PREDICTING AND UNPREDICTING: ANCHOR CHOICE IN JOINT INFERENCE

Anonymous authors
Paper under double-blind review

ABSTRACT

How can dynamics learned by prediction help explain hidden states from later evidence? We reuse a frozen library of locally invertible mechanisms to propose trajectories on unseen compositions, then weight them under a joint posterior when innovations and observations are uncertain. A conditional bound separates inverse execution, model error and uncertainty in the missing inputs. In a predeclared 24-world held-out study with both readings, moving the proposal anchor to the newly observed midpoint lowers effective sample size and full-path gate attainment in expanding chains, but improves both in contracting chains. Both proposals target the same learned posterior. At matched full-path posterior quality, a development-selected direct-inverse policy lowers complete cost against continuous-grid smoothing in the two expanding, two-evidence strata at both tested horizons. The result identifies where direct recovery helps joint inference and where its unresolved uncertainty still governs the answer.

1 LEARNING BY PREDICTING, INFERENCE BY UNPREDICTING

World models learn how an environment changes from experience, allowing agents to anticipate future outcomes and explain unobserved states from available evidence. Dreamer illustrates the value of imagined futures (Hafner et al., 2023); LeWorldModel (LeWM) pairs a compact predictive model with action search (Maes et al., 2026). Both uses depend on learned relationships, but they ask different questions of them.

Consider a water tank: its learned rule predicts the next water level from the current level and a recorded addition. Given a later level and the addition, the same rule can recover an earlier level. A noisy reading or an unrecorded addition leaves several plausible explanations; the inverse supplies a state for each proposed missing value, while the observation model determines their probabilities.

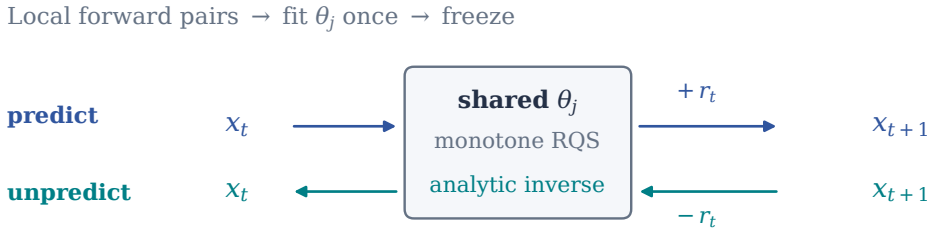

Figure 1: One learned transition answers two local queries. Forward pairs fit the shared spline θ_j; prediction evaluates it, while state recovery evaluates its inverse after a supplied disturbance is removed.

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.

A conventional feedforward dynamics predictor maps a current state and action to a successor state. Given a later outcome, however, explaining an earlier state requires a different computation. The learned dynamics remain useful, but the predictor's forward evaluation does not directly answer this reverse query. The issue is the interface between a learned relationship and the questions we ask of it: which reverse operations does the implementation make available, and what additional inference do they require?

Our starting point is that a relationship learned through prediction can also help explain observations. We describe these two uses as *learning by predicting, inference by unpredicting*; here, *unpredicting* means Bayesian inference over hidden causes through the shared learned model. Structured smoothers can already sample an interior state and complete the past and future (Briers et al., 2010). With dynamics learned only from forward pairs and then frozen, we ask which observed-time latent state should seed a finite-budget proposal for the *complete* trajectory. We evaluate terminal anchoring with terminal evidence alone. After adding a noisy midpoint reading, we compare terminal- and midpoint-anchored proposals in expanding and contracting chains. Both target the same learned posterior, so the comparison isolates how these implemented starting proposals cover it; it also tests whether direct inverse execution yields lower complete cost at matched full-path quality.

Reverse queries also admit optimization or sampling through a predictor, amortized inference, conditional generation, and structured smoothing (Maes et al., 2026; Zhu et al., 2025; Levine, 2018; Rybkin et al., 2021; Bari et al., 2022). We call the first route *soft inversion*. Our route uses direct local inversion to construct joint proposals, followed by Bayesian weighting for unresolved uncertainty.

Within unpredicting, a *local hard inverse* recovers the preceding physical state from a supplied successor and disturbance. Its known forward-trained parameters are reused across a new mechanism sequence. Unknown disturbances and noisy observations still require a joint conditional law. An anchor selects where a path proposal starts; Bayesian weighting still accounts for the unresolved quantities and determines its posterior quality.

We answer with three connected results. A conditional error account identifies where local prediction error, inverse gain and uncertainty in missing inputs enter composed recovery. A fresh root-only confirmation shows that frozen local mechanisms can answer unseen reverse compositions. The decisive joint experiment then adds an unknown innovation at every step and a second observation. Its predeclared terminal-versus-midpoint anchor reversal spans expanding and contracting dynamics; a matched continuous-grid smoother identifies a bounded cost advantage and clear failures. Figures 1 and 2 show the local operation and joint anchor choices; Figure 3 shows the central contrast. These controlled scalar-state queries leave higher-dimensional state recovery and action planning open.

Sections 2–3 define the operation and its error conditions. Section 4 establishes local predictive learning, and Section 5 tests joint inference with changing evidence. Appendix D retains the earlier root-only cost comparison. The supporting context studies in Appendices E–I use a separate inference implementation.

2 Local prediction and direct state inversion

The water-tank example changes which quantities are supplied without changing the underlying rule. We now specify the operation needed to do this with a learned transition. The distinction matters when moving from an invertible representation to actual recovery of an unknown physical state. For the tank, the forward rule is $v = .9x + u$, with recorded addition u; given exact contents v, recovery computes $x = (v - u)/.9$. A noisy final reading instead leaves a range of possible contents, and therefore a range of possible initial amounts. The inverse converts each candidate into an explanation; the prior and measurement model determine its probability. Figure 1 shows the local operation; Figure 2 shows its two uses in the joint query.

Hard inverse operations and probabilistic uncertainty. We use *hard inverse* to mean a structural inverse of a declared augmented map, not a single maximum a posteriori (MAP) explanation. An augmented construction has the form

$$T_\theta(x, \epsilon) = (y, r), \qquad (x, \epsilon) = T_\theta^{-1}(y, r), \tag{1}$$

with matched total dimensions and a bijection on its declared domain. The remainder r denotes the additional variables needed to invert the map; it retains information absent from y. For a specified (y, r), the inverse satisfies this learned relation up to numerical precision without an inner consistency optimizer. For observed y alone, however, r must be sampled or integrated under its appropriate conditional law. A generic base sample is not automatically that conditional; unknown priors, noisy observations and chain boundaries can still require inference. Structural inversion and probabilistic conditioning are complementary operations.

The state-recovery construction studied here uses a particular inverse: with the action or context supplied, a single transition has the form

$$x_{t+1} = f_{\theta,t}(x_t, r_t), \qquad x_t = g_{\theta,t}(x_{t+1}, r_t), \qquad g_{\theta,t}(f_{\theta,t}(x, r), r) = x. \tag{2}$$

For each supplied r, the state map is a bijection on its declared domain, and $g_{\theta,t}$ is its structural inverse. The augmented map is $(x, r) \mapsto (f_{\theta,t}(x, r), r)$: it retains the additional variables r, not a copy of the state x being recovered. Thus a call to $g_{\theta,t}$ requires both its successor state and its remainder. With a noisy observation of the successor, or an unknown remainder, their joint conditional distribution is part of the inference problem. A noise inverse conditional on an already known x does not provide this state-recovery operation.

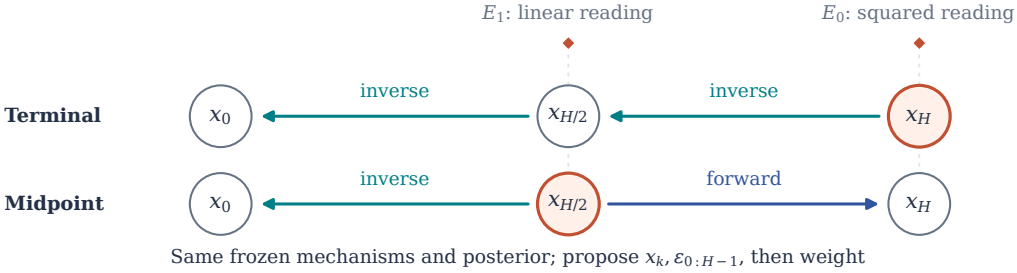

Figure 2: **Two evidence-site anchors for one joint posterior.** The observation sites, frozen mechanism library and unknown innovations are the same. Terminal anchoring executes all local inverses; midpoint anchoring executes inverses to the left and forward maps to the right. The highlighted anchor is a proposed latent state, not the noisy reading. Both proposals require importance correction against the same learned-model target.

For a complete path, the local inverse is one step inside a proposal. With noisy readings at the middle and end of a short chain, terminal anchoring draws a possible final state and the unknown innovation at each transition, then walks backward through the learned inverses. Midpoint anchoring draws a possible middle state and the innovations, walks backward to the root, and uses the forward maps to reach the end. Each path is then weighted using the initial prior, both observation likelihoods, and the exact proposal and change-of-variables factors. Thus the reading nearest an anchor helps propose a state, but only the weighted paths determine how well either proposal covers their common joint posterior.

This construction uses established invertible modeling (Ardizzone et al., 2019). In the scalar experiments, rational-quadratic splines (RQS) provide the state inverse through neural spline flow (NSF) machinery (Durkan et al., 2019). More concretely, each local mechanism j has one fitted parameter set θ_j (spline widths, heights and derivatives, plus an affine scale and shift). With additive disturbance r_t the implementation evaluates

$$x_{t+1} = F_{\theta_{j_t}}(x_t) + r_t, \qquad x_t = F^{-1}_{\theta_{j_t}}(x_{t+1} - r_t).$$

The deterministic experiments set $r_t = 0$; in the stochastic study, $r_t = \sigma\epsilon_t$ is unknown and proposed at inference time. The monotone spline's inverse is evaluated analytically with the same fitted parameters and no separately trained inverse head. A numerical alternative is to hold v, r fixed and solve $f_{\theta,t}(x, r) = v$ for x. Where the state inverse exists, an accurate solve recovers the same answer. The explicit inverse changes how that recovery is executed. Broader numerical inference methods need not perform such a solve at each step.

Learning the operation from local experience. In the scalar study, training pairs provide an input state and a noisy successor for each supplied mechanism. A squared predictive error, scaled by the known observation variance, fits its spline parameters; local validation selects the checkpoint. There is no separately fitted backward predictor for each sequence. After fitting, a requested mechanism sequence selects the modules and their reverse execution order. Thus the shared object is a library of learned local relations, not necessarily one identical transition at every time step. This is learning from paired local transitions; learning the physical inverse from arbitrary incomplete encounters remains a further problem.

From a noisy outcome to a distribution of explanations. In the earlier deterministic terminal query, a Gaussian proposal draws latent terminal states around the noisy reading; local inverses recover candidate roots, which are weighted by the root prior and inverse Jacobian. The Gaussian likelihood cancels against that particular proposal density. The output is a weighted distribution of possible roots, not the inverse of one noisy reading. In the main stochastic query, every innovation is unknown and the terminal reading is squared; its two signed branches require a different proposal and the full target-to-proposal ratio. Section 5 specifies that joint construction.

3 Composition error and remaining uncertainty

A direct inverse specifies how to recover one state. Composing such operations also raises questions about accumulated model error and the unknown quantities passed between steps. We first bound recovery with specified inputs, then account for their joint uncertainty. This analysis distinguishes an executable inverse from a complete inference procedure.

From local predictive error to composed recovery. We first consider recovery given a terminal state and remainders, before treating their uncertainty. Fix a horizon $H \geq 1$, a supplied sequence of mechanisms and known actions/context. Use the state inverse in Equation (2), with constant state dimension d, and abbreviate $g_{\theta,t}$ as g_t. All norms below are Euclidean, with corresponding induced operator norms. The true path follows $x^*_{t+1} = f^*_t(x^*_t, r^*_t)$; the true and learned models use the same state and remainder coordinates. The remainder may represent a process innovation, and is absent in a deterministic transition.

For every participating step, assume common constants $0 < \rho < 1$ and $0 \leq B < \infty$ satisfy

$$\|g_t(v, r) - g_t(v', r)\| \leq \rho\|v - v'\|, \qquad \|g_t(v, r) - g_t(v, r')\| \leq B\|r - r'\|.$$

The first bound says that differences between supplied successor states shrink through one inverse with the remainder fixed; the second limits the effect of getting that remainder wrong. For a scalar affine rule $v = ax + u$ with known u and $a > 1$, the inverse gain is $1/a < 1$. This expanding forward rule illustrates the favorable regime. The leaking tank motivates inverse use but has the opposite sensitivity, so it cannot illustrate this contraction guarantee. The domains must contain the reconstructed states, the true successor states, the learned predictions at true inputs, and the mixed state/remainder arguments used in these comparisons. The inverse identities and bounds must hold there, not only at observed training points. These conditions concern recovery of the state with r fixed; they do not require the augmented map retaining r to contract in all its arguments.

Let approximate execution start at $\hat{x}_H = \hat{h}$ and satisfy $\|\hat{x}_t - g_t(\hat{x}_{t+1}, \hat{r}_t)\| \leq \eta_t$ for $t = 0, \ldots, H - 1$. Define the recovery errors $e_t = \|\hat{x}_t - x^*_t\|$, including $e_H = \|\hat{h} - x^*_H\|$, and the local model discrepancies $\delta_t = \|f_{\theta,t}(x^*_t, r^*_t) - f^*_t(x^*_t, r^*_t)\|$. Thus η_t measures inverse execution error, whereas δ_t measures forward-model error at the true path input.

Proposition 1 (Conditional composition bound)**.** *Under the preceding domain and inverse-gain assumptions, backward execution satisfies*

$$e_0 \leq \rho^H e_H + \sum_{t=0}^{H-1} \rho^t \big(B\|\hat{r}_t - r^*_t\| + \rho\delta_t + \eta_t\big). \tag{3}$$

For $\hat{h} = x^*_H$, $\hat{r}_t = r^*_t$, $\eta_t = 0$ and $\delta_t \leq \delta$ at every step, this gives $e_0 \leq \rho(1 - \rho^H)\delta/(1 - \rho)$.

Proof. Compare the computed $\hat{x}_t$ first with $g_t(\hat{x}_{t+1}, \hat{r}_t)$, then with $g_t(\hat{x}_{t+1}, r_t^*)$, and finally with $g_t(f_{\theta,t}(x_t^*, r_t^*), r_t^*) = x_t^*$. The inverse-gain assumptions and triangle inequality give $e_t \le \rho e_{t+1} + B\|\hat{r}_t - r_t^*\| + \rho\delta_t + \eta_t$. Unrolling yields the bound and its geometric-sum special case. □

For a population version, fix the learned mechanisms and a law of true query paths. Let P_t denote the training input distribution for the mechanism used at step t, and ν_t the corresponding query-path input marginal, both on (x_t^*, r_t^*). Suppose $\nu_t \ll P_t$ with $d\nu_t/dP_t \le M$ almost everywhere, and $\mathbb{E}_{P_t}\|f_{\theta,t}(X,R) - f_t^*(X,R)\|^2 \le \epsilon^2$ for every step. The pathwise domain and inverse-gain assumptions must also hold almost surely under the query law. Then $\mathbb{E}_{\nu_t}\delta_t^2 \le M\epsilon^2$, so Minkowski's inequality gives

$$(\mathbb{E}e_0^2)^{1/2} \le \frac{\rho(1-\rho^H)}{1-\rho}\sqrt{M}\,\epsilon \tag{4}$$

under the same correct-terminal/remainder and exact-execution conditions. The expectation is over query paths, conditional on the fitted mechanisms; no independence among step errors is needed. A horizon-uniform bound requires $\rho < 1$, M and ϵ to hold uniformly as H changes. The model discrepancy is measured against the true transition, not a noisy training target; empirical training loss alone does not certify these conditions. This gives a route from local prediction to untrained composed recovery under the stated gain and coverage conditions. Accurate numerical inversion of the same model inherits the bound. Inverse instability in learned invertible models provides a complementary perspective (Behrmann et al., 2021). The water-tank illustration has inverse gain $1/.9 > 1$: it supports reverse execution but lies outside this contraction regime. Similarly, invertibility of a learned spline alone does not certify the required gain and coverage.

From local recovery to posterior inference. Write $r = (r_0, \dots, r_{H-1})$, $h = x_H$, and $G_\theta(h, r) = x_0$ for the exactly executed inverse chain; $x_j(h, r)$ denotes an intermediate reconstructed state. At differentiable points, the preceding assumptions imply $\|D_h G_\theta\| \le \rho^H$ and $\|D_{r_t} G_\theta\| \le B\rho^t$. These sensitivities hold with other arguments fixed. New evidence can still change the joint conditional law of all remainders.

To express that law as a density, additionally assume that, for each r, the map $x \mapsto f_{\theta,t}(x, r)$ is a continuously differentiable bijection with a continuously differentiable inverse on the relevant domains, and that $g_t(v, r)$ is jointly continuously differentiable in its arguments. Let the initial state have density p_0 and innovations have densities q_t, independent across steps and of the initial state. Let $E = \{(j, y_j) : j \in \mathcal{O}\}$ be observations conditionally independent given the trajectory, with likelihoods $L_j(y_j \mid x_j)$. Parameterizing by (h, r), called the terminal chart, gives

$$\mu_\theta(h, r \mid E) \propto p_0(G_\theta(h, r)) \prod_t q_t(r_t) \prod_j L_j(y_j \mid x_j(h, r)) \prod_t |\det D_v g_t(x_{t+1}, r_t)|. \tag{5}$$

Products over observations run over $j \in \mathcal{O}$ and transition products over $t = 0, \dots, H-1$. The expression is restricted to the transformed support and assumes a positive finite normalizer. Its Jacobian is block triangular because r is retained; the state chain rule supplies the product of determinants. This density change requires invertibility and smoothness, but not contraction. Correlated or state-dependent innovations require their joint generative density in place of $\prod_t q_t(r_t)$.

