COORDINATEWISE PRECISION ADAPTATION FOR FROZEN WORLD MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Local sensor faults change the reliability of observations used to infer a world model's latent state. We combine a calibrated diagonal base with causal innovation statistics for coordinatewise precision adaptation on frozen world models. We evaluate forty predictive coding world model (PCWM) and Gaussian recurrent state-space model (RSSM) checkpoints in two low-dimensional environments. Under local variance faults, one-step error decreases by 5.56–17.81% relative to independently calibrated static weights. Same-base controls retain 6.29–15.43% improvements on PCWM. A prospective comparison gives static and adaptive refinement the same filtering-time cap and four optimizer families, keeping their precision calibration fixed. Validation-selected adaptive configurations reduce local-fault error by 5.60–24.90% across all four settings relative to validation-selected static configurations. A separate multistep development evaluation yields 5.3–22.1% lower fault-condition error than static refinement at five transitions, with gains persisting at twenty transitions. These findings support coordinatewise state refinement for frozen world models under localized sensor corruption.

1 INTRODUCTION

World models infer latent states and predict action consequences (Ha & Schmidhuber, 2018; Hafner et al., 2019; 2020). Their predictions depend on how observations initialize those states. Local sensor faults change the reliability of individual coordinates while the learned dynamics can remain useful. State refinement offers a targeted deployment adjustment at this interface.

Classical adaptive filtering couples state and noise estimation (Särkkä & Nummenmaa, 2009; Särkkä & Svensson, 2023). Frozen-model residuals also contain representation and nonlinear decoding errors. Precision adaptation must combine learned encoder proposals and prior variances with calibrated residual scales.

Our controller retains the trained PCWM and Gaussian RSSM backbones (Figure 1). Causal innovation statistics modulate a calibrated diagonal base. Common temporal adaptation (T) changes a shared multiplier, while coordinatewise adaptation (C) changes individual weights. Disabling adaptation recovers the static update exactly.

We compare calibrated static, Huber, and inverse-multiquadric (IMQ) references under common noise switches and local faults. Our contributions connect coordinatewise control, same-base comparisons, and prediction under a common filtering-time allowance. Optimizer selection tests whether additional static optimization recovers the gains. Multistep forecasts assess how corrected states affect longer predictions.

2 RELATED WORK

World models and their use. Ensembles represent dynamics uncertainty (Lakshminarayanan et al., 2017; Chua et al., 2018), and short rollouts limit its impact during policy learning (Janner et al., 2019). Visual prediction and control span diverse architectures (Kaiser et al., 2020; Hafner et al., 2021; Micheli et al., 2023; Alonso et al., 2024). Reward and value predictions support planning (Schrittwieser et al., 2020; Hansen et al., 2022; 2024). Extensions include scalable imagination

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.

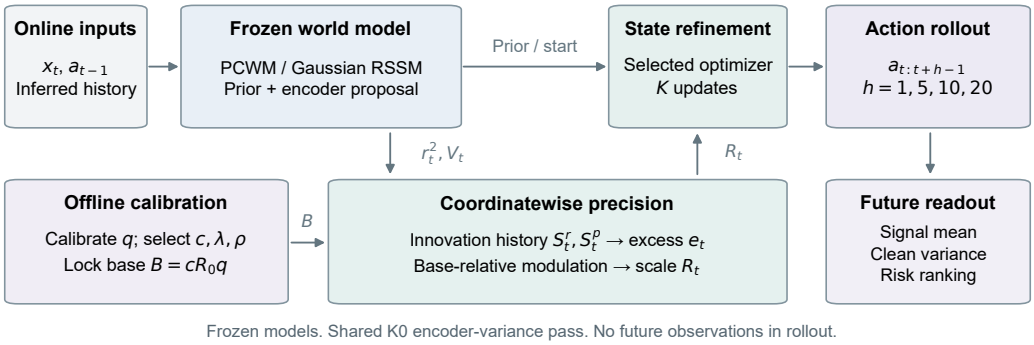

Figure 1: **From changing observations to future predictions.** Offline calibration and validation fix the precision controller. Online innovations modulate coordinate weights before a latent-state update. The corrected state initializes the next observed transition and the action-conditioned rollout. Decoded particles provide clean-signal means and variance scores. The encoder, transition model, and decoder remain frozen. Methods share the numerical K0 encoder variance and within-backbone prediction random streams.

(Hafner et al., 2025), temporal abstraction (Shaj et al., 2023), and language-conditioned learning (Hansen et al., 2026).

Latent-state inference. Recurrent models and structured variational inference represent latent-state uncertainty (Chung et al., 2015; Fraccaro et al., 2016; Krishnan et al., 2017; Karl et al., 2017; Fraccaro et al., 2017). Learned filtering includes factorized latent covariance updates (Becker et al., 2019), KalmanNet (Revach et al., 2022), and differentiable particle filters (Jonschkowski et al., 2018). Semi-amortized and iterative amortized inference refine initial proposals (Kim et al., 2018; Marino et al., 2018b), and amortized variational filtering learns iterative sequential updates (Marino et al., 2018a). Predictive coding connects precision-weighted errors to latent-state updates (Rao & Ballard, 1999; Bogacz, 2017).

Adaptive and robust filtering. Gaussian filters balance prior uncertainty and observation noise (Kalman, 1960; Särkkä & Svensson, 2023). Adaptive methods estimate noise through innovation covariance matching (Mehra, 1970) or coupled state/noise posteriors (Särkkä & Nummenmaa, 2009). Iterated filtering connects such state updates to Gauss–Newton optimization (Bell & Cathey, 1993). Robust estimation reduces the influence of large residuals (Huber, 1964). Generalized Bayesian updating admits loss-based observation contributions (Bissiri et al., 2016), and WoLF uses likelihood weighting with extended or ensemble Kalman filtering (Duran-Martin et al., 2024). Our frozen-model interface compares accumulated statistics and instantaneous robust weights. Encoder proposals initialize latent refinement, with residual scales held fixed during inner iterations. Decoder curvature or prior variance supplies diagonal preconditioning.

