$

For $q = 2$, take $p_1 = b_1/\sqrt{q} + 1/2$ and $\eta_{\mathrm{alg}} = \min\{1, (2K^2 b_1)^{-1}\}$. In either case,

$$\begin{aligned} \kappa &= K^2\chi, & r_1 &= (H+1)p_1 + K^3, \\ u_1 &= 8\kappa(1+\kappa)r_1/\gamma, & a_1 &= \sqrt{q}(p_1 + 2K^3 u_1), \\ \eta_0 &= \min\{\eta_{\mathrm{alg}}, \gamma/[16\kappa(1+\kappa)r_1], (2K^3 a_1)^{-1}\}. \end{aligned} \tag{28}$$

The factors are polynomial and thresholds inverse-polynomial in $q, K, H, \chi, \gamma^{-1}$; they are conservative sufficient bounds.

Proof of Theorem 4.2. Since $\|\mu\|_F \le M$ and $\eta \le 1$, contraction and inverse perturbation give $\|\widehat{B} - B\|_F \le b_1\eta$, $\|B^{-1}\| \le K^2$, $\|\widehat{B}^{-1}\| \le 2K^2$ and $\|\widehat{B}^{-1} - B^{-1}\| \le 2K^4 b_1\eta$. For $q \ge 3$, the unit invariant $T = T_S/\|T_S\|_F$ has measured residual at most 3η. Minimality gives $\|\mathcal{L}_\mu \widehat{T}\| \le 6\eta$; Lemma A.5, after one sign alignment, yields $\|\widehat{T} - T\| \le t_1\eta$. For $J = B^{-1}T$ and $s = \|T_S\|_F^{-2}$, the inverse bound gives $\|\widetilde{J} - J\|_F \le j_1\eta$ and $\|J\|_F \le K^2$. Summing the trace identity with the orthogonal transpose map gives $\|T_0\|_F^2 = nc_q$. Pullback then yields $s_0 \le s \le K^6/(nc_q)$ and $\|B\|_F \ge \sqrt{n}/K^2$. For $C(x,y) = \operatorname{tr}(J_x J_y) = sc_q B(x,y)$ and $j_1\eta \le 1$,

$$\|\widehat{C} - C\|_F \le c_1\eta, \qquad |\widehat{s} - s| \le \frac{\|\widehat{C} - C\|_F + sc_q\|\widehat{B} - B\|_F}{c_q\|\widehat{B}\|_F} \le s_1\eta.$$

As $c_q \ge 5/6$, $\widehat{s} \ge s_0/2 > 0$. The derivative bound for $s^{-1/2}$ gives $\|\widehat{J} - J_\pm\|_F \le j_2\eta$. Thus (25) has tensor error at most $p_1\eta$ from one exact product. For $q = 2$, omit precisely these cubic steps.

The true product has bilinear norm at most K^3, so $\|\widehat{R} - R\| \le r_1\eta$. Bauer–Fike confines perturbed eigenvalues to disks of radius $\kappa r_1\eta$ about the real anchor eigenvalues. These disks remain disjoint along $R + t(\widehat{R} - R)$ and each contains q roots. Sorting real parts selects the leftmost cluster, with $|\widehat{\theta} - \theta_{\min}| \le \kappa r_1\eta$; individual eigenvectors inside that cluster need not be stable. Put $D = R - \theta_{\min} I$ and $\widehat{D} = \widehat{R} - \widehat{\theta} I$. Then $\|\widehat{D} - D\| \le (1+\kappa)r_1\eta$. Testing on $\ker D$ bounds the selected singular values, hence $\|D\widehat{U}\| \le 2(1+\kappa)r_1\eta$ and $\|(I - UU^\top)\widehat{U}\| \le 2\kappa(1+\kappa)r_1\eta/\gamma$. Aligned orthonormal bases satisfy $\|\widehat{U} - U\| \le u_1\eta$ inside the half-gap ensured by (28). Since $\|L_x\| \le K^3\|x\|$, compression and trace-free projection give $\|\widehat{\rho}(x) - \rho(x)\|_F \le \sqrt{q}(p_1 + 2K^3 u_1)\eta\|x\|$. The true K^3 cap and inverse perturbation prove (6).

For candidate selection, write $E = \widehat{\rho} - \rho$, $a = a_1\eta \le (2K^3)^{-1}$. The identity $\mu_{\widehat{\rho}} - \mu = \widehat{\rho}^{-1}([E\cdot, \rho\cdot] + [\rho\cdot, E\cdot] + [E\cdot, E\cdot] - E\mu)$ and $\|[X,Y]\|_F \le 2\|X\|_F\|Y\|_F$ give

$$\|\mu_{\widehat{\rho}} - \mu\|_{\mathrm{op}} \le 16K^6 a_1\eta. \tag{29}$$

Thus an invertible output induces an exact Lie bracket even when the input violates Jacobi. □

Finite computation. The invariance matrix has $d_3 = \binom{n+2}{3}$ columns and nd_3 rows, giving conservative dense costs $O(n^{10})$ arithmetic and $O(n^7)$ storage before precision factors. For rational

input of length b, contractions, solves, square roots, and symmetric singular-subspace approximations have polynomial bit cost at prescribed accuracy. The proof supplies inverse-polynomial gaps for the invariant line, selected ideal, and output inverse.

Approximate all characteristic roots of $\widehat{R}$ with multiplicity (Pan, 2002), rather than cluster eigenvectors. In dimension $d = q^2$, root error τ may require coefficient error τ^d, still polynomial working precision in $d, b, \log(1/\tau)$; coefficient magnitudes add $O(d \log(1 + \|\widehat{R}\|))$ bits. The derived gaps then give $\|\widehat{\rho}_{\text{finite}} - \rho\| \le a_1\eta + \varepsilon_{\text{num}}$ at bit cost polynomial in $q, K, H, \chi, \gamma^{-1}, b, \log(1/\varepsilon_{\text{num}})$, also $\log(1/\eta)$ for relative precision. The bound is modulo the common automorphism and ideal-basis rotation, not individual singular vectors. Additional error $\le (4K^3)^{-1}$ enlarges caps to $4K^3$; include it in Appendix D's allocation.

Corollary A.6 (Selection without a supplied anchor name)**.** *Let a measured catalogue $\widehat{h}_1, \dots, \widehat{h}_L$ contain a valid anchor with the common envelopes of Theorem 4.2, at unknown index, and coordinate errors at most η. Form the invariant product once and extract on each candidate. Discard undefined, singular, or cap-infeasible outputs, using $K_c = 3K^3$. Select within η of the minimum full-tensor residual $\|\mu_{\widehat{\rho}} - \widehat{\mu}\|_F$. With $R_0 = 1 + 16nK^6a_1$ and $\eta_{\text{cat}} = \min\{\eta_0, [256q^2K_c^3(R_0+2)]^{-1}\}$, $0 < \eta \le \eta_{\text{cat}}$ gives $\inf_\alpha \|\widehat{\rho} - \alpha S\|_{\text{op}} \le 4qK_c^6(R_0+2)\eta$. Computation is polynomial with at most L extraction calls.*

Proof. The good candidate lies inside the cap and has tensor residual at most $R_0\eta$, by (29) and $\|E\|_F \le n\|E\|_{\text{op}}$ for bilinear tensors. Selection and comparison with truth add η each. Proposition A.4 gives the bound regardless of the selected candidate's spectral interpretation.

In finite arithmetic, use precision fixed by the good-candidate bounds. Singular-value intervals certify caps; discard ambiguous outputs. The slack between $2K^3$ and $3K^3$ preserves the good candidate. Split the objective tolerance among its rounding, residual evaluation, and selection. For cap-feasible maps and inverses, output perturbation ϵ changes the tensor bracket by at most $n(6K_c^2 + 2K_c^4)\epsilon$; choose ϵ below η divided by a fixed multiple of this factor. Bad candidates need no isolated internal singular vectors: their approximate outputs are judged by caps and residuals. □

Native compatibility is separate: neither Theorem 4.2 nor its one-mode extension below supplies it. Recovery from $\widehat{\mu}$ and K alone is not claimed. Conversely, $\text{diag}(-1, 0, 1)$ has repeated adjoint eigenvalue differences, which this construction does not separate. The cubic comes from (23), not $\text{tr}((\text{ad}\, x)^3)$, which vanishes in type A.

A.5 Calibration from one separated real mode

This extension relaxes the spectral promise of Theorem 4.2 to one separated real mode, whose complement may have repeated, complex, or defective eigenvalues but must satisfy a quantitative condition. The complete-law and native-coordinate contracts are unchanged, and the experiments, including Section F.1, test only the original algorithm, not this extension.

One-mode condition. For some true catalogue element h, $\|h\| \le H$, suppose that real V, real θ, and real $(q-1) \times (q-1)$ matrix C satisfy

$$Sh = V \begin{pmatrix} \theta & 0 \\ 0 & C \end{pmatrix} V^{-1}, \qquad \text{cond}\, V \le \chi, \qquad \frac{C + C^\top}{2} \succeq (\theta + \gamma)I, \quad \gamma > 0. \tag{30}$$

Thus θ is simple and strictly leftmost in real part. These are input promises, not conditions certified by a small fitted residual. The original real-simple spectrum satisfies (30) by choosing its least eigenvalue and a diagonal complement.

Lemma A.7 (Selection without diagonalizing the complement)**.** *Suppose R is similar, by a matrix of condition at most κ, to the direct sum of θI_q and q copies of $C^\top$ satisfying* (30)*. Let $\|\widehat{R} - R\| \le e$ and $0 < e \le \gamma/[16\kappa(1 + 2\kappa)]$. The q eigenvalues of $\widehat{R}$ with smallest real parts lie within $2\kappa e$ of θ. Their mean real part $\widehat{\theta}$ obeys the same bound. If $U, \widehat{U}$ are orthonormal bases of $\ker(R - \theta I)$ and the q smallest right singular subspace of $\widehat{R} - \widehat{\theta} I$, respectively, they can be aligned so that*

$$\|\widehat{U} - U\| \le \frac{8\kappa(1 + 2\kappa)e}{\gamma}. \tag{31}$$

Proof. For $\operatorname{Re} z < \theta + \gamma$, the numerical-range inequality gives $\|(zI - C)^{-1}\| \le (\theta + \gamma - \operatorname{Re} z)^{-1}$: the real part of $u^*(C - zI)u$ has absolute value at least $(\theta + \gamma - \operatorname{Re} z)\|u\|^2$, and Cauchy–Schwarz bounds the least singular value. On $\operatorname{Re} z = \theta + \gamma/2$ the resolvent of R is therefore bounded by $2\kappa/\gamma$. The Neumann series prevents any eigenvalue of $R + s(\widehat{R} - R)$, $0 \le s \le 1$, from crossing this line, and by continuity of polynomial roots the number on its left stays q.

Set $r = 2\kappa e \le \gamma/2$. On $|z - \theta| = r$ the scalar block has inverse norm $1/r$ and the other blocks at most $1/(\gamma - r) \le 1/r$, so $\|(zI - R)^{-1}\| \le \kappa/r = 1/(2e)$ and the same homotopy has exactly q roots in the disk. The disk lies left of the separating line, proving the selection and mean bounds.

Put $D = R - \theta I$; in similarity coordinates its nonzero block has least singular value at least γ. If $R = ZR_0Z^{-1}$ and $Z^{-1}v = (v_0, v_1)$, then $\|Dv\| \ge \sigma_{\min}(Z)\gamma\|v_1\|$ and $\operatorname{dist}(v, \ker D) \le \|Z\|\|v_1\|$, so $\|Dv\| \ge (\gamma/\kappa)\operatorname{dist}(v, \ker D)$. Let $\delta = (1 + 2\kappa)e$. Testing $\widehat{D} = \widehat{R} - \widehat{\theta} I$ on $\ker D$ shows that its q selected singular values are at most δ; as in the proof of Theorem 4.2, $\|D\widehat{U}\| \le 2\delta$ and $\|(I - UU^\top)\widehat{U}\| \le 2\kappa\delta/\gamma < 1/2$. Aligning principal vectors gives an orthonormal-basis distance at most twice this sine bound, hence (31). □

Theorem A.8 (Stable one-mode calibration)**.** *Retain Theorem 4.2's bracket, norm and error assumptions, including $\|\widehat{h} - h\| \le \eta$, but replace its full real-simple spectrum by* (30)*. Construct* (22)*–*(25) *once. Run* (26) *on both $\widehat{h}$ and $-\widehat{h}$. Write p_1, η_{alg} for the product constants of* (27) *and $\kappa = K^2\chi$, $r_1 = (H + 1)p_1 + K^3$ as in* (28)*, and set*

$$u = 8\kappa(1 + 2\kappa)r_1/\gamma, \qquad a = \sqrt{q}\,(p_1 + 2K^3u), \tag{32}$$
$$\eta_* = \min\{\eta_{\mathrm{alg}}, \gamma/[16\kappa(1 + 2\kappa)r_1], (2K^3a)^{-1}\}.$$

For $0 < \eta \le \eta_*$, at least one output is within $a\eta$ of an automorphism of S and has norm and inverse norm at most $2K^3$.

The useful sign need not be supplied. For a catalogue of L anchors containing such an h at an unknown index, scan its signed catalogue, filter using $K_c = 3K^3$, and select within η of the minimum pulled-back full-tensor bracket residual. With $R_* = 1 + 16nK^6a$ and $\eta_{\mathrm{cat},*} = \min\{\eta_*, [256q^2K_c^3(R_* + 2)]^{-1}\}$, $0 < \eta \le \eta_{\mathrm{cat},*}$ implies

$$\inf_{\alpha \in \mathrm{Aut}(\mathfrak{sl}_q)} \|\widehat{\rho} - \alpha S\|_{\mathrm{op}} \le 4qK_c^6(R_* + 2)\eta. \tag{33}$$

The scan uses at most $2L$ offline extractions and no extra terminal queries. The polynomial numerical-tolerance implementation of Section A.4 applies with these gap and precision parameters.

Proof. The invariant-cubic, inverse and normalization steps in Lemma A.5 and (27) do not use an anchor. They produce tensor error at most $p_1\eta$ from one exact product transported by F_+ or F_-, with both transport norms at most K and product bilinear norm at most K^3. The exact product for F_+ uses the anchor Sh; for F_- the input $-h$ uses $(Sh)^\top$. The latter has similarity $V^{-\top}$ and complement $C^\top$, with the same numerical-range and condition bounds. Consequently one of the signed candidates is covered; silently using the same sign for both products would be invalid.

Vectorized canonical right multiplication has the block form in Lemma A.7; transport to measured product coordinates costs at most K^2, so its similarity condition is at most κ. Since $\eta \le 1$ and $\|h\| \le H$, $\|\widehat{R} - R\| \le (H+1)p_1\eta + K^3\eta = r_1\eta$, and the lemma gives $\|\widehat{U} - U\| \le u\eta$. The selected exact space is the left ideal $\{vw^\top : v \in \mathbb{R}^q\}$ for a unit left eigenvector w of the represented anchor; the ideal's transport and inverse have norms at most K, hence restriction of left multiplication gives $\rho = \alpha S$ with $\|\rho\|, \|\rho^{-1}\| \le K^3$ exactly as in Lemma A.5. Since $\|L_x\| \le K^3\|x\|$ and trace-free projection is Frobenius-contracting, $\|\widehat{\rho}(x) - \rho(x)\|_F \le \sqrt{q}\,(p_1 + 2K^3u)\eta\|x\| = a\eta\|x\|$. The last threshold in (32) gives the $2K^3$ caps by inverse perturbation.

For $E = \widehat{\rho} - \rho$, the pulled-back bracket identity preceding (29) gives operator error at most $16K^6a\eta$, so tensor error at most $16nK^6a\eta$ and measured residual at most $R_*\eta$ for the good candidate. The rest is the proof of Corollary A.6: selection and comparison with the true table add η each, and Proposition A.4 with map and inverse caps K_c gives (33) regardless of which candidate is selected.

For finite arithmetic, follow Section A.4 with (32) in place of its constants: approximate characteristic roots with multiplicity, average the selected real parts, and approximate the singular subspace to a precision fixed by the inverse-polynomial invariant-line, selected-kernel and output-inverse gaps supplied by this proof; no eigenvectors of the defective complement are requested. Coefficient precision, root-tolerance conversion, cap slack and residual rounding are as there, so bit cost stays polynomial in $q, K, H, \chi, \gamma^{-1}$, input length, and logarithms of requested tolerances, with numerical error an additional output tolerance. □

Strict extension and retained boundary. For example, the block-diagonal matrices $\mathrm{diag}(-2, B)$ with $B = I_2$, $\left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, or $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ are traceless and satisfy (30) with $V = I$, $\theta = -2$, and $\gamma = 2$. They have respectively a repeated, nonreal, and defective unselected spectrum, all excluded by the original full real-simple condition. A single real eigenvalue without the quantitative block separation is not the theorem's assumption. Scanning $-\widehat{h}$ negates offline coordinates, not physical duration. Classical extraction from exact matrix algebras is prior work (Ivanyos et al., 2012); the added part is this noisy selection and bound with a potentially defective complement.

An anchor chooses a copy, not a different action law. Suppose the bracket and product are exact and only the anchor coordinates change, with $\delta_h = \|\widehat{h} - h\| < \gamma/(8\kappa K^3)$. For the good sign, exact right multiplication by the perturbed algebra element has exactly q roots in the separated left cluster by the same resolvent argument. Each eigenvalue of its $q \times q$ matrix occurs q times in right multiplication. Hence this cluster corresponds to one simple eigenvalue of that matrix; it is real, since otherwise its conjugate would also lie in the cluster. Its kernel is an exact q-dimensional left ideal. Exact extraction, and exact minimum-residual selection over cap-feasible candidates, therefore still give an automorphism of S, with zero intrinsic representation error. This needs exact arithmetic and exact residual minimization; it does not say anchor gaps are irrelevant when multiplication is noisy, or when a native frame rather than an equivalent representation is required.

A.6 State decoding and native coordinates

In the correct representation, (5) gives $M = s\varphi^\top$, $\varphi^\top s = 0$. The automorphism branches send it to PMP^{-1} and $-PM^\top P^{-1}$. For $q \geq 3$, orthonormalize two independent same-state responses in Frobenius norm: $M_j = s\varphi_j^\top$, $\|s\| = 1$, $\varphi_j \perp s$, $\langle \varphi_j, \varphi_k \rangle = \delta_{jk}$. Then

$$\sum_{j=1}^{2} M_j M_j^\top = 2ss^\top, \qquad \inf_{\|v\|=1} \sum_{j=1}^{2} \|(I - vv^\top)M_j^\top\|_F^2 = 2 - \lambda_{\max}\Big(\sum_j \varphi_j \varphi_j^\top\Big) = 1. \tag{34}$$

Exactly one branch has a common image. For $q = 2$, negative transpose is inner. Once oriented, the leading left singular vector of each nonzero M gives the state line. The pairing is $\mathrm{tr}(MA)$; in canonical orthonormal coordinates the response vector represents $M^\top$, not an indefinite Killing norm.

A complete response stack with R rows of norm at most M_0 and conorm γ supplies two rows with first norm and second orthogonal residual at least $\gamma/\sqrt{R}$. Their area is at least γ^2/R, so their smaller singular value is at least $\gamma^2/(2RM_0)$.

For a simplex-compatible anchor, let V contain its ordered learned native eigenvectors and $s^\star$ its learned uniform reference. The chart is

$$Q = \mathrm{diag}(V^{-1}s^\star)^{-1}V^{-1}, \qquad p = Qs/(\mathbf{1}^\top Qs). \tag{35}$$

Indeed $s \propto Gp$, $V = G\,\mathrm{diag}(t)$, and $s^\star \propto G\mathbf{1}$ give $Q \propto G^{-1}$. Arbitrary projective gauges do not preserve polynomial dictionaries, so eigengaps, reference coordinates, and denominators remain controlled.

Perturbation. Let S_0 be the aligned exact calibration, $\|S_0\|, \|S_0^{-1}\| \leq K_\rho$, and $a = \|\widehat{\rho} - S_0\| \leq (2K_\rho)^{-1}$. For response error e_r, inverse trace pairing gives

$$\|\widehat{M} - M\|_F \leq 2K_\rho e_r + 2K_\rho^2 a\|r\| =: \eta_M. \tag{36}$$

A known valid anchor uses $K_\rho = K^3$, $a = a_1\eta + \varepsilon_{\mathrm{num}}$. Catalogue selection uses $a = 4qK_c^6(R_0 + 2)\eta$ plus rounding and $K_\rho = 2K_c$: inverse perturbation bounds the nearby exact map and its inverse

by $2K_c$ when $a \le (2K_c)^{-1}$. These are checks on the joint confidence event, not deductions from a successful rollout.

For $\eta_M \le \|M\|_F/4$, the state-projector error is at most $2\eta_M/\|M\|_F$. Pair whitening pays its smaller singular value before (34) fixes orientation. Native-frame error is polynomial in the eigenvector, reference, and denominator gaps. For generator norms at most M, Duhamel gives $\|e^{t\widehat{A}} - e^{tA}\| \le te^{tM}\|\widehat{A} - A\|$; projective normalization pays its buffered denominator. This propagates states, actions, and their shared physical clock.

A single-state first response cannot orient $q \ge 3$: the fundamental world at $[e_1]$ with $f = s_2/s_1$ and dual world at $[e_2]$ with $g = -v_1/v_2$ both give value zero and response A_{21}.

Motion-law recovery. Decode complete outputs on the declared pre/post designs in chart (35). After removing redundant features, each projective row $B_i\Phi_i$ satisfies $y_{iq}(B_i\Phi_i)_p - y_{ip}(B_i\Phi_i)_q = 0$. The separated normalized design identifies its coefficient class, with the declared normalization and denominator/chamber bounds. Isolated group and effective-stage residuals reject missing parents; extra zero groups are deleted by minimum support. This identifies the effective dictionary, not arbitrary hidden factorizations. Scalar regression recovers the original score/offset; nonidentity outputs identify writes and nonzero input groups identify reads. Appendix D gives finite fitting and execution. These steps establish Section 6's motion-route execution guarantee under Definition A.1, with initialization, map, dependency and score costs charged.

A.7 Perturbation and query costs

For a full-column-rank equation $BJ = D$,

$$\widehat{J} - J = \widehat{B}^{\dagger}[(\widehat{D} - D) - (\widehat{B} - B)J], \qquad \|\widehat{J} - J\| \le \frac{2(e_D + e_B\|J\|)}{\gamma_B} \quad (e_B \le \gamma_B/2). \tag{37}$$

Apply (37) to the Hall profiles (14), then compute pivot brackets offline. The solve/centroid/support gaps and Section D.2's tensor conversion propagate error to calibration and (36). Section D.1 charges order-three Hall entries, discovery, failed candidates and full replay.

The inverse identity gives $\|(M+E)^{-1} - M^{-1}\| \le 2\|M^{-1}\|^2\|E\|$ when $\|M^{-1}E\| \le 1/2$. Apply these successively to (16) or (17), then (18). If true and fitted finite transports stay within $1/2$ of identity, the logarithm series has Lipschitz constant at most two; it still uses the public nonaliasing branch. If all L_a have norm at most M and suffix duration at most T, their conjugated adjoints have error at most $e^{2TM}(1 + 2TM)\max_a\|\widehat{L}_a - L_a\|$. Raw-coordinate recovery pays $\sigma_{\min}([\mathrm{vec}\, K_i]_i)$; factor and support isolation pay their declared gaps. The resulting bracket and anchor-coordinate errors feed Theorem 4.2, catalogue selection if needed, and (36).

For a confirmed first-duration design, use the centered four-node stencil at positive interior slots, $(\bar{Y}_{-2} - 8\bar{Y}_{-1} + 8\bar{Y}_1 - \bar{Y}_2)/(12r)$. Bias is at most $M_5 r^4/18$, noise proxy $65\sigma^2/(72Nr^2)$. Thus J entries with error ζ cost

$$O\left(\frac{J}{\pi_{\min}}\left[1 + \sigma^2\{r_{\mathrm{leg}}^{-2}\zeta^{-2} + M_5^{1/2}\zeta^{-5/2}\}\right]\log\frac{4J}{\delta}\right). \tag{38}$$

Each doubled-word endpoint adds $2L$ raw calls, or $2L + 2$ with insertions, beyond its history/continuation; state-response arms add L. Discovery and the optional mixed insertion derivatives retain their own entry counts. Second derivatives use the generic allocation in Appendix D, possibly $\widetilde{O}(\zeta^{-4})$ under C^5.

Finite action dimension does not bound exact terminal-mean rank. For the scalar flow $\dot{x} = -x^2$ from $x(0) = 1/2$, the terminal readout $h(x) = x$ gives $x(t) = (2+t)^{-1}$. For distinct positive prefix times $s_1, \dots, s_n$ and suffix times $t_1, \dots, t_n$, the terminal-mean matrix $H_{ij} = (2 + s_i + t_j)^{-1}$ is a Cauchy matrix, with nonzero determinant $\prod_{i<j}(s_j - s_i)(t_j - t_i)/\prod_{i,j}(2 + s_i + t_j)$. The flow has a one-dimensional generated action algebra, but the exact ranks are unbounded. This does not exclude useful finite-rank approximations or nonlinear predictive-state representations.