Here the initial state is unknown and has prior p_0. A supplied exact initial state would constrain (h, r) and require a different conditional treatment. Likewise, a noisy terminal measurement y_H contributes $L_H(y_H \mid h)$; it does not set $h = y_H$. Independent base draws for r need not follow the posterior in Equation (5).

For a posterior error bound, couple a draw of the full true path and remainders conditional on E with a draw $(\hat{h}, \hat{r}) \sim q$ used by the learned inverse. Call this one joint coupling γ. Assume the same remainder semantics in both models, the preceding inverse-gain and domain conditions on the coupled arguments, exact learned inverse execution, and finite second moments for the states and discrepancies below. Define

$$a = (\mathbb{E}_\gamma\|\hat{h} - x_H^*\|^2)^{1/2}, \quad b_t = (\mathbb{E}_\gamma\|\hat{r}_t - r_t^*\|^2)^{1/2}, \quad d_t = (\mathbb{E}_\gamma\delta_t^2)^{1/2}.$$

Let $P^*_{x_0|E}$ be the true initial-state posterior and $(G_\theta)_\# q$ the distribution of $G_\theta(\hat{h}, \hat{r})$. The quadratic Wasserstein distance W_2 is the infimum of root mean squared state distance over joint couplings

with those marginals. Applying Equation (3) under γ and then Minkowski's inequality gives

$$W_2(P^*_{x_0|E}, (G_\theta)_\# q) \le \rho^H a + B \sum_t \rho^t b_t + \sum_t \rho^{t+1} d_t. \tag{6}$$

The resulting state pair is one admissible coupling, so its root mean squared distance bounds W_2 from above. If the true dynamics are invertible, coupling terminal states and remainders already determines the true path. Otherwise, the full-path coupling above retains that missing uncertainty. Independently optimizing marginal couplings would not preserve the needed within-path dependence. Latent coordinates with different meanings require an explicit alignment before their discrepancies can be compared. The bound shows how posterior error propagates through recovery. Its size depends on how accurately the sampler represents the joint law of terminal states and remainders. These inequalities control recovery error under a specified coupling, not the variance of importance weights. They cannot rank terminal and midpoint effective sample sizes: proposal coverage, observation likelihoods, prior and Jacobian weights also enter that comparison. Contracting forward dynamics can amplify inverse errors, and even stable inverse maps need not yield a log-concave posterior. Sections 4–5 test composed reverse use, remaining inference and query cost; Appendix L reports earlier execution and coordinate diagnostics.

4 Locally learned mechanisms on unseen reverse compositions

Before joint inference, we check that local forward learning supplies a usable inverse library. Four monotone mechanisms per world are fitted from 1,024 noisy local forward pairs each; no composed inverse query is used in training. Ten fresh mechanism worlds supply unseen length-4/8 sequences with a known prior and terminal-noise law. The only unknown trajectory variable is the initial scalar state. The posterior score is the integrated squared difference between candidate and true root CDFs (excess CRPS/energy, in state units).

With one-dimensional reference inference for every learner, the eight-bin RQS attains mean true-posterior excess energy .002191, versus .005073 for a validation-selected monotone PCHIP and .103263 for the tested MLP. The RQS direct-inverse sampler lies 1.63×10^{-5} from the posterior of that *same learned RQS*, with ESS fraction .933. RQS also has the lowest local forward MSE, so this comparison establishes reuse of learned local relations, not an accuracy effect caused by invertibility. Appendix B gives the learners, uncertainty intervals, checks and raw amortized control.

5 Where to anchor joint unpredicting

The root-only query leaves no internal process uncertainty. We now ask whether direct state inversion remains useful when every step has an unknown innovation and evidence arrives at more than one time. Each frozen locally learned transition has the form $x_{t+1} = F_{\theta,j_t}(x_t) + \sigma\epsilon_t$, with $\epsilon_t \sim \mathcal{N}(0,1)$ and $\sigma = .02$; the sensor standard deviation is .05. The root prior p_0 is an equally weighted Gaussian mixture centered at $\pm.8$ with standard deviation .2, truncated to $[-2,2]$. The two readings are $y_H = x_H^2 + \varepsilon_H$ and $y_{H/2} = x_{H/2} + \varepsilon_{H/2}$. For an anchor $u = x_k$, the proposed $(u, \epsilon_{0:H-1})$ determines a path $x^{(k)}_{0:H}$: execute fitted inverses before k and fitted forward maps after k. Write ϕ for the standard-normal density and $J_k = \prod_{t<k} |F'_{\theta,j_t}(x_t^{(k)})|^{-1}$. The same learned-model joint target, its implemented proposal and the unnormalized importance weight are

$$\pi_{\theta,k}(u, \epsilon \mid E) \propto p_0(x_0^{(k)}) \prod_t \phi(\epsilon_t) \prod_{j \in \mathcal{O}} L_j(y_j \mid x_j^{(k)}) J_k,$$

$$Q_k(u, \epsilon \mid E) = a_k(u \mid E) \prod_t \phi(\epsilon_t), \qquad w_k \propto \frac{p_0(x_0^{(k)}) J_k \prod_{j \in \mathcal{O}} L_j(y_j \mid x_j^{(k)})}{a_k(u \mid E)}. \tag{7}$$

Here $L_H = \mathcal{N}(y_H; x_H^2, .05^2)$ and $L_{H/2} = \mathcal{N}(y_{H/2}; x_{H/2}, .05^2)$ when the second reading is present. Terminal anchoring uses $k = H$, $m = \sqrt{\max(y_H, 0)}$, and a proposal $a_H = .45\mathcal{N}(-m, s^2) + .45\mathcal{N}(m, s^2) + .10\mathcal{N}(0, 1.5^2)$, where $s = .10$ for $m < .2$ and $s = \max(.01, .05/(2m))$ otherwise. After the midpoint reading, $k = H/2$ instead uses $a_{H/2} =$

$.90\mathcal{N}(y_{H/2}, .05^2) + .10\mathcal{N}(0, 1.5^2)$. The Gaussian terms denote densities in u; the broad components preserve support. Each anchor density uses only its site's reading; the other reading enters the likelihood product in the weight. For E_0 alone, only terminal anchoring is evaluated. Both arms target the same posterior, but these anchor-specific proposal laws and inverse lengths differ. Merely changing coordinates while transforming the proposal with them would leave the importance weights unchanged. Setting $k = 0$ and $a_0 = p_0$ gives $J_0 = 1$ and the forward-prior proposal, whose weight is the product of the observation likelihoods.

Held-out comparison. The predeclared confirmation uses 24 fresh worlds, twelve with expanding and twelve with contracting forward dynamics. Each world has four teachers $F_j^*(x) = a_j x + .05\tanh(x) + b_j$, with independently drawn $b_j \sim \mathcal{U}[-.02, .02]$. Expanding draws $a_j \sim \mathcal{U}[1.045, 1.055]$, giving $F_j^{*\prime}(x) \in [1.045, 1.105]$ for all x; contracting draws $a_j \sim \mathcal{U}[.895, .905]$, giving $F_j^{*\prime}(x) \in [.895, .955]$. Each spline is fitted to 1,024 noisy forward pairs per mechanism with input $x \sim \mathcal{U}[-16, 16]$ and output-noise standard deviation .05; 256 independent pairs are used for validation. These teacher bounds are not global derivative guarantees for the learned splines. All generated held-out query paths stay within the training input interval (maximum $|x_t| = 8.74$); this does not certify the support of the full posterior. Each world then supplies unseen mechanism compositions of lengths $H = 16, 32$, a noisy squared terminal reading E_0, and an added noisy linear midpoint reading E_1. There is one generated query per world and horizon and three sampler repeats; the independent unit for intervals is the world.

All inferencers receive the same frozen learned transitions, prior, graph and likelihoods. The continuous-grid forward-filtering/ backward-sampling (FFBS) comparator and direct-inverse arms output full trajectory measures. True mechanisms and high-resolution learned/true posterior references are evaluator-only. The preregistered full-path quality gate combines multiscale RFF $\text{MMD}^2 \leq 2 \times 10^{-5}$, maximum root/midpoint/terminal CDF KS $\leq .03$, and root-sign mass error $\leq .05$. The MMD standardizes each time coordinate by its learned prior-predictive scale; Appendix A gives its kernel and feature count. All 96 learned/true reference cells and 288 analytic/numeric inverse checks pass their specified integrity gates.

At $N = 8192$ for $H = 16$ and $N = 16384$ for $H = 32$, midpoint/terminal ESS ratios are .232 and .137 in expanding dynamics, then 6.06 and 17.8 in contracting dynamics. The four paired world-bootstrap intervals exclude one, and the full-path quality-gate ordering reverses in the same direction (Figure 3). Median learned-posterior-path log inverse gain is negative in expansion and positive in contraction. This is consistent with the stability account in Section 3; anchor-matched proposal densities, evidence geometry, prior and Jacobian weights also change, so inverse gain is not isolated as the sole cause. In particular, the newly observed midpoint and the shorter reverse segment do not by themselves select the better finite-budget proposal in expansion.

Matched joint quality and complete cost. A policy chosen on separate development worlds freezes the anchor, sample budget and grid resolution for each regime, horizon and evidence state. The regime is supplied by the controlled experiment, not inferred by a learned selector. Figure 3 compares anchors at equal N; Table 1 evaluates this separately selected policy. A positive cost conclusion requires each method separately to pass the joint gate in at least 80% of the 36 held-out cells and a positive world-bootstrap interval for log(grid/direct complete CPU cost) on jointly passed cells. Complete cost charges setup, proposal preparation, execution, weighting and sampling. The ratio does not include failure retries. Table 1 reports all eight strata; cost ratios are descriptive whenever the reliability condition fails.

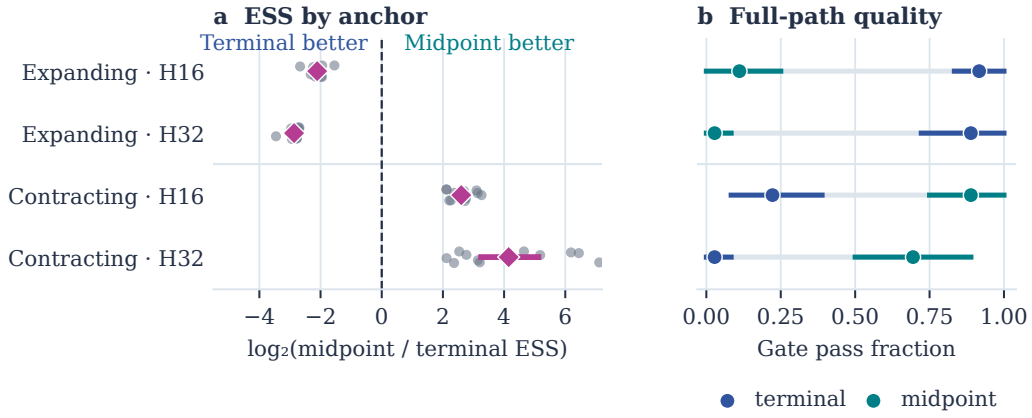

Figure 3: **The preferred inverse anchor reverses with dynamics.** At equal N and with both observations, points in (a) are world means of $\log_2(\mathrm{ESS}_{\mathrm{midpoint}}/\mathrm{ESS}_{\mathrm{terminal}})$; diamonds and bars are means and 95% world-bootstrap intervals. Panel (b) gives full-path gate pass fractions over 36 world-repeat cells, with world-bootstrap intervals. Both panels use $N = 8192$ at $H = 16$ and $N = 16384$ at $H = 32$. Terminal anchoring leads in expansion; midpoint anchoring leads in contraction. Both target the same learned posterior.

Table 1: Frozen-policy full-path comparison on 24 fresh worlds. Pass counts are out of 36 world-repeat cells; "both" is the joint pass count. Ratios exponentiate world-mean log(grid/direct complete cost) on jointly passed cells. Only bold rows satisfy the predeclared matched-quality cost criterion.

Dynamics	Evidence	H	Direct	Grid	Both	Cost ratio [95% CI]
Expanding	E_0	16	23	21	15	5.98 [5.92,6.02]
Expanding	$\boldsymbol{E_0 + E_1}$	**16**	**33**	**30**	**28**	**1.53 [1.52,1.55]**
Expanding	E_0	32	28	19	15	0.84 [0.83,0.85]
Expanding	$\boldsymbol{E_0 + E_1}$	**32**	**32**	**33**	**29**	**1.57 [1.56,1.58]**
Contracting	E_0	16	27	28	20	2.02 [1.47,2.62]
Contracting	$E_0 + E_1$	16	32	20	19	0.54 [0.48,0.61]
Contracting	E_0	32	10	29	6	3.63 [2.19,4.70]
Contracting	$E_0 + E_1$	32	25	26	18	1.73 [1.44,1.99]

The two expanding, two-evidence strata meet this criterion, with grid/direct cost ratios 1.53 and 1.57. They establish a bounded complete-cost benefit for this development-selected direct-inverse policy. Contracting $H = 32, E_0$ exposes a different boundary: the frozen terminal inverse passes only 10/36 cells while forward-prior importance passes 35/36. The other policy strata do not establish a reliable cost advantage. The full budget curves in Appendix A distinguish frozen selected settings from post-hoc descriptive points.

Execution and model fidelity. An identical-proposal analytic-versus-numerical inverse control passes all 288 path/weight parity checks and has median numerical/analytic *operator* time ratio 18.3, not an end-to-end gain. Mean learned-versus-true root-posterior CDF KS is .354/.168 in expanding/contracting worlds, distinct from learned-posterior inference error. Retraction of E_1 from a cache collapsed under E_1 fails for direct inversion and FFBS. The study has scalar state, one teacher/noise family, two evidence sites and one CPU evaluation setup.

6 Relation to Existing Work

Predictive world models support imagined planning, latent trajectory optimization and image-goal action search (Hafner et al., 2019; Hansen et al., 2024; Hafner et al., 2023; Zhou et al., 2024; Assran et al., 2025; Maes et al., 2026). Using a model trained for next-state prediction to estimate actions

at test time raises an interface question (Parthasarathy et al., 2025). Reverse-query routes include trajectory diffusion, multi-conditional world models and reverse use of video generators (Janner et al., 2022; Zhu et al., 2025; Alzayer et al., 2026); amortized planning and posterior inference reduce repeated query cost (Nguyen et al., 2026; Radev et al., 2020). Consistency and one-step policies offer fast action generation (Prasad et al., 2024; Wang et al., 2024), while WorldTest places hidden-state and dynamics-change queries alongside prediction (Warrier et al., 2025).

Our local inverse uses established invertible-network and neural-spline machinery (Ardizzone et al., 2019; Durkan et al., 2019); ACFlow and Simformer also support multiple conditional uses of a learned relationship (Li et al., 2020; Gloeckler et al., 2024). Control-as-inference, factor graphs, trajectory optimization, parallel smoothing and reusable search proposals retain temporal structure (Levine, 2018; Bari et al., 2022; Rybkin et al., 2021; Särkkä & García-Fernández, 2019; Pinneri et al., 2020). Two-filter smoothers can sample a midpoint state and complete both sides (Briers et al., 2010); centered and noncentered parameterizations show how inference cost can depend on model and data (Papaspiliopoulos et al., 2007). Here the frozen forward-learned inverse, innovation-prior proposal and exact importance weights make observation-site anchor choice a measurable finite-budget question. The held-out anchor reversal and matched-quality complete-cost boundary answer it for the tested chains.

7 Discussion

With the same learned joint posterior target, anchoring at the newly observed midpoint lowers ESS and full-path gate attainment in tested expanding chains, but raises both in contracting chains. The complete-cost benefit is conditional; terminal inversion in contracting $H = 32$ chains even fails where forward-prior importance succeeds. Inverse gain tracks the reversal but does not isolate its cause from proposal geometry, weighting and reverse execution length.

Direct execution, learned-model fidelity and posterior inference remain separate obligations. Matched numerical inversion reaches the same model answer at greater local operator cost, while the learned-versus-true gap remains material. The quality result covers scalar states, one unknown innovation per step and two evidence sites; higher state dimension, richer revision and task-scale planning remain untested. The older root-only cost reversal and exact affine prior-metric control rule out a universal speed claim.

Separate context studies (Appendices E–I) find that outcome-only experience improves later reverse and relational judgments in informative settings. A refreshed context mean can preserve the current single-encounter law yet lose on later relational risk. Appendix H gives a bounded preservation condition for declared future queries; learning such a state across unrestricted experience remains open.

The central joint-inference result cautions against choosing a proposal anchor from evidence location or reverse-chain length alone. The tested anchor-specific proposals change their ranking across the tested dynamics under this evidence geometry; their practical value must be judged at matched path quality and complete cost. The learned relationship remains shared across forward prediction and reverse queries; uncertainty over unexplained inputs remains explicit.

Reproducibility Statement

The anonymous supplement supplies the 24-world joint study's saved metrics, checks, frozen policy and primary/secondary analysis and figure-replay code. It excludes trajectories, checkpoints, campaign code and earlier-study artifacts; raw-sample gates and full experiments cannot be reproduced from it. Appendix S reports design, timing and limits.

AI use statement

AI tools contributed to conceptual development, literature review, mathematical derivation, experiment design and implementation, synthetic experiments, numerical analysis, figures, supplementary artifact preparation and writing.