3 Calibrated Precision Adaptation

3.1 A frozen world model and its refinement interface

Let $x_t \in \mathbb{R}^{D_x}$ and a_{t-1} denote the current observation and preceding action. A frozen world model supplies a latent prior mean μ_t, diagonal prior variance v_t, and encoder proposal $z_t^{(0)}$. Write the decoder mean as $g(z)$, conditional on RSSM's current recurrent state, and its learned residual variance as R_0. An effective residual scale R_t has variance units, with observation precision R_t^{-1}. For PCWM, the prior is formed by diagonal moment matching across five transition members. For RSSM, it is conditional on the recurrent state. The recurrent state is held fixed during the current stochastic-state update and evolves with the corrected state at the next transition.

Define the innovation and decoder-projected prior variance as

$$r_t = x_t - g(\mu_t), \qquad V_t = \operatorname{diag}\big[J_g(\mu_t)\operatorname{diag}(v_t)J_g(\mu_t)^\top\big]. \tag{1}$$

An observation-space residual scale R_t specifies the state objective

$$F_t(z) = \tfrac{1}{2}\|x_t - g(z)\|^2_{R_t^{-1}} + \tfrac{1}{2}\|z - \mu_t\|^2_{v_t^{-1}}, \qquad \|u\|_w^2 = \sum_j w_j u_j^2. \tag{2}$$

We initialize at the encoder proposal and use a diagonal Gauss–Newton preconditioner evaluated at that proposal,

$$D_t = v_t^{-1} + \operatorname{diag}\Big[J_g(z_t^{(0)})^\top \operatorname{diag}(R_t^{-1})J_g(z_t^{(0)})\Big], \tag{3}$$

$$z_t^{(k+1)} = \operatorname{clip}_{[-30,30]}\left(z_t^{(k)} - \eta D_t^{-1} \odot \nabla F_t(z_t^{(k)})\right). \tag{4}$$

The initial precision comparison fixes $K = 1$ and $\eta = 0.5$. The preconditioner remains fixed when evaluating larger iteration budgets. The optimizer comparison also evaluates prior-scaled gradient descent, Adam, and Gauss–Newton with backtracking (Section 4.2). Corrected states and controller statistics carry forward throughout each trajectory. The visual pilot estimates decoder curvature with randomized probes (Appendix E). Appendix D.3 uses a fixed linear decoder to explain step-size sensitivity.

3.2 Calibrate a base, then modulate its precision

Calibration supplies a coordinate shape q from encoder-only histories. For checkpoint s, normalize excess squared innovations by its learned variance $R_{0,s}$. The shape is

$$\begin{aligned} b_j &= \left\langle \operatorname{clip}\left(\frac{r_{t,j}^2 - V_{t,j}}{R_{0,s,j}}, 1, 16\right)\right\rangle_{\text{cal}}, \\ \ell_j &= \log b_j - \frac{1}{D_x}\sum_k \log b_k, \qquad q_j = \exp\left[\frac{\ell_j}{\max(1, \|\ell\|_\infty/\log 4)}\right]. \end{aligned} \tag{5}$$

Calibration assigns equal weight to nominal and target conditions. In the common-switch suite, both switching directions receive equal weight. We average transitions and trajectories within each condition, then average the ten checkpoints. Consequently, q has geometric mean one and entries in $[1/4, 4]$. Each backbone, environment, and suite shares one q. Write $Q_s = R_{0,s} \odot q$ and $B_s = cQ_s$, where validation selects c. We suppress the checkpoint index below.

Two exponential moving averages (EMAs) summarize residual and prior variation:

$$S_t^r = (1-\rho)S_{t-1}^r + \rho r_t^{\odot 2}, \qquad S_0^r = B, \tag{6}$$

$$S_t^p = (1-\rho)S_{t-1}^p + \rho V_t, \qquad S_0^p = 0, \tag{7}$$

$$e_t = \operatorname{clip}(S_t^r - S_t^p, Q/16, 128Q). \tag{8}$$

Past squared residuals receive weights $\rho(1-\rho)^i$ at lag i. This exponential memory smooths fluctuations and introduces a response lag after abrupt changes. The one-step study fixes $\rho = 0.2$. The multistep study selects ρ on its own validation split (Section 4.3). Common and coordinatewise modulation then use

$$\gamma_t = D_x^{-1}\sum_j e_{t,j}/B_j, \qquad R_{t,j}^{\mathrm{T}} = B_j\gamma_t^\lambda, \tag{9}$$

$$R_{t,j}^{\mathrm{C}} = B_j(e_{t,j}/B_j)^\lambda, \tag{10}$$

where $\lambda \in \{0, 0.25, 0.5, 1\}$ controls shrinkage towards the base. T preserves the coordinate ratios of B, while C adjusts them. Both recover static refinement at $\lambda = 0$. At $\lambda = 1$, C reduces to the clipped innovation covariance-matching statistic $R_t = e_t$ (Mehra, 1970). The update carries a corrected mean and obtains the next prior variance from the learned transition model. Clipping acts on e_t before modulation and gives $e_{t,j}/B_j \geq 1/(16c)$. Thus, $c \leq 1/16$ restricts both rules to $R_t \geq B$.