B RESET ISOLATION AND COMPLETE-LAW RECOVERY

B.1 CONSTANT-OVERWRITE CONTRASTS AND ADDITIVE ACQUISITION

Let R_r overwrite W_r by constants, $Z_i = \{r : i \notin W_r\}$, and compose R_S in a fixed public order without intervening raw calls. Overwritten coordinates forget the earlier history. For additive $h = c + \sum_i h_i(x_i)$, a raw-only continuation C and $g_i = h_i(C_i x_i(H)) - h_i(C_i x_i(H_0))$, the contrasts and their Boolean inversion are

$$\Delta_S = F(HR_SC) - F(H_0R_SC) = \sum_{i:S\subseteq Z_i} g_i, \qquad \sum_{S\supseteq Z} (-1)^{|S|-|Z|}\Delta_S = \sum_{i:Z_i=Z} g_i, \tag{39}$$

since $\sum_{Z\subseteq S\subseteq Z_i}(-1)^{|S|-|Z|}$ is one for $Z = Z_i$ and zero otherwise. The inversion is standard (Smith et al., 2022); constant erasure makes its inputs observable. Resets need not commute, but their pullbacks on additive functions modulo constants are commuting idempotents. Additivity is necessary: subtracting the score after resetting x_2 to r from $h = x_1x_2$ gives $x_1(x_2 - r)$, not one-record information. Copying also need not erase dependence.

Proposition B.1 (Additive reset acquisition)**.** *Let the unknown retention signatures Z_i be distinct and $h = c + \sum_i h_i$ analytic, each h_i future-active. Each record has a locally generating, effective split $\mathfrak{sl}_{q_i}$ action with line-parabolic isotropy, $q_i \geq 2$. Assume connected projected open patches, Lemma A.2's common language and compact probe coverage, a separated Hall or repeated-word entrance on every factor, and compatible native anchors/references, including at requested post-call states. All mask, positioning and pre/post-call words must be legal in the common buffer.*

Terminal contrasts recover signatures, native coordinates, raw supports and reset constants/targets up to permutation. They identify any complete affine call $P(x) = Tx + b$ (including singular T) and its functional reads/writes from projected spanning differences, without joint-open preparations. A finite additive score dictionary independent modulo constants recovers the original score and offset. Finite guarantees retain all local separation, design and deployment conditions; every mask is charged.

Proof. Enumerate $Z \subseteq [k]$ in (39). Its contrast is zero for unoccupied Z; for $Z = Z_i$, differentiating only the earlier history gives exactly the local profiles $D_{w(A_i)}(h_i \circ C_i)(x_i(H))$. Activity and projected openness detect every occupied signature under the isolated stopping schedule, not a plateau of zero estimates. Effectiveness separates local table columns even if records co-move. Use (14), or the shared-covector solves of Section A.2, then Theorems 4.1–4.2 and (35) factorwise. Compact probe coverage includes post-call states; r writes i precisely when $r \notin Z_i$, and decoding its poststate gives the reset constant.

For history pairs ℓ let $D_i = [x_i^{\ell,+} - x_i^{\ell,-}]_\ell$, $V_S^\ell = P(R_S x^{\ell,+}) - P(R_S x^{\ell,-})$, $Y_i = [\sum_{S\supseteq Z_i}(-1)^{|S|-|Z_i|}V_S^\ell]_\ell$ and $T_{:,i} = Y_i D_i^\dagger$. Indeed $V_S^\ell = \sum_{j:S\subseteq Z_j} T_{:,j}(x_j^{\ell,+} - x_j^{\ell,-})$: overwritten constants cancel. Boolean inversion gives $Y_i = T_{:,i}D_i$. Projected openness supplies $q_i - 1$ spanning differences per block, whose union has at most $\sum_i(q_i - 1)$ pairs; none need change only one record. One complete pair fixes b. Nonzero columns give reads and nonidentity rows give writes. The same contrast fits each declared h_i modulo a constant; independence of its analytic features gives a finite difference design. A reference terminal value fixes the aggregate offset. The common execution argument applies. □

For $\sigma_{\min}(D_i) \geq \nu_i$ and $\|\widehat{D}_i - D_i\| \leq \nu_i/2$, (37) (transposed) gives $\|\widehat{Y}_i\widehat{D}_i^\dagger - T_{:,i}\| \leq 2\nu_i^{-1}(\|\widehat{Y}_i - Y_i\| + \|T_{:,i}\|\|\widehat{D}_i - D_i\|)$. With L pairs and decoded Euclidean error η, the contrast error is at most $2^{k-|Z_i|+1}\sqrt{L}\eta$; scalar profiles incur the same signed-weight factor. Charge all 2^k masks, up to k resets per mask, and failed searches. The local bounds and Appendix D then give finite precision, not polynomial cost in k. This additive branch treats fixed affine complete calls; parameterized complete maps are covered by the degree-matched branch below. Raw durations remain editable. Affinity alone does not identify a sparse internal factorization.

B.2 BINARY CALIBRATION

Definition B.2 (Degree-matched reset contract)**.** Fix $1 \le d \le m-2$. Use the anonymous locked fields (9), native reference $x^\star = 1/2$, and unknown constant resets, each writing at least two records. Hidden retention signatures have public weight w and satisfy (10) through order d. Every nested calibration word, argument node, replay, and required post-call decoder is legal in $[a, 1-a]^m$, $a > 0$, including two final probe slots. The original signed polynomial score has degree at most d and is essential in every retained record. Complete enrolled maps $F_b(x, u)$ have joint degree at most d, $u \in [-1, 1]^{p_b}$, with $p_b \le p_0$ fixed. Deployment has its own buffered prefix class and influence bound. Coordinates, assignments, rates, targets, reset constants, and coefficients are hidden.

For $|K| \le d$, R_{S_K} with $S_K = \bigcap_{i \in K} Z_i$ retains exactly K: each exterior record is written by some name in the intersection. Exterior constants may depend on K and reset order.

Suppose the context bank has a visible response $b_e = f'_C(x_i)v_e(x_i)$ at every requested state and occupied signature. Use visible contexts at the reference and two separated small positive minus-flow states, enumerating names until the pilot is found. The response matrix is $B = \operatorname{diag}(f'_1, f'_2, f'_3)(v_e(x_j))_{j,e}$. Independence of the three quadratic fields and separated states give $\|B^{-1}\| \le C/\gamma$ at visible magnitude γ. Antisymmetrized ordered second derivatives give BT. Equation (37) with $J = T$, $D = BT$ and bounded T on the rate box gives rate/state error $C\zeta/\gamma$, not an extra inverse power of γ.

The Killing diagonal $(8\rho_H^2, 0, 0)$ identifies H, while $\kappa(-,+) = 4\rho_-\rho_+$. The minus name has the minority response sign, independently of f'_C. At the reference, $r = -b_-/b_+ = \rho_-/\rho_+$, and

$$\begin{aligned} \rho_H &= \sqrt{\kappa(H,H)/8}, \quad \rho_- = \sqrt{r\kappa(-,+)/4}, \qquad \rho_+ = \sqrt{\kappa(-,+)/(4r)}, \\ x_i &= u_i/(1+u_i), \qquad u_i = \sqrt{-\rho_+ b_-/(\rho_- b_+)}. \end{aligned} \tag{40}$$

Buffers separate roots and denominators. Decoding reset poststates gives targets and constants. Freeze only discrete assignments/signatures; re-estimate numerical rates and coordinates at final precision. Section B.3 supplies the contexts without prior coordinates or signatures.

Lemma B.3 (Conditional calibration from finite values)**.** *Suppose the two root names are correctly assigned and a genuine single-record context induces the same nonconstant rational $f = P/Q$, with $\deg P, \deg Q \le d$, after either preparation. Let $a = \rho_-$, $b = \rho_+$. If f has no pole at $1/2$ or on a common legal positive-time interval, $2d+1$ exact values of each of $r_-(t) = f((2+at)^{-1})$ and $r_+(t) = f(1-(2+bt)^{-1})$ determine a, b, f. Thus $4d+2$ values suffice after context and mode discovery.*

Proof. Each response is rational of numerator/denominator degree at most d. Equality at $2d+1$ non-pole nodes makes the cross-multiplied degree-$2d$ difference vanish. Interpolate reduced functions when degrees drop. Write $f(x) = c_0 + c_k(x-1/2)^k + c_{k+1}(x-1/2)^{k+1} + \cdots$. Nonconstancy and $Q(1/2) \neq 0$ give $1 \le k \le d$. Offline expansions $r_\pm(t) = c_0 + A_\pm t^k + B_\pm t^{k+1} + \cdots$ satisfy $A_- = c_k(-a/4)^k$, $A_+ = c_k(b/4)^k$, $B_-/A_- = -ka/2 - c_{k+1}a/(4c_k)$ and $B_+/A_+ = -kb/2 + c_{k+1}b/(4c_k)$. Consequently $\chi = |A_+/A_-|^{1/k} = b/a$, $a = -k^{-1}(B_-/A_- + B_+/(\chi A_+))$ and $b = \chi a$. Then $f(x) = r_-((x^{-1}-2)/a)$ as a rational identity, with removable singularities continued. All physical observations have positive durations. □

The lemma does not identify the H rate, contexts or assignments, and is not uniformly noise-stable: $f_\eta = c + \eta x$ loses rate information as $\eta \to 0$; response strength, poles, node design and branch isolation must still be controlled. The derivative pilot and its finite allocation remain the general guarantees.

B.3 NESTED PREPARATIONS AND POLYNOMIAL DETERMINATION

Put $\tau = T_{\mathrm{cal}}/(2d+1)$ and $\Delta_x = (4/5)a^2\tau/d$. For ordered $K = \{i_1, \ldots, i_s\}$, $s \le d$, and $K_j = \{i_1, \ldots, i_j\}$, execute in time order

$$\Phi_-^{t_1} R_{S_{K_1}} \cdots \Phi_-^{t_s} R_{S_{K_s}}. \tag{41}$$

Record i_j is last overwritten before layer j, then retained. With b_1 its initial value and other b_j the last constants, $x_{i_j} = b_j/(1+\rho_- b_j r_j)$ with $r_j = \sum_{\ell=j}^{s} t_\ell$. Independent cumulative nodes

$r_{j,n} = (2(s-j)+1+n/d)\tau$, $0 \le n \le d$, lie in ordered disjoint intervals. They give positive physical durations and coordinate gaps Δ_x, with exterior coordinates constant within each grid. The union G over all K, including the reference, has $Q_d = 1 + \sum_{s=1}^{d} \binom{m}{s}(d+1)^s$ points; with $N_d = \binom{m+d}{d}$ and $A_d = [(d+1)2^d\Delta_x^{-d}]^d$, every degree-$d$ polynomial obeys

$$\|\mathrm{coef}(p)\|_\infty \le (1 + N_d A_d)^{d+1} \max_G |p|. \tag{42}$$

Indeed interpolate a maximum-total-degree monomial on its support grid. A term with the same support exponents and an exterior variable would have larger degree, so exterior substitution preserves the coefficient. Recover the top degree, subtract on the actual grids, and descend. Univariate Lagrange numerators have coefficient sum at most 2^d and denominators at least Δ_x^d; summing nodes over at most d variables costs A_d, and each descent costs $1 + N_d A_d$. This allows different exterior constants and requires no well-conditioned square subsystem. Grids placing i first similarly norm degree-$(d-1)$ polynomials in d descents.

For blind discovery, enumerate ordered w-subsets $U_1, \dots, U_s$ of names, replace S_{K_j} by $\bigcap_{\ell \le j} U_\ell$ in (41), and enumerate all three raw names and the same nodes. Differentiate before the context. Equal weights imply that U_1 is contained in a true signature only when equal: an unoccupied U_1 erases all earlier input, whereas $U_1 = Z_i$ retains exactly input i. Thus false signatures give zero responses and true ones give one-record profiles. Define $V_d = 3\sum_{s=1}^{d} \binom{k}{w}^s (d+1)^s$ and $\kappa_i = \max_{\alpha_i \ge 0} |c_\alpha(h)|$. The derivative version of (42) gives a visible row of magnitude at least $\Gamma_i = (4/5)a^2\kappa_i/(1 + N_d A_d)^d$. The true tuple/minus name occurs in the bank and gives $v_-(z_i)\partial_i h(z)$; all earlier resets retain i. Each context costs $O(d(k+1))$ physical calls. Apply (40), measure the original score, and decode complete outputs on G times the $(d+1)^{p_b}$ argument grid. Equation (42) and argument interpolation identify coefficients and the functional scopes $P_b(j) = \{i : \partial_i F_{b,j} \not\equiv 0\}$ and $W_b = \{j : F_{b,j} - x_j \not\equiv 0\}$, with analogous argument groups. A constant reset may write without parents. This proves Theorem 5.1 for complete maps, not internal factorizations.

Thinness and its boundary. For prime $p \ge d+1$, index $m = p^2$ records by affine polynomials over $\mathbb{F}_p$. At $d+1$ positions, reset (t, c) writes the p records evaluating to c. If $f \notin K$, $|K| \le d$, then $\prod_{g \in K}(f - g)$ is a nonzero degree-d polynomial, so some position separates f from K. This standard cover-free code has $k = (d+1)p$ and $w = k - (d+1)$. Under locked PGL_2 motion, records sharing their last reset name, or no reset, move by one group element. Finitely many assignments give preparation dimension at most $3(k+1)$. For $d = 3, p = 17$, this is $207 < 289$, although (42) excludes every nonzero vanishing cubic. Nested resets reveal $(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)$ even though synchronized and simple equal-survivor states do not. Conversely, if each reset writes at least two equal constants, the last reset leaves some equal pair through the locked suffix. Thus $\prod_{i<j}(x_i - x_j)^2$ vanishes on all raw/reset preparations: bounded degree cannot be replaced by arbitrary analyticity.

B.4 The information lost by grouped reset queries

Query contract. Let $m \ge 5$ records start at $1/2$, with locked raw fields (9). All pair-reset names are available and overwrite their two records by $1/2$. A grouped word consists of a raw-only prefix, one contiguous block of pair resets, and a raw-only suffix. The buffer admits the small positive words below. Denote their endpoint set by $\mathcal{G}$, and the real polynomials of total degree at most d by $\mathcal{P}_{\le d}(m)$. Coordinates in this definition are analysis variables, not learner inputs.

Theorem B.4 (The complete cubic ambiguity)**.** *Under this query contract,*

$$\{p \in \mathcal{P}_{\le 3}(m) : p|_{\mathcal{G}} = 0\} = \mathrm{span}\{(x_i - x_m)(x_j - x_m)(x_i - x_j) : 1 \le i < j < m\}. \tag{43}$$

These elements are independent. There are $\binom{m-1}{2}$ invisible cubic directions and no nonzero invisible polynomial of degree at most two. The largest rank of a finite grouped cubic evaluation design is $\binom{m+3}{3} - \binom{m-1}{2}$: 50 among 56 features when $m = 5$. Available duration derivatives within the grouped language do not remove this ambiguity.

Proof. Locked motion preserves equality before the reset block. At its end, each coordinate is either unchanged or reset to $1/2$; the common suffix preserves the two equality classes. Thus every endpoint has at most two values.

Conversely, in every nontrivial bipartition choose a class of size at least two and overwrite exactly that class by pair resets within it. A small minus-flow prefix and plus-flow suffix, with their positive rates absorbed into u, v, produce $a(u, v) = (1+v+uv)/(2+u+v+uv)$ and $b(v) = (1+v)/(2+v)$; the Jacobian determinant $\det \partial(a, b)/\partial(u, v) = -(v + 2)^{-2}(uv + u + v + 2)^{-2}$ never vanishes. A small positive rectangle is legal and supplies a relatively open subset of each two-class plane. A polynomial vanishing on $\mathcal{G}$ therefore vanishes on every such plane.

Set $t = x_m$, $y_i = x_i - x_m$ for $i < m$, and write $p = \sum_{a+|\beta|\le 3} c_{a,\beta} t^a y^\beta$. For every $S \subseteq \{1, \dots, m-1\}$, p vanishes on the two-class plane $y = z\mathbf{1}_S$ with t, z independent, so each coefficient of $t^a z^k$, namely $\sum_{|\beta|=k,\ \operatorname{supp}\beta\subseteq S} c_{a,\beta}$, is zero. Boolean inversion gives a zero sum on each exact support. For $k \le 2$, each possible support has a unique monomial at that degree. For $k = 3$, singleton and triple supports also have a unique monomial, whereas $\{i, j\}$ has exactly $y_i^2 y_j$ and $y_i y_j^2$. Total degree three excludes any additional t factor. Hence the only free directions are their differences, as in (43); distinct pair supports give independence.

Subtracting the kernel dimension gives the stated rank. A finite design attaining it exists: unless selected evaluation functionals span the restricted space's dual, a nonzero restricted function vanishes on those points and is nonzero at another reachable point, which increases the rank. This does not guarantee the conditioning, cost, or maximal rank of every finite bank. Finally, invisible polynomials vanish identically on each grouped duration cell, so all derivatives within that language vanish. □

What is known, and what this adds. The factors are Vandermonde 3×3 minors. Watanabe & Yanagawa (2019, Theorem 1.1 and Proposition 2.1) establish the general ideal and its radicality. The proof above supplies the additional access-to-variety correspondence and an elementary degree-three specialization. With $K_{ij} = (x_i - x_5)(x_j - x_5)(x_i - x_j)$, one explicit three-record example is $(x_1 - x_2)(x_2 - x_3)(x_3 - x_1) = -K_{12} + K_{13} - K_{23}$, one element of a six-dimensional kernel. Every admissible cubic score has indistinguishable sufficiently small perturbations along that kernel, as long as the coefficient and essentiality conditions remain satisfied. Under its own contract, Theorem 5.1's nested preparation bank has zero kernel at the same degree. This is a distinction between query languages, not a universal superiority claim for designed over random interleaving.

Complete transcripts and imperfectly balanced menus. For a fixed common two-candidate loss menu $\ell_{\pm,j} = b_j \pm v_j$, write $d_b = b_1 - b_2$ and $d_v = v_1 - v_2$. Let P_+, P_- denote the laws of all information given to a decision rule, including discovery, calibration, selection and randomness. Its equal-prior regret obeys

$$\mathcal{R} \ge \frac{(|d_v| - |d_b|)_+}{2}\big(1 - \mathrm{TV}(P_+, P_-)\big). \tag{44}$$

For $|d_b| \ge |d_v|$ this is trivial. Otherwise the optima reverse and both gaps are at least $|d_v| - |d_b|$. The sum of the two error probabilities is at least $1 - \mathrm{TV}(P_+, P_-)$, proving the result. With identical transcripts and $d_b = 0$, the original $\Delta/2$ identity holds for every rule, $\Delta = |d_v|$. With unequal bases, the sharp identical-transcript Bayes value is $(|d_v| - |d_b|)_+/2$, attained by choosing the better base.

Suppose all world-dependent information consists of at most N adaptive Gaussian observations, with the same predictable conditional variance $\sigma_t^2 > 0$ in both worlds; auxiliary randomness and side information have a common law. Let g_t be the conditional mean difference at the same history and query. The chain rule and Gaussian likelihood ratio give $\mathrm{KL}(P_+\|P_-) = \mathbb{E}_+ \sum_{t=1}^N g_t^2/(2\sigma_t^2) \le \Lambda_{\mathrm{KL}}$. Stopped records can be padded with ignored common observations. By Pinsker's inequality $\mathrm{TV} \le \sqrt{\mathrm{KL}\,/2}$ (Tsybakov, 2009), Equation (44) then holds with factor $(1 - \sqrt{\Lambda_{\mathrm{KL}}/2})_+$.

Finite-duration coefficients. The statistic in Section F.2 is $c_{\mathcal{T}} = \sum_t t^2 D_{ab}(t)/\sum_t t^4$. If $D_{ab}(t) = ct^2 + r(t)$ and $|r(t)| \le Mt^3$ on the legal interval, direct substitution gives $|c_{\mathcal{T}} - c| \le M \sum t^5/\sum t^4 \le M \max \mathcal{T}$. If both paired-word predictions have error at most ϵ, their fitted-statistic error is at most $2\epsilon \sum t^2/\sum t^4$.

B.5 FINITE PRECISION AND COMPLETE-LAW COST

Let $M \geq 1$ bound score/map coefficients, $\kappa_h \leq \min_i \kappa_i$, and let $\kappa_s > 0$ lower-bound the largest coefficient in every nonzero input/argument group or nonidentity row $F_{b,j} - x_j$. Define

$$\begin{aligned} N &= \binom{m + p_0 + d}{d}, \qquad B_u = \sum_b (d+1)^{p_b}, \qquad A_u = [(d+1)d^d]^{p_0}, \\ L_d &= A_u(1 + NA_d)^{d+1}, \quad \Gamma = \min\{1, c_a\kappa_h/L_d\}, \quad K_0 = C_a/\Gamma, \\ \Lambda &= dNM, \qquad K_c = C_a L_d(1 + NM)K_0, \quad K_m = NK_c. \end{aligned} \tag{45}$$

Constants retain fixed degrees, rate/buffer boxes, calibration time, and anchor separation. At elementary jet/mean error ζ and sufficiently small $L_d N K_0 \zeta$, Appendix D's errors-in-variables bound gives coefficient error $K_c\zeta$ and map/score error $K_m\zeta$. Threshold at $\kappa_s/2$, refit absent groups to zero and unwritten rows to identity; truth remains feasible. Selected responses remain bounded below by a constant times Γ.

Starting at $R_0 = K_0$ (zero at the exact reference), let $R_\star$ bound every admitted prefix and partial-duration instance of

$$R^+ = \begin{cases} e^{C_a t}R + C_a K_0 t e^{C_a t}, & \text{raw duration } t, \\ \Lambda R + K_m, & \text{complete call}, \\ R + K_0, & \text{reset}. \end{cases} \tag{46}$$

The conservative reset bound permits arbitrary mixtures. Require

$$\zeta \leq c_a \min\{\Gamma, (L_d N K_0)^{-1}, \kappa_s/K_c, a/R_\star, \varepsilon/[K_m + (1+\Lambda)R_\star]\}. \tag{47}$$

An independent discrete pilot and fresh refinement use $J \leq C_a V_d(1 + k + B_u Q_d) + Q_d$ and $N_{\text{end}} \leq N_2(J, \zeta)$, with (62)'s mean envelope $C_a NM$. Physical work is $O(d(k+1)N_{\text{end}})$ raw/reset occurrences plus complete-call internals, including decoding of every raw/reset/pre/post state. Coordinate error is $K_0\zeta$, design error is absorbed in (45), and state error is $R_\star\zeta$. Evaluate current score/cache errors once from this state. Equation (47) preserves buffers and separated scopes and yields operational error ε with failure $\delta_0 + \delta$. The possibly exponential $R_\star$ precludes a general practical or dimension-optimal acquisition claim.

C LOCAL-LAW RECOVERY FROM MIXED PREPARATIONS

Retain Definition B.2's locked binary fields, unknown constant resets, reference, replay, and legal pre/post probes. The original degree-d_h score is essential in every retained record. The unknown program $P(x) = y$ of Section 5.2 satisfies (11) on the full local unit cube, with coefficient bound M and outputs in $[a_\star, 1 - a_\star]$. Parents are essential polynomial arguments. Put $t = 2\Delta + 1$ and require equal-weight separation (10) through $\max\{d_h, t, \max_b d_b\} \leq m - 2$. Here d_b bounds separately enrolled complete-map payloads, not the composed degree of P. Appendix B constructs native coordinates at degree d_h; coordinates, order, parents, and intermediate stages are not supplied.

C.1 MIXED INVERTIBILITY AND TARGETED DESIGNS

For $|K| = t$, $i \in K$, $U = K \setminus \{i\}$, and fixed exterior sources, let $\Phi(x_K) = (x_i, P_U(x))$. In the true topological order, used only for analysis, set $D_{jj} = \partial_{x_j} f_j$ and $B_{j\ell} = \partial_{y_\ell} f_j$. Then $J = DP = (I - B)^{-1}D$ is triangular. Replacing row i by the source identity and taking the principal minor gives $\det D\Phi = \prod_{j \in U} D_{jj}$ and $\|D\Phi\|_\infty \leq L := \max\{1, A/(1-\beta)\}$. The map is injective: traverse the true order with exterior sources and x_i fixed; each retained output is strictly increasing in its own source after predecessors are determined.