REFERENCES

Masaki Adachi, Satoshi Hayakawa, Harald Oberhauser, Martin Jørgensen, and Michael A. Osborne. Fast Bayesian inference with batch Bayesian quadrature via kernel recombination. In *Advances in Neural Information Processing Systems*, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/697200c9d1710c2799720b660abd11bb-Paper-Conference.pdf.

Amr Alexandari, Anshul Kundaje, and Avanti Shrikumar. Maximum likelihood with bias-corrected calibration is hard-to-beat at label shift adaptation. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 222–232, 2020. URL https://proceedings.mlr.press/v119/alexandari20a.html.

Hadi Alzayer et al. Masked visual actions for unified world modeling. *arXiv preprint arXiv:2607.19343*, 2026. URL https://arxiv.org/abs/2607.19343.

Lynton Ardizzone, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W. Pellegrini, Ralf S. Klessen, Lena Maier-Hein, Carsten Rother, and Ullrich Köthe. Analyzing inverse problems with invertible neural networks. In *International Conference on Learning Representations*, 2019. URL https://arxiv.org/abs/1808.04730.

Mido Assran, Adrien Bardes, David Fan, et al. V-JEPA 2: Self-supervised video models enable understanding, prediction and planning. *arXiv preprint arXiv:2506.09985*, 2025. URL https://arxiv.org/abs/2506.09985.

Salman Bari, Xiagong Wang, Ahmad Schoha Haidari, and Dirk Wollherr. Factor graph-based planning as inference for autonomous vehicle racing. *arXiv preprint arXiv:2203.03224*, 2022. URL https://arxiv.org/abs/2203.03224.

Christian Bayer and Josef Teichmann. The proof of Tchakaloff's theorem. *arXiv preprint math/0502473*, 2005. URL https://arxiv.org/abs/math/0502473.

Jens Behrmann, Paul Vicol, Kuan-Chieh Wang, Roger Grosse, and Jörn-Henrik Jacobsen. Understanding and mitigating exploding inverses in invertible neural networks. In *International Conference on Artificial Intelligence and Statistics*, 2021. URL https://proceedings.mlr.press/v130/behrmann21a.html.

Mark Briers, Arnaud Doucet, and Simon Maskell. Smoothing algorithms for state–space models. *Annals of the Institute of Statistical Mathematics*, 62(1):61–89, 2010. doi: 10.1007/s10463-009-0236-2. URL https://www.stats.ox.ac.uk/~doucet/briers_doucet_maskell_smoothingstatespacemodels.pdf.

Adnan Darwiche. A differential approach to inference in Bayesian networks. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, pp. 123–132, 2000. URL https://arxiv.org/abs/1301.3847.

Conor Durkan, Artur Bekasov, Iain Murray, and George Papamakarios. Neural spline flows. In *Advances in Neural Information Processing Systems*, 2019. URL https://arxiv.org/abs/1906.04032.

Bradley Efron. Two modeling strategies for empirical Bayes estimation. *Statistical Science*, 29(2): 285–301, 2014. URL https://arxiv.org/abs/1409.2677.

Fabian Falck, Ziyu Wang, and Chris Holmes. Is in-context learning in large language models Bayesian? a martingale perspective. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12784–12805, 2024. URL https://proceedings.mlr.press/v235/falck24a.html.

Saurabh Garg, Yifan Wu, Sivaraman Balakrishnan, and Zachary C. Lipton. A unified view of label shift estimation. In *Advances in Neural Information Processing Systems*, 2020. URL https://arxiv.org/abs/2003.07554.

Tao Ge, Jing Hu, Lei Wang, Xun Wang, Si-Qing Chen, and Furu Wei. In-context autoencoder for context compression in a large language model. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/0b276510ec2d3f6613a8b60c41ff0438-Paper-Conference.pdf.

Manuel Gloeckler, Michael Deistler, Christian Dietrich Weilbach, Frank Wood, and Jakob H. Macke. All-in-one simulation-based inference. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 15735–15766, 2024. URL https://proceedings.mlr.press/v235/gloeckler24a.html.

Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 2555–2565, 2019. URL https://proceedings.mlr.press/v97/hafner19a.html.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023. URL https://arxiv.org/abs/2301.04104.

Nicklas Hansen, Hao Su, and Xiaolong Wang. TD-MPC2: Scalable, robust world models for continuous control. In *International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=Oxh5CstDJU.

Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022. URL https://openaccess.thecvf.com/content/CVPR2022/html/He_Masked_Autoencoders_Are_Scalable_Vision_Learners_CVPR_2022_paper.html.

Michael Janner, Yilun Du, Joshua B. Tenenbaum, and Sergey Levine. Planning with diffusion for flexible behavior synthesis. *arXiv preprint arXiv:2205.09991*, 2022. URL https://arxiv.org/abs/2205.09991.

Jiun Jeong, Byeongwoo An, Gi-Soo Kim, and Kyubo Shin. Calibration-aware online adaptation under label shift. In *Proceedings of the Forty-Second Conference on Uncertainty in Artificial Intelligence*, volume 337 of *Proceedings of Machine Learning Research*, pp. 2480–2516, 2026. URL https://proceedings.mlr.press/v337/jeong26a.html.

Diederik P. Kingma and Prafulla Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. In *Advances in Neural Information Processing Systems*, 2018. URL https://arxiv.org/abs/1807.03039.

Brenden M. Lake, Ruslan Salakhutdinov, and Joshua B. Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332–1338, 2015. doi: 10.1126/science.aab3050.

Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv preprint arXiv:1805.00909*, 2018. URL https://arxiv.org/abs/1805.00909.

Yang Li, Shoaib Akbar, and Junier B. Oliva. ACFlow: Flow models for arbitrary conditional likelihoods. In *Proceedings of the 37th International Conference on Machine Learning*, 2020. URL https://proceedings.mlr.press/v119/li20a.html.

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/file/1e4d36177d71bbb3558e43af9577d70e-Paper.pdf.

Lucas Maes, Quentin Le Lidec, Damien Scieur, Yann LeCun, and Randall Balestriero. Leworldmodel: Stable end-to-end joint-embedding predictive architecture from pixels. *arXiv preprint arXiv:2603.19312*, 2026. URL https://arxiv.org/abs/2603.19312.

Samuel Müller, Arik Reuter, Noah Hollmann, David Rügamer, and Frank Hutter. Position: The future of Bayesian prediction is prior-fitted. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 81861–81875, 2025. URL https://proceedings.mlr.press/v267/muller25d.html.

National Institute of Standards and Technology. NIST Digital Library of Mathematical Functions, section 3.5(v): Gauss quadrature, 2026. URL https://dlmf.nist.gov/3.5#v. Accessed 14 September 2026.

Hoang Nguyen, Xiaohao Xu, and Xiaonan Huang. Latent geometry beyond search: Amortizing planning in world models. *arXiv preprint arXiv:2605.08732*, 2026. URL https://arxiv.org/abs/2605.08732.

Omiros Papaspiliopoulos, Gareth O. Roberts, and Martin Sköld. A general framework for the parametrization of hierarchical models. *Statistical Science*, 22(1):59–73, 2007. doi: 10.1214/088342307000000014.

Arjun Parthasarathy, Nimit Kalra, Rohun Agrawal, Yann LeCun, Oumayma Bounou, Pavel Izmailov, and Micah Goldblum. Closing the train–test gap in world models for gradient-based planning. *arXiv preprint arXiv:2512.09929*, 2025. URL https://arxiv.org/abs/2512.09929.

Cristina Pinneri et al. Sample-efficient cross-entropy method for real-time planning. *arXiv preprint arXiv:2008.06389*, 2020. URL https://arxiv.org/abs/2008.06389.

Aaditya Prasad, Kevin Lin, Jimmy Wu, Linqi Zhou, and Jeannette Bohg. Consistency policy: Accelerated visuomotor policies via consistency distillation. In *Robotics: Science and Systems*, 2024. URL https://arxiv.org/abs/2405.07503.

Stefan T. Radev, Ulf K. Mertens, Andreas Voss, Lynton Ardizzone, and Ullrich Köthe. BayesFlow: Learning complex stochastic models with invertible neural networks. *arXiv preprint arXiv:2003.06281*, 2020. URL https://arxiv.org/abs/2003.06281.

Oleh Rybkin, Chuning Zhu, Anusha Nagabandi, Kostas Daniilidis, Igor Mordatch, and Sergey Levine. Model-based reinforcement learning via latent-space collocation. In *International Conference on Machine Learning*, 2021.

Marco Saerens, Patrice Latinne, and Christine Decaestecker. Adjusting the outputs of a classifier to new a priori probabilities: A simple procedure. *Neural Computation*, 14(1):21–41, 2002. doi: 10.1162/089976602753284446.

Simo Särkkä and Ángel F. García-Fernández. Temporal parallelization of bayesian smoothers. *arXiv preprint arXiv:1905.13002*, 2019. URL https://arxiv.org/abs/1905.13002.

Maxime Vandegar, Michael Kagan, Antoine Wehenkel, and Gilles Louppe. Neural empirical Bayes: Source distribution estimation and its applications to simulation-based inference. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 2107–2115, 2021.

Zhendong Wang et al. One-step diffusion policy: Fast visuomotor policies via diffusion distillation. *arXiv preprint arXiv:2410.21257*, 2024. URL https://arxiv.org/abs/2410.21257.

Archana Warrier, Dat Nyugen, Michelangelo Naim, Moksh Jain, Yichao Liang, Karen Schroeder, Cambridge Yang, Joshua B. Tenenbaum, Sebastian Vollmer, Kevin Ellis, and Zenna Tavares. Benchmarking world-model learning with environment-level queries. *arXiv preprint arXiv:2510.19788*, 2025. URL https://arxiv.org/abs/2510.19788.

Gaoyue Zhou, Hengkai Pan, Yann LeCun, and Lerrel Pinto. DINO-WM: World models on pre-trained visual features enable zero-shot planning. *arXiv preprint arXiv:2411.04983*, 2024. URL https://arxiv.org/abs/2411.04983.

Chuning Zhu, Raymond Yu, Siyuan Feng, Benjamin Burchfiel, Paarth Shah, and Abhishek Gupta. Unified world models: Coupling video and action diffusion for pretraining on large robotic datasets. *arXiv preprint arXiv:2504.02792*, 2025. URL https://arxiv.org/abs/2504.02792.

A HELD-OUT JOINT EXPERIMENT: FULL DIAGNOSTICS

The comparisons, quality thresholds and selected policy were fixed using separate development worlds before evaluating the 24 held-out worlds. The selected configurations, complete costs, world-level uncertainty and reference checks are reported below. The direct-inverse, grid/FFBS and forward-prior methods see the same learned transition law and observations. At fixed N, changing the anchor changes proposal geometry and inverse length while preserving the posterior target; the comparison is not an analytic-versus-numerical inverse ablation.

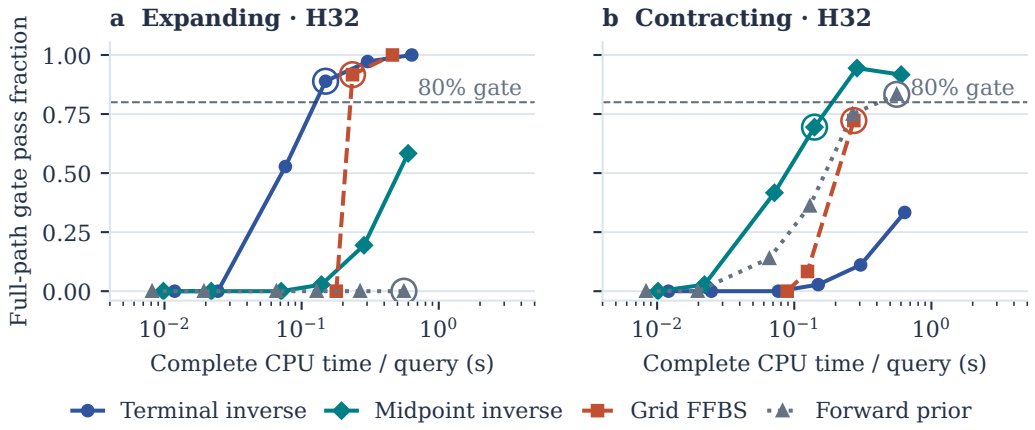

Figure 4: H32 with both readings ($E_0 + E_1$): full-path gate pass fraction versus complete CPU time per query. Hollow rings mark the development-selected settings; all other observed portfolio points and the connecting lines are post-hoc descriptive. The grid curves keep their selected resolutions fixed (16,385 cells in expansion, 8,193 in contraction). The attractive contracting midpoint $N = 32768$ point was not the frozen selected policy and does not establish a confirmed advantage.

The selected settings are below. "T" and "M" denote terminal and midpoint anchors; each grid entry gives grid cells and output paths. The regime is a known experimental stratum. The final column gives separate median complete costs among cells where both methods pass, in seconds; Table 1 uses a world-level mean log cost ratio, so it need not equal the ratio of these two medians.

Table 2: Development-selected policy and observed complete costs. All settings were fixed before held-out evaluation. N is the number of proposed or sampled full paths. Median costs are conditional on joint quality-gate success.

Dynamics	Evidence	H	Direct anchor, N	Grid cells, N	Median D/G cost (s)
Expanding	E_0	16	T, 8192	16385, 8192	.0387 / .2338
Expanding	$E_0 + E_1$	16	T, 8192	8193, 2048	.0390 / .0598
Expanding	E_0	32	T, 32768	8193, 8192	.3035 / .2556
Expanding	$E_0 + E_1$	32	T, 16384	16385, 2048	.1497 / .2347
Contracting	E_0	16	T, 65536	32769, 16384	.3182 / .8349
Contracting	$E_0 + E_1$	16	M, 8192	2049, 2048	.0362 / .0220
Contracting	E_0	32	T, 65536	32769, 32768	1.4305 / 3.4537
Contracting	$E_0 + E_1$	32	M, 16384	8193, 8192	.1439 / .2735

For the full-path MMD^2, each coordinate of a path is divided by its world-specific learned prior-predictive standard deviation, estimated from 8,192 paths without query evidence and held fixed across methods. With $d = H + 1$ standardized coordinates, the equal-weight Gaussian kernel mixture has bandwidths $\sqrt{d/4}$, $\sqrt{d}$ and $2\sqrt{d}$. Each of two independent seeded random-Fourier maps uses 256 features per bandwidth; their two squared mean-embedding distances are averaged. Reference paths are root-sign poststratified to match the grid reference's branch mass. This finite-feature

score is paired with the three marginal CDF and root-sign checks in the gate. The same threshold is applied within each horizon/regime stratum after this normalization; no raw unnormalized path distances are compared across horizons.

The quality gate is a finite-feature full-path diagnostic, paired with root/midpoint/terminal marginal CDF and root-sign checks. Reference-bank MMD floors, grid refinement, truncated-window mass, and inverse parity passed the frozen tolerances. The analytical operator-time ratio compares the same proposal and samples and excludes the remaining joint inference work. Retraction after a cache has collapsed under E_1 passes the joint gate in 0/36 cells in every stratum for both hard inverse and FFBS; this is a common importance-support limitation. The learned-versus-true root posterior gaps remain separate from the learned-law solver comparisons.

B EARLIER ROOT-ONLY COMPOSITION CONFIRMATION

The theory identifies sufficient conditions for recovery. We next ask whether mechanisms learned from noisy local pairs can answer complete reverse queries absent from training. Comparing mechanism classes helps locate how predictive fit affects those answers. This confirmation was frozen after development on four mechanism worlds and uses ten fresh worlds as independent units.

What the posterior score measures. For a candidate initial-state law Q and reference law P, define

$$\mathcal{R}(Q, P) = \int_{\mathbb{R}} (F_Q(u) - F_P(u))^2 \, du, \tag{8}$$

where F_Q and F_P are their cumulative distribution functions. For any water-amount threshold, each CDF is the probability that the initial amount was at most u; the score aggregates disagreement over those thresholds. This evaluates the distribution of explanations, not only its most likely value. For finite first moments this is expected excess continuous ranked probability score (CRPS), equivalently the scalar excess energy score under the convention with a one-half within-law distance term. We report it as excess energy, in the root's physical units. The earlier run in Appendix L additionally standardizes by the initial-prior scale. In the learner comparison, P is the posterior under the true simulator. In proposal comparisons, P is the reference posterior of the same learned model; that score measures numerical inference error. The two references answer different questions, and their scores are not interchangeable or an additive error decomposition. Each reported mean follows the stated world-level aggregation. The CDF score concerns the initial-state marginal, not the full trajectory posterior or the W_2 bound above.

B.1 LOCALLY TRAINED MECHANISMS, UNSEEN REVERSE COMPOSITIONS

Each world contains four monotone local teachers $f_j(x) = a_j x + b_j \tanh(c_j x) + d_j$, with $a_j \in [.8, 1.1]$, $b_j \in [0, .2]$, $c_j \in [.5, 1.5]$ and $d_j \in [-.2, .2]$. Each learner receives 1,024 noisy local pairs per mechanism, with inputs uniform on $[-6, 6]$ and observation standard deviation $.05$. Queries draw a root from an equally weighted, truncated mixture centered at $-.8$ and $.8$, compose a supplied mechanism sequence, and observe only the terminal value with standard deviation $.05$. Training contains no composed inverse query. Horizons four and eight are primary; every evaluated trajectory remains inside the local calibration interval. The query prior, mechanism sequence and observation-noise law are supplied. There is no internal process noise in this confirmation: only the root is unknown before observing the noisy terminal value. Thus the remainder vector in the preceding formulation is empty. These teachers are not restricted to the uniformly expanding regime. The study therefore probes composed reverse use beyond the theorem's sufficient condition.