Algorithm 1 Calibrated adaptation on a frozen world model

Require: Frozen backbone, calibration histories, and labeled validation histories
Calibrate q and lock precision and optimizer settings using the study's validation rule.
for each new observed history **do**
  Run the shared K0 history to cache encoder variances at prediction anchors.
  Set $B_s = cQ_s$, $S_0^r = B_s$, $S_0^p = 0$ and initialize from x_0.
  for each transition $t = 1, \dots, T$ with x_t and a_{t-1} **do**
    Compute the latent prior and encoder proposal from this arm's history.
    Set $R_t = B_s$.
    if $\lambda > 0$ **then**
      Compute (r_t, V_t) and update both accumulators and e_t.
      Set R_t with the selected rule in Eqs. 9–10.
    end if
    Apply the study's selected state optimizer and carry the corrected state forward.
    At prediction anchors, forecast with future actions and shared K0 initial variance.
  end for
end for
Ensure: Corrected latent states, clean-signal predictions, and rollout variance scores

Online updates use the current observation, preceding action, and inferred history. Clean signals provide offline validation targets and final scores.

Write relative precision as $w_{t,j} = B_j / R_{t,j}$. T assigns every coordinate the multiplier $\gamma_t^{-\lambda}$. For $\lambda > 0$, increasing one excess statistic therefore downweights every coordinate. C assigns $(e_{t,j}/B_j)^{-\lambda}$ separately, preserving other weights when their statistics stay fixed. Shared latent dynamics can still couple coordinates during recursive filtering.

3.3 Prediction and matched alternatives

K0 denotes encoder-only filtering with $K = 0$. K1 denotes a configuration with one state update. At prediction anchors, particles use the corrected state mean and diagonal encoder variance from the K0 history. An additional K0 filtering pass supplies these shared encoder variances. The transition-prior variance v_t enters refinement. Methods share particle random streams. The one-step study uses $P = 150$, allocating thirty particles to each PCWM transition member or propagating 150 RSSM recurrent particle states. The transition model receives future actions, and the decoded particle mean predicts the clean signal. The multistep study also scores the variance of decoded particles, excluding the decoder observation variance R_0.

The static reference uses $R_t = B$ throughout each trajectory. The robust references use

$$u_{t,j} = \frac{r_{t,j}}{\sqrt{V_{t,j} + B_j}}, \qquad w_{\mathrm{H},t,j} = \max\{10^{-4}, \min(1, \tau_{\mathrm{H}}/|u_{t,j}|)\},$$
$$w_{\mathrm{I},t} = \max\left\{10^{-6}, \frac{\tau_{\mathrm{I}}^2}{\tau_{\mathrm{I}}^2 + \sum_j r_{t,j}^2}\right\}. \tag{11}$$

Huber uses weight one at $u_{t,j} = 0$ and acts coordinatewise. IMQ uses one multiplier based on the innovation norm in observation units. Both set $R_t = B/w_t$ in the same refinement objective. These references transplant robust weight functions into the shared latent update. Validation selects the robust reference from the two families before test generation. An exploratory follow-up replaces the IMQ innovation norm with $\sum_j u_{t,j}^2$ (Appendix A.8).

T and C additionally maintain the two innovation accumulators. Their effective residual scales can move above or below an interior base. Huber and IMQ yield $R_t \geq B$. Algorithm 1 gives the deployment sequence.

4 EVALUATION UNDER SENSOR VARIATION

4.1 ONE-STEP EVALUATION

Backbones and data. We evaluate ten PCWM and ten Gaussian RSSM checkpoints in each of two custom stochastic swing-up environments, Pendulum and Cartpole. Both backbones use 32-dimensional stochastic states and two-layer, width-256 decoders. PCWM has five transition members. RSSM uses a 128-dimensional recurrent state. Each checkpoint used 180 training episodes and 6,000 optimization steps, with batches of 128 sixteen-transition subsequences. The training losses and environment dynamics appear in Appendices D.1 and D.

Clean signals are $(\cos\theta, \sin\theta, \dot{\theta}/8)$ for Pendulum and $(p/2.4, \dot{p}/5, \cos\theta, \sin\theta, \dot{\theta}/10)$ for Cartpole, where p is cart position and θ is pole angle. These signal units define noise injection and scoring.

Each suite has eight calibration, eight validation, and twenty test physical trajectories per environment and checkpoint seed. The two backbones share the corresponding physical trajectories. Every trajectory has 160 transitions, with predictions at anchors $10, 20, \ldots, 150$. Actions follow $a_t = \mathrm{clip}_{[-1,1]}(0.8a_{t-1} + 0.6\epsilon_t)$, where $a_{-1} = 0$ and $\epsilon_t \sim \mathcal{N}(0, 1)$. Filtering proceeds continuously from the initial observation. The open-loop forecast uses the next recorded action and targets the clean signal of the realized stochastic state.

Calibration and validation represent the target corruption families. Test event timing, affected coordinates, and severity are unknown during filtering.

Two prespecified target suites. The common noise switch suite alternates observation-noise standard deviations 0.05 and 0.20. Its two switch times are uniform integers in $[35, 55]$ and $[95, 115]$. It pairs schedules starting at the low and high noise levels on the same physical trajectories. T is the prespecified primary adaptive arm for their equally weighted whole-trajectory endpoint. The local suite changes the noise of two distinct randomly chosen observation coordinates during nonoverlapping intervals. Random onset and recovery times define two successive fault intervals (Appendix A.2). Faulty-coordinate standard deviations are sampled uniformly from $[0.20, 0.40)$. Healthy coordinates retain 0.05. C is its prespecified primary arm. Both controllers are evaluated on both suites. Nominal and stationary-high sensing provide matched secondary conditions. The local suite also includes stuck-sensor and bias-drift extrapolations. Appendix A details the transfer constructions.

Selection before test generation. The static grid contains $c = 2^{k/4}$ for $k = -16, \ldots, 28$. T and C search $c = 2^k$ for $k = -4, \ldots, 7$, crossed with the four λ values. Huber and IMQ use these twelve scales with thresholds $\tau_{\mathrm{H}} \in \{0.5, 1, 2, 4\}$ and $\tau_{\mathrm{I}} \in \{0.1, 0.3, 1, 3\}$, respectively. Each adaptive or robust family receives 48 candidates and the selected fine-grid static configuration as a shared fallback.