For $O = [m] \setminus U$, including i, Schur elimination retains all exterior paths:

$$\begin{aligned} (J_{UU})^{-1} &= D_U^{-1}\{I - B_{UU} - B_{UO}(I - B_{OO})^{-1}B_{OU}\}, \\ (J_{UU})^{-1}J_{Ui} &= D_U^{-1}B_{UO}(I - B_{OO})^{-1}e_i D_{ii}. \end{aligned} \tag{48}$$

The mixed inverse has rows $(1, 0)$ and $(-(J_{UU})^{-1}J_{Ui}, (J_{UU})^{-1})$. To bound it, let $C = |B|$ and $F = C_{UU} + C_{UO}(I - C_{OO})^{-1}C_{OU}$. Since C is nilpotent with row sums at most β, interpret it

as transition weights with missing mass sent to a cemetery state. The entries of F count disjoint first-return paths from U to U, of total probability at most the first-edge probability β. Likewise $C_{UO}(I - C_{OO})^{-1}e_i$ counts paths first stopped at i without another visit to U; acyclicity excludes revisiting i, so its row sums are at most β. Taking absolute values in (48) gives

$$\|(D\Phi)^{-1}\|_\infty \le B_\partial := \max\{1, [1 + (1 + A)\beta]/\alpha\}. \tag{49}$$

This improves the separate-resolvent bound $\max\{1, (1 + A\beta)/(\alpha(1 - \beta))\}$.

Equation (41) prepares a source rectangle containing $c + [-\ell, \ell]^K$, where $\tau = T_{\mathrm{cal}}/(2t + 1)$ and $\ell = (2/5)a^2\tau$. Its mixed image contains $\Phi(c) + [-b, b]^t$, $b = \ell/(2B_\partial)$. To see this, lift $\Phi(c) + sv$ by $x'(s) = (D\Phi(x(s)))^{-1}v$, $x(0) = c$. For $\|v\|_\infty \le b$, the lift moves by at most $\ell/2$; compactness and nonsingularity extend it to $s = 1$. This inverse is analytical, not a physical call.

Set $N_t = \binom{t+r}{r}$, $C_0 = (r + 1)^t(r/b)^{rt}$, $v_{\max} = 6/5$, $\delta_x = \min\{\ell, (4C_0rN_tL)^{-1}\}$, and $Q_{\mathrm{loc}} = \binom{m}{t}[2 + \lceil \tau v_{\max}/\delta_x \rceil]^t$. Rational cumulative-duration meshes of width $\delta_x/v_{\max}$ prepare state sets G_K using at most Q_{loc} preparations. Tensor interpolation on the mixed cube has coefficient norm at most C_0: univariate numerator coefficient sum is at most 2^r and denominator at least $(2b/r)^r$. For $\deg p \le r$, $\mathrm{Lip}_\infty(p) \le rN_t\|\mathrm{coef}\, p\|_\infty$. Lifting and rounding interpolation nodes therefore loses at most $rN_tL\delta_x\|\mathrm{coef}\, p\|_\infty$, which is absorbed to give

$$\|\mathrm{coef}\, p\|_\infty \le 2C_0 \max_{x \in G_K} |p(x_i, P_U(x))| \qquad (\deg p \le r). \tag{50}$$

C.2 A SHARED PREPARATION POOL

Fix rational $0 < T \le 1$, let $v_{\min} = (4/5)a^2$, and define

$$\begin{aligned} &H = \min\{h \ge 0 : B_\partial A v_{\max}\beta^{h+1}/[v_{\min}(1 - \beta)] \le 1/2\}, \\ &s_0 = 1 + (t - 1)\sum_{j=0}^{H} \Delta^j, \qquad s_{\mathrm{pool}} = \max\{d_h, t, s_0 - 1, \max_b d_b\} \le m - 2. \end{aligned} \tag{51}$$

Require (10) through s_{pool} and legality in the common buffer of every all-reset permutation with the following intervening durations, every prefix, replayed complete call, and decoder probe, including continuous duration limits. Put $\tau_p = T/(2k + 1)$, $B_q = 2B_\partial/v_{\min}$, $L_q = v_{\max}L$, $b_p = \min\{\tau_p/(8B_q), a/2, a_\star/2, 1/2\}$, $\lambda_c = \binom{k}{t}^{-1}[2b_p/(\tau_pL_q)]^t b_p^{2r}/(144^rN_t)$, and $\lambda = \lambda_c/2$. Sample a uniform reset permutation π. Independently sample cumulative times q_j on an equal-width rational midpoint grid in $I_j = [(2(k - j) + 1)\tau_p, (2(k - j) + 2)\tau_p]$, with $h = \tau_p/N \le \min\{\tau_p/4, \lambda_c/(8rN_tL_q)\}$. Execute $R_{\pi(1)}, \Phi_-^{q_1 - q_2}, \ldots, R_{\pi(k)}, \Phi_-^{q_k}$. All durations are positive and sum to $q_1 < T$. The sampler knows the pilot's minus name, but not targets, constants, graph, or a successful permutation.

Proposition C.1 (Shared mixed covariance)**.** *For each fixed admissible program, let ϕ_r contain the N_t monomials of total degree at most r in t variables, and let $M_t = m\binom{m-1}{t-1}$. Independent repetitions with $Q_{\mathrm{pool}} \ge (8N_t/\lambda)\log(2N_tM_t/\delta_{\mathrm{design}})$ satisfy, simultaneously for every $i \in K$, $|K| = t$,*

$$\frac{1}{Q_{\mathrm{pool}}}\sum_a \phi_r(z_a)\phi_r(z_a)^\top \succeq \frac{\lambda}{2}I, \qquad z_a = (x_i^{(a)}, P_{K\setminus\{i\}}(x^{(a)})), \tag{52}$$

with probability at least $1 - \delta_{\mathrm{design}}$.

Proof. First use continuous uniform times in the interiors of I_j. Every record has a writing reset. If $\ell(j)$ is its last writer's position, then $x_j = a_{\pi(\ell(j)),j}/[1 + \rho_- a_{\pi(\ell(j)),j}q_{\ell(j)}]$. Each source depends on one time, with derivative $-\rho_-x_j^2$ of magnitude in $[v_{\min}, v_{\max}]$; the full source Jacobian has row-sum norm at most $v_{\max}$.

Fix i, K, U and set $S = \{i\} \cup \mathrm{Anc}_{\le H}(U)$, including length-zero paths, so $|S| \le s_0$. Separation supplies distinct names h_j for $j \in K$, each writing j and no other member of S. They occupy the last t positions with probability $\binom{k}{t}^{-1}$. Condition on such a permutation and all other times. Targets have distinct free times and sources in $S \setminus K$ are fixed, while exterior ancestors may co-move. Every path from S^c to U has length at least $H + 1$; the finite expansion $J = \sum_{j \ge 0} B^j D$

gives $\|J_{U,S^c}\|_\infty \le A\beta^{H+1}/(1-\beta)$. For the actual mixed map $\Psi(q_K) = (x_i, P_U(x))$ and the fixed-exterior map Φ at the same state, $D\Psi = D\Phi\,\mathrm{diag}(\nu_j : j \in K) + E$ with $v_{\min} \le |\nu_j| \le v_{\max}$ and $\|E\|_\infty \le Av_{\max}\beta^{H+1}/(1-\beta)$. Equations (49) and (51) imply $\|(D\Psi)^{-1}\|_\infty \le B_q$ by a Neumann bound, while $\|D\Psi\|_\infty \le L_q$. Segment lifting from the selected-time center covers a mixed cube of halfwidth b_p: displacement is at most $B_q b_p \le \tau_p/8$, inside the central half. Global injectivity is unnecessary.

On any cube $z_0 + [-b_p, b_p]^t \subset [0,1]^t$, tensor Legendre expansion gives

$$\frac{1}{(2b_p)^t}\int_{z_0+[-b_p,b_p]^t} p(z)^2\,dz \ge \frac{b_p^{2r}}{144^r N_t}\|\mathrm{coef}\,p\|_2^2. \tag{53}$$

The degree-j Legendre recurrence bounds coefficient ℓ_1 norm by 3^j; normalization contributes at most $3^{j/2}$. Each tensor basis element of total degree at most r thus has coefficient norm at most 6^r. Cauchy–Schwarz costs $\sqrt{N_t}$ and substitution $(z - z_0)/b_p$ costs $(2/b_p)^r$; squaring gives (53).

The row bound yields $|\det D\Psi| \le L_q^t$. By the area formula, with multiplicity at least one on the covered cube, the time integral of $p(\Psi)^2$ is at least L_q^{-t} times its cube integral. Multiply by time density τ_p^{-t} and probability $\binom{k}{t}^{-1}$, then average the conditioned times: covariance is at least $\lambda_c I$, uniformly in those values.

Round times to rational cell midpoints. For $\|\mathrm{coef}\,p\|_2 = 1$, $|p| \le \sqrt{N_t}$ and $\mathrm{Lip}_\infty(p) \le r\sqrt{N_t}$. The mixed state changes by at most $L_q h$ and its squared response by at most $2rN_tL_qh \le \lambda_c/4$. Hence the implemented sampler has covariance at least λI. Since $\|\phi_r\|^2 \le N_t$, matrix Chernoff (Tropp, 2012) bounds failure for one set by $N_t \exp[-Q_{\mathrm{pool}}\lambda/(8N_t)]$. A union bound over M_t sets proves the claim. □

At fixed $\Delta, r, \alpha, A, \beta, a, a_\star, T$, the quantities H, s_0 are fixed, $b_p = \Theta(k^{-1})$, and $\lambda_c = \Theta(k^{-(t+2r)})$, giving (13). This counts decoded complete-call pairs for each fixed unknown admissible program, with simultaneous coverage of all candidate differences under this section's ancestor-separation and all-permutation legality conditions. It yields the same coefficient, parent and execution guarantees as the targeted design after scalar decoding and precision costs, which remain additional. Fixed-strength s codes with logarithmic alphabet follow from L_0 independent Bernoulli writing bits of probability $1/(s+1)$ and their complements. Signatures have weight L_0, $k = 2L_0$, and separation of a designated record from s others has probability at least $1/[4(s+1)]$ per position. A union bound over m^{s+1} choices gives $L_0 = O(s^2 \log m)$. Binomial tails exclude masks writing fewer than two or more than $m-2$ records for large m (Idalino & Moura, 2026). This does not establish permutation legality or cheap blind decoding. For $k = \Theta(\sqrt{m})$, $\Delta = 1$, $r = 2$, the pool bound $O(m^{7/2}\log m)$ is worse than targeted $O(m^3)$.

Ancestor protection is needed in this proof, though not asserted necessary for every shared design. In centered variables take $v_4 = z_4/2$, $v_2 = z_2/8 - v_4/4$, $v_3 = z_3/4$. If records $2, 4$ share last writer and constant, then $v_2 = 0$ despite positive own-source derivatives, contractive parents, and distinct last writers for targets $\{1,2,3\}$.

C.3 Fitting, finite precision, and execution

For both designs,

$$\|\mathrm{coef}\,p\|_\infty \le L_E\|p\|_E, \tag{54}$$

with maximum residual and $L_E = 2C_0$ for targeted grids, or empirical RMS and $L_E = \sqrt{2/\lambda}$ on the pool's design event. For each row i, enumerate $C \subset [m]\setminus\{i\}$, $|C| \le \Delta$, and fit degree-r $g(x_i, y_C)$ to y_i. A zero-residual candidate differs from truth on one source and at most 2Δ outputs. Equation (54) forces every difference coefficient to vanish, so minimum cardinality selects exactly the essential parents. Their graph gives a topological order.

Let $n_{\mathrm{loc}} = \binom{\Delta+1+r}{r}$ and decoded error be at most e. Clip fitting data, not executed polynomials, to $[0,1]$. Accept coefficient-bounded fits of residual at most $R_e = (1 + rn_{\mathrm{loc}}M)e$. Section D.2 with $N = n_{\mathrm{loc}}$ gives $\|\mathrm{coef}(g - f_i)\|_\infty \le C_\theta e$, where $C_\theta = 2L_E(1 + rn_{\mathrm{loc}}M)$; set $\eta_f = n_{\mathrm{loc}}C_\theta e$. Embed coefficients in the union-parent basis. If each essential group contains a coefficient of magnitude κ_s

and $C_\theta e < \kappa_s/2$, no accepted set omits it. Require $\eta_f \le a_\star/2$ and $r n_{\mathrm{loc}} C_\theta e \le \frac{1}{2}\min\{\alpha, 1-\beta\}$ to retain the buffer, positive own-source derivatives, and contractive parents. Traversing the true derivatives gives within-call bounds, not temporal stability:

$$\|\widehat{P}(x) - P(x)\|_\infty \le \frac{\eta_f}{1-\beta}, \qquad \|\widehat{P}(\widehat{x}) - P(x)\|_\infty \le \frac{A\|\widehat{x} - x\|_\infty + \eta_f}{1-\beta}. \tag{55}$$

With score coefficient bound $M_h \ge 1$, activity $\kappa_h > 0$, and Appendix B's $A_{d_h}, V_{d_h}, Q_{d_h}$, set

$$\begin{aligned} N_h &= \binom{m+d_h}{d_h}, & L_h &= (1 + N_h A_{d_h})^{d_h+1}, \\ \Gamma_h &= \min\{1, c\kappa_h/L_h\}, & K_{\mathrm{dec}} &= C/\Gamma_h, \\ K_h &= 4N_h L_h (1 + d_h N_h M_h) K_{\mathrm{dec}}, & L_h^{\mathrm{score}} &= d_h N_h M_h, \\ e &= K_{\mathrm{dec}}\zeta, & K_P &= n_{\mathrm{loc}} C_\theta K_{\mathrm{dec}}/(1-\beta), \\ J &\le C V_{d_h}(1 + k + Q_{\mathrm{design}} + Q_{d_h}) + Q_{d_h}, & N_{\mathrm{end}} &\le N_2(J, \zeta). \end{aligned} \tag{56}$$

Here $Q_{\mathrm{design}} \in \{Q_{\mathrm{loc}}, Q_{\mathrm{pool}}\}$ and the scalar envelope is $C N_h M_h$. Decoder slots follow the complete call, so derivatives avoid its expanded degree. Charge every replay: $O((t+d_h)(k+1)N_{\mathrm{end}})$ raw/reset occurrences plus internal program work. Freeze the pool before fresh estimates. Conditional scalar failure is δ_{data}; total failure is $\delta_0 + \delta_{\mathrm{data}} + \delta_{\mathrm{design}}$, or $\delta_0 + \delta_{\mathrm{data}}$ for targeted grids. The bank size V_{d_h} contains $\binom{k}{w}$, potentially large even for $k = O(\log m)$.

Use (46) with $K_0 = K_{\mathrm{dec}}$ for raw/reset steps and $R^+ = A(1-\beta)^{-1}R + K_P$ for programs. For a uniform prefix/partial-duration bound $R_\star$, impose the preceding group and buffer conditions and

$$\zeta \le c\min\{\Gamma_h, a/R_\star, \varepsilon/[K_h + K_P + (1 + L_h^{\mathrm{score}} + A/(1-\beta))R_\star]\}. \tag{57}$$

State, score, and current-cache errors are then at most ε; cache defects are charged once. Storage is $O(m(n_{\mathrm{loc}} + \Delta))$ and full evaluation $O(m(r n_{\mathrm{loc}} + \Delta))$. Parent search uses $m\sum_{j=0}^{\Delta}\binom{m-1}{j}$ fits; execution retains versions, features, incoming edges, queues, and score reductions.

Examples and boundaries. For $z_i = x_i - 1/2$, $v_i = y_i - 1/2$, the chain $v_1 = z_1/4 + z_1^2/32$, $v_j = z_j/4 + z_j^2/32 + (v_{j-1} + v_{j-1}^2)/16$ has $r = 2$, $\Delta = 1$, $\alpha = 7/32$, $A = 9/32$, $\beta = 1/8$, outputs in $[47/128, 87/128]$, and exact complete degree 2^j, with depth-independent local bounds.

This chain has $\mathrm{Lip}(P) \le 9/28$; high exact degree is not an approximation lower bound. In contrast, $f(x) = 1/2 + (6/5)(x - 1/2) - (7/5)(x-1/2)^3$ has no parents but expands near $1/2$, so temporal error retains all raw/payload Lipschitz factors. Finally, remove own-source nondegeneracy and set $v_1 = (x_1 - 1/2)/4$, $v_2 = v_1^2/4$, $v_3 = (x_3 - 1/2)/4 + c(v_2 - v_1^2/4)$, $|c| \le 1/8$. These buffered contractive quadratic recursions have identical complete calls but different essential parents at $c = 0$ and $c \ne 0$. Identical tools preserve every terminal observation. Extra affine records and an essential score embed them in a larger reset code: complete-map recovery remains possible, uniform local-recursion recovery does not.

D Estimation, execution, and excitation

Costs include discovery and rejected attempts; finite guarantees apply on the stated isolated, buffered branch.

D.1 Positive-node estimation

An independent pilot fixes J entries $\partial^{\alpha_j} f_j(t_j^\star)$ with $1 \le s_j \le 3$ variable durations, $k_j = |\alpha_j| \le 3$, and linear assembly of row-sum norm A_{asm}. Each mean is legal on the positive interior of $[b_j, b_j + a_j]^{s_j}$, $b_j \ge 0$, and has a public holomorphic bound $|f_j| \le M_j$, $M_j \ge 1$, on its tensor Bernstein ellipse $E_2^{s_j}$. Targets may lie at a boundary such as zero, without executing zero or negative durations. Real analyticity alone does not give the width or envelope (Demanet & Townsend, 2019). Numerical bases and coordinates are re-estimated after the pilot.

For one entry write s, k, a, M and $C_{s,k} = s(1 + (2k)!)2^{5k+3s+1}$. Choose the least $n \ge \max\{1, k\}$ satisfying

$$C_{s,k} M a^{-k} (n+1)^{2k+s} 2^{-n} \le \eta/4. \tag{58}$$

Tensor Chebyshev–Gauss interpolation uses $(n+1)^s$ interior nodes and derivative weights

$$\partial^\alpha I_n f(t^\star) = \sum_\nu w_\nu f(t_\nu), \qquad W := \sum_\nu |w_\nu| \le (2/a)^k n^{2k} [2(n+1)]^s. \tag{59}$$

Its bias is bounded by (58)'s left side. Omit zero weights and allocate fresh samples

$$N_\star = 1 \vee \left\lceil 32\sigma^2 W^2 \eta^{-2} \log(4J/\delta) \right\rceil, \qquad N_\nu = \lceil N_\star |w_\nu| / W \rceil. \tag{60}$$

Across all entries let $K_q = \sum_{j,\nu} N_{j,\nu}$. With logged conditional exposure at least $\pi_{\min}$, total failure is at most δ after at most

$$\frac{2}{\pi_{\min}} [K_q + 4\log(2/\delta)] \quad \text{attempts}, \qquad K_q \le \sum_j [(n_j+1)^{s_j} + N_{\star,j}]. \tag{61}$$

Unfilled quotas abstain and still count; a direct mean has one unit weight.

The Gauss cardinal polynomial is $[1+2\sum_{l=1}^n \cos(l\theta_\nu) T_l(x)]/(n+1)$, of norm below two. Markov's inequality and rescaling give (59). Analytic coefficients satisfy $|c_l| \le 2^s M 2^{-|l|_1}$. To bound the tail, use $\|T_l^{(u)}\|_\infty \le l^{2u}$ and $(1+l)^u \le u! \binom{l+u}{u}$: summing gives $\sum_{l\ge 0} 2^{-l}(1+l)^u \le u! 2^{u+1}$ and $\sum_{l>n} 2^{-l}(1+l)^u \le 2^{-(n+1)}(n+2)^u u! 2^{u+1}$. Sum over a coordinate beyond n and apply the cardinal bound to its interpolant. This yields (58) with the displayed $C_{s,k}$; $2k+s \le 9$ makes n logarithmic in Ma^{-k}/η.

The full-past conditional sub-Gaussian law holds at accepted-branch stopping times. The weighted noise proxy is at most $\sigma^2 W^2 / N_\star$, so (60) bounds all errors by $\eta/4$. Conditional probabilities at least $\pi_{\min}$ give the Bernoulli domination in (61), without paired-counterfactual noise. Round each node by at most $a\eta/(32sMW)$ and half its boundary distance. Cauchy's bound $8M/a$ on a disk inside the ellipse makes rounding error at most $\eta/4$; computing weights to the same error keeps the total at most η.

For J direct-arm entries with at most two slots and total order two, write

$$N_2(J,\zeta) = N_0 + C\{J[(n+1)^2 + 1 + \sigma^2 a_t^{-4} \zeta^{-2} (n+1)^{12} \log(4J/\delta)] + \log(2/\delta)\}. \tag{62}$$

Here $a_t \le 1$ is the minimum legal analytic width, $n = O(1 + \log(M_{\mathrm{arm}} a_t^{-2}/\zeta))$, and N_0, δ_0 are independent pilot cost and failure, yielding total failure $\delta_0 + \delta$. Binary rational flows on fixed rate/buffer boxes with bounded context times admit this complex neighborhood. Resets substitute bounded constants, and a score with N coefficients bounded by M has envelope CNM. Decoder slots follow fixed complete calls, avoiding derivatives of expanded compositions.

Without a holomorphic envelope, use legal positive C^5 Taylor stencils. If $D_r = B + rC + O(C_k r^2)$ is order k with absolute weight sum at most $A_k r^{-k}$, then $2D_{r/2} - D_r$ has bias $O(C_k r^2)$. Choosing permitted $r \asymp \sqrt{\zeta / C_k}$ yields sufficient cost

$$\widetilde{O}\big(V_{\mathrm{end}} A_k^2 C_k^k \pi_{\min}^{-1} \zeta^{-(k+2)} \log(1/\delta)\big), \quad k \le 3; \qquad N_{\mathrm{mean}} = \widetilde{O}\big(V_{\mathrm{end}} \pi_{\min}^{-1} \zeta^{-2}\big). \tag{63}$$

All legal-radius and duration-division factors, and full replay, remain charged.

D.2 Fitting and simultaneous error control

Let the degree-r candidates have N coefficients bounded by M, and let $\|\mathrm{coef}\, p\|_\infty \le L_E \|p\|_E$ norm their differences. Decoded input/output error e gives truth residual $R_e = (1 + rNM)e$. Any bounded-coefficient fit with residual at most R_e therefore satisfies $\|g - f\|_E \le 2R_e$ and $\|\mathrm{coef}(g-f)\|_\infty \le 2L_E(1+rNM)e$. This works for maximum or empirical RMS residual. Minimum cardinality removes only zero groups once separated groups cannot be omitted. For all-grid fitting the matrix perturbation is CNe; absorbing $CNeL_E \le 1/2$ gives $4L_E(1+rNM)e$ without a conditioned square subset. Remove redundant features before whole-group tests.

In an aligned orthonormal factor of dimension $n = q^2 - 1$, a bilinear operator error converts to tensor error with factor n; entrywise error ζ_μ costs $n^{3/2}\zeta_\mu$. Include basis/projector errors in $K_{\mathrm{tab}}\zeta$, a bound on both table and anchor errors. Set $(C_{\mathrm{alg}}, \eta_{\mathrm{use}}) = (a_1, \eta_0)$ for a known anchor and $(4qK_c^6(R_0+2), \eta_{\mathrm{cat}})$ for a catalogue. With $K_{\mathrm{tab}}\zeta \le \eta_{\mathrm{use}}$ and $\varepsilon_{\mathrm{num}} \le K_{\mathrm{tab}}\zeta$, Corollary A.6's

tolerance split gives $a = C_{\mathrm{alg}} K_{\mathrm{tab}} \zeta + \varepsilon_{\mathrm{num}}$. Equations (37), (36) and the frame/law bounds give amplification K_{asm}, including $(C_{\mathrm{alg}} + 1)K_{\mathrm{tab}}$ and the absolute norm/inverse factors. The inverse, pair, visibility, native-frame and buffer tests remain mandatory.

For structural margin $\Gamma_\star$ and deployment gain K_{op} on the fixed prefix-closed buffered class, choose

$$\zeta \le c \min\{\eta_{\mathrm{use}}/K_{\mathrm{tab}}, \Gamma_\star/K_{\mathrm{asm}}, \varepsilon/(K_{\mathrm{op}} K_{\mathrm{asm}})\}. \tag{64}$$

Branches using their binary decoder omit the calibration threshold. Allocate summable failure over searches, epochs and fresh confirmations. On that joint event isolation resolves structure and deterministic propagation gives (3) at most ε; calibration adds no independent failure event. Count all distinct measured entries and separate pilot/confirmation samples.

For a confirmed analytic design independent of accuracy, put $\eta = \zeta/A_{\mathrm{asm}}$ and $a_t = \min\{1, \min_j a_j\}$. Section D.1 gives

$$N_0 + O\big(\pi_{\min}^{-1}\big\{J[(n+1)^3 + 1 + \sigma^2 a_t^{-6} \eta^{-2} (n+1)^{18} \log(4J/\delta)] + \log(2/\delta)\big\}\big). \tag{65}$$

Thus refinement is $N_0 + \widetilde{O}(\varepsilon^{-2})$ with failure $\delta_0 + \delta$ at fixed conditioning. The calibration runtime excludes dictionary search and processing the ε^{-2} samples.

D.3 Frozen execution and true-system error

Section 6's induction proves exact maintenance of the frozen model on its versioned DAG, which here also records source/output versions, call/argument ports, numerical rules and score reductions. Constant resets write without parents; cancellation can make the safe cone nonminimal or dense.