We compare an eight-bin monotone RQS, a factorized two-layer multilayer perceptron (MLP), and a monotone piecewise cubic Hermite interpolator (PCHIP) whose smoothing resolution is selected using local validation only. All three receive one-dimensional numerical integration over the initial state, with the reference checks reported below. Labels such as "exact root inference" denote these numerical reference calculations rather than closed-form exact arithmetic. This reduces solver approximation as a confound in the learner comparison, subject to those checks. We also execute the RQS through its analytic local inverses with 2,048 corrected importance samples. A variable-length sequence-conditioned NSF supplies an amortized control: one arm uses only local pairs and an-

other adds 8,192 learned-model H1/H2 examples; both are evaluated raw and with a defensive prior mixture and exact importance correction. H4/H8 remain unseen by these arms.

Table 3: Confirmation on ten fresh mechanism worlds. Values are mean true-posterior excess energy on H4/H8; intervals bootstrap worlds. Differences are paired against exact inference in the learned RQS.

Method	Excess energy (95% CI)	Difference from RQS (95% CI)
RQS, exact root inference	.002191 [.001559,.002887]	—
RQS, hard-inverse samples	.002209 [.001551,.002904]	.000018 [.000014,.000021]
Monotone PCHIP, exact	.005073 [.004060,.006075]	.002882 [.001825,.003916]
MLP, exact	.103263 [.078718,.125921]	.101072 [.076914,.123927]
Corrected sequence flow, +8192	.002771 [.002007,.003589]	.000580 [.000429,.000739]
Raw sequence flow, +8192	.112184 [.080052,.145401]	.109993 [.078978,.144558]

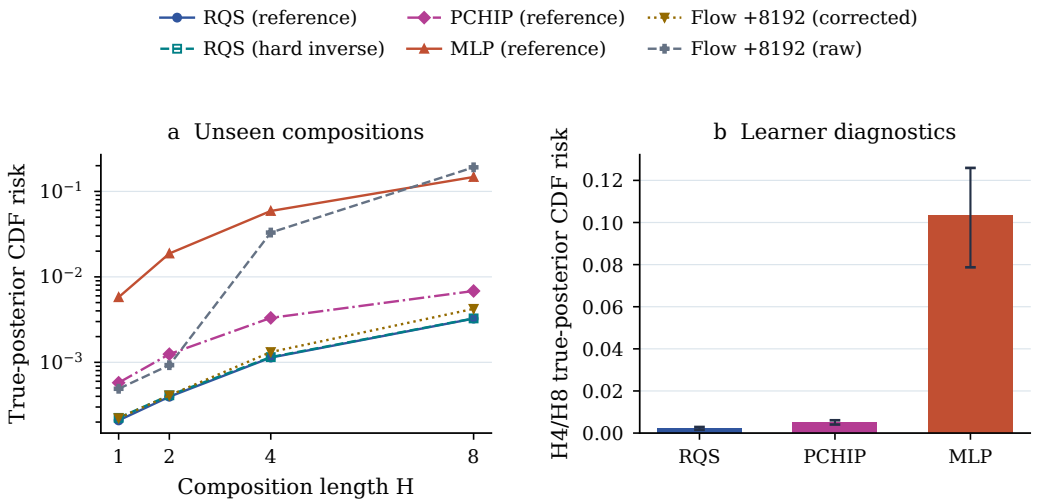

Figure 5: Untrained composition confirmation. Left: posterior risk grows with composition length; exact RQS inference and direct hard-inverse inference remain close, while raw amortized extrapolation degrades and importance correction repairs much of the gap. Right: the H4/H8 model-class comparison after giving every learner exact root inference. Error bars are world-bootstrap intervals.

The hard-inverse sampler's mean excess relative to the learned-RQS posterior is 1.63×10^{-5}, with mean effective sample size (ESS) fraction .933 and no failures. It therefore closely approximates the posterior of its learned model in this regime. Against the true simulator, RQS has lower local forward mean squared error (MSE) (5.32×10^{-5}) than PCHIP (2.05×10^{-4}) and MLP (3.57×10^{-4}). The learner comparison reflects predictive fit as well as reverse implementation.

All forty fixed reference-refinement checks pass, with maximum CDF change 1.21×10^{-5}. Nominal 90% interval coverage is 87.45% for RQS, 84.68% for PCHIP and 24.70% for MLP. These results show useful composed reverse inference with imperfect calibration. The experiment characterizes how established RQS/NSF machinery behaves when locally learned transitions answer unseen multi-step queries.

B.2 Paired learner diagnostics

On 32 previously evaluated worlds, we compare locally fitted learners under the same root-space numerical inference and posterior-refinement checks. The residual MLP uses least-squares initialization and development-selected hyperparameters; RQS and ordinary MLP reuse their checkpoints.

These paired diagnostics complement the fresh-world confirmation by measuring how the learned transition affects composed reverse risk. The revised residual MLP does not improve on its earlier identity-initialized variant (risk .018953).

Table 4: Paired learner comparison on 32 previously evaluated worlds. H4/H8 true-posterior excess energy; intervals bootstrap mechanism worlds.

Local predictive learner	Excess energy (95% CI)
RQS	.002997 [.002399,.003692]
Monotone PCHIP	.005053 [.004465,.005680]
Residual MLP, LS initialization	.053225 [.040868,.066730]
MLP	.021727 [.017919,.025762]

These studies establish that locally learned mechanisms can answer the tested reverse compositions. They also separate learned-model error from approximation to that model's posterior. So far only the initial state is unknown. We next test whether the same local operation remains useful when every transition also has an unobserved disturbance.

C Earlier stochastic-chain bridge

The earlier comparison keeps only the initial scalar state unknown. We now let each transition have an unobserved innovation, $x_{t+1} = f_{\theta,j_t}(x_t) + r_t$, with $r_t \sim \mathcal{N}(0, .02^2)$. The direct operation is still local: $x_t = f^{-1}_{\theta,j_t}(x_{t+1} - r_t)$ for a *proposed* r_t. It never receives the simulator's actual innovations. At length H, a query thus contains an unknown root and H coupled innovations. A noisy squared terminal reading supplies the first evidence E_0; a noisy linear midpoint reading is subsequently added as E_1. A squared terminal reading can leave both signs of the initial state plausible. The midpoint reading constrains an internal state and thereby changes which joint explanations are likely.

Eight fresh worlds provide locally learned monotone spline mechanisms, fitted only from noisy forward pairs. Each world supplies one fixed, unseen mechanism sequence and query at each of $H = 16, 32$. At 2,048 particles, we compare a terminal-anchored direct inverse proposal, an independently generated midpoint-anchored proposal after E_1, reweighting the stored terminal proposal, and forward-prior importance sampling. All methods use the same frozen learned law and exact importance correction. An independent backward-message grid computes the learned root posterior; a separate true-mechanism grid measures model error. Table 5 reports the maximum absolute difference between each weighted root CDF and its learned-model reference, averaged over the eight worlds. The final column compares the numerical learned and true root-posterior references, independently of proposal choice.

Table 5: Root-posterior CDF KS in fresh stochastic chains ($N = 2048$). E_0 is the terminal observation; E_1 adds midpoint evidence. Direct anchors the new proposal at the terminal state; middle anchors at the midpoint; warm reweights stored E_0 particles. The first four columns measure sampling error against the *learned* law. The last is learned-versus-true posterior mismatch. Dashes indicate proposals that do not apply.

H	Evidence	Direct	Middle	Warm	Prior	Model gap
16	E_0	.0188	—	—	.2705	.245
32	E_0	.0319	—	—	.3740	.280
16	$E_0 + E_1$	.0353	.0794	.0348	.3355	.258
32	$E_0 + E_1$	.0447	.1257	.0460	.5398	.431

For the terminal-only query, direct inversion makes a useful joint proposal: its mean effective sample sizes are 1,452 and 877 at $H = 16, 32$, compared with 18 and 6 for forward-prior importance. Its root-CDF errors to the learned law are .0188 and .0319. The paired world-bootstrap intervals for direct-minus-prior CDF KS are $[-.3623, -.1792]$ and $[-.4437, -.2489]$, respectively. These measure a proposal comparison, rather than an exclusive ability to invert.

The separate fresh length-4/8 stochastic bridge gives a useful point of comparison. Averaged over its 32 queries after adding midpoint evidence, a fresh midpoint proposal has learned-law root-CDF KS .0348, versus .0464 for a terminal proposal and .1117 for forward-prior importance; its learned-versus-true posterior gap is .0931. These empirical-CDF values were recomputed on the same eight worlds after correcting interpolation between weighted particles. Because the short and long bridges use different worlds and query sets, their contrasting anchor ordering is descriptive rather than an isolated horizon effect.

Adding E_1 changes the result. Terminal-anchor ESS falls to 549/329, while the fresh midpoint anchor reaches only 144/46 at $H = 16/32$. Reweighting stored terminal particles nearly matches a fresh terminal proposal here. The midpoint proposal's root-CDF error rises to .0794/.1257, worse than the terminal proposal despite using the new observation as its anchor. A long unconditioned segment from the midpoint to the terminal state is a plausible reason, but this experiment does not isolate that mechanism. The midpoint sensor mostly resolves the sign ambiguity; restoring the previous posterior only restores stored, unresampled particles. Thus the test demonstrates evidence addition, not robust recovery after a lost branch.

There is a second, larger limit. Local held-out forward RMSE is .00765, yet the learned-versus-true root-posterior gaps reach .245/.280 under E_0 and .258/.431 after E_1. In an evaluator-only diagnostic that supplies the true terminal state and innovations, the composed inverse gain along the observed paths averages .367/.155 at $H = 16/32$, but recovered-root absolute error still grows from .033 to .060. Small local errors can accumulate even when terminal perturbations attenuate. These pathwise gains do not certify the uniform contraction assumption of Equation (3). The experiment therefore links the bound's error terms to an observed posterior mismatch without claiming to verify all its sufficient conditions.

A post-confirmation structured-solver check provides perspective. On the previously evaluated $H = 32, E_0 + E_1$ worlds, a 4,097-point forward/backward grid has mean root-CDF KS .0051 in 22.5 ms; in a separate exploratory rerun, 2,048 weighted direct-inverse trajectories have root KS .0562 in 23.7 ms. For a root marginal in this scalar setting, structured smoothing is therefore the stronger tested interface. Direct inversion returns weighted *joint* trajectories, but this earlier bridge's grid/FFBS comparison did not match output representations or quality; the later matched comparison is in Section 5. The direct operation is useful here as a posterior proposal; the evidence does not establish a general speed or accuracy advantage over structured inference.

These results make two dependencies visible: the posterior over unresolved innovations changes when evidence is added, and a well-executed learned inverse still inherits errors in the learned mechanism. We next return to the simpler root-only query to isolate the complete cost of direct execution from those additional sources of difficulty.

D Earlier root-only proposal and geometry controls

Here there is no internal process noise, so the initial scalar state is the only unknown trajectory variable. The comparison holds the learned relation, prior, observations and output task fixed. The direct proposal draws a terminal state from the observation-likelihood proposal, composes local state inverses, and corrects the prior and Jacobian weights. The Laplace proposals use the same learned posterior target with their own normalized proposal densities. With normalized weights w_i, the effective sample size fraction is $1/(N\sum_{i=1}^{N} w_i^2)$ for N samples.

D.1 Corrected proposals and complete cost

The cost comparison freezes the original RQS checkpoints on twelve paired worlds. It distinguishes a grid-assisted Laplace mixture from a continuous three-start L-BFGS-B Laplace mixture. Both use normalized truncated components, a defensive prior mixture with its actual integer allocation, and full importance correction. The optimizer receives no posterior-reference grid or analytic inverse. One of 480 optimizer batches reports nonconvergence; its approximate proposal remains in the results with density correction. At 2,048 samples, hard inverse and both Laplace proposals pass the .001 learned-posterior risk threshold on every H4/H8 query. These errors measure inference on the learned relation, separately from the true-world errors in Table 4.

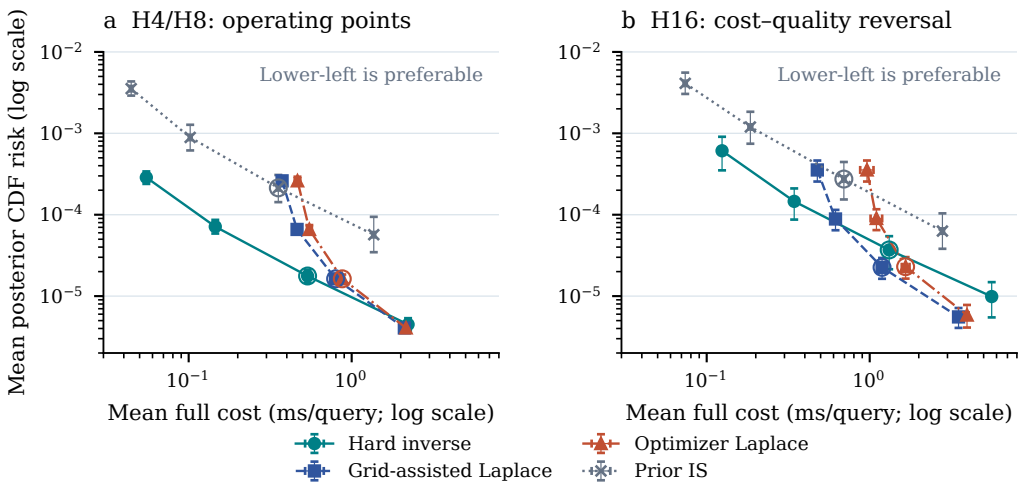

Figure 6: **The useful operating range depends on budget and horizon.** The same twelve previously evaluated worlds at 128, 512, 2,048 and 8,192 samples; lines connect these measured budgets, and open rings mark 2,048. Both axes are logarithmic. Each point is a world-averaged full CPU cost and learned-posterior CDF risk; bars are pointwise 95% world-bootstrap intervals. H4/H8 are equally weighted within worlds; H16 is the prespecified stress point. Shared model, observations, priors and cached runtime are fixed. Grid evaluation, mode setup, sampling and correction are charged; preparation is charged once per displayed batch even when reused across Monte Carlo repeats.

Table 6: Corrected proposals on twelve previously evaluated worlds, $N = 2048$. Risk entries are multiplied by 10^5. Cost is median CPU time per query within a batch of 16 at H4/H8, including proposal setup and correction; it is not single-query latency. All methods share the same cached RQS implementation.

Proposal	H4/H8 risk	H16 risk	ms/query
Hard inverse	1.766	3.693	0.588
Grid-assisted Laplace IS	1.632	2.257	0.824
Optimizer Laplace IS	1.631	2.296	0.894
Prior IS	21.391	27.426	0.370

Table 6 charges the grid baseline's model-grid evaluation and each Laplace method's mode preparation, in addition to sampling and correction. Preparation is charged anew for each reported batch cost, although physically reused across Monte Carlo repeats; sampling-only costs are not used for the complete-cost comparison. Evaluator integration is excluded equally. Timings use float64, one CPU thread and fixed affinity on a shared host. At H4/H8, hard inverse is less expensive than both Laplace implementations. At H16, grid-assisted Laplace instead has lower mean risk and lower measured cost (1.171 versus 1.308 ms/query), and higher ESS (0.924 versus 0.840). The favorable cost ordering therefore changes with horizon. Figure 6 shows the full retained sample-budget sweep: the tradeoff cannot be summarized by a single fixed-budget speed ratio. These curves use world-mean costs, whereas Table 6 reports medians of the measured batch-amortized costs.

D.2 Coordinate geometry and structured inference

To separate execution cost from posterior geometry, 120 exact affine-Gaussian cells compare forward, state and inverse coordinates. Raw condition numbers differ, but after each chart receives its full prior metric the forward/inverse spectral gap is below 1.2×10^{-12} in log condition number. A Kalman/Rauch–Tung–Striebel (RTS) mixture smoother additionally agrees with independent dense

conditioning in means, covariances, likelihoods and mixture weights to within 2.6×10^{-13}. Core marginal smoothing is linear in horizon; the reference check explicitly materializes a quadratic-size covariance matrix. This exact-model comparison explains how a suitable prior metric can remove an apparent coordinate advantage. In the scalar terminal-query experiments, the unknown root dimension remains one as horizon grows. The stochastic-chain experiment in Section C adds one unknown innovation per step, but still has a scalar physical state. Appendix Figure 11 displays the raw-coordinate and full-prior-metric comparison side by side.

The choice between explicit state inversion and a structured alternative thus depends on the mechanism, proposal and query. Together with the stochastic result, this places direct execution among several possible reverse-query interfaces rather than making it the whole inference procedure.

E SUPPORTING STUDY: ONE RELATIONSHIP AND A CHANGING CONTEXT

These supporting studies examine how outcome-only experience changes a context belief used by the same predictive likelihood. Their reversible cause/noise map retains the unknown cause; conditioning still integrates over that cause. They therefore provide evidence for shared probabilistic reuse, separately from the physical state-inverse construction in Section 2.

E.1 EXPERIENCE AND EXPLANATION IN THE SUPPORTING SETTING

Suppose a learner first sees paired causes and consequences and learns to predict what follows from a possible cause. Later, the causes disappear from view: target experience contains only consequences, while the learner is asked to explain likely hidden causes and relations among several encounters. Can those partial encounters improve the explanations without a new target cause-label stream or a separately trained inverse predictor? This earlier controlled setting tests a complementary use of the shared relationship: adapting a context belief from incomplete experience.