Let $L^{\mathrm{val}}_{s,k}(a)$ be validation clean-signal mean squared error (MSE) for checkpoint s, condition k, and method a. The normalized objective is

$$
\begin{aligned}
N_k(a) &= \frac{1}{10} \sum_s \frac{L^{\mathrm{val}}_{s,k}(a)}{L^{\mathrm{val}}_{s,k}(\mathrm{K0})}, \\
J(a) &= \tfrac{1}{2} N_{\mathrm{nominal}}(a) + \frac{1}{2|\mathcal{K}|} \sum_{k \in \mathcal{K}} N_k(a).
\end{aligned}
\tag{12}
$$

Static selection minimizes J over its fine grid. The set $\mathcal{K}$ contains the two noise-switch schedules or the single local variance-fault condition. Each other family minimizes J subject to $N_{\mathrm{nominal}}(a) \leq 1.05 N_{\mathrm{nominal}}(\mathrm{static})$. The robust reference is the better validation winner of Huber and IMQ. Its combined pool has 96 family-specific candidates. The static, adaptive, and robust references each select their own scale. All selections and calibration-derived controls are locked before test generation. All eight selected static scales are interior to their grid. The four RSSM primary adaptive scales equal the lower bound $1/16$. Their accumulated residuals therefore act in the downweighting direction.

Endpoint and statistical unit. One-step clean-signal MSE averages squared prediction error across observation coordinates. Within each checkpoint, we then average anchors, physical trajectories, and the specified target conditions with equal weights. For method a and reference b, the paired effect is

$$\Delta(a, b) = \frac{100}{10} \sum_{s=1}^{10} \left[\frac{L_s(a)}{L_s(b)} - 1 \right]. \tag{13}$$

Negative $\Delta(a, b)$ indicates improvement. The primary family contains 2 suites $\times$ 2 backbones $\times$ 2 environments $\times$ 2 references, giving sixteen comparisons with one Holm correction (Holm, 1979). Two-sided sign-flip tests enumerate all 2^{10} sign assignments of checkpoint-paired effects. Their exactness requires sign symmetry under the paired null (Hemerik & Goeman, 2018). We report pointwise 95% Student-t intervals. The intervals and tests are conditional on the existing checkpoints and locked selection. Eight separately corrected TOSTs assess stationary-high equivalence within a prespecified ± 2 percentage-point margin (Schuirmann, 1987).

4.2 Optimizer comparison under a common time cap

We test local-fault prediction when both precision rules can select their state optimizer. The original calibration shape, scales, λ, and ρ remain fixed for each rule. Static and C each search 48 candidates from four equally sized families: fixed diagonal Gauss–Newton, prior-scaled gradient descent, Adam (Kingma & Ba, 2015), and Gauss–Newton with Armijo backtracking. Each family crosses $K \in \{1, 2, 4\}$ with four step sizes. All retain Eq. 2, the encoder initialization, and latent clipping. Each transition holds R_t, the prior, and RSSM's recurrent context fixed during optimization. Each optimizer carries forward its own corrected states and causal innovation statistics.

Both rules receive the same filtering-time cap, 1.05 times the original C/GN1 median timing. Synchronized timing uses three rounds on the first validation trajectory of checkpoint zero. Timing includes diagnostics and backtracking, excluding the common K0 pass and future rollout. Validation minimizes Eq. 12 under this cap and the original static/GN1 nominal bound. The original static/GN1 remains a shared fallback. Thus, static refinement can spend its lower controller cost on additional optimization.

Selections precede generation of 400 fresh physical test trajectories, shared across backbones. The forty existing checkpoints retain the original anchors and $P = 150$ particle budget. Two C configurations, original GN1 and validation-selected, each compare with selected static across four settings. These eight comparisons share a separate Holm correction and the same checkpoint-level statistics. Appendix B gives the grids, selections, and measured time ratios.

4.3 Multistep development evaluation

We separately evaluate the forty checkpoints at $h \in \{1, 5, 10, 20\}$. Calibration, validation, and development holdout each contain eight physical trajectories per environment and checkpoint. Nominal and local-fault conditions share trajectories, actions, and measurement draws across backbones. Anchors $10, 20, \ldots, 140$ on 160-transition histories serve all horizons.

The study recalibrates each controller on its own splits and uses $P = 60$ particles with the common K0 encoder variance. Validation selects (c, λ, ρ) with $\rho \in \{0.05, 0.2, 0.5, 1\}$ for C and T. Every family shares the static fallback and 5% nominal validation constraint. Selections precede holdout generation (Appendix C).

The multistep risk averages nominal and fault MSE at $h = 5, 10, 20$. Each term uses its checkpoint's independent calibration K0 MSE as denominator, floored at 10^{-8}. We average checkpoint-paired percentage changes within each setting. These descriptive development estimates use separate configurations and sampling budgets.