For nodes V, edges E, node cost $c(v)$, edited ports P_j, visited nodes V_j, and traversed edges E_j, charge

$$W_0 = O\Big(|V| + |E| + \sum_{v \in V} c(v)\Big), \quad W = W_0 + O\Big(\sum_j \Big[|P_j| + |V_j| + |E_j| + \sum_{v \in V_j} c(v)\Big]\Big). \tag{66}$$

Charge implementation overhead, solves, indexing, copies, and score reductions, with a logarithm for heap scheduling. Streaming pays traversed incoming edges and affected score features. Store maps, reverse lists, versions, and caches. Rebuild after name changes, schedule insertion/deletion, or model changes. Retained state guards require uniform margins.

Section 6's truth-error recurrence $r^+ \le L_a r + d_a$, with current cache error evaluated once and score error equal to fitted defect plus Lipschitz-amplified state error, yields Appendices B–C's recurrences; motion includes influence, denominator, and chart bounds in K_{op}. The fixed-class supremum covers adaptive admitted arguments.

D.4 A restricted-language excitation lower bound

Theorem D.1. *For $m = 2^r$, $r \ge 3$, two replayable worlds have m projective two-dimensional factors, three bounded nonzero raw components generating $\mathfrak{sl}_3^{\oplus m}$, a fixed-gap real-simple anchor, and pairwise Frobenius max-distance at least $m^{-2}/30000$ from every automorphism lock. Every positive-time word of total raw duration at most 10^{-6} and at most two program calls is legal in both, with mean gap at most $4000\, 2^{-m}$. Yet r repeated programs have scores $\pm 1/2$. With independent $N(0, \sigma^2)$ noise,*

$$N \le \sigma^2 4^m / 64000000 \quad \Longrightarrow \quad \inf_{\widehat{F}} \max_{\theta \in \{+,-\}} \Pr_\theta\{|\widehat{F}(P^r) - F_\theta(P^r)| > 1/4\} \ge 3/8. \tag{67}$$

The learner may be adaptive and randomized and know supports, dimensions, templates, and the unordered candidate pair, but not the chosen world/reference.

Proof. Let $z_i^\pm$ be the roots of $T_m(z) = \pm 1/2$ and $\mu_\pm = m^{-1} \sum_i \delta_{z_i^\pm}$. The monic polynomials $2^{1-m}(T_m \mp 1/2)$ differ only in their constants. Newton's identities give equal moments below degree m and degree-m gap 2^{1-m}. All roots lie in $(-1, 1)$ and their combined separation is at least m^{-2}, using $\|T_m'\|_\infty \le m^2$ between levels and intervening extrema.

Let $\bar{p} = \mathbf{1}/3$, $u = (1,-1,0)^\top/\sqrt{2}$, $v = (1,1,-2)^\top/\sqrt{6}$, $a = 1/100$, and $p_i(0) = \bar{p} + az_iu$. For normalized p, write $(x,y) = a^{-1}(u^\top p, v^\top p)$ and set $R = vu^\top - uv^\top$, $A_i = (3/2 + z_i/2)R$, $B = \operatorname{diag}(-4/5,-1/5,1)$ and $C = E_{12} + E_{23} + E_{31}$. Raw maps are normalized exponentials. Their Frobenius norms are respectively in $[\sqrt{2}, 2\sqrt{2}]$, $\sqrt{42}/5$, and $\sqrt{3}$; the anchor gap is $3/5$. Take $P(x,y) = (2x^2 - 1, 0)$ and $h = m^{-1}\sum_i x_i$. In homogeneous coordinates, for $b = \mathbf{1}^\top p$, $P([p]) = [b^2(\bar{p} - au) + (2/a)(u^\top p)^2 u]$, whose coordinate sum is b^2. This is a local noninvertible quadratic, not a positive-coefficient family member.

For $\lambda = (-3/5, -6/5, 9/5)$, the operators $U_j = \prod_{k\neq j}(\lambda_j - \lambda_k)^{-1}(\operatorname{ad} B - \lambda_k I)\, C$ give E_{12}, E_{23}, E_{31}; brackets give opposite roots and diagonals. An automorphism fixing B, C is identity, so distinct A_i cannot lock; (15) gives generation. If B, C errors are at most $\eta \le 1/60$, errors in $C, [B,C], [B,[B,C]]$ are at most $\eta, 10\eta, 76\eta$. The descending coefficient triples of U_j are $(-25/36, 5/12, 3/2)$, $(5/9, -2/3, -3/5)$, and $(5/36, 1/4, 1/10)$. Root, opposite-root, and diagonal errors are consequently at most $60\eta, 360\eta, 1560\eta$. The six off-diagonal units are orthonormal and the diagonal Gram has eigenvalues $1, 3$, giving $\|\varphi - I\|_{F\to F} < 5000\eta$. Also $\|\varphi A_i - A_j\|_F \le \eta$ implies $|z_i - z_j|/\sqrt{2} \le 15001\eta$, proving the lock gap; $\eta > 1/60$ is immediate.

For $|z| \le 2$, $\|R\|_1 = 2/\sqrt{3}$, $\|A(z)\|_1 < 3$, and $\|B\|_1 = \|C\|_1 = 1$. Any raw block of total duration $T \le 10^{-6}$ obeys $\|G - I\|_1 \le e^{3T} - 1 \le 6T$. On normalized input of ℓ_1 norm at most 13, its denominator stays within $78T$ of one and normalization changes p by at most $6T\cdot 13\cdot 14/(1-78T) \le 2184T$; both x, y change by less than one. Through two programs the successive x bounds are $2 \to 3 \to 19 \to 20 \to 801 \to 802$, with $|y| \le 1$. Since $\|p\|_1 \le 1 + a(\sqrt{2}|x| + 4|y|/\sqrt{6}) < 13$, induction closes. Every one-factor response $f_w(z)$ is thus holomorphic near $|z| \le 2$ and bounded by 1000, uniformly over the language, as adaptive queries require. For real inputs the sharper change $24T$ in p gives x bounds $1.0024, 1.00961152, 1.01201152, 1.04834, 1.05074$ and $|y| \le .0024$, keeping simplex coordinates above .32 and proving legality.

Now $F_\pm(w) = \int f_w \, d\mu_\pm$. Moment matching and Cauchy's coefficient bound give $\sup_{w\in\mathcal{A}_{\text{train}}} |F_+(w) - F_-(w)| \le 2000 \sum_{k\ge m} 2^{-k} = 4000\, 2^{-m}$ over the training language $\mathcal{A}_{\text{train}}$. But $T_2 \circ T_k = T_{2k}$ gives $F_\pm(P^r) = \pm 1/2$ with positive deployment states. Adaptive Gaussian transcript KL is at most $N(4000\, 2^{-m})^2/(2\sigma^2) \le 1/8$. Pinsker gives testing error at least $3/8$ (Tsybakov, 2009), and $1/4$-accurate prediction identifies the sign. Include learner randomness in the transcript and pad stopped runs with ignored arms. □

The duration and program-count caps define the training language, not positive-time access generally; querying P^r removes the obstruction. Known-model evaluation costs $O(m \log m)$, and deployment perturbations on the tube amplify by at most m^2. Three names give at most six depth-two directions, versus dimension $8m$, so this is not an instance of Appendix E's uniform Hall design.

E Polynomial discovery in a sparse-output family

Fix $q \ge 2$, $D = q^2 - 1$, and $\bar{p} = \mathbf{1}/q$. The unknown simple linear hypergraph has edge sizes $2 \le |e| \le r$, pair intersections at most one, and degrees $d_i \in [2q - 1, \Delta_{\text{tool}}]$. Conditional on its skeleton and an anchor port, incidence parameters are independent on relative full-dimensional compact log-concave cells, with $a_A \le \|A_{e,i}\|_F \le b_A$ and $\max_{i,j\in S_e} \|A_{e,i}\|_F / \|A_{e,j}\|_F \le C_A$. Anchor cells are regular real semisimple with public spectral/frame gaps and compatibility (35). Other cells are strictly Metzler and contain perturbations of E_{a0}, E_{0a}; their modes are hidden.

Let Q be a public orthonormal basis of $\mathbf{1}^\perp$ and write $x_j = \bar{p} + Qz_j$. Nonidentity program and payload components have the affine-residual form

$$T(x_0, \dots, x_s) = \bar{p} + Q\left(b + B_0 z_0 + \sum_{j=1}^{s} B_j z_j\right), \qquad s \le \Delta_P. \tag{68}$$

Active parameter groups are independent; others are zero. Require $\sigma_{\min}(B_0) \ge \kappa_0$, $\kappa_P \le \|B_j\|_F \le K_P$ for active $j > 0$, and, for constants $\rho_P, \rho_{\text{fwd}}, \rho_{\text{back}} < 1$,

$$\max_i \sum_{j\in \text{pa}(i)} \|B_{i0}^{-1} B_{ij}\| \le \rho_P, \quad \max_i \sum_j \|B_{ij}\| \le \rho_{\text{fwd}}, \quad \max_j \sum_i \|B_{ij}\| \le \rho_{\text{back}}. \tag{69}$$

Nonidentity payload groups are separated from identity by κ_{pay} and have derivative bound $K_{\mathrm{pay}}^{\mathrm{pub}} = O_{q,\mathrm{loc}}(1)$. The task is $h_s(y) = c_h + \sum_{s\in S_{\mathrm{out}}} u_s^\top y_s$, with $u_s \in \mathbf{1}^\perp$ and $u_0 \le \|u_s\| \le u_1$. Set $A_\star = \mathrm{Anc}_{G_P}(S_{\mathrm{out}})$. Payloads writing there read only that subsystem; its legality/ledger updates depend only on it and the public ledger. With $D_0 = \mathrm{diag}(B_{i0})$, block strictly lower L_0, and $g_s = Q^\top u_s$ on sinks (zero elsewhere), $f_{\mathrm{id}} = h_s \circ P$ has gradient $c = D_0^\top (I - L_0)^{-\top} g$. We prove $\chi_\star = \min_{i\in A_\star} \|c_i\| > 0$ below.

Nondegenerate parameter law. Conditional on the discrete supports, the program, payload and nonzero score groups are independent of each other and of the raw parameters. Each has a full-dimensional rectangular support in all entries of its active matrix/vector blocks (scores use Q coordinates), of bounded center and side half-widths in $[s_0, s_1]$, and density between c_{den} and C_{den} times the uniform density there. The positive constants $s_0, s_1, c_{\mathrm{den}}, C_{\mathrm{den}}$ are fixed local parameters. The product support lies inside the norm, invertibility and contraction bounds above; draws are not subsequently conditioned on those bounds. Small boxes around strict feasible points give nonempty such laws. Raw incidence cells retain their stated regularity and witness conditions. Independence alone, singularly supported program laws, or cells shrinking arbitrarily with m do not give the quantitative conclusion.

Assume public semialgebraic chambers give centers and widths $\tau_{\mathrm{leg}}, \eta_{\mathrm{leg}} \ge c_{\mathrm{leg}} m^{-C_{\mathrm{leg}}}$ for Hall/divided, pre/post, program/payload/score, and P–raw–P–test arms and all prefixes, on one analytic branch with finite public ledger. Generation does not imply these permissions. Deployment permits raw length $L_{\mathrm{act}} \log(m+1)$ and $L_{\mathrm{pay}} \log(m+1)$ enrolled calls, with buffer $\rho_{\mathrm{buf}} \ge c_{\mathrm{buf}} m^{-C_{\mathrm{buf}}}$. Population statements allow any finite causal depth; rates fix $H_\star$.

Theorem E.1 (Task-relative discovery and precision)**.** *Almost surely the observable action is $\bigoplus_{i\in A_\star} \mathfrak{sl}_q$, with supports $S_e \cap A_\star$. Legal terminal data recover native coordinates, effective program parents/laws, payloads, and the original score without the skeleton, sinks, ports, or coefficients. Population execution is exact on finite admitted words. At fixed $q, r, \Delta_{\mathrm{tool}}, \Delta_P, H_\star$ and model cells, operational error is at most ε, with failure $\delta_{\mathrm{gen}} + \delta_{\mathrm{data}}$, using*

$$\begin{aligned} N_{\mathrm{nat},C^5} &\le \widetilde{O}_{q,\mathrm{loc},\mathrm{H}_\star}\left(V_{\mathrm{end}}(m/\delta_{\mathrm{gen}})^C \varepsilon^{-5} \log(2/\delta_{\mathrm{data}})\right), \\ N_{\mathrm{nat,an}} &\le N_{0,\mathrm{nat}} + \widetilde{O}_{q,\mathrm{loc},\mathrm{H}_\star}\left(V_{\mathrm{end}}(m/\delta_{\mathrm{gen}})^C \varepsilon^{-2} \log(2/\delta_{\mathrm{data}})\right). \end{aligned} \tag{70}$$

Discovery cost $N_{0,\mathrm{nat}}$ is polynomial and independent of ε; exponents may differ between bounds. All attempted endpoints and physical executions count.

Writing $E^\star = \{e : S_e \cap A_\star \ne \varnothing\}$ and $M^\star = |E^\star| \le |A_\star| \Delta_{\mathrm{tool}}$, a separate C^5 allocation first detects names and then refines only the active subsystem:

$$\begin{aligned} N_{\mathrm{det}} &\le \widetilde{O}(V_{\mathrm{end}} |E| (m|E|/\delta)^{C_{\mathrm{det}}} \log(|E|/\delta)), \\ N_{\mathrm{prec}} &\le \widetilde{O}(V_{\mathrm{end}} (M^\star |A_\star| \Delta_P/\delta)^C \rho_{\mathrm{buf}}^{-C} \varepsilon^{-5} \log(2/\delta)). \end{aligned} \tag{71}$$

E.1 Visibility and Hall excitation

For x uniform on $[-r_0, r_0]^{q-1}$, $0 < r_0 \le 1$, put $s = (e_0 + x)/\sqrt{1 + \|x\|^2}$ and $N_s = \{s\varphi^\top : \varphi^\top s = 0\}$. Then, on $\mathfrak{sl}_q$,

$$\mathbb{E}\,\mathrm{Proj}_{N_s} \succeq \frac{4r_0^4}{55(1 + (q-1)r_0^2)^2} I. \tag{72}$$

Indeed, write $A^\top = \left(\begin{smallmatrix} a & b^\top \\ c & B \end{smallmatrix}\right)$, $C = B - aI$, and $F_A(x) = c + Cx - x(b^\top x)$. Cube moments give $\mathbb{E}\|F_A\|^2 \ge \|c - r_0^2 b/3\|^2 + (r_0^2/3)\|C\|_F^2 + (4r_0^4/45)\|b\|^2$. Tracelessness implies $\|A\|_F^2 \le \|b\|^2 + \|c\|^2 + \|C\|_F^2$, while $\|c\|^2 \le 2\|c - r_0^2 b/3\|^2 + 2r_0^4\|b\|^2/9$. The wedge with $e_0 + x$ dominates F_A; normalization costs $(1 + (q-1)r_0^2)^2$, proving (72).

For the second bound, let u be uniform on $[-1, 1]^d$ and $x(u) = x_0 + \tau(Ju + R(u))$, where J is square with conorm κ, $0 < \tau\kappa \le 1$, and $\sup\|R\| \le \eta$ with $2^{6k+12} d^{k/2} \eta \le \kappa$. Every Hilbert-valued polynomial of degree at most k satisfies

$$\mathbb{E}\|p(x(u))\|^2 \ge \frac{c_k d^{-C_k} \tau^{2k} \kappa^{2k}}{(1 + \|x_0\|)^{2k}} \|\mathrm{coef}\, p\|^2. \tag{73}$$

For $p_0(u) = p(x_0 + \tau J u)$, fixed-degree tensor Legendre bounds give $\sup_{[-2,2]^d}\|Dp_0\| \le 2^{6k}d^{k/2}\|p_0\|_{L^2}$. Thus $E = J^{-1}R$ obeys $\|p_0(u + E(u)) - p_0(u)\|_{L^2} \le 2^{-12}\|p_0\|_{L^2}$. Translation costs $C_k(1 + \|x_0\|)^k$, inverse substitution costs $(\tau\kappa)^{-k}$, and coefficient/Legendre conversion is polynomial in d at fixed k. This proves (73); unused duration coordinates may be conditioned on when the bounds hold uniformly.

Parameter visibility and generation. The complete affine program has derivative norm at most $K_0/(1 - \rho_{\mathrm{fwd}})$, $K_0 = \max_i\|B_{i0}\|$, and inverse norm at most $K_{\mathrm{inv}}^{\mathrm{pub}} = \kappa_0^{-1} + \rho_P$. Its task gradient is

$$c_i = B_{i0}^\top \sum_{\pi: i \rightsquigarrow s} B_{i_1 i}^\top B_{i_2 i_1}^\top \cdots B_{s i_{k-1}}^\top g_s. \tag{74}$$

For an ancestor i, choose a longest path from i to an output, of length $\ell \le H_\star$. In one coordinate of c_i, a choice of matrix entries along that path, one entry of B_{i0} and one of g_s gives a degree-$D_i = \ell+2$ monomial with coefficient one. No other path uses that same monomial, and no longer path can extend it. Conditional on groups off this path, that coefficient is unchanged. The path groups have at most $O_{q,\mathrm{loc}}(H_\star + 2)$ free entries. Translation and rescaling to their unit boxes preserve a highest-degree coefficient of magnitude at least $s_0^{D_i}$. Finite-dimensional coefficient/L^1 norm equivalence on these fixed-degree boxes therefore gives a uniform positive L^1 lower bound. The density comparisons transfer both this bound and the polynomial sublevel estimate (Carbery & Wright, 2001, Theorem 2) from uniform measure. Consequently $\Pr\{\|c_i\| \le t \mid \text{off-path groups}\} \le C_{q,\mathrm{loc},\mathrm{H}_\star} t^{1/(H_\star+2)}$ for $0 < t \le 1$. Thus $c_i \ne 0$ almost surely, and a union bound gives $\chi_\star \ge c_{q,\mathrm{loc},\mathrm{H}_\star}(\delta/m)^{H_\star+2}$ with probability at least $1 - \delta$. The gradient is state-independent. Score, payload, and legality closure make every record outside $A_\star$ invisible to all admitted futures.

For each record, $(z_{ef}) \mapsto \sum_{e<f} z_{ef}[A_{e,i}, A_{f,i}]$ has a nonzero D-minor: with regular diagonal anchor H, $X_a = E_{a0}$ and $Y_a = E_{0a}$, the brackets $[H, X_a], [H, Y_a], [X_a, Y_b] = E_{ab} - \delta_{ab}E_{00}$ span $\mathfrak{sl}_q$. The witness persists under admissible cell perturbations. Linear overlap makes these columns singleton-supported. Fixed-degree minor anti-concentration and the catalogue union bound yield conorm $\kappa_{\mathrm{nat}} \ge c_0(\delta/m)^{C_0}$ at fixed local parameters. Depth-two words span the full latent direct sum, observationally restricted to $A_\star$.

Positive-duration designs. Let $M = |E| = O(m)$, $d = mD$, and $L_W = 3M(M-1)$. Concatenate all ordered-distinct e, f, e blocks into W, with every duration centered at $\tau > 0$. Let K_{ch} control group/state charts, B_{end} the first four duration derivatives, and B_{st} the first five derivatives of Hall/divided means. On the buffer, $B_{\mathrm{end}} \le M^{C_q}(1 + b_A)^{C_q}K_{\mathrm{ch}}^{C_q}$ and $B_{\mathrm{st}} \le m^{C_{q,\mathrm{loc}}}K_{\mathrm{st}}^{C_{q,\mathrm{loc}}}$, where $K_{\mathrm{st}} = 1 + b_A + K_{\mathrm{ch}} + K_{\mathrm{inv}}^{\mathrm{pub}} + (1 - \rho_{\mathrm{fwd}})^{-1} + (1 - \rho_{\mathrm{back}})^{-1}$. The first/last right-trivialized column difference of an e, f, e block is $\tau[A_f, A_e] + O_q(b_A^3\tau^2)$, up to common invertible suffix transport. Aggregate error is at most $C_qM^3(1 + b_A)^3\tau^2$. Choose a dyadic $\tau \in (\bar\tau/2, \bar\tau]$, where $\bar\tau = c_q \min\{1, \tau_{\mathrm{leg}}, \kappa_{\mathrm{nat}}/[M^3(1 + b_A)^3]\}$ and $\gamma_{\mathrm{rect}} \ge c_q\tau\kappa_{\mathrm{nat}}$, and set $r_{\mathrm{nat}} = c_q\gamma_{\mathrm{sub}}\rho_{\mathrm{sub}}$ with $\gamma_{\mathrm{sub}} = \min\{1, \gamma_{\mathrm{rect}}/\sqrt{1 + 4d(L_W - d)}\}$ and $\rho_{\mathrm{sub}} = c_q \min\{1, \tau, \eta_{\mathrm{leg}}, \rho_{\mathrm{marg}}, \gamma_{\mathrm{sub}}^4/[240d^4B_{\mathrm{end}}(1 + B_{\mathrm{end}})^3]\}$. Weyl, rank-revealing QR on actual duration columns, and quantitative inversion give a selected d-column cube covering a product-state ball of radius r_{nat}. Selection is proof-only; the learner samples the full public catalogue. Choose divided duration below $\min\{\tau, \eta_{\mathrm{leg}}, (8b_A)^{-1}, \rho_{\mathrm{sub}}, c\kappa_{\mathrm{nat}}/[M^C(1 + B_{\mathrm{st}})^C]\}$ and charge its reciprocal. For each fixed-size state or union-parent marginal, QR gives square Jacobian conorm κ_I and derivative bound B_I. Set $\rho_{\mathrm{marg}} = \min_I 2^{-6k_I-11}\kappa_I/(B_I d_I^{1+k_I/2})$ with $k_I \le 4$; the derived conorms control (73)'s remainder uniformly in unused coordinates.

At a fixed state put $\beta_H(t) = d(f_{\mathrm{id}} \circ W(t))$ and let B_β bound its third duration derivative. Given nonzero cotangent β and displacement v, set $C = \beta v^\top/\|\beta\|^2$. The traceless isotropy matrix with lower block $C - \mathrm{tr}(C)I/q$, upper-left entry $-\mathrm{tr}(C)/q$, and zero off-diagonal blocks fixes the state and moves the covector by $C^\top\beta = v$. Its norm is at most $C_q\|v\|/\|\beta\|$. Applied factorwise, fixing the endpoint cancels Hessian and cross-factor terms. The group-map right inverse and (74) therefore give $\sigma_{|A_\star|(q-1)}(D_t\beta_H(t^\circ)) \ge a_{\mathrm{joint}} := c_qK_{\mathrm{ch}}^{-C_q}\chi_\star\gamma_{\mathrm{sub}}$. Choose dyadic $\eta < \min\{\eta_{\mathrm{leg}}, \tau/2, \rho_{\mathrm{sub}}, (a_{\mathrm{joint}}/(8B_\beta\sqrt{L_W}))^{1/2}\}$. The genuine bank consists of $W(t^\circ)T_{\mathrm{id}}$ and $W(t^\circ \pm \eta e_j)T_{\mathrm{id}}$, all at positive durations. Central-difference remainders are below half the conorm. Averaging yields $\lambda_{\mathrm{cyc}} = \eta^2 a_{\mathrm{joint}}^2/[4K_{\mathrm{ch}}^2(2L_W + 1)]$ and $\lambda_{\mathrm{state}} = c_qm^{-C_q}\kappa_{\mathrm{nat}}^2 r_{\mathrm{nat}}^4 K_{\mathrm{ch}}^{-C_q}$. The

latter follows from (73) and the coefficient-injective quadratic projective field, equivalently (72). Profile covariance is at least $\lambda_{\mathrm{prof}} I$, with $\lambda_{\mathrm{prof}} = \lambda_{\mathrm{cyc}}\lambda_{\mathrm{state}}$. Conditional response covariance first bounds $\lambda_{\mathrm{cyc}} \sum_i \|X_{A_i}(x_i)\|^2$; averaging histories then gives the claim without independent states.

E.2 From profiles to geometry and laws

If $\mathbb{E} Z_j^\top Z_j \succeq \lambda_{\mathrm{prof}} I$ and $\|Z_j\| \le R_{\mathrm{prof}}$, matrix Chernoff (Tropp, 2012) gives normalized stack gap $\sqrt{\lambda_{\mathrm{prof}}/2}$ after $8R_{\mathrm{prof}}^2 \lambda_{\mathrm{prof}}^{-1} \log(d/\delta)$ samples. If a proof-only Hall basis has conorm κ_{Hall} among L columns, strong rank-revealing QR with $f = 2$ (Gu & Eisenstat, 1996) gives $\gamma_B = \kappa_{\mathrm{Hall}} \sqrt{\lambda_{\mathrm{prof}}/2} / \sqrt{1 + 4d(L-d)}$. Writing the selected stack $B = ZS$, we have $\|S\| \le \sqrt{d} M_{\mathrm{Hall}}$ and $\|S^{-1}\| \le R_{\mathrm{prof}}/\gamma_B$. Record $K_{\mathrm{act}} \ge \max\{1, \|S\|, \|S^{-1}\|, \operatorname{cond} S\}$. The centroid equations have nonzero singular gap at least $(16 d^{3/2} \sqrt{q} K_{\mathrm{act}}^6)^{-1}$. In canonical compact-form coordinates their squared operator is a sum of adjoint commutator squares, with kernel consisting of a scalar on each simple ideal. Every nonzero root-lattice Casimir eigenvalue is at least q^{-1}; the bracket norm is at most $2d^{3/2}$, and operator/quotient-angle transport loses at most $8K_{\mathrm{act}}^6$ (Humphreys, 1972; Isaev & Krivonos, 2021).