The forward conditional alone does not answer that question. Reverse inference also depends on which causes are plausible in the current context. A conservative baseline is therefore to keep the source context fixed and use the learned relationship only to invert each fresh query. That avoids treating unseen causes as labels, but it can discard information: when target outcomes identify the shared context through the learned cause–consequence relationship, they can revise which hidden causes are plausible even though no target causes are shown. Bayesian inversion is standard; the substantive question here is whether that outcome-only context signal can continue to improve reverse use without ever revealing its missing causes.

In this supporting context-adaptation setting, unpredicting uses the learned relationship together with the current context belief. Bayesian conditioning, smoothing, message passing, and planning-as-inference are established machinery; the implementation target is to combine structural inverse operations with conditioning through the same predictive factors, retaining uncertainty and making the remaining numerical inference explicit. We use *bidirectional association* for this reuse of one uncertain relationship rather than separate forward and reverse predictors. The relationship supports anticipation—what follows from a possible state—and explanation— which states or missing pieces could account for an observation. Ambiguity should yield several plausible explanations with probabilities. Generative concept models illustrate the value of representations that support several uses of the same knowledge (Lake et al., 2015).

The target updates in this paper use ordinary self-supervision from incomplete encounters: an observed consequence constrains the context even though its cause is hidden. We call this broader source of correction *super natural supervision*. Other examples could include aligned measurements, neighboring moments, another sensory channel, or an action's consequence; masked reconstruction is a familiar instance (He et al., 2022). The intended loop is experience $\rightarrow$ prediction $\rightarrow$ observation $\rightarrow$ plausible explanations $\rightarrow$ model update. It still obeys information limits: an observation cannot teach a distinction that it erases.

The corresponding coherence requirement is that changing what is observed or queried still means conditioning or marginalizing one shared probabilistic object. This is the precise sense in which the underlying distribution is "commutable" across directions of use; below we use the standard term *probabilistic coherence*, and no algebraic commutativity is assumed.

We study this loop under prior shift: the learned consequence mechanism stays fixed while target experience changes which causes are common. Outcome-only observations therefore have a specific possible role—they can revise the shared context belief without supplying missing cause labels. The same revised belief is then used for ordinary reverse inference and supplied multi-encounter relations that were absent from source fitting. Symmetric and weak-information conditions test when the observation side fails to identify the hidden distinction, while calibration controls expose the separate effect of an imperfect learned channel.

The shared task is reuse under changing information: source encounters expose a relationship, later observations reveal only part of it, and subsequent questions require explanations or relations among encounters. Our controlled setting lets us evaluate hidden explanations against known truth while varying observation informativeness and retained context uncertainty. Supplied component families and question semantics make those comparisons interpretable; the consequence relationship is still fitted from encounters. This separates evidence for useful reuse within the setting from the further question of learning its structure.

Generative modeling supplies machinery for this idea. Flows couple generation and density evaluation in an invertible representation (Kingma & Dhariwal, 2018); invertible inverse models retain latent variables for ambiguous explanations (Ardizzone et al., 2019). ACFlow and Simformer support different conditional queries within shared modeling frameworks (Li et al., 2020; Gloeckler et al., 2024). Yet full-state invertibility alone does not say which uncertainty and dependencies must survive for later joint questions or evidence updates.

Several literatures expose parts of that requirement. Predictive state representations define state through future tests (Littman et al., 2001); probabilistic compilation supports queries and evidence changes through polynomial structure (Darwiche, 2000). ICAE compresses context for later prompts (Ge et al., 2024), while martingale analyses ask whether in-context predictions evolve coherently with evidence (Falck et al., 2024). We ask when one belief representation can replace another for a declared family of future joint queries and revisions.

This bidirectional program also creates a supporting preservation problem. If the same learned experience must later answer a different conditional, combine several encounters, or revise an explanation after new evidence, what can its representation safely discard? Agreement on today's single-encounter answers need not be sufficient. Neural Empirical Bayes learns a likelihood and an unknown source distribution for continuous inverse inference (Vandegar et al., 2021); calibrated label-shift work emphasizes calibration and identification (Alexandari et al., 2020; Garg et al., 2020). We therefore measure three different things separately: whether outcome-only experience improves the target answer, whether a replacement state preserves the fitted model's later answers, and whether those answers are accurate against target truth.

E.2 Shared context model

Let X be a cause and Y its observed consequence. Source encounters fit a conditional density $\widehat{K}_\theta(y \mid x)$. A supplied family of component cause densities $g_c(x)$ and context weights $w \in \Delta_{C-1}$ define

$$q_w(x) = \sum_{c=1}^{C} w_c g_c(x), \qquad J_w(x, y) = q_w(x)\widehat{K}_\theta(y \mid x). \tag{9}$$

Here q_w says how often different causes occur in the current context, while $\widehat{K}_\theta$ says what consequences those causes tend to produce according to the learned relationship. Their product J_w is the shared object that can be read forward for prediction or normalized backward for explanation. Prediction uses $\widehat{K}_\theta(Y \mid X)$, while reverse inference—our *unpredicting* direction—normalizes the same joint to obtain $J_w(X \mid Y)$. After a declared context change, the mechanism remains fixed and the cause frequencies change. All studied component priors have overlapping support. This is out-of-distribution (OOD) adaptation under prior shift.

Let the shared target context $W \in \Delta_{C-1}$ have prior $\pi(dw)$. For target history $D = (y_1, \ldots, y_n)$, *outcome-only adaptation* means updating from these observed Y values while their corresponding target causes X remain unobserved. Define $H_c(y) = \int g_c(x)\widehat{K}_\theta(y \mid x)dx$ and $H_w(y) =$

$\sum_c w_c H_c(y)$. Assuming the denominator below is positive and finite,

$$\mu_D(dw) = \frac{\pi(dw)\prod_{j=1}^n H_w(y_j)}{\int_{\Delta_{C-1}} \prod_{j=1}^n H_u(y_j)\,\pi(du)}. \tag{10}$$

For an integrable function u, write $\mu_D(u) = \int u(w)\mu_D(dw)$. Each target cause is marginalized inside $H_w(y_j)$ and is not observed by this update. Mechanistically, each observed consequence y_j favors contexts whose current mixture of causes would have made that consequence more likely. The history therefore changes which causes are considered common without supplying their missing identities. For a fresh query outcome y, reverse inference after this history is

$$p_D(x \mid y) = \frac{\int q_w(x)\widehat{K}_\theta(y \mid x)\,\mu_D(dw)}{\int H_w(y)\,\mu_D(dw)}, \tag{11}$$

whenever the denominator is positive and finite. This is the concrete *unpredicting* operation studied below: the learned consequence channel stays fixed, while outcome-only history changes the context belief used to explain a new consequence. The two steps have distinct roles: D learns the current context, while the fresh y asks how plausible each hidden cause is for one new observation under that learned context.

Different encounters are independent conditional on W, but their posterior predictions can be dependent after integrating the uncertain shared context:

$$J_D(v_{1:k}) = \int \prod_{i=1}^k J_w(v_i)\,\mu_D(dw), \qquad v_i = (x_i, y_i). \tag{12}$$

In the simulated world w_* is fixed, so this dependence represents the learner's context uncertainty, rather than a physical interaction between causes. This is why a relation among several encounters can carry information that no single-encounter marginal exposes: the encounters share uncertainty about which context generated them.

Reversibility and cognitive interpretation. The learned conditional is Gaussian with mean $f_\theta(x)$ and positive scale σ_θ. The map $(x, y) \leftrightarrow (x, [y - f_\theta(x)]/\sigma_\theta)$ retains cause and residual noise and is invertible. Conditioning still requires integration over missing values. This is shared density machinery of the kind used by flows (Kingma & Dhariwal, 2018), rather than an automatic inverse for a partial observation. The cognitive interpretation is one relationship used to anticipate, explain and revise. The experiments do not isolate an accuracy gain caused by invertibility, or test equivalence to human cognition.

E.3 A SUPPLIED LANGUAGE FOR UNTRAINED RELATIONAL QUESTIONS

Here $y_{1:k}$ denotes fresh query outcomes, distinct from the outcome-only history D used to form μ_D. For a binary cause property $s(x)$, write $B_{i1}(w) = \sum_c w_c \int s(x) g_c(x) \widehat{K}_\theta(y_i \mid x) dx$ and $B_{i0}(w) = \sum_c w_c H_c(y_i) - B_{i1}(w)$. A useful reading is that B_{i1} is the joint evidence for observing y_i and having the queried property true, while B_{i0} is the corresponding evidence for it being false. Conditional on one context w, multiplying these terms combines encounters; summing assignments that satisfy h selects the requested relation. A Boolean relation $h(s_1, \ldots, s_k)$ has joint numerator

$$N_h(w) = \sum_{s_{1:k}: h(s_{1:k})=1} \prod_i B_{i,s_i}(w), \qquad L(w) = \prod_i [B_{i0}(w) + B_{i1}(w)]. \tag{13}$$

The answer is $\mu_D(N_h)/\mu_D(L)$, with positive finite denominator. The implementation computes these quantities by a count recurrence. The properties are $X > 0$ and $|X| > .75$; the relations are all, strict majority and agreement (all true or all false). The learned mechanism supplies the likelihoods, while the supplied grammar specifies how properties compose. No relation-specific predictor is trained.

If b later outcome-only observations, separate from D and the query outcomes, have likelihood product $A(w)$, the revised answer is $\mu_D(AN_h)/\mu_D(AL)$. Equivalently, for a fixed integrable function f and a belief μ that already accounts for the relevant query evidence,

$$\mu_A(f) - \mu(f) = \frac{\mathrm{Cov}_\mu(f, A)}{\mu(A)}. \tag{14}$$

Here $\mu_A(dz) = A(z)\mu(dz)/\mu(A)$, with $A \geq 0$, $0 < \mu(A) < \infty$ and $\mu(|f|A) < \infty$. For the relational query, take $\mu(dw) = L(w)\mu_D(dw)/\mu_D(L)$ and $f(w) = N_h(w)/L(w)$ wherever $L(w) > 0$. This first incorporates the fresh query outcomes, then the later evidence A; averaging the within-context answer under μ_D alone would omit the query's update to context. New experience changes expectations through its association with possible contexts. If contexts that make f large also make the new evidence A likely, the covariance is positive and the revised answer rises; negative association pushes it down. This is the local mechanism by which another ordinary encounter can revise an earlier explanation. Saving only the current expectation need not preserve this update.

F SUPPORTING STUDY: RELATIONAL OOD REUSE

What the controlled setting tests. Fixing the cause-component family and relation grammar lets all comparators use the same questions and information. Context adaptation tests what new outcome-only experience contributes; replacing the context state tests what uncertainty later uses require. Known simulator probabilities separately measure accuracy against the world. These controls expose distinctions that a forward prediction score alone cannot resolve. They do not test learning the component family or discovering the question semantics.

Design. The simulator draws X from three Gaussian components with means $-1, 0, 1$ and standard deviation .35, and sets $Y = X^2 + a\sin(3X) + \mathcal{N}(0, .25^2)$. The learner receives the component shapes, a regression basis and a Gaussian conditional family, but not this simulator formula. For each $a \in \{.03, .6\}$, it fits an intercept, a linear feature and 20 fixed Gaussian radial-basis features by ridge regression from 1,024 paired source encounters; residual variance is fitted from the same data. Source component weights are equal. The two target contexts are $(.8, .1, .1)$ and $(.1, .1, .8)$, each with 64 outcome-only observations and two additional observations. The context prior is uniform Dirichlet$(1, 1, 1)$. There is no reverse network or separately trained compressor.

Each target has 32 fresh groups of four encounters. The same groups are evaluated at both history sizes, using prefixes of length 1, 2 and 4. A relational aggregate contains 384 query slots. The all and strict-majority events coincide for $k = 2$; their duplicate weights remain in the declared grammar. All 3,584 single and relational slots share two source fits and four target histories. In particular, the four informative target/stage readouts share one fitted channel and two histories; they are not four independent training replications.

Comparators and estimands. Frozen Bayesian context retains $\pi(dw)$ and frozen equal weights retain $w = (1/3, 1/3, 1/3)$. Both receive the query outcomes and the complete relational interface, but do not use the target outcomes to revise which causes are thought common in the current context. Their contrast with adaptation therefore isolates what is gained by learning context from incomplete target experience. Full Bayes, the learned-product and generic compressed states receive identical history and query information. Ablations retain three atoms matching only first moments, a single atom at the *base* posterior mean, or products of exact individual marginals under the full query-conditioned belief. Reweighting the base atom does not move it after new observations; it is not a refreshed posterior-mean baseline at the second stage.

For an event with true conditional probability p_*, the expected binary Brier score of $\widehat{p}$ is $(\widehat{p} - p_*)^2 + p_*(1 - p_*)$. We report its excess term $R = \mathbb{E}[(\widehat{p} - p_*)^2]$, averaging the fixed query slots. Truth is used only for evaluation. Reverse negative log-likelihood (NLL) is averaged over the 128 fresh target encounters. The prespecified informative-channel utility criterion requires an absolute excess-risk reduction of at least .001 and a relative reduction of at least 10% against each frozen comparator, together with reverse gain above .05 nats. The weak channel was an ungated diagnostic.

Outcome-only experience improves later hidden-cause inference. The frozen Bayesian comparator makes the conservative choice above: it uses the same learned channel and fresh query outcomes, but it never lets the preceding outcome-only history revise which causes are common. If hiding all target causes also removed the information needed to improve their later explanation, adaptation would collapse to this baseline. In the informative channel it does not. All four informative target/stage readouts satisfy the criterion. Adapted relational excess risk is 0.000169–0.001404, compared with .02983–.04583 for frozen Bayesian context. The corresponding reduction is 96.9–99.4%; against the frozen equal-weight point it is 97.5–99.7%. These percentages concern

excess conditional risk, not total Brier loss or accuracy. Ordinary reverse inference improves by 0.229–0.256 nats per encounter. The learned-product and generic compressed states preserve these readouts to numerical precision. Thus target causes are never supervised, yet consequence-only history changes their later explanations by identifying shared context through the learned relationship. This effect is conditional: weak-information and symmetric controls show the boundary when the observed side does not identify the relevant hidden distinction.

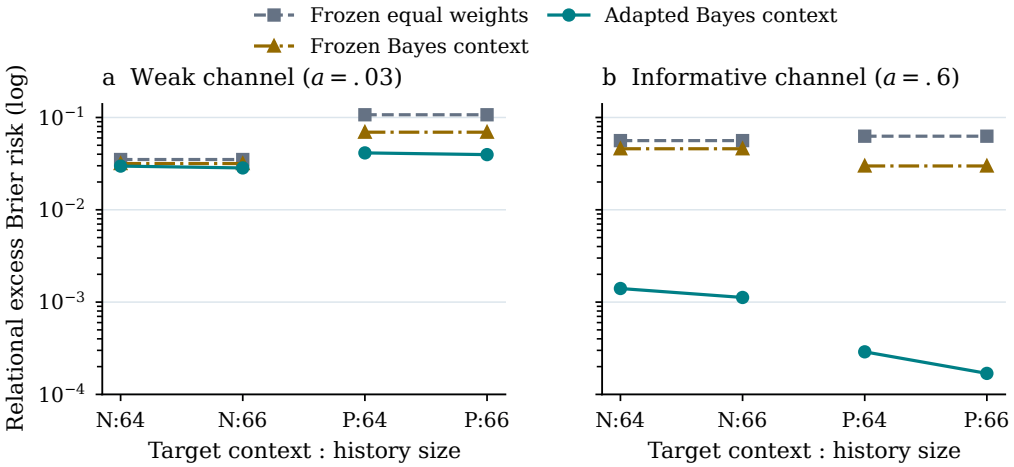

Figure 7: Outcome-only context adaptation improves aggregate relational excess Brier risk in both studied channels. N and P denote negative- and positive-component-dominant targets. The informative channel passes the declared utility criterion at both history sizes in both targets. The weak N:64 cell improves only 6.15% over frozen Bayesian context. Lines connect paired descriptive readouts; no target causes are observed during adaptation, and the frozen Bayesian comparator sees the same query outcomes but does not use preceding target outcomes to revise context. No independent-replication uncertainty is implied. Lines join the two stages within each target. The vertical axis is logarithmic.

Table 7: Context representation and maximum probability discrepancy from full fitted-model Bayes over all recorded queries and stages. The two degree-six states achieve the same atom reduction. Atom count is not a measured total memory or runtime frontier. The three-atom ablation preserves the base single-encounter law but can lose later answers.

Context representation	Atoms	Maximum answer error
Full posterior rule	1,296	0
Learned affine products	28	3.86×10^{-13}
Generic polynomials	28	6.01×10^{-14}
First-moment state	3	4.66×10^{-01}
Base mean atom	1	2.69×10^{-01}

What compression preserves. Both degree-six states retain 28 positive atoms and agree with full Bayes to 3.86×10^{-13} across the record (Table 7). Their ordinary reverse-score discrepancies are at most 6.03×10^{-14} nats. The first-moment ablation matches the base single-encounter law, but after two observations its context-mean discrepancy reaches .19176. This local comparison illustrates the distinction in Proposition 2. The construction and matched generic control establish faithful bounded reuse, with no distinctive compression or computational advantage attributed to the learned coordinates.