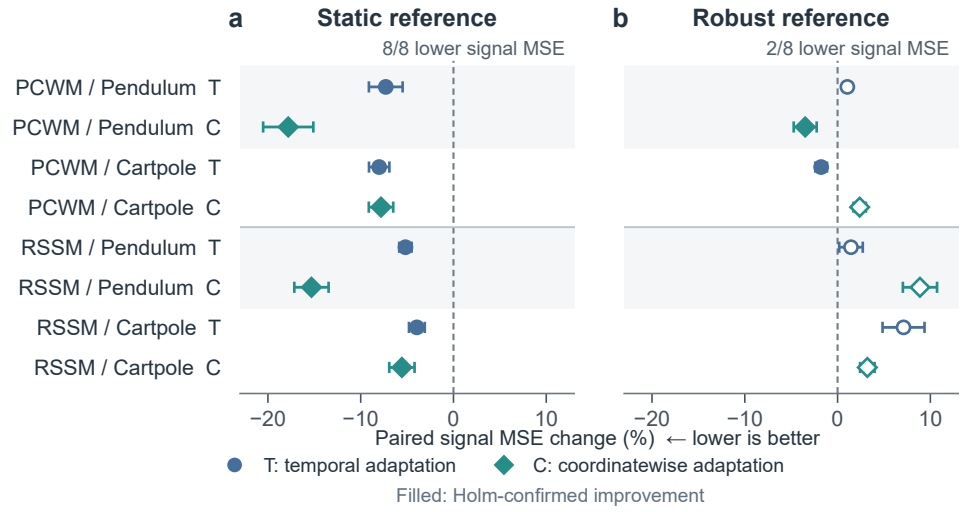

Figure 2: **One-step prediction under changing sensor conditions.** All sixteen prespecified comparisons use validation-selected references. T targets common switches and C targets local faults. Points show mean checkpoint-paired MSE changes with pointwise 95% t intervals over ten checkpoints. Filled markers indicate improvement with Holm16 $p \leq 0.05$.

5 Results

5.1 Coordinatewise adaptation improves local-fault prediction

Under local variance faults, C reduced one-step clean-signal MSE by 5.56–17.81% relative to independently calibrated static weighting. The PCWM reductions were 7.81% on Cartpole and 17.81% on Pendulum. Across the one-step evaluation, all eight prespecified adaptive/static comparisons passed the unified Holm correction (Figure 2). The four T configurations reduced common-switch error by 3.94–8.01%. Table 2 reports absolute errors for every setting.

PCWM/Pendulum C also reduced error by 3.49% relative to selected Huber weighting, with a 95% interval of $[-4.73, -2.25]\%$. PCWM/Cartpole T improved by 1.77% over Huber under common switches. Both comparisons had Holm16 $p = 0.03125$, while the robust reference had lower error in the other six settings. An exploratory standardized-IMQ comparison retained the Pendulum C advantage, while Cartpole T had 2.09% higher error than IMQ (Appendix A.8).

5.2 Adaptation retains gains under a common time cap

After optimizer selection for both rules, C reduced local-fault error by 24.90% and 5.60% on PCWM/Pendulum and PCWM/Cartpole. RSSM reductions were 15.56% and 6.16%, respectively. All four comparisons passed Holm8 correction ($p = 0.015625$). Figure 3 shows their intervals and measured filtering times.

Both PCWM C configurations selected two prior-scaled gradient steps at $\eta = 0.3$. RSSM C retained one Gauss–Newton step at $\eta = 0.5$. Original C/GN1 retained significant gains against selected static in three settings. Its PCWM/Cartpole change was -0.47% ($[-2.11, 1.18]\%$). The gains therefore depend on precision and optimizer choice within the shared time allowance. All four nominal-change intervals spanned zero. Stationary-high PCWM/Cartpole error increased by 3.15% (Appendix B).

5.3 Online modulation adds to the selected fixed base

The same-base control retains the adaptive configuration's base B and sets $\lambda = 0$. Relative to this control, PCWM C reduced local-fault error by 6.29% on Cartpole and 15.43% on Pendulum. Their pointwise 95% reduction intervals were $[5.03, 7.54]\%$ and $[11.26, 19.60]\%$, respectively. Figure 5 compares these configurations on each checkpoint's K0 scale.

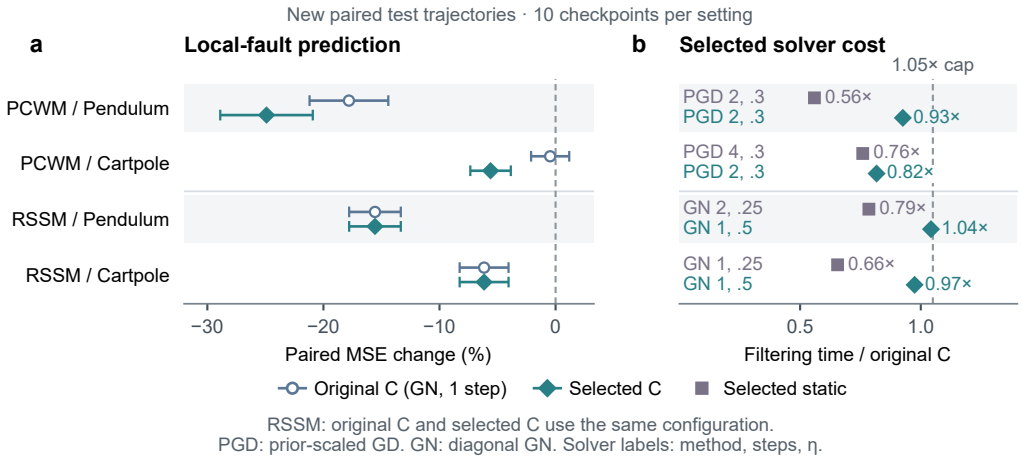

Figure 3: **Local-fault prediction under a shared filtering-time cap.** **a** Original and validation-selected C compare with validation-selected static refinement. Bars show pointwise 95% t intervals across ten checkpoint-paired effects. The eight comparisons share Holm correction. RSSM's two C configurations coincide. **b** Selected static and C filtering times are normalized by original C/GN1 timing. The dashed line marks their common 1.05 cap. Both rules search 48 optimizer candidates with their original precision calibration fixed.

With both rules selected within each suite, C reduced local-fault error relative to T by 3.89–14.89% across the four settings (Table 6).

During the first 21 fault observations, C retained 86.2–97.0% of its paired nominal healthy-coordinate precision, compared with 41.6–59.8% for T. These descriptive geometric summaries cover ten checkpoints per setting (Appendix A.6).