Canonical centroid idempotents are orthonormal for $(q^2-1)^{-1} \operatorname{tr}(TU)$. A Gaussian combination has all eigenvalue gaps at least $1/(8m^2)$ with probability greater than $15/16$, by a union bound on Gaussian differences. Repeat $O(\log(1/\delta))$ times and test separated intervals. Transport costs at most K_{act}^2; the raw support gap is at least $a_A/(2K_{\mathrm{act}})$.

For an orthonormal factor inclusion U_i, SU_i and its inverse onto the canonical block are bounded by K_{act}. The centroid gaps control perturbed inclusions. Each raw-name catalogue contains an unlabeled native anchor with the prescribed spectral/frame gaps and coordinate norm at most Kb_A. Theorem 4.2 and Corollary A.6, after Section D.2's tensor conversion, therefore give polynomial-time calibration at fixed q and cells: every new error factor, inverse tolerance and catalogue size is polynomial in the derived bounds. No new Hall word or derivative is needed. The same response stack gives the orientation pair by Section A.6, and visible pre/post responses; the native reference and anchor then give (35). This proves polynomial geometric bounds, not polynomial runtime for arbitrary dictionary fitting.

For $w = b + D_0 z + L_0 w$, let $T = (I - L_0)^{-1} D_0$ and $H = T^{-1}$. Complete-call pairs give $B_{i0} = H_{ii}^{-1}$, $B_{ij} = -H_{ii}^{-1} H_{ij}$, and $b_i = H_{ii}^{-1}(Hw(0))_i$. The fixed-degree marginal pullback yields union-parent feature Gram $\gamma_{\mathrm{aff}} I$ with inverse-polynomial γ_{aff}. Missing active B_j costs squared residual $\ge \gamma_{\mathrm{aff}} \|B_j\|_F^2$; minimum support removes zero extra groups. In $W = FA^\top$, design gap γ and error $e_F \le \gamma/2$ give coefficient error $2(e_W + \|A\| e_F)/\gamma$. Inverse perturbation transfers it to the recovered coefficients. Program/payload groups and scores retain their specified gaps and anti-concentration bounds.

E.3 Complex refinement and discovery

Use $\|p\|_{\infty,1} = \max_i \|p_i\|_1$. Consider an arm with $s \le 3$ variable slots and $b \le 4$ raw runs separated by affine maps of norm at most G, with real positive normalized intermediate states, total raw generator norm at most L, and maximum generator norm B. Set

$$Q_L = 4e^{2L+1}, \qquad H_c = \max\{2, G, Q_L\}, \qquad r_c = [8bs(1+B)H_c^{2b+1}]^{-1}. \tag{75}$$

For a real raw product, $\|M^{\pm 1}\|_1 \le e^L$ and positivity gives $\mathbf{1}^\top M p \ge e^{-L}$. A complex duration change $\delta = \sum B|\Delta t_j| \le sBr_c$ perturbs M by at most $\delta e^{L+\delta}$. Thus incoming error e becomes $e^{L+\delta}(e+\delta)$ before normalization and at most $Q_L(e+\delta)$ afterwards; affine steps multiply by G. Iteration yields $e \le bH_c^{2b}\delta$ and $(bH_c^{2b}+1)\delta \le (4H_c)^{-1}$. These bounds close the denominator induction, also for partial runs, and keep complex states below norm two. The Bernstein ellipse on $[0,a]$ is within $a/2$ of that interval, so $a \le r_c$ suffices.

Hall means $H\Sigma C_j$ and $HB_\star \Sigma C_j$ have raw H, at most three varied slots, one raw/enrolled $B_\star$ and two programs. Therefore $b \le 4$, $L = O_q(M^2 b_A \tau) = O_{q,\mathrm{loc}}(1)$, $B \le \sqrt{q} b_A$, and $G = \max\{\sqrt{q} K_0/(1-\rho_{\mathrm{fwd}}), G_{\mathrm{payload}}\}$. The score envelope is $|c_h| + 2mu_1 \le B_Y + 3mu_1$. Taking a below every legal width and (75) bounds a^{-1} and all envelopes by $(m/\delta_{\mathrm{gen}})^{C_{q,\mathrm{loc},\mathrm{H}_\star}}$.

Let $a_\natural$ be the minimum of 1, κ_{nat}, $\chi_\star$, γ_{anc}, γ_{ch}, $\pi_{\min}$, ρ_{buf}, K_{ch}^{-1}, τ_{leg}, η_{leg}, κ_0, κ_P, κ_{pay}, u_0, $1-\rho_P$, $1-\rho_{\mathrm{fwd}}$ and $1-\rho_{\mathrm{back}}$. All gaps/radii are at least $cm^{-C}a_\natural^C$, with upper envelopes bounded by its reciprocal. The public history/continuation catalogue has size at most $m^{C_q}[1+\log(m/a_\natural)]^4$. Parameter anti-concentration and a union bound give $a_\natural \geq c(\delta_{\mathrm{gen}}/m)^C$ with probability $1-\delta_{\mathrm{gen}}$. On this event, summable-confidence dyadic refinement resolves ranks and groups without labeled minors; polynomial coarse cost includes failed arms and is independent of ε. Fresh refinement uses (63) or (65). Spatial gain is controlled by $K_{\mathrm{inv}}^{\mathrm{pub}}$ and contractions; temporal gain is at most $\sum_{j=0}^{N_{\mathrm{pay}}}(K_{\mathrm{pay}}^{\mathrm{pub}})^j \leq m^{C_{q,\mathrm{loc}}}$. Buffer/frame bounds complete (70).

Discovery versus active-subsystem precision. For (71), set $g_e(t) = f_{\mathrm{id}}(T_e^t s^\star) - f_{\mathrm{id}}(s^\star)$. Closure makes it identically zero for inactive names. For an active name, $\beta_e = g_e'(0) = \sum_{j\in S_e\cap A_\star}\langle c_j, X_{A_{e,j}}(s_j^\star)\rangle$ has a unique path monomial of degree at most $H_\star+3$, witnessed by identity path blocks, a sink covector and an incident root direction. Anti-concentration with failure $\delta/4$ gives $\min_{e\in E^\star}|\beta_e| \geq \chi_{\det} = c_{\det}(\delta/(4|E|))^{C_{\det}}$. A public second-derivative bound is $B_{\det} = [m(1+b_A)K_{\mathrm{inv}}^{\mathrm{pub}}(1-\rho_{\mathrm{fwd}})^{-1}]^{C_{q,\mathrm{loc},H_\star}}$. Refine dyadic $t_j = 2^{-j}t_{\mathrm{div}}$ until $t_j B_{\det} \leq \chi_{\det}/2$; the first such step is at least half the smaller of t_{div} and $\chi_{\det}/(2B_{\det})$. Only permitted P–raw–P–test arms and prefixes are issued. Simultaneous radius $t_j\chi_{\det}/8$ distinguishes zero from active $|g_e(t_j)| \geq t_j\chi_{\det}/2$ at threshold $t_j\chi_{\det}/4$. Each name costs $O(V_{\mathrm{end}}(t_j\chi_{\det})^{-2}\log(|E|(j+1)^2/\delta))$, including rejected scales, giving the detection bound.

Restriction to active records preserves every active incidence, the local degree, singleton minors and conditional independence; payload/ledger closure preserves observable semantics. Substitute (mD, M, S_e) by $(|A_\star|D, M^\star, S_e\cap A_\star)$ in the same C^5 proof, retain ρ_{buf} and allocate failure across parameter, detection, design and precision events. This gives the active refinement bound. No inactive state accuracy, new inverse-square active bound, or bound on deleting individual couplings is inferred.

F Experimental protocols, controls and scope

The studies test distinct mechanisms: calibration from noisy tables (Section F.1), positive-time scalar acquisition and frozen-model reuse (Section F.2), and reset-based identification (Section F.4). They do not jointly implement the full discovery and confidence guarantees. Hidden states and structure are evaluator-only except in identified component checks. Queries, primitive calls and wall time are charged separately; query-free deployment follows acquisition and initialization. Independent fitting is the main comparison; earlier coupled-fit controls remain explicitly labelled. Reset, caching and native-family studies retain their own access conditions and failure denominators.

F.1 Direct validation of the constructive calibration

We test Theorem 4.2 with a representation $S : W \to \mathfrak{sl}_q$ satisfying $\|S\| = \|S^{-1}\| = K$, and an anchor with real spectral gap γ and eigenvector condition χ. Perturbations of the table and anchor have Frobenius error η, without Jacobi projection (median cyclic defect .14 at $\eta = 10^{-2}$). The learner receives only $(\widehat{\mu}, \widehat{h}), q, K$; S, its spectrum and eigenvectors are evaluator-only.

For each branch $A_a = Sa$ or $A_a = -(Sa)^\top$, the evaluator obtains unit-Frobenius P minimizing $\sum_a \|\widehat{\rho}(a)P - PA_a\|_F^2$ by a smallest-singular-vector solve. Branches with $|\det P| < 10^{-14}$ are discarded; the reported error is $E_{\mathrm{align}} = \min_{\text{retained branches}}\|\widehat{\rho} - PA.P^{-1}\|_{\mathrm{op}}$. This feasible alignment upper-bounds the infimum in (6), rather than optimizing globally over $\mathrm{Aut}(\mathfrak{sl}_q)$. SVD uses double precision; alignment condition numbers were not retained, and the determinant threshold is not a conditioning certificate.

The 360-run grid uses $q \in \{3,4\}$, $K \in \{2,3\}$, six seeds and fifteen noise levels from 10^{-14} to 1; every run returns. Noiseless errors at $q = 2, 3, 4$ are below $4\,10^{-15}$. The output norm caps $2K^3$ hold in 359/360 noisy runs; the exception at $\eta = 1$ has inverse norm 279 against cap 54. All runs with $\eta \leq 10^{-1}$ satisfy both caps. Observed stability extends well beyond the sufficient radius (Figure 4). At $(q, K, H, \chi, \gamma) = (3, 2, 4, 2, 1)$, (28) gives $\eta_0 = 5.52\,10^{-18}$ and $a_1 = 1.13\,10^{16}$. Error remains proportional to η over 10^{-14}–10^{-2}, with median amplification .18 in this configuration and .12–.53 across the four (q, K) pairs. Thus the tested range extends fifteen orders beyond the sufficient

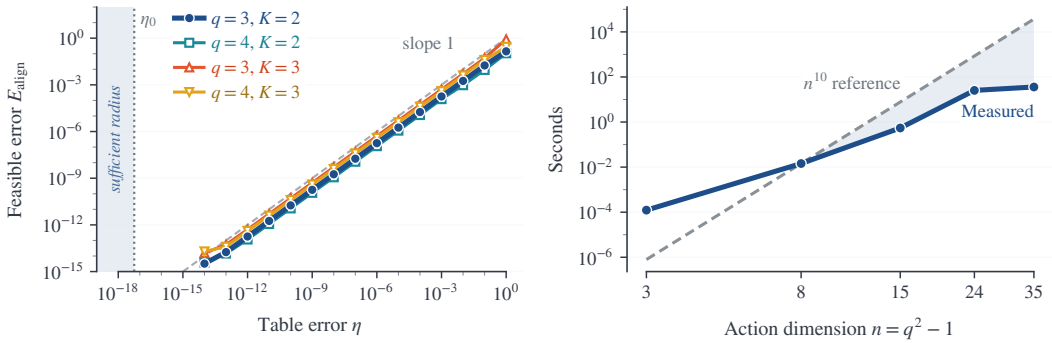


Figure 4: **Calibration on perturbed table inputs.** Left: feasible alignment error versus table error, including cases outside the sufficient radius ($q = 3, K = 2$ for the marked radius). Right: wall time through $q = 6$, with a normalized n^{10} growth reference. The algorithm receives the noisy table, anchor, dimension and norm bound, not the true representation. These panels isolate calibration from scalar acquisition; Section F.2 tests the latter with positive-duration access.

radius, with amplification seventeen orders below its bound. At $\eta = 1$, amplification ranges from .08 to 5.4.

At $\eta = 10^{-8}$, sweeps of $K = 1.5$–6, $\gamma = .125$–2 and $\chi = 1$–16 give empirical log–log exponents $K^{2.31}, K^{2.89}$ and $\chi^{.21}, \chi^{.43}$ at $q = 3, 4$, versus bound exponents $K^{30.9}, K^{31.0}$ and $\chi^{2.22}, \chi^{2.57}$. Error varies by only $5\,10^{-7}$ relatively across the sixteenfold gap range, whereas the bound scales as $\gamma^{-.78}$. Increasing K from four to six changes amplification from .72 to 3.3 at $q = 3$ and from .88 to 3.7 at $q = 4$. These are sampled parameter sensitivities, not improved uniform bounds.

Fitting and computational cost. A least-squares structure-constant fit receives the same tables and public constants, with no near-truth initialization. At $q = 3, \eta = 10^{-8}$ it is more accurate ($7\,10^{-10}$ versus $1.8\,10^{-9}$) but slower (28 versus .04 seconds). At $q = 4$, it fits 225 parameters to 3,375 residuals using six independent Gaussian-SVD orthogonal starts, Levenberg–Marquardt, at most 20,000 residual evaluations and sixty seconds per start, and a 400-second cell cap. Timeouts count as nonconvergence. At $\eta = 10^{-8}$, two of three $K = 2$ instances retain residuals $8.17, 9.29$ despite reported convergence from all starts; one of three $K = 3$ instances retains 12.24. These full-table residuals establish fitting failures. Construction is deterministic and two to three orders faster at these sizes; this comparator is not the theoretical exhaustive search.

The invariance operator is assembled sparsely on $W^* \otimes \mathrm{Sym}^3 W^*$ using its last-three-slot symmetry. Against the full-tensor solve at $q = 3, 4$, errors agree within 27%; the invariant line's arbitrary sign can select opposite branches. Median time ranges from below one millisecond at $q = 2$ to 36 seconds at $q = 6$ ($n = 35$), with peak resident memory 5.4 GiB and error $5.0\,10^{-10}$ at input error 10^{-8}. The timing exponent over $q = 3, \ldots, 6$ is 5.7, excluding sub-resolution $q = 2$. The normalized n^{10} reference depicts a conservative bound, not predicted runtime. Error/input ratios fall from .38 at $q = 2$ to .05 at $q = 6$ in these instances only.

Anchor selection. Following Corollary A.6, we construct multiplication once, extract on each candidate, discard undefined or uncapped outputs, and minimize $\|\mu_{\widehat{\rho}} - \widehat{\mu}\|_F$. Catalogues of $L = 4, 8, 16$ contain one canonical anchor and decoys with repeated minimal eigenvalues, imaginary spectra or gap 10^{-9}. Across $q = 3, 4$, five noise levels and five seeds, all 150 cells return, with median amplification .12. Candidate survival is 94%; canonical-anchor selection is 60%. Other selections are valid when a decoy's minimal eigenvalue is simple: the criterion controls representation error, not anchor identity. This tests the catalogue rule, not Theorem A.8's signed scan. Figure 5 summarizes the parameter, fitting and anchor comparisons.

F.2 Positive-time acquisition, fitting and reuse

The common setup below supports the main independent-restart comparisons: initialization (Section F.2.1) and new-objective transfer (Section F.2.2). Section F.2.3 retains the original coupled-restart controls. Subsequent subsections examine scale, alternative readouts, dimension discovery

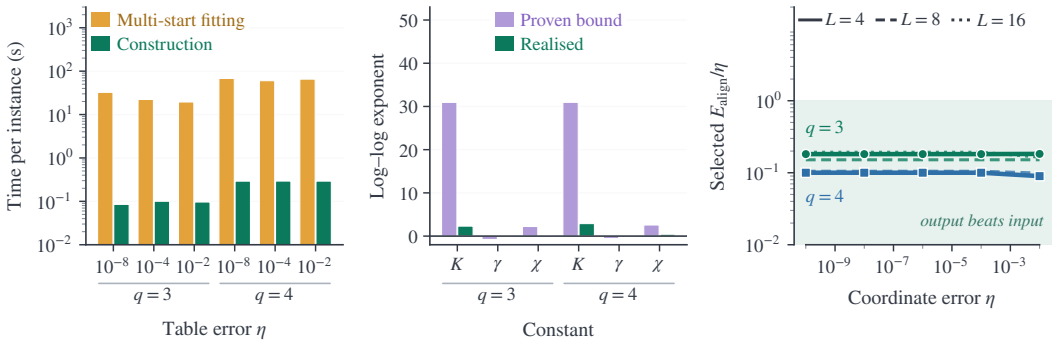


Figure 5: **Calibration cost and quantitative bounds.** Left: median seconds per instance for multi-start fitting and construction on the same tables. Centre: log–log exponents of the amplification in the representation norm, the anchor gap and the eigenvector condition: proven from (28), realised at $\eta = 10^{-8}$. Right: error of the output that Corollary A.6's residual rule selects, relative to the coordinate error it was given, with one canonical anchor hidden among decoys; the shaded band is where the selected output is more accurate than its own input, and every cell lands there.

and composition, followed by bracket and physical-time diagnostics. Each extension states its own access and comparison conditions.

System and access. We test two known two-dimensional projective factors ($q = 3$), 18 anonymous names and an unknown additive native-linear score. Eight names act on each factor and two on both. Per factor, two diagonal generators have spectra $(-.65, 0, .65)$ and $(.35, -.65, .30)$; six others have off-diagonal entries $U[.015, .07]$, one distinct entry increased by .55, and centered diagonal entries $U[-.18, .18]$. Reference states are Dirichlet$(12, 12, 12)$ and score coefficients are $U[-.8, .8]$. Shared tools combine independently permuted factor catalogues with weights $(.4, .6)$ and $(.3, .4)$; a final hidden permutation anonymizes names. Learner priors are the dimensions, spanning-raw design and readout class, not targets, matrices, states, basis or anchor identity. Queries have positive durations, total duration at most 4.5, and simplex coordinates above 10^{-7}; simulator legality checks reveal no state.

Acquisition and estimation. Sixty-four history/continuation pairs of lengths four/two and durations $\{.1, .25, .4\}$ provide responses at nine positive Chebyshev nodes in $(0, .6)$. A degree-$d \le 8$ least-squares fit defines the first-derivative functional w_d at zero; mixed derivatives use $w_d \otimes w_d$ on the 9×9 grid. The interpolation rule ($d = 8$) is polynomial-exact through degree eight, but retains smooth-response bias:

$$\left| \sum_j w_{8,j} f(t_j) - f'(0) \right| \le M_9 \sum_j |w_{8,j}| t_j^9/9!, \qquad |f^{(9)}| \le M_9.$$

Mixed derivatives likewise retain approximation error; this fixed grid does not implement Section D.1's adaptive bias guarantee. Bracket noise amplification is $\|w_8\|^2 = 2.2 \times 10^4$, versus 874 at $d = 4$ and 42 at $d = 2$.

After other acquisition, 256 repeated reads estimate $\hat{\sigma}$ by sample standard deviation. Separately for first responses and brackets, pooled Lepski selection chooses the lowest $d \in \{2, \dots, 8\}$ agreeing with every higher e within three difference-noise RMS values: $\hat{\sigma}\|w_d - w_e\|$ or $\sqrt{2}\hat{\sigma}\|w_d \otimes w_d - w_e \otimes w_e\|$. The rule was fixed on six development worlds. At zero estimated noise it selects interpolation; acquisition order and the shared observation prefix are unchanged in all 240 evaluation cells. Early-stopped interpolation acquisitions require completion reads before the noise probe.

A pivoted response solve selects 16 raw directions and estimates their bracket. The adjoint commutant and real Schur split identify two eight-dimensional factors. Calibration scans unknown raw anchors with cap $K = 10$; independent same-state responses select fundamental orientation. Sixty-four additional histories fit an unconstrained bilinear ratio in the recovered rays by homogeneous least squares. All negative weights and inverse operations are offline. This raw-span study does not test dimension discovery, deeper Hall words, native payload dictionaries or program recovery.

Cohorts, controls and costs. Twenty-four evaluation worlds are independent of development. Endpoints have independent Gaussian noise, $R = 1$, with one noiseless fit and three repeats at each $\sigma = 10^{-10}, 10^{-8}, 10^{-6}$: 240 cells, not independent worlds. Construction uses 1,266,800 endpoints (1,266,544 plus 256 noise probes) and 10,079,056 primitive occurrences. Both arms receive 512 validation endpoints, equally split across lengths $4, 8, 16, 32$, adding 7,680 occurrences; construction does not use validation. These are logical simulator queries. Nonzero-noise estimates satisfy $\hat{\sigma}/\sigma = .905$–$1.099$ (median .994), with common standardized draws across noise levels; noiseless estimates are below 4×10^{-16}. Selected first/bracket degrees are $8/8$ without noise; $6/5, 7/5$ or $7/6$ at 10^{-10}; $5/4$ at 10^{-8}; and 3–4/2–3 at 10^{-6}.

Direct fitting estimates generators, states and additive coefficients with the same known factor dimensions, generator cap ten and score cap .8. Squared off-diagonal parameters admit zero actions in the closed Metzler family. Eight random starts use float64 Adam for 12,000 steps, batches of 1,024, cosine step size .012 to .0003, then at most 500 L-BFGS iterations on a fixed 4,096-row subset. Validation MSE selects checkpoint and start. Two preliminary strictly positive parameterizations were excluded before test inspection; no evaluation worlds were screened. Direct fitting receives the original acquisition prefix and validation: 1,267,056 endpoints in 192 cells, 256 fewer than construction plus validation; 1,255,040 in 48 high-noise early-stop cells, 12,272 fewer (12,016 completion reads plus 256 noise probes). Primitive counts range from 10,033,536 to 10,085,200. Total observations and compute are therefore unequal. Two fitted cohorts share this setting: independent restarts use separate gradient clipping and L-BFGS solves (Sections F.2.1–F.2.2; main Figures 2–3); coupled restarts share clipping and line search and are retained as separately labelled controls.

Median noiseless interpolation-construction time, including acquisition, is 5.85 CPU seconds; direct optimization costs 34.41 RTX-4090 seconds excluding acquisition. Serial comparisons on five cells give $-.14\%$ to $+.42\%$ timing changes for noise-matched versus interpolation construction. These heterogeneous timings do not establish physical-service speedup.

Tasks and metrics. Each model scores 48 unseen menus: eight for each length $8, 16, 32$ and each of two tasks, call permutation or single-duration editing. Each menu has 32 distinct candidates; base durations are $.025, .05, .075$, and edited durations $.01, .03, .06, .09, .12$. Test words are disjoint from training/validation, with no negative or illegal calls or post-freeze learner queries in all 240 cells (minimum prefix coordinate .0328).

Normalized regret is excess true score above the menu minimum divided by its range. Ranges are $1.02\,10^{-4}$–$.0510$ (median .00935), all above 10^{-12}. Averages first combine menus and repeats within worlds, then worlds; Figure 7 uses 2,000 world-bootstrap resamples. Completion means finite predictions, not accuracy-gated release.

Structural metric and inference. For a gauge-invariant structural metric, vectorized generators form $M \in \mathbb{R}^{9\times 18}$ per factor. Its rank-eight row space records linear relations among public names and is unchanged by $M \mapsto GM$ for invertible G, including conjugation. Linear-relation error (*catalogue-relation error* in the plots) is the spectral distance between learned/true row-space projectors, maximized over factors after linear-assignment matching, in $[0, 1]$. Truth is evaluator-only and both arms use the same metric. This does not measure every Lie relation or physical scale: doubling all generators preserves the projector but changes durations and nonzero brackets.

Inference averages repeats within each of 24 worlds, then uses paired two-sided Wilcoxon tests (exact without tied absolute differences; otherwise normal with tie correction) and 10,000-resample world-bootstrap 95% intervals.

F.2.1 Independent fitting and constructive initialization

We test whether construction improves the same fitting procedure. A truth-free gauge search maps recovered generators, rays and rational readout into the fit's constrained parameterization: positive simplex state, nonnegative off-diagonal generators of norm at most ten, and additive coefficients bounded by .8. Negative off-diagonal entries are floored and changes recorded; this is not exact coordinate transport.

Mapping succeeds in $24/24, 72/72, 72/72, 69/72$ cells across increasing noise. Three high-noise failures have minimum state coordinates $-.008$ to $-.062$; one also violates the coefficient cap. They remain failures, not clipped successes. A constrained gauge follow-up finds feasible states but

changes generators by 11–18%; subsequent fitting selects a common random start in all three, adding no benefit or eligible cases. Median floored off-diagonal mass is 1.7×10^{-6}, 7.4×10^{-5}, .0020, .058; median generator shifts are 3.5×10^{-6}, 5.5×10^{-5}, .0015, .044. Clipping without gauge search gives validation RMSE .24–.28. Mapping therefore contributes substantive operations and cost.