F.1 ACCURATE COMPUTATION AND ACCURATE EXPLANATIONS CAN DISAGREE

Table 8: All eight relational aggregates. Risk columns report excess Brier; reverse gain is single-encounter NLL improvement in nats from the frozen equal-weight source model. Lower risk and higher gain are better. The base mean atom stays fixed at the second stage. In the informative positive target, approximations improve truth-risk despite disagreeing with full fitted-model Bayes. Rows share fits, histories and query groups as specified in the design.

a	Context	n	Full Bayes	Base mean atom	Marginal product	Reverse gain
0.03	N	64	0.029740	0.032002	0.032017	0.075
0.03	N	66	0.028409	0.032002	0.030938	0.087
0.03	P	64	0.041332	0.072519	0.072171	0.128
0.03	P	66	0.039596	0.072519	0.069459	0.137
0.60	N	64	0.001404	0.002042	0.001827	0.229
0.60	N	66	0.001123	0.002042	0.001502	0.234
0.60	P	64	0.000289	0.000124	0.000173	0.256
0.60	P	66	0.000169	0.000124	0.000097	0.254

Table 8 retains a useful reversal. In the informative positive target with 64 observations, full fitted-model Bayes has risk $.000289$, while the mean atom and product of exact marginals score $.000124$ and $.000173$. Both approximations also have lower risk after the additions. For any two answers a, b, the identity

$$R(a) - R(b) = \mathbb{E}[(a-b)^2] + 2\mathbb{E}[(a-b)(b-p_*)] \tag{15}$$

separates a nonnegative disagreement term from its signed association with the reference's error. For the base mean atom against full Bayes in this cell, these terms are $.0000860$ and $-.0002511$. Thus a less faithful answer can be closer to the true answer. The evaluation fixes the world's context and the realized finite history; the source likelihood is also learned. There is no true-channel, finite-history Bayesian arm in this probe to separate these contributions. The reversal neither contradicts Bayesian optimality under its own joint distribution nor selects a uniformly better approximation.

Aggregate transfer also need not benefit every question. In the weak negative target, four-cause positive-majority risk worsens against frozen equal weights by $.04266$ at the base stage and $.03628$ after two observations. Four-cause all-positive risk worsens against frozen Bayesian context by $.00330$ and $.00179$. These are retained descriptive strata of the supplied grammar, including its duplicate two-cause events, rather than new success criteria. Even in the informative positive target, the additional observations improve relational aggregate risk while slightly reducing ordinary reverse gain from $.25559$ to $.25442$ nats. Individual experience need not improve every use monotonically.

These experiments establish the bounded reuse result before asking a narrower representation question. Because the same adapted context belief supports single-encounter explanations, multi-encounter relations and later evidence, Appendix H asks what uncertainty a replacement state must retain to preserve those fitted-model uses.

G SUPPORTING STUDY: ESTIMATED SOURCE STRUCTURE AND GEOMETRIC SENSING

The first study supplies the component shapes and uses one scalar observation family. The next experiment tests two dependencies of that result: whether outcome-only reuse survives estimating those shapes from source encounters, and whether it remains useful with a different observation geometry. It also adds independent source fits and competent point-estimation controls. These are extensions of the same predicting/unpredicting task; question semantics remain supplied.

Less supplied structure, different observations. The scalar settings retain the informative and weak channels above. In two additional settings, $X \in \mathbb{R}^2$ is a hidden position and Y consists of noisy distances to the origin, or to both the origin and $(1, 0)$. The cause components have means $(-1, 0)$, $(0, .8)$ and $(1, 0)$ and covariance $.35^2 I$; each sensor adds independent Gaussian noise with standard deviation $.25$.

The range-only channel cannot distinguish the reflected outer components, while the second distance breaks that reflection. Different sensor settings use separate fits and draws; their empirical comparison is not a paired sensor-addition ablation.

Each fit receives 1,024 paired source encounters and estimates the consequence mean and diagonal residual variance. A separate three-component diagonal Gaussian mixture fits source X without component identities. Its estimated shapes are then frozen before outcome-only target adaptation. Target component weights use the same two shifts as the first study. We cross supplied versus estimated shapes with fitted versus true channels; supplied true shapes and channels are privileged diagnostic arms. The estimated-shape/fitted-channel arm learns both numerical relationships from source measurements. It still receives the component count, distribution families and property grammar. For vector observations, the cause/residual representation of Appendix E applies componentwise.

Repeated evidence and matched controls. There are twelve independent source fits per setting, four target histories per direction per fit, and 32 fresh four-encounter query groups per history. Histories have 64 observations and are also evaluated after two additions. Every arm uses the same history and query observations within a comparison. We compare full Bayesian adaptation with frozen Bayesian context, a point at the posterior mean refreshed from each complete history, the original fixed-base mean, and maximum likelihood (ML) mixture weights. All methods answer the same relational queries. The refreshed mean receives the computation needed to reconstruct the full-history posterior; it is a query-state ablation, not an efficient standalone memory.

We report both target directions separately in the accompanying local record and average histories and directions within each source fit for the main table. Intervals resample source fits, not query slots. The primary relational aggregate counts distinct events, removing the duplicate two-cause majority/all event; the original slot-weighted aggregate remains recorded for comparison. Continuous reverse scores are evaluated against hidden target X only after inference. The primary estimated-shape/fitted-channel readout is shown below.

Table 9: Extension with source-estimated component shapes and a fitted channel, after 64 outcome-only encounters. Positive gains indicate improvement over frozen Bayesian context. Intervals are descriptive 95% fit-cluster bootstrap intervals over twelve independent fits per setting; they are not adjusted for multiple comparisons. Risk uses distinct relational events.

Setting	Excess Brier	Brier gain [95% CI]	Reverse gain [95% CI]
Scalar, weak	0.0460	0.0016 [-0.0022, 0.0053]	0.023 [-0.005, 0.050]
Scalar, informative	0.0078	0.0308 [0.0284, 0.0329]	0.222 [0.207, 0.236]
2D, one sensor	0.0542	0.0060 [0.0024, 0.0097]	0.056 [0.030, 0.081]
2D, two sensors	0.0014	0.0041 [0.0036, 0.0046]	0.098 [0.090, 0.106]

Reuse survives both extensions in the informative settings. With estimated shapes and a fitted channel, the informative scalar and two-sensor settings improve both relational excess Brier and reverse NLL at every source-fit mean, averaging each fit's histories and two target directions. Mean excess-risk reductions are 79.8% and 74.6%, respectively (Table 9). These are reductions of excess risk, not accuracy percentages. Some individual histories still worsen. The expanded averages are smaller than the original probe's 96.9–99.4% readouts; new fits, histories, estimated shapes and distinct-event weighting define a separate comparison. The first result remains a bounded observation, not the expected gain of every later run.

Single-encounter agreement can conceal relational loss. The refreshed mean and full Bayes have the same complete current single-encounter law and hence the same ordinary reverse NLL. Nevertheless, the mean state has higher average relational excess risk in all four settings (Figure 8). The differences are largest in the weak scalar and single-sensor settings: 0.0195 and 0.0204, respectively. Thus the distinction isolated by the preservation argument has a measured consequence here, even after correcting the earlier fixed-base-mean limitation. This is an average comparison in the declared settings, not a claim that every approximation is worse on every query.

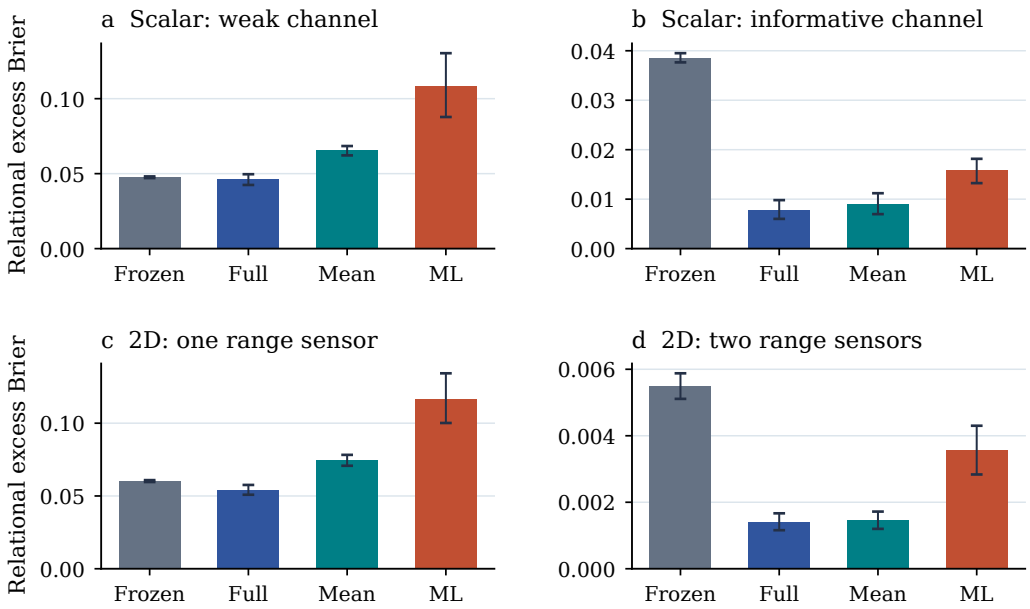

Figure 8: Matched inference controls with estimated source shapes, after 64 outcome-only observations. Mean denotes the refreshed posterior-mean context. Error bars resample source fits. Each panel retains its own risk scale; lower risk is better. Fitted-model fidelity and truth-risk remain different criteria.

Information boundaries remain visible. The weak scalar gain has an interval spanning zero; its apparently positive aggregate does not establish reliable improvement. The single-sensor setting improves the aggregate, but some sign-related questions worsen. Its true observation law distinguishes central from outer causes without distinguishing the two reflected outer components. Learning some context information therefore does not identify every queried distinction. Across 96 history/target cells, Bayesian adaptation increases relational risk in 41 weak-scalar cells and 26 single-sensor cells. ML adaptation is more harmful on average in these two settings despite satisfying the likelihood optimality check. The source-shape/channel crossings in Appendix R retain the corresponding oracle diagnostics.

Together these observations extend useful bidirectional reuse to source-estimated distribution shapes and a distinct observation geometry. They also expose a practical evaluation issue: a reverse-score check alone can miss a loss on multi-encounter questions. The tested component family, question semantics and announced prior shift remain explicit assumptions. Long-horizon continual inference is outside this experiment.

H Supporting result: preserving later queries and revisions

Here *preservation* has a narrow technical meaning: replacing one context belief by another leaves a declared family of fitted-model query answers and evidence revisions exactly unchanged. It does not mean lower target-world risk or general usefulness.

The general expectation interface. This abstracts the mechanism above rather than introducing a new one. In the relational experiment, Z is the shared context W, N_q and L_q play the roles of the relational numerator and query-evidence likelihood, and A is the likelihood contributed by later outcome-only encounters. The abstraction asks which expectations a replacement belief must retain so that the same predict–unpredict–revise computation returns the same declared answers. Let μ be a belief over a shared latent state Z, and write $\mu(f) = \int f(z)\mu(dz)$. Queries are indexed by $q \in \mathcal{Q}$.

Conditional on Z, N_q is the joint likelihood of the query event and its query observations, while L_q is the likelihood of those observations alone. Let $A \in \mathcal{A}$ be the nonnegative likelihood function of additional observations. Assume these observations are conditionally independent of the query event and query observations given Z. The revised answer is

$$T_{A,q}(\mu) = \frac{\mu(AN_q)}{\mu(AL_q)}, \qquad \mathcal{F}_{\mathcal{Q},\mathcal{A}} = \operatorname{span}\{1, AN_q, AL_q : q \in \mathcal{Q}, A \in \mathcal{A}\}. \tag{16}$$

For each compared belief and declared (q, A), assume the displayed expectations exist, $0 < \mu(A) < \infty$, $0 < \mu(AL_q) < \infty$, and $0 \le N_q \le L_q$. If a replacement belief agrees with μ on this function space, it preserves the declared answers and revisions. This is a sufficient condition from Bayesian normalization, not a minimal-memory claim. It applies beyond mixtures, but unrestricted question or evidence families need not admit a finite-dimensional space. The affine mixture below supplies a particular finite closure, using standard moment identities.

The degree bound has a direct operational reading in the mixture model. One encounter contributes one affine factor in the context weights, so a query over k encounters has degree at most k. Each of b later evidence factors can increase the degree by at most one. Preserving moments through degree $b + k$ therefore preserves every numerator and normalizer used by the declared query-and-revision budget.

Proposition 2 (Current-law inclusion and bounded revision)**.** *Fix any finite collection of component encounter laws J_c on a common measurable space and let $J_w = \sum_c w_c J_c$. Encounters are independent conditional on the same context W. Let μ and ν be two context distributions on Δ_{C-1}, with integers $b \ge 0$ and $k \ge 1$. If $\mathbb{E}_\mu[W_c] = \mathbb{E}_\nu[W_c]$ for every component, they induce the same complete single-encounter law. If they also agree on the expectation of every polynomial in W of total degree at most $b + k$, they induce the same normalized k-encounter law after the same b nonnegative affine encounter-evidence factors, provided the update normalizers are positive and finite. Their conditional event probabilities agree wherever the common query evidence has positive finite density, or almost everywhere under its evidence law.*

Linearity gives $\int J_w(v)d\mu(w) = \sum_c \mathbb{E}_\mu[W_c]J_c(v)$. Products for k encounters and b additions have degree at most $b + k$; matching their expectations preserves both numerators and normalizers. The proof in Appendix N covers discrete or continuous fixed component laws, any finite component count and measurable events of the encounters. It does not require our Gaussian family or Boolean grammar. It requires fixed affine factors and a bounded future-use budget; arbitrary shared parameter uncertainty is outside this finite-degree argument.

For a concrete witness, let a shared scalar context W be the success probability of conditionally independent Bernoulli encounters. Compare $\mu = \delta_{1/2}$ with $\nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Both have $\mathbb{E}[W] = 1/2$, so they induce exactly the same single-encounter Bernoulli law. Their pair-success probabilities are nevertheless $\mathbb{E}_\mu[W^2] = 1/4$ and $\mathbb{E}_\nu[W^2] = 1/2$; after one success, the next-success probabilities $\mathbb{E}[W^2]/\mathbb{E}[W]$ are respectively $1/2$ and 1. For two encounters conditionally independent given W, let S_i be the event that encounter i satisfies the same query, with context-dependent probability $r(W)$. Then

$$\Pr(S_1 \cap S_2) - \Pr(S_1)\Pr(S_2) = \operatorname{Var}_\mu[r(W)]. \tag{17}$$

Thus preserving the entire current law can discard uncertainty needed for another use. This analytical witness involves neither a learned-channel error nor an empirical claim about how often a particular compression fails. The earlier single-encounter interface is retained inside the richer contract; no lower bound on every smaller representation is asserted.

For representation evaluation, the consequence is specific: matching all current single-encounter answers does not certify matching future joint answers or evidence revisions. The witness establishes this insufficiency, while the experiment measures its consequences for particular learned models. Neither establishes how frequently a learned memory will discard the relevant uncertainty.

Positive context representations. We use classical positive quadrature and finite-feature recombination (National Institute of Standards and Technology, 2026; Bayer & Teichmann, 2005; Adachi et al., 2022). In a scalar feasibility study, six quadrature atoms preserved 54 declared Bayesian query/revision checks against a 513-point reference to 1.9×10^{-15}. That study measured fitted-model fidelity only.

The subsequent three-component study uses a 1,296-node context rule and positive recombination preserving degree-six functions, covering $k \leq 4$ and $b \leq 2$. Operationally, the compressed state is a smaller discrete belief over possible contexts whose weighted expectations match the functions needed by those later uses. Positivity keeps the retained atom weights interpretable as probabilities, rather than as signed correction coefficients.

The tested learned-product features are two affine functions obtained from the fitted likelihoods. Together with a constant they have rank three in both fits. Their complete degree-six closure therefore spans the same $\binom{8}{2} = 28$ dimensional polynomial space as ordinary simplex coordinates. Both representations use the same support, objective and positive linear program. An invertible change of feature coordinates gives the same feasible constraints, so the generic construction recovers this particular operation. This is stronger evidence than agreement of output scores alone. For C components and degree d, the full polynomial space has dimension $\binom{C-1+d}{d}$. Its growth and the cost of component integration limit tractability; the theorem's scope is not a general efficiency guarantee.

I SUPPORTING STUDIES: EARLIER TRANSFER AND THE INFORMATION BOUNDARY

The relational construction extends a sequence of completed single-encounter experiments. We retain their distinct models, data and estimands rather than pooling them as relational replications.

The first construction fits a supplied quadratic triangular flow from 4,096 paired encounters. All three seeds improve forward and reverse held-out NLL over initialization, while preserving positive and negative explanations for $Y = 1$. A subsequent neural model, with two 64-unit SiLU hidden layers, learns a nonquadratic mean from raw X. In all three seeds it improves both scores over initialization and a fitted quadratic control. After 512 cause-only observations under a shift from $\mathcal{N}(0, 1)$ to $\mathcal{N}(1, .8^2)$, adapting its Gaussian cause prior improves reverse NLL by 0.379–0.402 nats without changing the forward conditional. This measures learned OOD transfer, not an advantage over an equally expressive conditional density combined with the same prior.

The outcome-only experiment then fits ten neural channels, five seeds each for $Y = X^2 + a\sin(3X) + \mathcal{N}(0, .25^2)$ with $a = 0$ or .6. Target causes are mixtures of $\mathcal{N}(-1, .45^2)$ and $\mathcal{N}(1, .45^2)$ with positive-component weight .1 or .9. At 4,096 outcome-only encounters, Bayesian weight adaptation improves reverse NLL by 0.230–0.272 nats in all ten informative seed/shift cells. The target causes are hidden during adaptation.