PCWM/Pendulum C also improved by 7.35% ($[2.35, 12.35]\%$) over a ten-transition causal delay. The delayed controller uses weights from its own evolving history. Appendix A.4 gives the weighting, timing, and replay controls. For PCWM/Cartpole common switches, the same-base effect is -0.91% ($[-2.81, 0.98]\%$), compared with -8.01% against the separately selected static base. Independently selected static comparisons combine calibration and online modulation.

5.4 Prediction gains extend across rollout horizons

Under local variance faults, C had lower mean paired error than static refinement at every evaluated horizon in all four settings (Figure 4). Five-step reductions ranged from 5.3% to 22.1%. At twenty steps, reductions ranged from 0.63% to 6.34%. On PCWM, fault-condition errors were also 0.12–2.00% lower than the selected robust reference across the six environment/horizon combinations at $h = 5, 10, 20$.

The aggregate endpoint combines all three longer horizons and both sensing conditions. C reduced this risk by 2.27–10.04% relative to static refinement (Table 23). The four-setting mean change relative to selected robust weighting was -0.18%.

Reading risk from the rollout. We rank forecasts by coordinate-averaged clean-signal particle variance. Normalized area under the risk–coverage curve (nAURC) divides mean retained risk across coverage levels by full-set risk. Constant scores give nAURC one. At $h = 20$, C variance gave mean nAURC of 0.201–0.463. Static and K0 variance gave similar rankings on the same C errors (Appendix C).

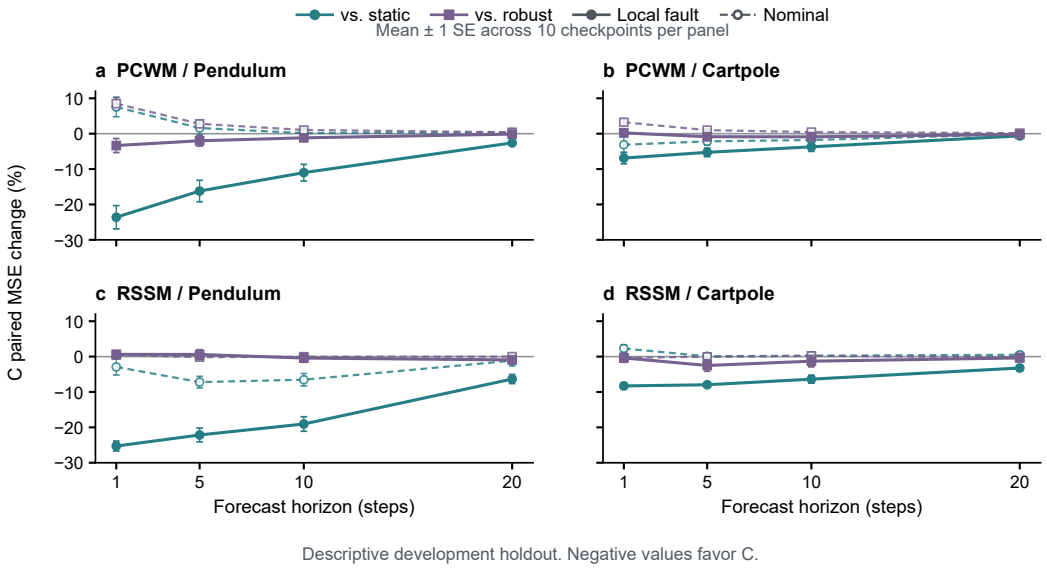

Figure 4: **State refinement across prediction horizons.** Each curve shows mean checkpoint-paired error changes for C relative to calibrated static or validation-selected robust refinement. Nominal conditions and local variance faults are displayed separately. All ten existing checkpoints contribute to each backbone/environment panel. Error bars show one standard error across checkpoints. These descriptive development estimates use newly selected multistep configurations. Negative changes favor C.

5.5 Spatial weighting in image observations

A separate pilot uses five existing PCWM checkpoints with 32×32 PixelPendulum images and two local patch-noise events per history. Validation selected the C family's static fallback, $c = 4$ and $\lambda = 0$. A fixed active C ($c = 64$, $\lambda = 0.5$) yielded mean B/R ratios of 2.40 in faulty patches and 10.85 on healthy pixels. Its risk changed by $-0.221 \pm 0.145\%$ against same-base static and $+0.266 \pm 0.050\%$ against selected static (mean $\pm$ one SE). Separate exact-curvature and projected-variance diagnostics appear in Appendix E.4.

5.6 Deployment cost and operating conditions

The controller adds two D_x-dimensional accumulators to the frozen model. Static K1 filtering took 2.68–3.22 ms per transition on an RTX 5090. The original K1 adaptive/static ratios were 1.60–1.67, excluding the common K0 pass and future rollout. Figure 3 reports the selected optimizers' costs. Appendices A.5 and A.7 give depth and robust-reference timing comparisons.

6 Discussion and conclusion

Calibrated coordinatewise precision provides an inference adjustment for frozen world models under localized sensor corruption. Same-base comparisons identify a predictive contribution from online modulation on PCWM. Under a common filtering-time cap, adaptive configurations retain gains across both backbones after each precision rule selects its optimizer. This result holds at the original calibrated scales. Separate development evaluations extend the gains over static weighting to longer horizons.

The primary evaluations cover two low-dimensional systems with representative fault families available for calibration and labeled validation. The pixel pilot retains Pendulum dynamics and a 32-dimensional latent state. Broader visual tasks and closed-loop decisions are important extensions. The current controller adjusts residual scales, while persistent mean offsets require additional modeling.