Table 1: **Construction as an initializer with independent restarts.** One of eight random starts is replaced by the mapped construction; each start has its own clipping and L-BFGS solve. Both fitting arms use the same observation bank, but construction has additional acquisition costs. Score error is the within-cell median of menu maximum absolute errors, then the median over paired cells; relation error is the median over paired cells, and top-1 pools their menus. The three mapping failures remain in the attempted denominator.

	Eligible	Test-score error		Relation error		Top-1 (%)	
σ	/attempted	Random	Constr.	Random	Constr.	Random	Constr.
0	24/24	$7.33\,10^{-5}$	$6.56\,10^{-8}$	.337	$7.23\,10^{-8}$	98.2	100.0
10^{-10}	72/72	$6.82\,10^{-5}$	$3.06\,10^{-6}$	.318	$1.98\,10^{-4}$	98.1	100.0
10^{-8}	72/72	$7.18\,10^{-5}$	$7.15\,10^{-6}$	.305	.022	98.1	99.9
10^{-6}	69/72	$6.48\,10^{-5}$	$2.01\,10^{-5}$	.329	.061	98.3	99.4

Paired fitting. Both arms reuse all 240 training/validation banks without new fitting queries. The random arm uses eight pre-generated starts; the constructive arm replaces one with the mapped initialization. Each start has 12,000 float64 Adam steps under Section F.2's batch/step schedule, its own norm-20 clipping, and separate L-BFGS (at most 500 iterations on the same 4,096-row subset). The seven common trajectories are computed once; all starts share a separately seeded minibatch stream, not gradient norms or line search.

Validation MSE selects among initialization, every 500 Adam steps and the final L-BFGS iterate, before test evaluation. All optimization runs complete; missing mappings are separate failures. Saved checkpoints reproduce all selections, and independent prediction checks agree to 1.1×10^{-15}.

The prespecified primary error (within-cell median of menu maximum absolute errors), validation RMSE and relation error improve in 231/237 eligible cells and tie in six, where a common random start is selected. World-averaged primary error improves in all 24 worlds through 10^{-8} ($p = 1.2\times10^{-7}$) and in 23 with one tie at 10^{-6} ($p = 2.7\times10^{-5}$). Missing mappings are not imputed. Worst-menu error improves/ties/worsens in $230/6/1$ cells and top-1 in $109/124/4$; regressions occur only at 10^{-6}, with top-1 losses in two worlds. Pooled menu accuracy in Table 1 is descriptive, not an independent-sample test. The selected constructive checkpoint is its initialization in every noiseless cell and 65/72 cells at 10^{-10}: optimization is not uniformly beneficial. The comparison removes cross-start coupling, but not structural information carried by projected zeros, where the squared parameterization has zero derivative. It does not isolate pure basin selection or unrestricted optimizer superiority.

Total cost. The mapped start additionally pays construction, gauge search and projection, including the extra reads charged above. Equal fitting schedules do not equalize total cost; acquisition, mapping and fitting were not jointly timed.

F.2.2 New objectives with frozen dynamics

To test reuse beyond the acquisition objective, each world's score coefficients are redrawn three times from $U[-.8,.8]$, leaving tools and states unchanged. Frozen models refit only their readout from nested banks of $K = 16, 32, 64, 128, 256$ new noisy observations, then score the same 48 menus under the new objective. New words are equally split across lengths $4, 8, 16, 32$ with durations $\{.025, .05, .075\}$.

Construction refits nine numerator coefficients and retains the rational denominator, which represents state normalization and is generically identifiable up to scale by numerator/denominator coprimality. Random and constructively initialized direct fits retain generators/states and refit six additive coefficients. Linear PSR, GRU and Transformer retain their states or penultimate features and refit a linear output. All readouts use ridge regression with exact leave-one-out selection.

Independent-restart transfer and readout capacity. Independent, validation-selected models from Section F.2.1 use the same worlds, objectives, noise streams, budgets, menus and readout selection as the coupled-fit control. Dynamics remain frozen. Table 2 retains all attempted cells, including three unavailable high-noise initializations. Means first aggregate objectives/repeats within worlds; initialized and paired comparisons restrict both arms to the same eligible cells.

To test readout capacity, we additionally refit the nine products of the fitted models' normalized record coordinates. This contains the six additive features: for $\mathbf{1}^\top p = \mathbf{1}^\top q = 1$, $p_i = \sum_j p_i q_j$ and $q_j = \sum_i p_i q_j$. Thus a correct native representation can express the true additive objective family in either class. Construction instead uses bilinear features over its learned, frozen denominator. The control expands fitted-model readouts without new observations; it does not assume the learned coordinates or denominator are exact.

Table 2: **Objective transfer with independent restarts.** Top-1 (%) after K new reads; dynamics remain frozen. Direct and initialized fits use additive (six coefficients) or expanded bilinear (nine) readouts. All 24 worlds contribute; cell counts are 24, 72, 72, 72 across noise levels, except 69 eligible initialized cells at 10^{-6}. Means first aggregate objectives/repeats within worlds. Acquisition costs are separate and unequal. Figure 2 includes the intermediate budgets.

			Direct fit		Initialized fit	
σ	K	Construction	Additive	Bilinear	Additive	Bilinear
0	16	100.0	84.4	84.4	100.0	100.0
0	256	100.0	84.5	84.7	100.0	100.0
10^{-10}	16	100.0	84.7	84.0	99.8	99.8
10^{-10}	256	100.0	84.8	84.8	99.9	99.9
10^{-8}	16	99.4	85.2	84.6	98.6	98.6
10^{-8}	256	99.6	85.2	85.3	98.7	98.7
10^{-6}	16	90.3	85.0	84.8	93.8	93.5
10^{-6}	256	92.1	85.0	85.0	93.9	93.9

Construction beats both independent random-fit readouts in every world at every budget through 10^{-8}. At sixteen reads, its paired top-1 advantage over the additive fit is 15.6, 15.3, 14.1 and 5.3 percentage points across the four noise levels; paired world-bootstrap 95% intervals are $[12.6, 19.1]$, $[12.2, 18.8]$, $[11.4, 17.3]$ and $[2.4, 8.5]$. At 10^{-6} it loses to this fit in five worlds. Initialization improves high-noise top-1 over construction in 21/24 paired worlds, but not every error summary: mean normalized regret is .00424 versus .00353 for construction on the same eligible cells. This is conditional reuse evidence, not uniform dominance. Absolute errors also retain the high-noise reversal. At sixteen reads, median world-averaged menu-maximum errors for construction/direct-additive are $7.12\,10^{-10}/5.96\,10^{-4}$, $2.76\,10^{-6}/6.18\,10^{-4}$, $6.66\,10^{-5}/5.73\,10^{-4}$ and $.00109/6.10\,10^{-4}$. Corresponding mean normalized regrets are below $10^{-14}/.0148$, below $10^{-14}/.0142$, $3.67\,10^{-5}/.0127$ and $.00356/.0136$.

Original fitted cohort and broader controls. The following comparisons retain the original coupled fits, allowing direct comparison with the saved structure-free baselines and fixed-time composition study; they are not the independent-restart results above.

Table 3: Top-1 accuracy (%) under three new objectives per world after refitting only a readout, in the 48 cells with saved structure-free models (24 worlds, $\sigma = 0$ and 10^{-8}, one repeat). Fitted controls retain the original coupled restarts.

New reads K	16	32	64	128	256
Construction	99.7	99.7	99.8	99.9	99.9
Construction, then fit	99.0	99.0	99.0	99.0	99.0
Direct fit	79.2	79.2	79.0	79.1	79.1
Linear PSR	58.6	61.7	63.3	63.9	64.1
GRU	20.9	25.3	28.4	34.8	38.5
Transformer	11.4	10.5	13.3	14.8	17.7

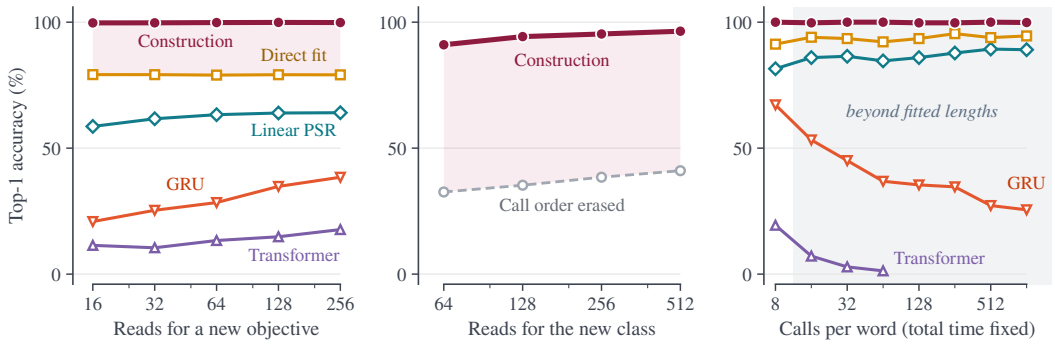


Figure 6: **Reuse controls under the original fitted protocol.** Left: new-objective readout fits with coupled-restart controls and saved structure-free models. Centre: new CNN classes, with synthetic tools. Right: longer words at fixed total duration, not extended physical horizons. These are the original cohorts; main Figure 2 isolates independent-restart objective transfer.

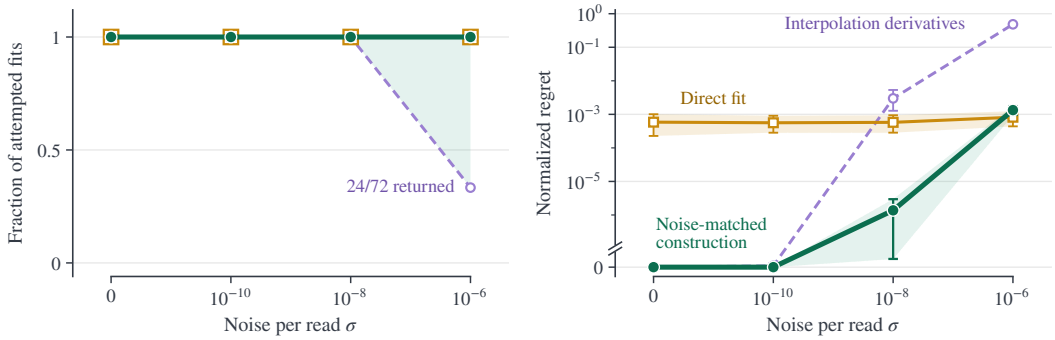


Figure 7: **Completion and regret with coupled fitting.** Left: finite predictions on every menu, over 24 worlds. Right: normalized regret with world-bootstrap 95% intervals and a linear scale below 10^{-7}. Dashed: interpolation derivatives, whose high-noise regret covers 24 fits in 17 worlds. The observation prefix is shared; construction adds reads and compute is not matched (Section F.2).

In the coupled cohort, sixteen reads give construction top-1 $100\%, 100\%, 99.4\%$ through 10^{-8} (median worst-menu errors $7.0 \times 10^{-10}, 2.3 \times 10^{-6}, 6.0 \times 10^{-5}$); direct fitting stays at 78.8–79.7% across budgets and noise levels. Construction wins in all 24 worlds at every budget through 10^{-8} on decisions ($p < 2 \times 10^{-5}$) and prediction ($p = 1.2 \times 10^{-7}$). At 10^{-6}, construction reaches 91.4–92.1% with at least 32 reads, beating direct fitting in every world. Initialized fitting reaches 93.7–93.9% on its 69 eligible cells and wins in every world at every budget; with sixteen reads its median worst-menu error is 2.3×10^{-4} versus 9.3×10^{-4} for direct fitting on those cells. Table 3's structure-free representations were selected for the original score; construction beats each on decisions in all 24 worlds at every budget ($p < 2 \times 10^{-5}$).

New classifier objectives. For Section F.2.5's twelve worlds, the same acquired dynamics are reused for the other nine class log-probabilities. Only the kernel-ridge readout is refitted, using 64–512 new observations per class. The original relation errors are reproduced. At 512 reads (.04% of the original 1,267,312-endpoint acquisition), top-1 is 96.4% over 5,184 menus, versus 91.0% at 64 reads; every world exceeds 92.6%. A readout on per-tool total durations, which erases order, achieves 41.0% on the same reads. Construction wins in all twelve worlds ($p = .0022$), with normalized regret .0011 versus .27. Refitting the original true-class readout on 512 such reads gives 94.6%.

F.2.3 ACQUISITION-SCORE CONTROLS WITH COUPLED FITTING

These acquisition-score controls use coupled restarts, unlike main Figure 3. Coupled constructive initialization improved primary error in 236/237 cells and relation error in 235, but worsened relation error relative to construction alone in 68.

Table 4: Positive-time closed loop with coupled fitting, 24 independent worlds. Constructive uses the noise-matched derivative functionals; Interpolation is the frozen study's interpolation functional on the same observations. Completion means finite predictions on all menus, not acceptance under an accuracy gate. Regret is the world mean over completed fits; Constructive and Direct cover all 72 fits in 24 worlds at every nonzero level.

Noise/read	All-menu completion			Normalized regret	
σ	Constructive	Interpolation	Direct	Constructive	Direct
0	24/24	24/24	24/24	$9.11\,10^{-16}$	$5.85\,10^{-4}$
10^{-10}	72/72	72/72	72/72	$3.90\,10^{-16}$	$5.61\,10^{-4}$
10^{-8}	72/72	72/72	72/72	$1.37\,10^{-6}$	$5.77\,10^{-4}$
10^{-6}	72/72	24/72	72/72	$1.34\,10^{-3}$	$8.21\,10^{-4}$

Prediction and structural recovery. Construction has MAEs $1.06\,10^{-9}$, $2.10\,10^{-6}$, $5.04\,10^{-5}$ and $1.23\,10^{-3}$ and worst errors $2.58\,10^{-8}$, $2.09\,10^{-5}$, $4.74\,10^{-4}$ and $.0144$ across increasing noise. Direct-fit MAEs are 8.48–$8.95\,10^{-5}$, worst $.00114$: construction predicts better through 10^{-8}, but worse at 10^{-6}. Interpolation on the same observations has nonzero-noise MAEs $1.05\,10^{-5}, 1.27\,10^{-3}, .225$ and 48 high-noise non-returns: 31 factor-split, one Schur-ordering and 16 calibration-cap failures. Noise-matched estimation removes these failures, with completion/probe costs as above. Endpoint noise σ is not the bracket error η in Theorem 4.2; these returns do not establish its input conditions or uniform d_{op} guarantee.

The unconstrained readout is broader than the additive model:

$$\frac{a^\top s}{c^\top s} + \frac{b^\top t}{d^\top t} = \frac{s^\top (ad^\top + cb^\top)t}{s^\top (cd^\top)t}.$$

It imposes neither a rank-one denominator nor shared numerator factors; its contribution to high-noise error has not been isolated. Noiseless constant and order-erased controls have normalized regret $.500$ and $.244$.

Truth/self and factor-permutation controls give zero; random models give median $.997$ and true-support random models $.999$. Re-expression with conditions $10, 10^3, 10^6$ gives errors $4.0\,10^{-15}, 1.7\,10^{-11}, 1.1\,10^{-5}$. Eighth singular-value medians are .28–.34 in both arms, with the ninth at machine zero (a fifteen-order gap).

A separate support diagnostic thresholds generator Frobenius norms at 10^{-3} after factor permutation; it is gauge-dependent, not a confidence test. Construction matches all memberships in the 96 cells at $0, 10^{-10}$ and none at higher noise. Nevertheless, acting names outrank inactive names in every factor through 10^{-8} (median gaps .32–.33), and in 138/144 factors at 10^{-6} (median/mean AUC $1/.998$, gap .25). Direct fitting orders them correctly in 351/480 factors, with median AUC one, mean .95–.96, and acting/inactive scores near $.34/.14$. Its fixed threshold misclassifies every inactive name despite these rankings. Thus relation error, support ordering and threshold agreement answer different questions.

Table 5 separates structure from decisions. Construction has smaller relation error in 235/240 cells, including 67/72 at 10^{-6}. Direct fitting's $.440$–$.469$ errors indicate imperfect relations rather than absent support information. Constructive top-1 is higher/tied/lower in $114/53/1$ of the 168 cells through 10^{-8}, but $9/20/43$ at 10^{-6}, where direct fitting decides better despite larger structural error. Both arms select an exact optimum in at least one menu in every cell and have median normalized regret zero. Across all 1,920 direct-fit restarts, the minimum relation error is $.140$; no cell contains a restart below .1. This rules out validation selection as the sole explanation within the tested optimizer budget, not identification by other direct methods.

Finite-duration order effects. For each of 153 name pairs, six held-out histories and $t = .02, .04, .08$, define $D_{ab}(t) = y(Hab) - y(Hba)$ and $c_{\mathcal{T}} = \sum_t t^2 D_{ab}(t) / \sum_t t^4$. This through-

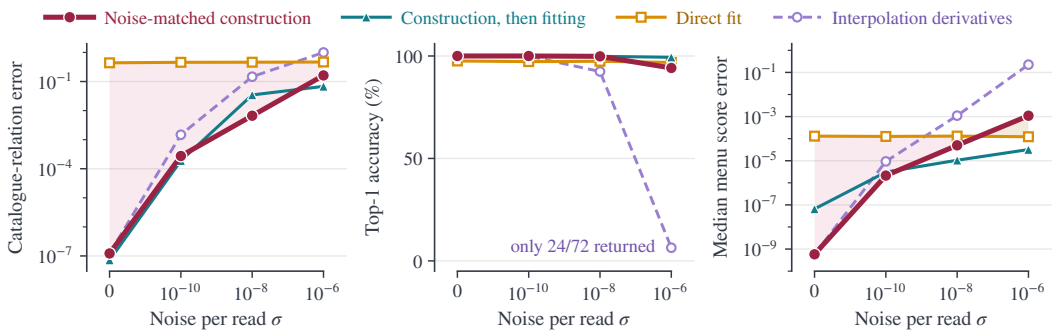


Figure 8: **Original estimator and fitted-cohort controls.** Left: catalogue-relation error; centre: menu top-1; right: median menu-maximum error across cells. Dashed interpolation returns in only 24/72 high-noise cells; noise-matched construction returns throughout. Fitted arms use the original coupled restarts. These complete curves retain the earlier comparison; main Figure 3 uses independent restarts.

origin contrast has Taylor bias (Section B.4); it is not the derivative-limit coefficient. Both scorers follow (14)'s left-action convention. Of 918 contrasts per cell, 427–498 per world exceed 10^{-3}; sign agreement excludes the remainder by convention.

At $\sigma = 0, 10^{-10}, 10^{-8}, 10^{-6}$, median relative errors are $4.0\,10^{-8}, 1.7\,10^{-4}, 3.7\,10^{-3}, 9.8\,10^{-2}$ for construction; 4.5–$4.8\,10^{-2}$ for direct fitting; and $4.2\,10^{-7}, 1.6\,10^{-4}, 4.5\,10^{-3}, 1.0\,10^{-2}$ for initialized fitting. Corresponding sign agreement is 1.000 through 10^{-8} and .989 at 10^{-6} for construction, .997–.998 for direct fitting, and 1.000 throughout for initialized fitting. Relation and contrast errors correlate across 717 returned models (log–log correlation $+.977$; within-family $+.992, +.694, +.985$). Binned contrast-error medians are $1.1\,10^{-7}, 1.6\,10^{-4}, 3.0\,10^{-3}, 6.5\,10^{-3}, 4.5\,10^{-2}, 7.2\,10^{-2}$ for relation-error bins below 10^{-6}, to $10^{-3}, 10^{-2}, .1, .5$, and above. These shared-world associations neither constitute 717 independent replications nor give a sufficient order-effect guarantee.

World-level comparisons. Construction improves relation error in all worlds through 10^{-8} ($p = 1.2 \times 10^{-7}$) and 23 at 10^{-6} ($p = 3.6 \times 10^{-7}$). Mean log-error differences are $-6.56, -3.21, -1.81, -.45$ decades, with intervals $[-6.72, -6.40], [-3.28, -3.13], [-1.89, -1.73], [-.55, -.35]$. Top-1 improves/ties/worsens in $14/10/0, 20/4/0, 19/4/1$ worlds through 10^{-8} ($p = 8.3 \times 10^{-4}, 8.6 \times 10^{-5}, 1.1 \times 10^{-4}$); it worsens in twenty at 10^{-6} ($p = 1.2 \times 10^{-4}$). Median worst-menu error improves in 24, 24 and 23 worlds at the first three levels and worsens in all 24 at the last.

Table 5: Action relations versus decisions with coupled fitting, over the same 240 cells as Table 4. Relation error is the median over cells that returned a model; a random model scores about one. The interpolation entry at 10^{-6} covers its 24 returned models. Top-1 is the fraction of that arm's own scored menus whose selected candidate is the true optimum. The constructive and direct arms return a model in all 240 cells.

Noise/read	Relation error			Menus decided top-1	
σ	Constructive	Interpolation	Direct	Constructive	Direct
0	$1.22\,10^{-7}$	$1.22\,10^{-7}$	.440	100.0%	97.6%
10^{-10}	$2.74\,10^{-4}$	$1.48\,10^{-3}$	.458	100.0%	97.3%
10^{-8}	$6.66\,10^{-3}$	.146	.462	99.8%	97.4%
10^{-6}	.162	.999	.469	94.2%	96.8%

F.2.4 More factors and larger factors

The dimension-generic learner uses the same pipeline with a commutant of dimension m: eigenvalue clustering of a fixed generic traceless element yields blocks of public sizes $q_f^2 - 1$. The readout is

a multilinear ratio in the Kronecker product of recovered rays. For $n = \sum_f (q_f^2 - 1)$, the design uses $\max(64, 3n)$ contexts and $\max(64, 4\prod_f q_f)$ readout histories. Each factor has $q - 1$ diagonal generators with traceless spectra in $[-.65, .65]$ and $q(q-1)$ Metzler generators drawn as in Section F.2; one co-moving name joins each adjacent factor pair. Names are randomly permuted. The derivative rule is unchanged; eight fresh worlds per configuration and noise level follow one smoke test, without evaluation-world tuning.

Table 6: Construction with more or larger factors, eight worlds per row and noise level; every world returns a model. Endpoints include the 256 noise-probe reads.

Factors	n	Names	Endpoints	Relation error, $\sigma = 0$		Relation error, $\sigma = 10^{-10}$	
				median	worst	median	worst
$3, 3$	16	17	1,266,080	$1.3\,10^{-6}$	$5.5\,10^{-6}$	$4.5\,10^{-3}$	$9.1\,10^{-3}$
$3, 4$	23	24	2,861,681	$1.8\,10^{-7}$	$3.9\,10^{-6}$	$1.1\,10^{-3}$	.013
$3, 3, 3$	24	26	3,265,636	$4.6\,10^{-7}$	$9.2\,10^{-5}$	$6.5\,10^{-3}$	.225
$4, 4$	30	31	6,393,650	$3.1\,10^{-7}$	$2.0\,10^{-6}$	$4.7\,10^{-3}$	.019
$3, 3, 3, 3$	32	35	7,848,292	$1.6\,10^{-6}$	$2.3\,10^{-5}$	$9.5\,10^{-3}$	.041

All 80 constructions at $\sigma = 0, 10^{-10}$ return (Table 6); mixed-size relation errors use each factor's rank $q_f^2 - 1$. Noiseless top-1 is perfect on all 48 menus per world. At 10^{-10}, mean top-1 in table order is $1.000, 1.000, .893, .997, .961$; one three-factor world with relation error .225 lowers the third entry. The dominant bracket budget is $162\max(64, 3n)\binom{n}{2}$, scaling as n^3 at fixed conditioning. At 10^{-8}, returns are $8, 8, 6, 7, 3$ of eight, with median relation errors $.084, .040, .12, .17, .39$. Quadrupling contexts for $(3, 3, 3)$ and $(4, 4)$ costs 12,973,972 and 25,495,880 endpoints, increases returns to seven and eight, lowers median errors to $.068, .076$, and improves returned-model top-1 from $.903, .914$ to $.949, .969$. The roughly halved errors are consistent with averaging. The equal-size $(3, 3)$ scaling family is harder than the original family ($.084$ versus $.0067$ at 10^{-8}), so size alone does not explain difficulty.

F.2.5 A TRAINED CLASSIFIER AS THE HIDDEN SCORE

A frozen CIFAR-10 classifier supplies the scalar readout: five 3×3 GELU convolutions and average pooling, 140,714 parameters, twelve training epochs, and 75.5% test accuracy. Each new world couples a fresh tool catalogue to a test image with unmodified true-class probability in $(.3, .9)$. Two colour-balance rays control the central 16×16 region and border through gains $3s/(\mathbf{1}^\top s)$; the uniform ray is identity. Tools act as before; the learner sees only float64 log-probabilities, without added noise.