What the observation erases. At $a = 0$, the two exact component outcome densities coincide. No number of square-only observations identifies their frequency, although a query invariant to that frequency may remain identifiable. A noisy channel $C = X + \mathcal{N}(0, .8^2)$, available independently with probability $1/8$ during adaptation, restores reverse gains of 0.334–0.380 nats in every symmetric seed/shift cell. The individual test query still observes only Y. This isolates the benefit of informative experience with a changed sensor budget.

Errors in a learned channel can instead create a false distinction. In the symmetric outcome-only condition at 4,096 observations, marginal maximum likelihood (ML) harms reverse NLL in seven of ten cells, by as much as 8.275 nats, while its largest observed marginal-likelihood improvement is only $.0003822$ nats. Bayesian integration reduces the largest harm to $.567$ nats but still harms four cells. This illustrates a known tension in empirical Bayes and shift estimation (Efron, 2014; Garg et al., 2020; Vandegar et al., 2021): fitting an observed distribution and recovering a useful hidden explanation are different objectives.

I.1 CALIBRATING CONTINUOUS EXPLANATIONS FROM OBSERVED MEASUREMENTS

A separate calibration experiment uses twenty fresh neural models across $a \in \{0, .03, .1, .6\}$ and five seeds. It tests raw inference, affine Gaussian likelihood calibration, two uses of bias-corrected temperature scaling (BCTS) (Alexandari et al., 2020), and the exact mechanism. All fitted remedies share 4,096 source calibration pairs from a known balanced component mixture; an independent 2,048 pairs assess source generalization. Target adaptation sees only Y. The full record contains 400 cells across two target counts and two shifts.

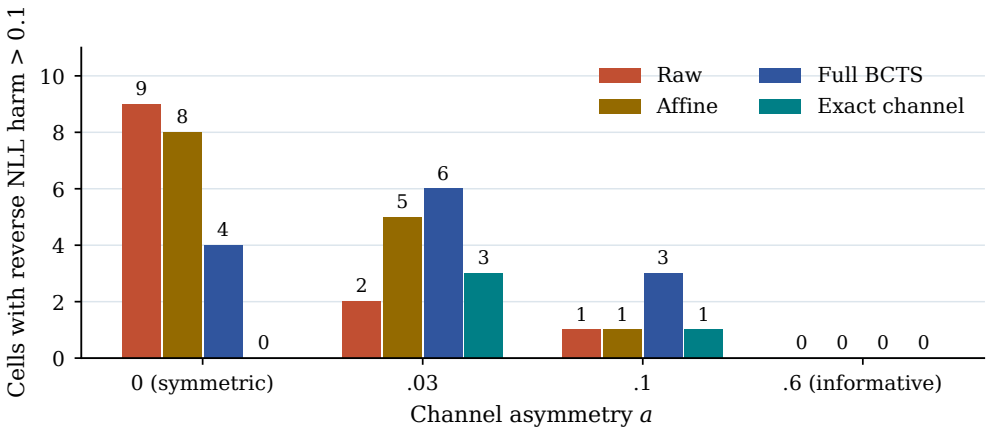

Figure 9: Earlier calibration study, separate from the relational experiment: material ML reverse harm after 4,096 outcome-only observations. Each bar counts all five seeds and both target shifts; change is relative to the interface's own frozen output. Full BCTS reduces symmetric failures but does not remove near-boundary harm. The exact channel also has finite-sample failures under weak information. Counts compare adaptation behavior, not absolute model quality across different frozen interfaces.

The implementation bridge is to calibrate and reuse component probabilities without observing component identities. For the supplied $Z \to X \to Y$ model, source measurements supply $r(X) = \Pr(Z = + \mid X)$, giving, for any fixed predictor,

$$\mathbb{E}[\mathrm{CE}(1\{Z = +\}, a(Y))] = \mathbb{E}[\mathrm{CE}(r(X), a(Y))]. \tag{18}$$

This elementary population identity licenses the soft calibration objective. It is not a finite-sample calibration guarantee and depends on the known component family. BCTS corrects component odds and the full interface combines the corrected probabilities with continuous within-component inverse densities:

$$p_{\mathrm{full}}(x \mid y) = \sum_z \rho_{w,z}(y) h_z(x \mid y), \qquad h_z(x \mid y) = \frac{g_z(x)\widehat{K}_\theta(y \mid x)}{H_z(y)}. \tag{19}$$

This realizes the published label-shift correction as a normalized continuous explanation. It differs from inserting a corrected weight into the original joint and need not retain that joint's forward conditional. The exact contract is given in Appendix Q. It has not been extended or tested as a calibration repair for the three-component relational experiment.

In the informative channel, full-BCTS Bayesian reverse gains are 0.217–0.283 nats in all ten cells. At $a \in \{0, .03\}$, however, ML harm above .1 nats remains in 13/20 affine and 10/20 full-BCTS cells, spanning all five seeds for each remedy (Figure 9). Full BCTS sets its slope to zero in three of five symmetric models, suppressing an unreliable distinction. This improves restraint without identifying the hidden frequency. At $a = .03$, even exact-channel ML harms three of ten cells above the threshold, so finite-sample weak identification also matters. BCTS improves its source fit loss in all twenty models, yet independent source cross entropy worsens in twelve. The experiment therefore tests a fitted remedy, not the guarantees for perfectly calibrated identifiable models, and does not settle other calibration methods.

J Related work for the supporting context studies

The following comparisons concern context adaptation and query preservation, complementing the implementation comparisons in Section 6.

Learned relationships and prior adaptation. Invertible inverse models, arbitrary conditionals and shared simulation-based inference already support multiple uses of learned relationships (Ardiz-

zone et al., 2019; Li et al., 2020; Gloeckler et al., 2024). Neural Empirical Bayes (NEB) (Vandegar et al., 2021), Sections 1–2 and 5.1–5.2, develops the paired-likelihood, unknown-prior and continuous-posterior pipeline. It also compares observed and hidden distribution errors, including symmetry and regularization. We examine supplied multi-encounter relations, bounded context compression and revision alongside truth-risk in a fully specified learned mixture.

Calibration and downstream error. Unlabeled prior adjustment (Saerens et al., 2002), calibrated ML adaptation (Alexandari et al., 2020), label-shift error analysis (Garg et al., 2020) and empirical Bayes modeling distinctions (Efron, 2014) motivate the earlier controls. Recent calibration-aware online adaptation explicitly couples calibration with downstream excess cross entropy (Jeong et al., 2026). Its auxiliary calibration procedure was not evaluated here. Our soft-responsibility and continuous-density interface adapts these components to the available measurements.

Representing future questions. Predictive state representations encode sufficient future tests (Littman et al., 2001); probabilistic compilation supports many queries and evidence changes through polynomial structure (Darwiche, 2000). Positive moment preservation is classical (Bayer & Teichmann, 2005; National Institute of Standards and Technology, 2026). The recombination subroutine described by Adachi et al. (2022), Section 2, is direct machinery for the finite-feature compression used here. These connections explain why the degree-six control is decisive for the declared future queries and updates.

K Interpretation of the supporting context results

The supporting context studies give another test for reusable generative knowledge. Forward prediction alone does not show whether a learned relationship remains useful when the direction of observation changes. Both adapted and frozen models can perform reverse Bayesian inference; the distinction exposed here is whether consequence-only history can revise the shared belief used by that inference. In the bounded setting studied here, it can do so when observations identify the shared context, and reuse reaches an information boundary when they do not. The connections below identify research targets, not results on language models or unrestricted continual learning.

When does new experience produce coherent belief revision? For a reusable model, the relevant question is not only whether it can evaluate a reverse conditional once, but whether later evidence changes that conditional in a way consistent with the same underlying belief. Our outcome-only result gives a small positive example: observed consequences revise context, and that revised context improves later hidden-cause explanations without target cause labels. Falck et al. (2024) study Bayesian coherence of in-context learning on exchangeable data through the martingale property. Their Sections 2.3–2.4 already distinguish single-step predictions from epistemic information in sample paths and use integration over shared uncertainty. Prior-fitted prediction makes the assumed distribution of tasks central to amortized Bayesian inference (Müller et al., 2025). Our shared context law connects anticipation, unpredicting and revision, while the information controls show why an uninformative observation cannot identify a hidden frequency. The open extension is a learned interface that preserves joint uncertainty as both questions and context change. An operational test would compare joint queries and evidence-consistent updates under the same model, separately from target accuracy. We test explicit Bayesian integration, not whether an amortized neural or language-model interface passes those tests. Exchangeability here concerns encounters within a context, not arbitrary token order or observations across an unmodeled context change.

Supporting question: what should compressed memory preserve for future use? ICAE learns compact context slots for reconstruction, continuation and later instruction prompts (Ge et al., 2024). Predictive state representations instead define state through sufficient future tests (Littman et al., 2001). Equation (16) specifies the question/update functions relevant to such sufficiency; Proposition 2 supplies an exact finite-mixture realization. The resulting design question is which future queries and evidence updates a memory must support. Here these uses are declared in advance and the degree bound supplies a sufficient state; the experiment does not learn that choice. Learning or approximating this space for changing query families remains open. Testing it requires queries held out from representation training, repeated updates, and a measured memory–accuracy cost. The present $b \leq 2$ revision test does not establish unrestricted continual inference. The contract and

witness give an evaluation target for such memories, without demonstrating that existing context compressors fail it.

Which inference accuracy matters under distribution shift? Calibration-aware adaptation studies downstream error as well as prior estimation (Jeong et al., 2026). Our matched compression and truth-risk measurements separate exact preservation of a fitted belief from accuracy against the world. The open question is when approximation, calibration or belief revision yields a dependable downstream benefit under misspecification and finite evidence. The current reversals establish that fidelity alone does not select the best answer in the recorded cells; a true-channel finite-history comparison and fresh fitted channels would be needed to isolate their cause and stability. The earlier calibration findings remain complementary evidence with a distinct interface, rather than a tested repair of the relational model.

The supporting context results demonstrate bounded bidirectional reuse from incomplete experience. A relationship learned for forward prediction can be combined with a changing context belief so that consequence-only target experience improves later explanations of hidden causes and supplied relations among them. The effect is conditional on what the observed side identifies and on the quality of the learned channel. The supporting preservation contract then states what uncertainty a replacement context state must retain for the declared later queries and updates, while the truth-risk controls show that exact fitted-model preservation is a different objective from being correct in the shifted world. The expansion retains informative gains when component shapes are estimated from source experience and observations describe two-dimensional geometry. Its refreshed-mean control further shows why matching current single-encounter answers can miss relational loss. These gains retain the supplied grammar, fixed mechanism and overlapping-support prior shift; neither arbitrary semantic inference nor a causal benefit of invertibility is established. These boundaries suggest a concrete evaluation principle: test whether learned knowledge stays coherent and useful as observation direction, evidence, and query change, rather than evaluating forward prediction alone.

L EARLIER DIRECT-INVERSE EXECUTION AND COORDINATE DIAGNOSTICS

This study of locally learned transitions and held-out compositions provides additional execution and coordinate diagnostics. Its posterior reference has unresolved resolution checks, so the main cost comparison uses the separately verified proposals in Appendix D.

Learned nonlinear mechanisms. Each of twelve held-out libraries contains four scalar transitions

$$f_j(x) = a_j x + b_j \tanh(c_j x) + d_j, \quad a_j \in [1.04, 1.08],\ b_j \in [.01, .03]. \tag{20}$$

For each mechanism, 1,024 noisy local pairs train a 16-bin monotone rational quadratic spline with affine output and tails, using the established analytic inverse of neural spline flows (Durkan et al., 2019). Two initializations are retained; the learning rate is selected on three separate development libraries. Held-out queries compose random mechanism sequences of horizons $1, 4, 8, 16, 32$, unseen as composed inverse tasks during training. Each library trains its own local models. Dense-grid derivative estimates with exact affine-tail checks exceed 1.020 on the checked domain.

Given a teacher-generated terminal state, direct state recovery composes the learned local spline inverses. Safeguarded local and global bisection instead use the same learned forward maps as numerical oracles. At $H = 32$, all three methods have mean absolute recovery error (MAE) .00865 and zero failures; the maximum MAE disagreement between analytic and numerical execution is 6.51×10^{-15}. The error is therefore learned-model composition error rather than a solver discrepancy. The execution timings used different batch sizes (32 versus eight queries) and 48 fixed bisection steps. We therefore use them only to diagnose the cost of these implementations; Appendix D reports the matched full-query comparison.

Coordinate geometry. To separate wall-clock implementation from posterior geometry, we fit linear maps to four local expanding affine mechanisms in dimensions one and two, then compute exact Gaussian posterior precision matrices. In forward coordinates a terminal observation contains products $A_{H-1} \cdots A_t$; in terminal-inverse coordinates the corresponding products use $A_t^{-1} \cdots A_{H-1}^{-1}$. At $d = 2, H = 64$, the median learned forward-coordinate condition number is 1.35×10^{10}, versus

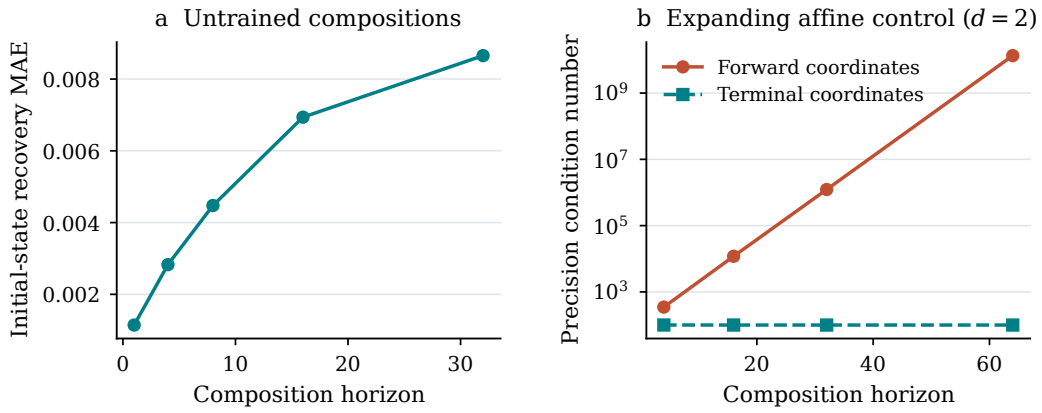

Figure 10: Held-out hard-inverse experiment. Left: learned composition error grows with horizon, so direct inversion does not erase forward-model error. Right: in expanding affine mechanisms, forward-coordinate conditioning grows with horizon while terminal-inverse conditioning remains bounded in this family. The left panel averages two initializations and twelve libraries; the right reports medians across learned affine cases in twelve libraries.

100 in terminal coordinates; the median ratio is 7.42×10^{-9}. These are precision-matrix calculations. Prior whitening can eliminate the contrast, and contracting dynamics can reverse its direction.

Noisy posterior queries. We next add process noise .02 and terminal noise .05. A dense-grid backward-message recursion supplies a provisional one-dimensional reference. The hard-inverse proposal samples a terminal value around the observation and local process innovations, then traverses the same learned transitions backward with exact Jacobian weights. We compare it with forward-prior importance sampling, the identical proposal executed through numerical inversion, and a full-state dense Laplace importance proposal using four MAP starts and an exact target/proposal correction, implemented with dense matrix operations. No tested method fails. The reference checker compares domain extension at identical grid spacing, omitting the computed coarse-to-fine comparison; its true flags therefore do not establish resolution convergence. Its initial-state tail check also does not certify intermediate-state truncation.

At the primary $H = 16$ and 128 particles, direct inference has mean standardized excess continuous ranked probability score (CRPS) .000964 and mean effective sample size (ESS) fraction .966, versus .02158 and .0623 for forward-prior importance. This earlier run divides integrated squared CDF error by the initial-prior standard deviation; ESS fraction is $1/(N\sum_{i=1}^{N} w_i^2)$ for normalized importance weights w_i. The library-level paired CRPS difference is $-.0206$ with bootstrap CI $[-.0238, -.0177]$. On the prespecified first query of each sequence, dense Laplace has CRPS .000871 and direct inference .000900; their paired difference CI $[-.000273, .000351]$ includes zero. Analytic and numerical execution of the same terminal proposal agree in CRPS within 3.91×10^{-12}. All these posterior metrics are conditional on the unresolved reference check. The prespecified marginal quality gate also requires relative standard-deviation error $\leq .1$ and sign-probability error $\leq .03$, alongside CRPS $\leq .005$. At 128 particles direct passes 80.9% of all 768 queries, below the 90% target; 512 particles raises this to 96.6%. On the matched 96-query subset, direct and Laplace pass 85.4% and 78.1% at 128, and both pass 97.9% at 512. Joint-posterior quality has not been measured. The original timers include unequal weighting and setup work and lack warm-cache controls; the matched cost analysis is reported in Appendix D.

These diagnostics show agreement between analytic and numerical state recovery and a coordinate effect with a clear prior-metric boundary. Section 4 tests true-posterior risk on fresh untrained compositions; the unresolved reference check above limits posterior conclusions from this earlier study.