REPRODUCIBILITY STATEMENT

The accompanying code includes environments, training configurations, evaluation runners, tests, and checkpoint-level results with provenance hashes. Appendices A, B, and C specify the precision, optimizer, and multistep evaluations. Appendix D details backbone training and the environments.

AI USE DISCLOSURE

AI tools supported method development and evaluation, as well as manuscript refinement. Empirical results were verified against saved outputs, and the authors take responsibility for the final work.

REFERENCES

Eloi Alonso, Adam Jelley, Vincent Micheli, Anssi Kanervisto, Amos Storkey, Tim Pearce, and François Fleuret. Diffusion for world modeling: Visual details matter in Atari. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-1873. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/6bdde0373d53d4a501249547084bed43-Abstract-Conference.html.

Philipp Becker, Harit Pandya, Gregor Gebhardt, Cheng Zhao, C. James Taylor, and Gerhard Neumann. Recurrent Kalman networks: Factorized inference in high-dimensional deep feature spaces. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 544–552. PMLR, 2019. URL https://proceedings.mlr.press/v97/becker19a.html.

Bradley M. Bell and Frederick W. Cathey. The iterated Kalman filter update as a Gauss–Newton method. *IEEE Transactions on Automatic Control*, 38(2):294–297, 1993. doi: 10.1109/9.250476. URL https://doi.org/10.1109/9.250476.

Pier Giovanni Bissiri, Chris C. Holmes, and Stephen G. Walker. A general framework for updating belief distributions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78(5):1103–1130, 2016. doi: 10.1111/rssb.12158. URL https://academic.oup.com/jrsssb/article/78/5/1103/7040623.

Rafal Bogacz. A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, 76:198–211, 2017. doi: 10.1016/j.jmp.2015.11.003. URL https://www.sciencedirect.com/science/article/pii/S0022249615000759.

Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/hash/3de568f8597b94bda53149c7d7f5958c-Abstract.html.

Junyoung Chung, Kyle Kastner, Laurent Dinh, Kratarth Goel, Aaron C. Courville, and Yoshua Bengio. A recurrent latent variable model for sequential data. In *Advances in Neural Information Processing Systems*, volume 28, pp. 2980–2988, 2015. URL https://papers.nips.cc/paper/2015/hash/b618c3210e934362ac261db280128c22-Abstract.html.

Gerardo Duran-Martin, Matias Altamirano, Alex Shestopaloff, Leandro Sánchez-Betancourt, Jeremias Knoblauch, Matt Jones, Francois-Xavier Briol, and Kevin Patrick Murphy. Outlier-robust Kalman filtering through generalised Bayes. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12138–12171. PMLR, 2024. URL https://proceedings.mlr.press/v235/duran-martin24a.html.

Marco Fraccaro, Søren Kaae Sønderby, Ulrich Paquet, and Ole Winther. Sequential neural models with stochastic layers. In *Advances in Neural Information Processing Systems*, volume 29, pp. 2199–2207, 2016. URL https://papers.nips.cc/paper_files/paper/2016/hash/208e43f0e45c4c78cafadb83d2888cb6-Abstract.html.

Marco Fraccaro, Simon Kamronn, Ulrich Paquet, and Ole Winther. A disentangled recognition and nonlinear dynamics model for unsupervised learning. In *Advances in Neural Information Processing Systems*, volume 30, pp. 3601–3610, 2017. URL https://papers.nips.cc/paper_files/paper/2017/hash/7b7a53e239400a13bd6be6c91c4f6c4e-Abstract.html.

David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018. doi: 10.48550/arXiv.1803.10122. URL https://arxiv.org/abs/1803.10122.

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. PMLR, 2019. URL https://proceedings.mlr.press/v97/hafner19a.html.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/pdf?id=S1lOTC4tDS.

Danijar Hafner, Timothy Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering Atari with discrete world models. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=0oabwyZbOu.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse control tasks through world models. *Nature*, 640(8059):647–653, 2025. doi: 10.1038/s41586-025-08744-2. URL https://www.nature.com/articles/s41586-025-08744-2.

Nicklas Hansen, Xiaolong Wang, and Hao Su. Temporal difference learning for model predictive control. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8387–8406. PMLR, 2022. URL https://proceedings.mlr.press/v162/hansen22a.html.

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.

Nicklas Hansen, Hao Su, and Xiaolong Wang. Learning massively multitask world models for continuous control. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/092359ce5cf60a80e882378944bf1be4-Abstract-Conference.html.

Jesse Hemerik and Jelle Goeman. Exact testing with random permutations. *TEST*, 27(4):811–825, 2018. doi: 10.1007/s11749-017-0571-1. URL https://link.springer.com/article/10.1007/s11749-017-0571-1.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979. URL https://www.jstor.org/stable/4615733.

Peter J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964. doi: 10.1214/aoms/1177703732. URL https://doi.org/10.1214/aoms/1177703732.

Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. When to trust your model: Model-based policy optimization. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/5faf461eff3099671ad63c6f3f094f7f-Abstract.html.

Rico Jonschkowski, Divyam Rastogi, and Oliver Brock. Differentiable particle filters: End-to-end learning with algorithmic priors. In *Proceedings of Robotics: Science and Systems*, Pittsburgh, Pennsylvania, 2018. doi: 10.15607/RSS.2018.XIV.001. URL https://roboticsproceedings.org/rss14/p01.html.

Łukasz Kaiser, Mohammad Babaeizadeh, Piotr Miłos, Błażej Osiński, Roy H. Campbell, Konrad Czechowski, Dumitru Erhan, Chelsea Finn, Piotr Kozakowski, Sergey Levine, Afroz Mohiuddin, Ryan Sepassi, George Tucker, and Henryk Michalewski. Model based reinforcement learning for Atari. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1xCPJHtDB.

Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960. doi: 10.1115/1.3662552. URL https://doi.org/10.1115/1.3662552.