The derivative window is $.0125$, selected on two development worlds: $.3$ gave relation error near one, while $.6$ failed in the one world tested; error fell below 10^{-6} at the smallest tested window $.0125$. Other acquisition steps are unchanged. A Gaussian-kernel ridge readout uses 512 further legal histories of lengths four–32, upper-triangular $ss^\top$ features per record, median-heuristic bandwidth and leave-one-out ridge selection. Total cost is 1,267,312 endpoints per world. All twelve evaluation worlds return, with median/worst relation error $1.0 \times 10^{-6}/4.3 \times 10^{-6}$. On 576 menus, top-1 is 96.9%, versus 45.0% with order erased and 3.1% for random choice; normalized regret is $.0018$ versus $.263$. Median worst-menu error is $.0028$ against median range $.048$, consistent with readout regression limiting prediction. Top-1 improves in all worlds (Wilcoxon $p = .0021$; world-bootstrap 95% interval $[95.5\%, 98.1\%]$). The readout is a trained network, but the tools remain simulated colour transformations.

F.2.6 DISCOVERING THE DIMENSIONS

With only anonymous names and the type-A contract supplied, the learner infers n, m, q_f before the corresponding acquisition steps. The first-order profile over 64 contexts determines n by its largest consecutive singular-value ratio, accepted if at least ten and the residual spectrum is below $3\hat{\sigma}\|w\|(\sqrt{64} + \sqrt{N})$. Here N is catalogue size and $\hat{\sigma}$ comes from degree-seven residuals. The adjoint commutant determines m after removing the identity, requiring ratio at least five; generic

commutant eigenvalue clusters must have sizes $q_f^2 - 1$. Each stage may abstain. Thresholds were development-fixed; correct dimensions leave subsequent queries unchanged.

Table 7: Cells in which n, m and every q_f are all returned and correct; every other cell abstains, and no cell returns a wrong set of dimensions.

Worlds	$\sigma = 0$	10^{-10}	10^{-8}	10^{-6}
Closed loop, factors $3, 3$	$24/24$	$72/72$	$72/72$	$7/72$
Scaling, factors $3, 3$	$8/8$	$8/8$	$3/8$	
Scaling, factors $3, 4$	$8/8$	$8/8$	$5/8$	
Scaling, factors $3, 3, 3$	$7/8$	$7/8$	$0/8$	
Scaling, factors $4, 4$	$8/8$	$7/8$	$2/8$	
Scaling, factors $3, 3, 3, 3$	$8/8$	$7/8$	$0/8$	
Trained classifier, factors $3, 3$	$12/12$			

No complete output has incorrect dimensions (Table 7). Through 10^{-8}, all closed-loop cells succeed, with accepted ratios at least 255 for n and 11.9 for m; at 10^{-6}, 13 abstain at n and 52 at m, despite successful supplied-dimension construction. Scaling succeeds in 39/40 noiseless and 37/40 low-noise cells. Quadrupling contexts at 10^{-8} raises $(3, 3, 3)$ and $(4, 4)$ successes from zero/two to three/four of eight. Two intermediate errors occur in one $(3, 3, 3)$ world: $n = 23$ instead of 24 at 10^{-8} under either budget, which shares the first 64 contexts, and $m = 2$ instead of three without noise. Both subsequently abstain. Replayed acquisition observations match the original 240 cells exactly.

Dimension labels versus factor extraction. The identity $n = \sum_f (q_f^2 - 1)$, $q_f \geq 2$, can determine factor sizes when its unordered decomposition is unique; $n = 16$ permits only $8 + 8$. Otherwise the original gates remain. Across 388 checks this rule adds 57 correct label sets and no incorrect ones, raising high-noise closed-loop labels from 7/72 to 59/72; thirteen still fail the unchanged n gate. Labels alone supply neither a stable factor basis nor a qualified executor.

F.2.7 Structure-free baselines

Three model families receive the same original observations and 512 validation endpoints, without record, factor or projective priors: a two-layer GRU and three-layer Transformer trained for 6,000 Adam steps with batches of 4,096 words; and a bilinear continuous-time predictive-state model $y = c^\top \exp(t_L G_{e_L}) \cdots \exp(t_1 G_{e_1}) z_0 + b$ of dimension nine or sixteen, trained for 4,000 steps. Each keeps its validation-best checkpoint; validation selects the linear PSR in all 96 cells (one repeat per world and noise level). These models do not return action relations. Table 8 reports held-out menu accuracy.

Table 8: Top-1 accuracy (%) on the 48 unseen menus of each of 24 worlds, one noise repeat per world. Linear PSR is the validation-selected structure-free model; fitted controls use the original coupled restarts.

σ	Construction	Construction, then fit	Direct fit	Linear PSR	GRU	Transformer
0	100.0	100.0	97.6	94.1	65.0	37.5
10^{-10}	100.0	100.0	97.0	94.1	64.8	37.0
10^{-8}	99.7	99.8	97.4	94.1	64.8	36.8
10^{-6}	94.2	99.2	96.8	93.3	61.7	36.2

Through 10^{-8}, construction beats the selected structure-free model on decisions in 23/24 worlds, tying the other (Wilcoxon $p \leq 2.5 \times 10^{-5}$), and on prediction in all worlds ($p = 1.2 \times 10^{-7}$). Median worst-menu errors are $5.8 \times 10^{-10}, 2.3 \times 10^{-6}, 5.0 \times 10^{-5}$ versus 3.8×10^{-4}. At 10^{-6}, decision wins/ties/losses are $13/2/9$ ($p = .18$), while the structure-free model predicts better in 23 worlds. Constructive initialization then improves decisions in 21 worlds, tying three ($p = 5.4 \times 10^{-5}$), and prediction in all worlds. The GRU/Transformer, trained on at most eight calls, achieve 62–65%/36–38% top-1 on eight–32 calls. Parametric direct fitting beats the selected structure-free model in sixteen–nineteen worlds and loses in three–four at each noise level.

F.2.8 LONGER AND FINER WORDS

Acquisition words contain at most eight calls and validation words at most 32. Frozen models now score $L = 8, 16, \ldots, 1024$ calls of duration $2/L$: total physical time stays two, while the shortest calls approach .002. Sixteen menus per length reorder eight segments, each containing $L/8$ calls from its own name pair; 32 distinct candidates retain median score ranges .014–.020. The Transformer's 64-position limit restricts its evaluation to $L \leq 64$. Table 9 reports menu accuracy; Figure 6 combines this study with objective transfer and classifier results, excluding initialized-fit rows.

Table 9: Top-1 accuracy (%) on menus of reordered segments at fixed total time, in the 48 cells of Table 3; no model saw a word longer than 32 calls. Fitted controls retain the original coupled restarts.

Calls L	8	16	32	64	128	256	512	1024
Construction	100.0	99.7	100.0	100.0	99.7	99.7	100.0	99.9
Construction, then fit	100.0	99.7	99.6	99.9	99.9	99.9	100.0	99.5
Direct fit	91.3	94.0	93.5	92.2	93.5	95.4	93.9	94.5
Linear PSR	81.5	85.9	86.5	84.6	85.9	87.8	89.3	89.1
GRU	67.2	53.3	45.1	36.8	35.4	34.6	27.2	25.5
Transformer	19.4	7.2	2.9	1.3				

Across 240 cells, construction's median worst-menu error at eight versus 1,024 calls is 1.3×10^{-9} versus 1.1×10^{-9} without noise and 1.4×10^{-4} versus 9.6×10^{-5} at 10^{-8}. Top-1 remains 99.0–100% through 10^{-8}. At 10^{-6}, initialized fitting gives 97.7–99.3% against direct fitting's 92.0–95.9% on the same 69 eligible cells. Continuous-time direct and PSR models also retain short-word accuracy; the GRU declines from 67.2% to 25.5%, and the Transformer from 19.4% to 1.3% at 64 calls. Construction beats both token models in all worlds at every tested length ($p < 2 \times 10^{-5}$), and PSR in nineteen–24 worlds, losing in at most one. These results concern finer composition at fixed physical time, not longer-horizon stability.

F.2.9 REUSING A FROZEN ACTION REPRESENTATION

Three tasks test new-tool enrollment, substitution and duration optimization using the same 24 acquisition worlds. States and targets remain evaluator-only, and original acquisition costs remain charged.

New-tool enrollment. A new complete call $.45I + .55Q$ acts on one unknown factor, with independent Dirichlet$(3, 3, 3)$ columns of Q. Its target and native coordinates are hidden. Frozen generators, rays and readout predict pre-call states; terminal responses decode post-call states using 18 names at nine positive nodes in $(0, .6)$, the empty suffix and two single-call suffixes (486 endpoints per history). Cross-product least squares fits a homogeneous map on each factor from 16 or 64 nested four-call histories. The dictionary is projective-gauge closed; cross-factor/nonlinear maps and singular resets are not tested.

Warm and cold controls jointly refit old generators, rays, readout and new maps from the constructed model or random parameters, with identity new maps. Both use identical new scalars and all old rows as eligible samples: 1,200 float64 Adam steps, batches of 256 old/256 new rows, cosine step size .002 to .0001, cap-20 shared gradient clipping, and validation every 200 steps including initialization. Selection minimizes equally weighted old/new validation MSE. Shared GPU batching couples the controls; independently clipped comparisons are reported below. An independent 128-observation validation bank is charged once, giving 7,904/31,232 new endpoints at $R = 1$. Evaluation uses 96 unseen words with one/two new calls among $4, 8, 16$ old calls, plus 48 old-only words. Preparation durations are $.04, .12, .24$; evaluation durations $.025, .05, .075$. All banks are disjoint. One replicate per world at $\sigma = 0, 10^{-8}, 10^{-6}$ gives 72 cells; both stages complete without deployment queries. Noiseless bases equal the original models; noisy bases use noise-matched acquisition with the same enrollment data.

In the original comparison (Table 10), frozen enrollment beats both coupled refits in every world at $0, 10^{-8}$ and both budgets; worst errors with 64 histories are $4.11 \times 10^{-8}, 4.87 \times 10^{-4}$. At 10^{-6},

Table 10: Original new-tool enrollment: fixed-degree decoding and unnormalized homogeneous fitting, with coupled-gradient refits. Entries are medians across 24 worlds of maximum score error on 96 unseen mixed words. All methods share new scalars; budgets exclude the separately charged original acquisition. Table 11 tests conditioned fitting and independently clipped controls.

Noise	Base models	Histories	Frozen	Warm joint	Cold joint
0	$24/24$	16	$1.25\,10^{-9}$	.00595	.0132
0	$24/24$	64	$1.25\,10^{-9}$	.00200	.0158
10^{-8}	$24/24$	16	$1.09\,10^{-4}$	.00468	.0133
10^{-8}	$24/24$	64	$1.05\,10^{-4}$	.00246	.0159
10^{-6}	$24/24$	16	.0195	.00475	.0131
10^{-6}	$24/24$	64	.00494	.00269	.0158

warm refitting wins in 19/24 worlds, and one frozen map reaches error 8.12. Thus preserving old outputs does not establish accuracy, and finite return is not release qualification.

Representation and decoder controls. A fourth arm freezes the same representation but learns its new map by gradient descent in the coupled optimizer. All four arms return at $8, 16, 32, 64$ histories in all 72 cells. Frozen arms change old parameters/outputs by at most 3.3×10^{-16}. At 64 histories, median maximum errors for decoder/frozen-gradient/warm/cold are $(1.3\times10^{-9}, .0013, .0020, .016)$ without noise, $(1.05 \times 10^{-4}, .0012, .0025, .016)$ at 10^{-8}, and $(.0049, .0037, .0027, .016)$ at 10^{-6}. Decoding wins in every world at eight and 64 histories at the first two levels; warm fitting is best in fifteen high-noise worlds.

Eight histories cost 4,016 endpoints and give decoder errors $1.9 \times 10^{-9}, 1.6 \times 10^{-4}$ at $0, 10^{-8}$, with little subsequent change. Warm fitting improves from .0074 to .0020 without noise as histories increase. At eight histories, warm/cold fitting increases old-word error from 1.3×10^{-9} to $.00058/.0046$ without noise; at 10^{-8} warm gives .00062 versus .00011 frozen. At 10^{-6}, however, warm improves old-word error from .0027 to .00052. Original decoding costs .6–4.9 CPU seconds versus 67.4–67.5 seconds for the shared GPU optimizer; noisy reruns were untimed. These are neither per-method matched nor end-to-end timings.

Conditioned map estimation. To reduce learned-gauge sensitivity, input/output rays are sphered by empirical second moments, normalized, fitted homogeneously and transformed back. Moments with eigenvalue ratio at most 10^{-12} are rejected. Refinement minimizes cross products of unit predicted/observed rays, fixing the largest matrix entry to one (200 evaluations, tolerance 10^{-10}). Both factors are fitted without supplied targets, truth coordinates or positivity clipping. Degree-eight and acquisition-selected decoders use the same 64 histories/31,232 endpoints, without deployment-based selection. This adds offline computation, not a gauge-invariant release certificate.

Six variants return in all 168 cells: the original 72, 48 independent high-noise base/enrollment repeats, and 48 fresh enrollment-noise cells sharing the original high-noise bases. Fixed-degree refinement reduces the original-bank worst error from 8.12 to .0127. Generalization is checked on three additional 96-word banks per world, disjoint from original test/validation words, using unchanged lengths/durations and shared across arms and repeats. These remain 24 worlds.

On the 48 fresh cells, selected-degree refinement beats independently clipped warm/cold/frozen-gradient controls in $34/48, 48/48, 46/48$ cases; fixed-degree refinement also beats warm in $34/48$. Neither degree selection nor sphering alone explains the improvement. Every refined fresh-cell maximum exceeds 10^{-3}, and the independent base/noise repeats have worst new-word error .0117, precluding a uniform 10^{-2} claim. Validation-selected history count in another 48-stream check raises worst error from .688 to 1.29. Conditioning improves map estimation, not the base model's high-noise accuracy or total acquisition cost.

Substitution and duration edits. These tests retain the original coupled-restart direct and initialized fits, not the independent-restart cohort of Section F.2.1. For each anonymous name, fit a nonnegative combination of two other names with weights summing to at most 1.5. A first-order control uses 64 contexts and a pivoted expansion in 16 directions, without bracket calibration. Both orders are tested at durations $.1, .3, .6$ in twelve held-out prefix/suffix contexts (72 comparisons per target). Success

Table 11: **New-tool fitting with fresh enrollment noise and unseen words.** Two noise streams per each of 24 existing worlds, $\sigma = 10^{-6}$, $R = 1$, 288 new words per cell. Every arm returns. Controls use the same observations, 1,200-step schedules and validation selection, now with per-arm gradient clipping. Threshold counts are evaluator summaries, not release tests; base acquisition remains separately charged.

	Maximum error per cell		Cells below threshold	
Estimator	Median	Worst	10^{-3}	10^{-2}
Homogeneous, degree eight	.00439	1.47	0/48	29/48
Homogeneous, selected degree	.00411	17.9	0/48	36/48
Sphered only, degree eight	.00376	5.37	0/48	42/48
Sphered only, selected degree	.00401	.0319	0/48	44/48
Sphered + refined, degree eight	.00309	.00839	0/48	48/48
Sphered + refined, selected degree	.00303	.00758	0/48	48/48
Warm joint refit	.00574	.0946	2/48	28/48
Cold joint refit	.0175	.103	0/48	8/48
Frozen-map gradient fit	.00684	.121	0/48	28/48

requires score and evaluator-only native-state errors at most 10^{-3}. Proposal thresholds $10^{-3}, 10^{-2}$ apply to relative Euclidean representation residuals, not gauge-invariant certificates.

Construction, direct fitting, initialized fitting and the first-order control release 328, 0, 190 and 258 proposals at 10^{-3}, and 357, 0, 224 and 336 at 10^{-2}, respectively, with no false release. All 240×18 opportunities, including three absent initializations, remain in each denominator. Zero releases indicate failure of this rule, not general behavioral failure. Without noise, construction and first-order control both substitute all 48 shared targets; the latter uses 10,368 already acquired scalars (72,576 primitive occurrences). Substitution requires no new queries and does not demonstrate the necessity of full calibration.

For duration optimization, each frozen model updates two schedules at each length $8, 16, 32$, initialized uniformly in $[.035, .095]$. Twelve analytic-gradient projected steps of maximum coordinate size .01 optimize over $[.01, .12]$; a 24-candidate random search includes initialization. Costs are thirteen gradient versus 24 score evaluations, not equal compute. Construction and direct fitting improve true score in all 1,440 tasks each; initialized fitting does so in all 1,422 eligible tasks, at every noise level and without deployment queries. This establishes a finite-task use, not uniform gradient accuracy or superiority to direct fitting.

F.2.10 Bracket-acquisition diagnostic

This intervention uses the original interpolation construction, not the noise-matched models. Replace the estimated bracket by the noiseless output of the same pipeline, retaining sign, scale and finite-difference conventions; clean data are evaluator-only and noisy acquisition remains charged. At $\sigma = 10^{-10}, 10^{-8}, 10^{-6}$, median relation error changes from $.00148, .146, .999$ to $7.98 \times 10^{-6}, 7.98 \times 10^{-4}, .0861$. High-noise top-1 rises from .062 to .865 and all 48 non-returns disappear. Additional factor-split replacement changes nothing; the noiseless intervention is identity. Conditional errors have different return sets.

The intervention identifies bracket acquisition as consequential, not mixed differences as its unique cause: $\widehat{X} = \widehat{B}^{\dagger}\widehat{D}$ depends on estimated responses and differences. Replacing their combined output does not allocate residual error uniquely to calibration/readout, or provide a usable noiseless oracle.

F.2.11 Calls versus physical time

Using original interpolation/direct models, menus independently vary $L = 8, 16, 32, 64$ and $T = .5, 1, 2, 4$, with step duration T/L. Median worst error over increasing T is 4.06×10^{-5}–5.50×10^{-4} at $L = 8$ and 1.94×10^{-5}–1.51×10^{-4} at $L = 64$ for construction; direct-fit ranges are 9.24×10^{-5}–2.17×10^{-3} and 3.84×10^{-5}–4.49×10^{-4}. At $T = 4$, eightfold larger L reduces errors by factors 3.6 and 4.8. Holding step duration fixed instead increases error with L. Dependence on T is consistent with the telescoping bound; improvement with L is empirical, not predicted by it.

Subdivision satisfies $e^{tA} = (e^{tA/k})^k$. Of 5,184 model pairs, thirty exceed the preregistered 10^{-12} invariance threshold, all high-noise constructions with generator norm 418 and exponential condition 2.4×10^4. Other strata remain below 2.4×10^{-13}; truth changes by at most 1.1×10^{-15}. Maximum deviations grow as $3.2 \times 10^{-10}, 1.1 \times 10^{-9}, 2.4 \times 10^{-9}$ for $k = 2, 4, 8$, consistent with numerical accumulation and providing a truth-free sensitivity check.

A scorer-equivalent evaluation exposes rational denominators. The median minimum across each menu is .67–.72 through 10^{-8} and .0030–.0084 at 10^{-6}; the global minimum is 2.1×10^{-7}, without sign changes. Large scores thus accompany near-zero denominators, not observed pole crossings. At high noise the median denominator increases from .0030 at $L = 8$ to .0084 at $L = 64$, excluding increasing call count as its cause here. Direct fitting has no rational readout denominator; its state sum-normalization remains 1.22–1.73.

F.3 Allocating repetitions by the estimator's own weights

We test repetition allocation separately from degree selection, using the original interpolation estimator. For independent reads, $\mathrm{Var}(\sum_j w_j \bar{Y}_j) = \sum_j w_j^2 \sigma_j^2 / r_j$, so fixed-budget optimal quotas satisfy $r_j \propto |w_j| \sigma_j$, as in (60). Uniform and weighted arms match both endpoints and primitive calls, since words within an estimator have equal length. At one read per node, the implementation reproduces every original model exactly. Profile and bracket nodes are repeated explicitly in the simulator and averaged; readout targets remain single-read. Counts charge every simulated endpoint and primitive occurrence, not physical-service measurements.

At multipliers $M = 8, 30$, predicted bracket-variance ratios are $.385, .374$ (standard-deviation ratios $.621, .612$). Across 24 worlds, weighted/uniform median relation-error ratios are $.625, .557$ at $\sigma = 10^{-10}$ and $.640, .554$ at 10^{-8}; weighting improves 95/96 paired cases. At $M = 8$, world-median menu errors fall from 3.76×10^{-6} to 2.00×10^{-6} and from 3.77×10^{-4} to 1.95×10^{-4} at these noise levels. Uniform $M = 21$ matches weighted $M = 8$ in predicted variance and within 5.4% in relation error, but costs 26,596,144 versus 10,131,904 endpoints, a factor 2.62. All arms select identical pivot names. Noiseless relation errors agree to median relative 1.4×10^{-3}, as expected when allocation cannot affect statistical noise. Decision performance is saturated: median regret is zero throughout and top-1 is one, except .979 for both allocations at $\sigma = 10^{-8}, M = 8$. Weighting improves structural accuracy here, not decisions.

Reducing the node count offers a distinct bias–variance tradeoff. Chebyshev–Gauss Vandermonde condition is 1.414 at every tested count, and fewer nodes reduce $\sum_\nu |w_\nu|$. Relative to nine nodes, eight/seven/six/five nodes multiply derivative bias for e^t by $47, 1.9 \times 10^3, 6.6 \times 10^4, 1.8 \times 10^6$, while freeing $21\%, 40\%, 56\%, 69\%$ of the quadratic bracket budget. Weighted/uniform variance ratios range from .373 to .529. The allocation study increases repetitions without changing this truncation bias.

F.4 Reset isolation and thin preparations

Query-language witness. In 24 paired worlds, five records share locked binary motion and ten anonymous pair resets to $1/2$. Scores are $h_\pm(x) = .2 + .03 \sum_i x_i \pm (x_1 - x_2)(x_2 - x_3)(x_3 - x_1)$. Grouped and interleaved arms each use 384 words/2,688 primitive occurrences per service, sharing multisets of four raw calls and three resets. Grouped words place all resets between raw-only prefix/suffix segments starting at the common reference. This preserves at most two distinct record values, annihilating the cubic. Interleaving yields maximum paired mean differences .00201–.00968 per world versus exact zero when grouped. Evaluator-constructed record-swap menus reverse optima in all 24 pairs, with gaps 2.57×10^{-4}–4.73×10^{-4}. These means are noiseless and menus use evaluator-known symmetry.

Across 192 specificity-control cells with strengths $0, .01, .1, 1$ and either this cubic or random degree-at-most-three polynomials, neither arm separates zero-strength pairs, and both separate random nonzero polynomials in all worlds. The witness concerns this cancellation, not superiority over all query designs.

Blind acquisition and decisions. A common scalar-only learner compares grouped, interleaved and random-legal grammars without supplied masks, flow families, rates or score sign. Grouped/interleaved arms match every call multiset. The original random arm matches their dis-

tribution, but shared random streams alter individual multisets; the corrected control independently permutes exactly the same multisets. These are separate controls.

Möbius inversion over reset subsets identifies signatures. Ratios of second to squared first derivatives identify three flow families; isolating reset sets decode records, followed by rates, masks and a cubic readout. Post-freeze prediction uses learned dynamics without queries. The original degree-six estimator uses seven positive Chebyshev nodes in $(0, .15)$, at 8,431 endpoints and 53,529 primitive occurrences per complete $R = 1$ acquisition. All non-returns remain in attempted denominators.

Menus pair legal reorderings within a public multiset, symmetrically under world exchange, with base-score differences small relative to the cubic contribution. They need not have equal base scores or opposite optima in every pair. The complete-transcript/imbalance bound in Section B.4, rather than the exact balanced-menu special case, applies to their interpretation. Noiseless grouped accuracy is .522 (world-bootstrap interval $[.512, .531]$), versus 1.000 for interleaved and random-legal arms; normalized regrets are .478 and zero. The interval excludes .5. Design ranks are $50/56$ versus $56/56$, with smallest singular values 1.2×10^{-18} and 1.7×10^{-4}. The grouped language's six-dimensional cubic kernel is a theorem; finite design rank and complete discovery are distinct empirical checks. At $\sigma = 10^{-7}$, one/sixteen/256 reads per question give conditional accuracies $.883/.963/.990$ (interleaved), $.507/.504/.498$ (grouped) and $.866/.955/.991$ (distribution-matched random). Wordwise-matched random control returns 48/48 with perfect accuracy without noise and 46/48 with conditional accuracy .884 at 10^{-7}, versus .883 interleaved. It was tested only at one read, not sixteen/256; no comparison establishes designed interleaving as superior to random legal ordering.

On eight unfiltered reorderings, grouped accuracy is .986 without noise: the cubic controls only selected decisions. Original fixed-degree discovery also deteriorates equally across grammars, from five correctly isolated records at $\sigma \leq 10^{-7}$ to 2.75 at 10^{-6} and .50 at 10^{-5}. The probe interval was fixed by a completion scan before decision evaluation. The following comparison changes the estimator, not that interval or the acquisition questions.

Noise-aware derivatives. Pooled Lepski selection over degrees $2, \ldots, 6$ (multiplier three) estimates first/second derivatives from discovery curves and reuses them in decoding. Signatures, empirical gates, dynamics fitting and menus are unchanged. Both arms add 256 repeated legal reads for noise estimation, totaling 8,687 per complete acquisition at $R = 1$, without supplied noise, states or supports. The rule was fixed on original worlds, then tested on 24 new worlds, both paired score signs and three noise levels: 288 attempts including both estimators, not 288 independent worlds.