M VISUAL BOUNDARY CHECK FOR STRUCTURED INFERENCE

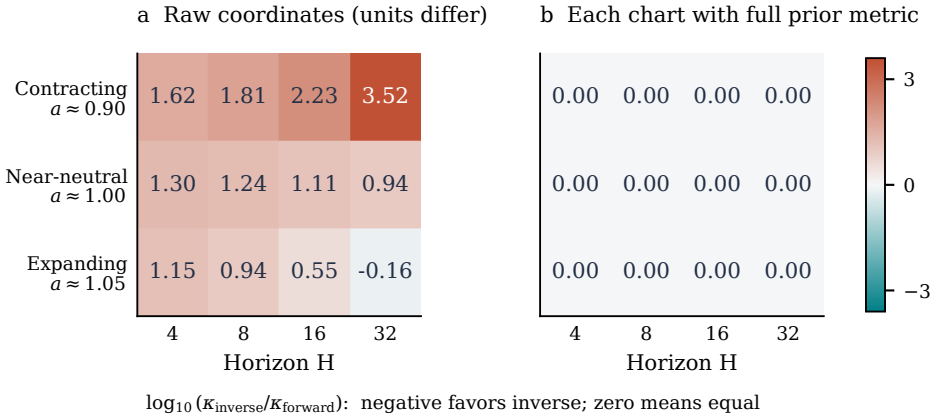

Figure 11: **Supplying the full prior metric removes the affine chart gap.** Exact affine-Gaussian calculations. Each entry is the median log condition-number ratio across four development worlds, with terminal observations and process/observation standard deviations .05. The two charts represent the same posterior. Raw coordinates have different units: the forward coordinates standardize prior innovations, whereas the terminal coordinate is unstandardized, so the raw ratio is not an intrinsic comparison. After each chart receives its full prior metric, all displayed ratios round to zero on the log scale. The shared color scale is centered at equality; negative values favor the inverse chart. Independent Kalman/RTS smoothing validates the reference algebra.

This view uses 48 terminal-observation cells from the 120-cell affine-Gaussian coordinate study. The other cells vary observation masks or process noise. The calculation qualifies raw conditioning illustrations. Section C separately tests a nonlinear stochastic chain; it does not inherit a coordinate advantage from this affine calculation.

N ALGEBRA OF PRESERVATION AND REVISION

Proof of Proposition 2. Use the common dominating measure $\lambda = \sum_c J_c$ to write component densities, or apply the same finite expansions directly to encounter measures. For a context law μ, let $\bar{w}_c = \int w_c \, d\mu(w)$. Linearity gives $J_\mu(v) = \sum_c \bar{w}_c J_c(v)$, so equal context means give identical encounter measures. All their marginals agree; conditional event probabilities agree almost everywhere with respect to the shared evidence law, or pointwise wherever the chosen density versions have positive finite denominators.

After b additional encounters, the unnormalized density of k future encounters is

$$\int A(w) \prod_{i=1}^{k} J_w(v_i) d\mu(w).$$

Each factor is affine, so this integrand is a polynomial of degree at most $b+k$. The update normalizer is $\mu(A)$, with degree at most b; event numerators have degree at most $b+k$. Under the stated positive finite normalizers, equal polynomial expectations therefore give identical normalized joint laws and relational conditionals. The Gaussian mixtures used here have positive finite relevant densities. □

This sufficient condition concerns the declared query and revision budget. Preserving moments through degree six does not preserve arbitrarily long future evidence products. Nor does it imply that every smaller representation fails: redundant questions, degeneracies and other encodings can reduce what must be retained.

For any integrable f and nonnegative evidence factor A with $0 < \mu(A) < \infty$, provided $\mu(|f|A) < \infty$,

$$\mu_A(f) - \mu(f) = \frac{\mu(fA) - \mu(f)\mu(A)}{\mu(A)} = \frac{\mathrm{Cov}_\mu(f, A)}{\mu(A)}.$$

If f is a query conditional within a context, the measure in this identity must already account for that query's observed outcomes. Integrating an already conditioned $p_w(x \mid y)$ under an unreweighted context posterior generally does not equal conditioning the integrated joint.

An added positive factor can be retracted from the same atom state by dividing its masses by that factor and renormalizing. This statement assumes the factor and atom support are retained. It is not a promise to recover arbitrary discarded history after repeated recompression. The saved diagnostics check addition-order invariance of compressed masses and restoration from a retained base state with an empty factor set. They do not numerically test factor division from an updated state or reevaluate restored query answers.

The first-encounter inclusion applies directly to the original affine joint interfaces. The earlier fitted Gaussian priors are retained as fixed point states; this proposition does not imply that a first moment suffices for arbitrary uncertainty over Gaussian parameters. Full BCTS introduces corrected component probabilities and its own normalization; its earlier empirical gains retain their own interface contract. They are not automatically laws of the new uncalibrated relational joint.

O Information that an observation can identify

For a known channel K and allowed prior class $\mathcal{P}$, define the observation operator $Tq = \int K(\cdot \mid x)q(dx)$. The prior family is globally identifiable from the population observation law precisely when $\ker T \cap (\mathcal{P} - \mathcal{P}) = \{0\}$. Identification of a particular prior requires its own observation fiber to be a singleton. A query $Q(q)$ is identified if it is constant on each feasible fiber $\{q \in \mathcal{P} : Tq = \nu\}$; identifying all of q is stronger.

For a bounded function h, let $T_h q = \int h(x)K(\cdot \mid x)q(dx)$. On a nonempty observation fiber, the conditional expectation of $h(X)$ is common to all feasible priors if and only if $T_h(q - p) = 0$ for every feasible difference in that fiber. The denominator measure is the same, so equality of the Radon–Nikodym derivative classes is equivalent to equality of the numerator measures.

For a finite latent space and the full probability simplex, a linear query $f^\top q$ is identified by Mq if and only if $f \in \mathrm{rowspan}([M; \mathbf{1}^\top])$. Sufficiency follows by writing the query as a linear combination of observed constraints and total mass. For necessity, any vector outside that row space has a nonzero pairing with some vector in its null space. Small positive and negative perturbations of an interior prior along that null vector remain feasible and observationally indistinguishable. Restricted prior families or boundary fibers require their actual feasible differences, so unrestricted null-space necessity cannot simply be carried over.

In the symmetric two-component experiment, $H_+(y) = H_-(y)$ because the cause components are reflections and the outcome depends on X^2. All weights therefore produce the same outcome law. A sign-sensitive cue changes the observation operator; its population mean identifies the mixture weight. Bayesian uncertainty represents this distinction but cannot create information absent from the observation.

P Relational implementation and numerical scope

The regression feature vector is $[1, x, \exp(-\frac{1}{2}((x - c_j)/.35)^2)_{j=1}^{20}]$, with centers equally spaced from -3 to 3. Ridge penalty is 10^{-3} on every coefficient except the intercept. The residual standard deviation is the square root of mean training squared error, without a degrees-of-freedom adjustment; its values are .24946 and .24952. The 128 separate validation pairs diagnose fit and reversibility and do not select a checkpoint or basis.

The context rule maps $(u, v) \in [0, 1]^2$ to $((1 - u)(1 - v), u, (1 - u)v)$. A uniform simplex prior has factor $2(1 - u)$ in these coordinates. Thirty-six Gauss–Legendre nodes on each axis integrate the degree-70 likelihood/query products and this Jacobian through degree 71 in each transformed variable. The resulting 1,296-point reference is therefore exact for these finite context polynomials

in exact arithmetic, conditional on the computed component integrals. It is not exact integration of the learned continuous cause densities.

For component integrals, the boundaries -5, $-.75$, 0, $.75$ and 5 define four intervals with 96 Gauss–Legendre points each. Saved 192-point and denser-context checks cover the first four query groups per fit/target. Their maximum answer discrepancies are 8.78×10^{-7} for the learned channel, 4.13×10^{-8} for truth and 2.67×10^{-15} for context resolution. The learned discrepancy is close to the frozen 10^{-6} tolerance. Twenty-two stored probabilities exceed one by at most 1.56×10^{-15} from rounding; they remain in the original record without post hoc clipping or a changed success gate.

Learned feature coordinates use the outcome likelihood at $Y = .25$ and the positive-cause joint likelihood at $Y = 1$. Their full affine rank makes their complete degree-six products an invertible basis change of generic polynomial coordinates. The same positive linear program, posterior support and deterministic sine objective are used in both constructions. Maximum recorded span-reconstruction residual is 8.89×10^{-12}. Each solution has 28 positive atoms. Reported atom counts omit basis storage, integration and compression cost and hence do not establish total-memory or runtime advantage.

The fitted conditional cause/noise map has maximum roundtrip error 4.44×10^{-16}. Reordering the two added factors leaves the checked compressed mass vectors unchanged at recorded precision. Restoring the retained base state with an empty factor set changes masses by at most 2.78×10^{-17}. These are mass and base-state reconstruction checks, not numerical tests of inverse-factor retraction or restored query answers. They do not establish density accuracy, human cognition or a benefit of invertibility by ablation.

Q Earlier calibration interface and full final-count table

The twenty neural models use 8,192 source pairs and 4,000 fixed optimization steps each. The source calibration mixture is balanced with component standard deviation .45, giving $r(x) = \mathrm{sigmoid}(2x/.45^2)$. BCTS fits $a(y) = \mathrm{sigmoid}(\tau \log[H_+(y)/H_-(y)] + b)$ with $\tau \geq 0$ and free b, using Eq. (18). Let p_0 be the mean calibrated source prediction. The target likelihood factor, up to a weight-independent multiplier, and corrected positive-component probability are

$$D_w(y) = \frac{(1-w)(1-a(y))}{1-p_0} + \frac{wa(y)}{p_0}, \qquad \rho_{w,+}(y) = \frac{wa(y)/p_0}{D_w(y)}.$$

ML optimizes the summed log factors; Bayesian adaptation integrates their product with a uniform weight prior. Calibration uncertainty is not integrated. The joint-only arm inserts an estimated weight into the original joint. The full conditional arm uses Eq. (19); its frozen reference has $w = p_0$. Raw, affine and exact-channel comparisons retain their corresponding original joint references. The saved record also contains absolute NLL and differences from a common raw reference, so arm-relative changes must not be read as absolute rankings.

At $a = .1$, exact-channel ML has one material-harm cell; at $a = 0$ it has none. Full-BCTS source cross entropy worsens on the independent source set in twelve of twenty models, including seven of ten near-boundary models. These controls limit attribution: a fitted correction can generalize poorly, and finite-sample weak identification can harm a downstream query even with an exact channel. Neither mechanism licenses a verdict on all calibrated Bayesian adaptation.

R Implementation of the context expansion

The expanded experiment uses new source fits, histories and query groups, analyzed separately from the original probe. There are 48 source fits and 384 target histories nested within them. The same query groups are reused at the 64- and 66-observation readouts. Source-fit resampling retains all associated histories, target directions and paired methods. The 4,000 bootstrap resamples give descriptive percentile intervals for the evaluated source-fit settings.

Table 10: Every calibration interface at the final count of 4,096. Each row includes five seeds and two shifts. Positive changes mean worse continuous reverse NLL in nats against the interface's own frozen output. The two BCTS uses have identical displayed changes in these saved aggregates but retain different density contracts and absolute references.

a	Interface	ML harm $> .1$	Mean ML change	Mean Bayes change
0.00	raw	9/10	5.002	0.047
0.00	affine	8/10	4.930	0.043
0.00	bcts_joint	4/10	1.838	0.054
0.00	bcts_conditional	4/10	1.838	0.054
0.00	oracle	0/10	0.000	-0.000
0.03	raw	2/10	-0.157	-0.294
0.03	affine	5/10	0.026	-0.255
0.03	bcts_joint	6/10	0.046	-0.258
0.03	bcts_conditional	6/10	0.046	-0.258
0.03	oracle	3/10	-0.107	-0.298
0.10	raw	1/10	-0.270	-0.329
0.10	affine	1/10	-0.268	-0.333
0.10	bcts_joint	3/10	-0.135	-0.327
0.10	bcts_conditional	3/10	-0.135	-0.327
0.10	oracle	1/10	-0.269	-0.335
0.60	raw	0/10	-0.249	-0.249
0.60	affine	0/10	-0.249	-0.249
0.60	bcts_joint	0/10	-0.249	-0.249
0.60	bcts_conditional	0/10	-0.249	-0.249
0.60	oracle	0/10	-0.249	-0.249

What is estimated. The scalar channel keeps the original fixed RBF features. The two-dimensional channel uses an intercept, the two coordinates and 49 Gaussian radial-basis features on a 7×7 grid over $[-2.5, 2.5]^2$, with width $.55$. Both fit ridge coefficients with penalty $.001$ except on the intercept, and estimate diagonal residual variances from source residuals. An additional 256 source pairs diagnose held-out fit without selecting a basis or checkpoint. The source mixture uses three diagonal Gaussian components, five fixed EM initializations, a 500-iteration cap, tolerance 10^{-5} and variance regularizer 10^{-4}. Training likelihood selects among the five starts. No target score selects a component model. Its source weights serve the mixture fit; the target context prior remains uniform Dirichlet$(1, 1, 1)$ for every arm. Uncertainty in fitted component shapes and channel parameters is not integrated.

Matched inference and numerical checks. Every method uses the same fitted component likelihoods and relational composition. The ML baseline maximizes the observed-outcome mixture likelihood over the three-component simplex. Its implementation checks active faces and an interior profile and verifies a floating-point first-order likelihood-gap bound of 10^{-12} nats per observation. Numerically tied candidates within $32\epsilon_{64}$ nats use a deterministic minimum-norm rule. Likelihood optimality does not guarantee that weakly identified weights are accurate.

The context rules use 36^2 and 44^2 points. Continuous cause integrals use piecewise Gaussian quadrature: 256 versus 512 points per interval in one dimension, and 128 versus 192 per interval on each axis in two dimensions. The first-coordinate boundaries are $-5, -.75, 0, .75, 5$; the second-coordinate boundaries are $-5, 0, 5$. All scored query groups and reverse scores receive coarse/fine checks, including recomputation of history updates. The maximum conditional likelihood times a Gaussian outside-box tail bound also bounds omitted component evidence relative to its computed density. We separately record fixed-weight ML query integration error, changes in optimized weights and the likelihood optimality certificates. End-to-end refinement sensitivity must not be attributed to quadrature alone in a nearly unidentified problem.

Source observations, fitted coefficients, mixture parameters, target histories, test encounters, numerical component integrals, context states and raw query probabilities support recalculation of the reported scores. Independent checks cover information access, normalization, scoring and the ML solver. Longer update streams and repeated recompression are left for future investigation.

Table 11: All shape/channel crossings for full Bayesian adaptation at 64 observations. Shapes and channel parameters labeled given/true are privileged diagnostics. Gains use each arm's own frozen Bayesian reference; the excess-risk column supplies absolute quality against the same world conditional. Source fits are shared within each setting's crossed comparison.

Setting	Shapes / channel	Excess Brier	Brier gain	Reverse gain
Scalar, weak	Given / fitted	0.04573	0.00204	0.025
Scalar, weak	Estimated / fitted	0.04604	0.00158	0.023
Scalar, weak	Given / true	0.04396	0.00348	0.034
Scalar, weak	Estimated / true	0.04355	0.00373	0.036
Scalar, informative	Given / fitted	0.00703	0.03123	0.221
Scalar, informative	Estimated / fitted	0.00780	0.03077	0.222
Scalar, informative	Given / true	0.00687	0.03118	0.222
Scalar, informative	Estimated / true	0.00770	0.03072	0.224
2D, one sensor	Given / fitted	0.05201	0.00811	0.066
2D, one sensor	Estimated / fitted	0.05420	0.00601	0.056
2D, one sensor	Given / true	0.05425	0.00665	0.054
2D, one sensor	Estimated / true	0.05685	0.00410	0.039
2D, two sensors	Given / fitted	0.00133	0.00424	0.099
2D, two sensors	Estimated / fitted	0.00140	0.00410	0.098
2D, two sensors	Given / true	0.00089	0.00436	0.099
2D, two sensors	Estimated / true	0.00091	0.00426	0.098

S NUMERICAL CHECKS AND REPRODUCIBILITY LIMITS

The tables and figures use saved experimental outcomes. For the primary 24-world joint study, the comparisons, thresholds and policy were fixed using separate development worlds before held-out evaluation. Each reference cell has two independently constructed reference banks. All 96 learned/true reference cells and 288 analytic/numeric operator checks pass the declared gates. The largest grid CDF refinement difference is .00519 against the .006 tolerance; the largest reference-bank MMD floor is 1.89×10^{-6} against 4×10^{-6}. World, rather than sampler repeat, is the unit of the reported bootstrap intervals. The same learned posterior target and complete-cost accounting apply to the direct-inverse and continuous-grid comparisons.

The earlier local-recovery study has a numerical qualification: its reference check extended the integration domain without refining grid spacing, so that check alone does not certify resolution or joint-posterior accuracy. Its learning rate was chosen on three development libraries before twelve held-out libraries were evaluated. The later 32-world learner diagnostics and twelve-world proposal comparison reuse previously evaluated worlds and are not fresh confirmations. Posterior refinement was checked for each reported learner and query in the 32-world comparison; one PCHIP cell needed a finer grid. The twelve-world proposal reference used 32,769- and 65,537-point grids, with maximum score change 7.31×10^{-8}. Paired intervals bootstrap worlds; Holm-adjusted sign tests address direction consistency rather than mean effect.

For the short stochastic-chain study, reported CDF values were recomputed on the same saved samples because the original weighted-particle CDF interpolated between samples. The longer stochastic-chain study uses eight fresh worlds, four mechanisms per world, 1,024 noisy forward pairs and 256 validation pairs per mechanism, a two-component Gaussian initial prior, process-noise standard deviation .02, and terminal/midpoint observation-noise standard deviation .05. Each world supplies one query at each length and both evidence states. All 40 prescribed apparatus/reference checks pass; the largest 32,769/65,537-grid root-CDF refinement change is .00443 and the largest independently checked true-law refinement change is .00011. A later root-only grid-cost sweep reused those same worlds. The study did not include matched sparse smoothing, HMC or particle-filter controls. Its 512-path FFBS reference does not resolve fine joint-trajectory scores; the table therefore reports root marginals.

For the initial relational probe, fitted coefficients, encounter datasets and posterior atom states were not retained, limiting independent reanalysis without rerunning its code.