Table 12: **Derivative selection on 24 fresh worlds.** Interleaved acquisition; two signs per world. Decision success averages the selected two-candidate menus and signs within each world, then averages worlds. Acquisition failures and menus with any nonfinite prediction count as failures, not successful decisions. Costs and empirical gates are identical between arms.

	Acquisitions returned /48		All-attempt decision success	
σ	Fixed degree	Selected degree	Fixed degree	Selected degree
10^{-7}	48	48	.865	.983
10^{-6}	38	48	.425	.911
10^{-5}	28	46	.293	.490

At 10^{-6}, all 48 selected-degree acquisitions isolate five records, every world improves decision success, and maximum menu error is $.00565$. At 10^{-5}, only twelve attempts isolate all five; conditional accuracy is .512 among 46 returns. Table 12 counts non-return and nonfinite predictions as decision failures. Degree selection shifts the useful range but supplies neither a uniform release guarantee nor general planning performance. The exact $\Delta/2$ regret conclusion requires identical adaptive transcripts, including discovery/calibration within the grouped language and the same conditional noise law, plus exactly balanced opposite-optimum menus. This empirical protocol is not asserted to meet that special case.

Signature and preparation components. These component checks of Theorem 5.1 use locked binary fields (9), hidden names/rates, and signed degree-d scores essential in every record. Composite reset constants depend on order. Out-of-buffer words are rejected, not clipped. Discovery uses

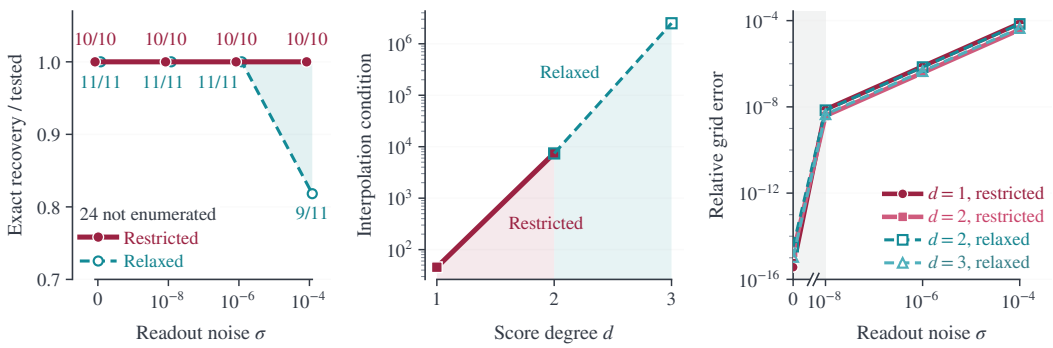


Figure 9: **Reset component checks with scope separated.** Solid/filled marks satisfy the degree and minimum-write-size restrictions; dashed/open marks use relaxed masks. Neither stratum implements the full acquisition theorem. Left: signature recovery fractions and blind-search denominators. Centre: median interpolation condition. Right: relative in-design errors using supplied coordinates, with a linear noise scale below 10^{-8}. All degree-three cells are relaxed.

terminal contrasts, but interpolation receives exact prepared coordinates and legality uses simulator states; native decoding, complete-map learning and deployment are untested.

Complementary masks produce weight-L_0 signatures among $2L_0$ resets. Of 108 runs, 48 satisfy minimum write size two and $d \le m-2$; sixty use relaxed masks, including singleton/empty writes. Every cubic case is relaxed. Blind enumeration is attempted only for $\binom{2L_0}{L_0} \le 40{,}000$: 40 admissible and 44 relaxed attempts recover signatures in 40/40 and 42/44, without false positives. Twenty-four skipped runs are not successes. Both misses are relaxed $\sigma = 10^{-4}$ cases, with unoccupied contrast 1.5×10^{-4} exceeding minimum occupied contrast 4.7×10^{-5}. All 21 attempts at tested $\sigma \le 10^{-6}$ recover exactly. Search sizes for $L_0 = 2, 3, 5$ at $m = 4$ and $L_0 = 3, 7, 11$ at $m = 6$ are $6, 20, 252$ and $20, 3432, 705432$. Intersections retain their intended sets in all 612 admissible and 1,452 relaxed checks. Across 108 runs, preparation Jacobian minimum singular values are .018–.28 and conditions one–fourteen. Supplied-coordinate in-design interpolation errors are 10^{-16}–10^{-14} without noise, about 5×10^{-9} at $\sigma = 10^{-8}$ and 5×10^{-5} at 10^{-4}; median conditions for degrees one/two/three are $45.5, 7397, 2.5 \times 10^6$. Figure 9 summarizes these component results; they are not held-out law errors or full-contract cubic evidence.

Higher-dimensional additive isolation. A component test of Proposition B.1 uses locked linear motion on $\mathbb{R}^n$, $n = 2, 3, 4$, with quadratic record readouts. Exact inclusion–exclusion coefficients select occupied signatures in every generated family; recovered-form errors are 1.7×10^{-14} at $n = 2$, 4.5×10^{-12} at $n = 4$, worst 1.3×10^{-10}. Four of 72 runs stop because isolation exceeds the legality radius at $m = n = 4$. The 2^{L_0} composites and two prefixes amplify noise by about 3×10^3 at $n = 2, 3$ and 6×10^4 at $n = 4$. Costs are 432–2,304 reads without repeats and sixteenfold with repeats. This tests isolation, not projective decoding or complete affine-call recovery, which requires its own retained-output design.

F.5 Five-record Brier service and acquisition

The 1,797-image digits dataset is stratified by odd/even labels into 718 classifier-training, 539 service and 540 evaluator examples. A standardized logistic classifier (regularization one) defines five training-probability quintiles. Each hidden coordinate is a group's probability; the service returns binary Brier loss (Brier, 1950), $h(x) = 539^{-1} \sum_{\ell \in \text{service}} (x_{g(\ell)} - y_\ell)^2$. Labels, groups and sufficient statistics are hidden. Data/loss are real; tools, programs and access restrictions are synthetic, not an industrial deployment or privacy mechanism.

The hidden-order program has root $y_i = (1 - a_i)x_i + a_i/2$ and subsequent rows $y_i = (1 - w_i)x_i + w_i y_{\text{pa}(i)}$, where $w_i = a_i + b_i(x_i - 1/2)$, $a_i \sim U[.1, .2]$, $b_i \sim U[-.04, .04]$. A hidden payload target receives $.75(.5 + .5u) + .125(x_j + x_k)$ from two other records. Raw rates are $U[.85, 1.15]$; ten anonymous pair resets write $1/2$, which the learner estimates. Public priors are native forms, five

records, signature weight six, local/payload degree two and indegree at most one; all assignments, coordinates, rates, constants and coefficients are hidden.

After scalar signature discovery, two isolating contexts per record give cubic readouts of its post-context coordinate. Variable projection (Golub & Pereyra, 1973) eliminates cubic coefficients while fitting rates in $[.8, 1.2]$ over six mode assignments/two starts, using eighty training and forty validation histories. Scalar residuals select the fit. Six decoding arms per record combine each context with no kick or a positive .35 kick from either learned root. A 161-point grid and bounded minimization search $[.015, .985]$. Empirical acceptance requires calibration RMS $\leq \max\{5\sigma/\sqrt{R}, 10^{-8}\}$ and decoder RMS $\leq \max\{10s, 10^{-7}\}$, where $s = \max\{\sigma/\sqrt{R}, \text{calibration RMS}, 10^{-9}\}$. Another minimum more than .04 away within $4s^2$ mean-square loss triggers abstention. These are not confidence certificates: Lemma B.3 assumes genuine context/mode discovery and does not certify this joint calibration, third rate or noisy gates. An erroneous release is empirical acceptance followed by violation of the study's evaluator accuracy threshold; abstention and post-freeze qualification are distinct.

Program identification uses four nodes per axis, 36 validation preparations and joint acyclic parent selection, releasing only a unique surviving DAG. Two hundred additional histories supply decoded features and independent scalar targets. Local/payload minimax fits have coefficient cap two; the cubic score cap is three; payload arguments are $-1, 0, 1$. Full acquisition costs 49,577 independent endpoints within a 50,000 cap, with $R = 1$ and $\sigma_{\text{mean}} = \sigma$. This finite-value procedure implements neither general derivative acquisition nor the shared-pool theorem. At $\sigma = 10^{-7}, 10^{-6}, 10^{-5}$, fixed-cubic acquisitions originally return $18, 19, 19$ of twenty. The extra low-noise failure is numerical: the bounded-coefficient minimax program admits zero coefficients and sufficiently large residual bound. Unchanged constraints with presolve disabled/alternative solvers give 19/20 at every level without changing the original 56 predictions. All 57 completions recover exact parents and payload writes. The persistent abstention is legal-buffer extrapolation: a coordinate near .077 lies outside the calibration range $[.290, .903]$.

Calibration-degree selection. Degrees one–three are selected by RMS on the existing 400 validation scalars, sharing 800 training scalars, mode/rate search, domain and guards. Two fresh noise streams for twenty worlds at $\sigma = 10^{-6}, 10^{-5}$ give eighty pairs/160 attempts; both arms use the numerical fallback. Fixed cubic returns 76/80 and selected degree 80/80, all with correct parents/writes, 49,577 endpoints and zero deployment queries. Calibration needs no additional queries, but downstream query paths can differ.

On 196,608 unseen words per case, worst score errors remain $.00326, .0288$ at the two noise levels; RMSE improves in only 11/38 and 12/38 jointly completed pairs. Selection reduces abstention, not uniformly error, and supplies no confidence certificate.

F.6 Plan tasks and equal-endpoint controls

Development selected two tasks from six Brier/signed-cubic types before evaluation on twenty independent worlds: choosing independently sampled plans, or reordering the same calls with the final program fixed. Each has eight contexts at lengths $8, 16, 32, 64$ and 1,024 candidates per context. The shared candidate bank defines finite-bank regret under original service loss, not regret over all plans. These comparisons and the cache study use the original $\sigma = 10^{-5}$ cohort (nineteen completions/one abstention), not selected-degree models.

Plan-aware predictors receive 40,000 training and 10,000 validation endpoints with disjoint call-name sequences across training/validation/test. Plans start raw, contain a middle payload and end with a program; sampling probabilities are .14 per raw name, .022 per reset and .18 each for program/payload. Arguments are uniform on $[-.8, .8]$, with raw duration at most six. Predictors can query complete deployment plans; compiler calibration has the two-program restriction. Equal endpoints do not equalize primitive calls, priors or computation. Validation selects a 128-unit GRU or two-layer 256-unit MLP, trained on standardized targets with AdamW (at most 120 epochs, step 10^{-3}, batch 1,024, weight decay 10^{-5}). A finite-memory control selects suffix length $4, 8, 16, 32$ and gradient-boosted trees using last occurrences/arguments, cumulative durations and counts, without hidden writes. Early stopping/selection are validation-only; neural training uses RTX 4090 and evaluation a common host.

Each query policy independently spends 50,000 endpoints across both tasks. Uniform screening samples 781 candidates per context without replacement and repeats with the remainder. Screen–remeasure samples 390, retains the best 48 noisy candidates and repeats. Table 13 averages 32 contexts within each of nineteen paired worlds. Finite-memory and screen–remeasure mean regrets (choice/order) are $(.00268, .00395)$ and $(.00467, .00226)$. Learned local laws beat both, the neural predictor and uniform screening in all nineteen completions, not all twenty attempts.

Table 13: Brier regret on 19 paired completions (one abstention in 20); 32 contexts and 1,024 candidates per task, 50,000 endpoints per method for both tasks, whole-map acquisition charged.

Model or query policy	Plan choice	Call order
Learned local laws	$1.16\,10^{-5}$	$4.22\,10^{-6}$
Whole map, same coordinates	$1.95\,10^{-5}$	$1.33\,10^{-5}$
Validation-selected GRU/MLP	$9.19\,10^{-4}$	$1.10\,10^{-3}$
Direct uniform screening	$1.22\,10^{-3}$	$5.57\,10^{-4}$

The whole-map ablation retains learned coordinates, raw/reset laws, payload and score, replacing only the program with a polynomial fitted to the same decoded pairs. Validation chooses degrees one–four and ridge $0, 10^{-12}, 10^{-9}, 10^{-6}, 10^{-3}$, selecting the fewest terms within 1.25 times best error with floor 10^{-10}. Selected degrees are one/two/three in two/sixteen/one fits. Both models pay 49,577 endpoints. Local/whole-map/tied world counts at task-mean tolerance 10^{-12} are $8/4/7$ for choice and $8/3/8$ for order. Together, these comparisons support decision-making with learned coordinates and laws; they do not isolate the added value of local parent recovery.

With 50,000 queries devoted to either task alone, enumeration covers 32×1024 candidates and repeats, achieving all-world mean regrets zero and 8.76×10^{-8}. These are not jointly a 50,000-query solution. Even after consecutive-reset canonicalization, the combined bank contains 64,643 distinct words; this is not an adaptive-query lower bound. An affine control's best-budget source error/span is .0951–.0976 over $(1, 2, 4, 8, 16)(n+1)$ observations, so its transfer failure also reflects underfitting.

F.7 Caching, truth errors and construction costs

The nineteen original high-noise models are frozen. Local-law/whole-map full replay, checkpoint replay and per-record propagation share a scheduler and initialized prefix. Checkpoints store the five-dimensional state after each call and restart at an edited argument, stopping only at exact equality with the old state. No approximate stopping or new pruning is allowed.

For each world and $L = 64, 256, 1024, 4096$, two plans concatenate 64-call blocks with raw duration rescaled to at most five. Twelve sequential argument edits in $[-.8, .8]$ cycle among the latest payload, the payload nearest 32 calls before the end and a uniformly selected earlier payload, without effect/speed selection. Three randomized-order timings use one pinned CPU worker and one BLAS thread, including bookkeeping, edges and score work but excluding evaluator checks. Distinct edits, not timing repeats, give 456 cases per length.

Table 14: Original update times in microseconds, conditional on true score change exceeding 10^{-6}. Take the median of three timings per edit, average within each world and then take the median across contributing worlds. Active counts are $186, 133, 137, 145$ of 456; there are eighteen contributing worlds at $L = 256$ and nineteen otherwise. All other edits remain in errors and construction costs.

	Full replay		Checkpoint		Propagation	
L	Local	Whole	Local	Whole	Local	Whole
64	1308	584	320	162	558	643
256	5159	2230	233	103	477	567
1024	20640	8653	203	102	451	559
4096	83176	34368	231	103	478	611

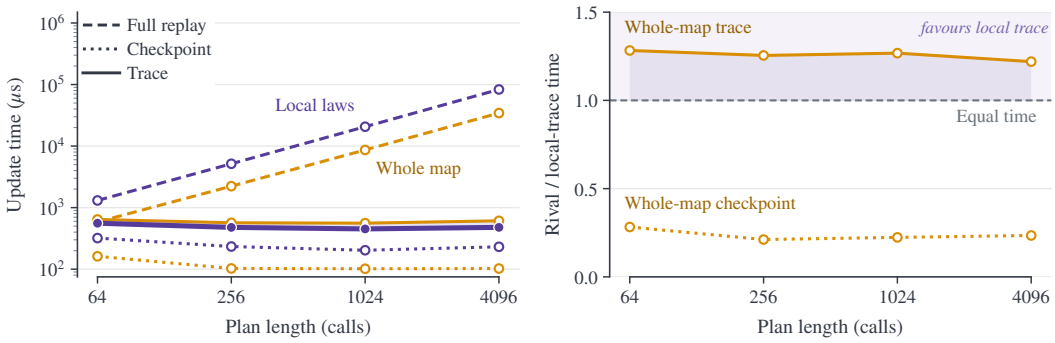


Figure 10: **Maintenance cost on frozen models.** Edits whose true score change exceeds 10^{-6}. Left: full replay, checkpoint replay and trace propagation for local and whole-map models. Right: rival time divided by local-trace time; values above one favour the latter. The local model is 1.22–1.28 times faster within the common trace scheduler, but whole-map checkpoints cost only .21–.28 of local-trace time. Local checkpoint and trace policies share the same fitted model; whole-map policies do not.

On the learned local model, both caching policies preserve accuracy, but checkpointing beats propagation. At $L = 4096$, active-edit costs are .231 ms (checkpoint), .478 ms (propagation) and 83.18 ms (full replay; Table 14). Within the common propagation scheduler, paired local/whole-map speedup is 1.22–1.28, aggregating per-edit ratios within/across worlds (Figure 10). These are maintenance savings after initialization, not a unique sparse-parent advantage.

All caches match their own full models within 2.78×10^{-16}, with no post-freeze queries or changes outside visited nodes. At evaluator threshold 10^{-3}, local/whole-map failures among all 456 edits are $17/23, 3/24, 9/15, 0/3$ as L grows; maxima are $(.00355, .00157), (.00185, .00244), (.00124, .00155), (.000823, .00180)$. Restricting to active, mutually accurate edits preserves checkpointing's advantage. At $L = 4096$, all 152 uniformly historical edits have measured zero effect, whereas 128/152 latest-payload edits exceed 10^{-6}. Maintaining overwritten changes cheaply is not task improvement.

Initial full/checkpoint/trace times at $L = 4096$ are $83.45/86.08/92.17$ ms for local laws and $34.69/36.76/196.09$ ms for whole maps. Checkpoint arrays occupy 163,880 bytes and traces median 15,728 nodes, not whole-process memory. Median noiseless search costs 265 versus 217 seconds under concurrent load, excluding acquisition. No search-runtime gain or physical-service break-even is established.

A separate 120-world study covers 23,040 edits, including 9,980 score-changing cases. Average node revisits are $80.7, 80.7, 80.1$ for local, dense learned-order and original-state memoization, versus 1570 for full evaluation. Original-state control has stronger access. These work counts are not timings and do not isolate a sparse-parent advantage.

F.8 Local-law recovery on restricted preparations

Two noiseless checks test identification without composition expansion. Targeted preparation uses five records, ten equal-value pair resets, a hidden-order quadratic chain and a signed cubic difference of two Vandermonde products plus offset. Complete degrees are $2, 4, 8, 16, 32$. Jets use six positive nodes $j10^{-4}$ and 48-digit arithmetic. All five signatures among 210 candidates and minimum parents are recovered. The 271 histories cost 106,298 terminal means, 97,560 program invocations and 2,278,549 primitive calls. Maximum program error on 64 held-out inputs in $[.2, .8]^5$ is 3.56×10^{-17}; sixteen unseen thirty-call sequences have state/score errors at most $8.50 \times 10^{-17}/1.96 \times 10^{-17}$ without post-freeze queries. A cubic whole-map fit on the same decoded pairs also approximates well (4.24×10^{-6}), despite higher exact degree.

Shared-pool preparation uses 128 random all-reset histories, 54,818 terminal means and 1,443,604 primitive calls to recover all five signatures and minimum parents. Maximum state/score error on sixteen unseen thirty-call sequences is 4.92×10^{-17}, without deployment queries. Pre-call states use

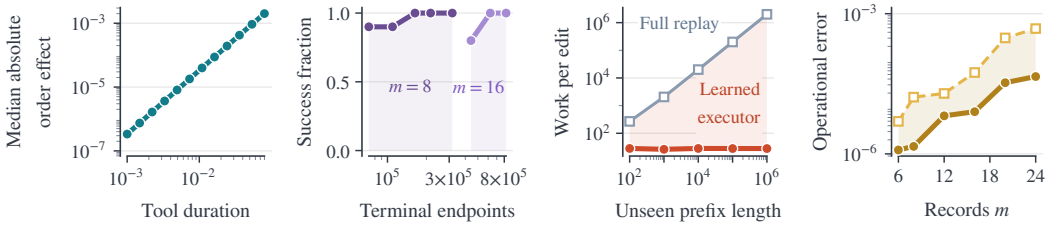


Figure 11: **Native compilation at population endpoints, not finite-noise learning.** Leftmost: absolute order effects; disjoint controls are exactly zero. Second from left: compiler success versus endpoint budget at $m = 8, 16$. Third from left: update work versus unseen prefix length; local work is flat while full replay grows. Rightmost: operational error versus record count, showing the median seed maximum (solid) and worst passing seed (dashed).

scalar-learned parameters, outputs use terminal probes and truth is evaluator-only. Repetition reproduces counts exactly and sequence error to ten digits; unreported float64 least-squares diagnostics vary by at most relative 4×10^{-4}. Parent coupling is .003 for the pool versus .0625 for targeted preparation. These instance-specific checks establish neither a matched sampling advantage nor the conservative pool-size bound, noisy release, payload recovery or argument-edit maintenance.

F.9 Population native compilation to twenty-four records

Population-endpoint acquisition tests Appendix E's sparse-output family under fixed sampling, analysis and stopping criteria. Learners receive anonymous names, durations, a merge token and one scalar per charged trial, not states, targets, overlap graph, local outputs, parents or coefficients. Worlds are statistical units; success intervals are Wilson intervals, while continuous errors/slopes use 10,000-resample bootstrap. This is not finite-noise evidence for Theorem E.1's budget.

Table 15: Fresh-seed scale study at population endpoints. Endpoints and error are medians. The last two columns are log–log slopes of update work against prefix length for the learned executor and for full replay, over prefixes 10^2, 10^3, 10^4, 10^5 and 10^6.

m	seeds	passed	success (95% CI)	endpoints	max op. error	EEI slope	replay slope
6	20	20	1.00 [.84,1]	119,888	$1.21\,10^{-6}$	−.00048	.972
8	20	20	1.00 [.84,1]	215,448	$1.44\,10^{-6}$	.00025	.969
12	16	15	.94 [.72,.99]	462,296	$6.52\,10^{-6}$	.00087	.965
16	30	26	.87 [.70,.95]	826,476	$7.99\,10^{-6}$	−.00055	.960
20	30	29	.97 [.83,.99]	1,468,550	$3.34\,10^{-5}$	−.00015	.956
24	30	27	.90 [.74,.97]	2,275,562	$4.52\,10^{-5}$	.00016	.951

Across 252 cells, 235 pass: scale studies contribute 146/137 attempted/passing cells (Table 15); budget studies at $m = 8, 16$ contribute 70/62; evaluator-side gauge stress over six conditioning stretches contributes 36/36. Order-effect analysis reuses these cells. Every cell has zero hidden-information access and post-freeze queries; passing discrete answers match full recomputation. Median endpoints increase nineteenfold, from 119,888 at $m = 6$ to 2,275,562 at $m = 24$. Over unseen prefixes 10^2–10^6, update-work slopes range from -5.5×10^{-4} to 8.7×10^{-4}, versus replay's .951–.972 (Figure 11). This measures affected-cone maintenance in the sampled family, not an advantage over all caching policies; Section F.7 reports a comparison where checkpointing is faster. Nineteen cells provide 1,216 noncommuting pairs. Order-gap slope against duration is 1.985 (bootstrap interval $[1.9849, 1.9856]$), with perfect sign accuracy. Median learned error is 1.06×10^{-10} versus 1.31×10^{-5} when commutators are suppressed in the same pipeline, isolating the role of order information.

All seventeen failures remain in denominators: twelve exceed the prespecified operational-error gate (worst .145 at $m = 16$ under the smallest budget); five fail numerically during compilation, two at the learned positive chamber and three near a conic projection pole. Failures cluster at small budgets and $m = 16$. The companion finite-noise implementation remains unstable under increased

repetition, specifically in conic parameterization/router compilation; no positive finite-noise result is claimed. Population results establish neither representation optimality nor real-system deployment.

F.10 Numerical checks of quantitative ingredients

At eighty-digit precision, all 120 monomial Grams satisfy (53) for $t = 3, 5$, $r = 2, 3, 4$, halfwidths $1/4$–$1/128$ and four placements. Fitted halfwidth exponents $4.001, 6.004, 8.007$ match $2r$, with minimum bound slack 1.4×10^3. Eigenvalues below 10^{-77} require precision beyond ordinary double arithmetic.

Uniform-cube sampling preserves half the population Gram eigenvalue in 90% of repetitions at 60–170 samples. Measured affine-basis-invariant leverage for $(t, r) = (3, 2), (3, 3), (5, 2)$ is $52, 190, 131$, constant over four halfwidths. This is not the legal random-reset distribution: Proposition C.1 additionally accounts for favorable permutations, ancestor co-motion, density, rounding and simultaneous candidates. These checks do not sharpen (13).

For Theorem D.1's Chebyshev pair at $m = 8, 16, \ldots, 256$, arithmetic checks reproduce equality of moments below m, degree-m difference 2^{1-m} and root separation above m^{-2} (slack 1.633–1.6449). With $g = 4000\, 2^{-m}$, the test $Ng^2/(2\sigma^2) \le 1/8$ yields budget $\sigma^2 4^m/64{,}000{,}000$. These numerical checks supplement, rather than replace, proofs.

Exploratory scope. Additional high-noise trials produced non-returns and incomplete recovered structure; simple fits could match or outperform the learner on single-argument tasks. The reported studies do not establish a cross-service advantage or make decision accuracy a certificate of structural recovery.