Maximilian Karl, Maximilian Soelch, Justin Bayer, and Patrick van der Smagt. Deep variational Bayes filters: Unsupervised learning of state space models from raw data. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=HyTqHL5xg.

Yoon Kim, Sam Wiseman, Andrew Miller, David Sontag, and Alexander Rush. Semi-amortized variational autoencoders. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 2678–2687. PMLR, 2018. URL https://proceedings.mlr.press/v80/kim18e.html.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015. URL https://arxiv.org/abs/1412.6980.

Diederik P. Kingma and Max Welling. Auto-encoding variational Bayes. In *International Conference on Learning Representations*, 2014. URL https://arxiv.org/abs/1312.6114.

Rahul G. Krishnan, Uri Shalit, and David Sontag. Structured inference networks for nonlinear state space models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, pp. 2101–2109, 2017. doi: 10.1609/aaai.v31i1.10779. URL https://ojs.aaai.org/index.php/AAAI/article/view/10779.

Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://papers.nips.cc/paper_files/paper/2017/hash/9ef2ed4b7fd2c810847ffa5fa85bce38-Abstract.html.

Joseph Marino, Milan Cvitkovic, and Yisong Yue. A general method for amortizing variational filtering. In *Advances in Neural Information Processing Systems*, volume 31, pp. 7857–7868. Curran Associates, Inc., 2018a. URL https://proceedings.neurips.cc/paper_files/paper/2018/hash/060afc8a563aaccd288f98b7c8723b61-Abstract.html.

Joseph Marino, Yisong Yue, and Stephan Mandt. Iterative amortized inference. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 3403–3412. PMLR, 2018b. URL https://proceedings.mlr.press/v80/marino18a.html.

Raman K. Mehra. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, 15(2):175–184, 1970. doi: 10.1109/TAC.1970.1099422. URL https://doi.org/10.1109/TAC.1970.1099422.

Vincent Micheli, Eloi Alonso, and François Fleuret. Transformers are sample-efficient world models. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=vhFu1Acb0xb.

Rajesh P. N. Rao and Dana H. Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87, 1999. doi: 10.1038/4580. URL https://www.nature.com/articles/nn0199_79.

Guy Revach, Nir Shlezinger, Xiaoyong Ni, Adrià López Escoriza, Ruud J. G. van Sloun, and Yonina C. Eldar. KalmanNet: Neural network aided Kalman filtering for partially known dynamics. *IEEE Transactions on Signal Processing*, 70:1532–1547, 2022. doi: 10.1109/TSP.2022.3158588. URL https://ieeexplore.ieee.org/document/9733186.

Simo Särkkä and Aapo Nummenmaa. Recursive noise adaptive Kalman filtering by variational Bayesian approximations. *IEEE Transactions on Automatic Control*, 54(3):596–600, 2009. doi: 10.1109/TAC.2008.2008348. URL https://users.aalto.fi/~ssarkka/pub/vb-akf-ieee.pdf.

Simo Särkkä and Lennart Svensson. *Bayesian Filtering and Smoothing*. Cambridge University Press, second edition, 2023. doi: 10.1017/9781108917407. URL https://doi.org/10.1017/9781108917407.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature*, 588:604–609, 2020. doi: 10.1038/s41586-020-03051-4. URL https://www.nature.com/articles/s41586-020-03051-4.

Donald J. Schuirmann. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, 15(6):657–680, 1987. doi: 10.1007/BF01068419. URL https://link.springer.com/article/10.1007/BF01068419.

Vaisakh Shaj, Saleh Gholam Zadeh, Ozan Demir, Luiz Ricardo Douat, and Gerhard Neumann. Multi time scale world models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 26764–26775. Curran Associates, Inc., 2023. doi: 10.52202/075280-1163. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/54d8aab579b5a9ed3395764c7341ebec-Abstract-Conference.html.

A PROSPECTIVE EVALUATION OF CALIBRATED PRECISION ADAPTATION

This evaluation measures clean-signal MSE under unknown changes in sensing conditions. The authors recorded the protocol locally and fixed the implementation before generating test trajectories. The calibration, validation, and test sets are newly generated. This appendix gives the complete primary and equivalence results.

A.1 ADDITIONAL CALIBRATION DETAILS

Equation 5 defines the calibration statistic using innovations r_t and decoder-projected prior variances V_t from complete K0 histories. For each checkpoint, we first average the clipped ratios over observed transitions. We then average paired schedules within each physical trajectory and trajectories within each condition. Nominal and target conditions receive equal weight, followed by equal weighting of the ten checkpoints. Each checkpoint retains its own $R_{0,s}$ when forming the shared shape q. The same histories later supply the scale-matched reference in Appendix A.4.

The coordinatewise rule also has the log-scale form

$$\log R^{\mathrm{C}}_{t,j} = (1-\lambda)\log B_j + \lambda \log e_{t,j}. \tag{14}$$

Partial adaptation interpolates between the base and online scale. Clipping precedes this interpolation, with bounds proportional to Q. For $c \le 1/16$, the clipping bound gives $R_t \ge B$ coordinatewise. The relative upper bound is $R_{t,j}/B_j \le (128/c)^{\lambda}$ for both rules.

A.2 ADDITIONAL DATA AND SELECTION DETAILS

We evaluate 40 existing checkpoints: PCWM and Gaussian RSSM on Pendulum and Cartpole, with ten training seeds per backbone/environment. Evaluating two suites gives 80 checkpoint/suite cells. Each cell has eight calibration, eight validation, and twenty test trajectories, all disjoint. Corresponding PCWM and RSSM cells share physical trajectories, giving 1,440 distinct trajectories (320 calibration, 320 validation, 800 test).

Common noise switches pair low/high/low and high/low/high schedules on the same physical trajectory, actions, standard-normal measurement draws, and switch times. The local suite selects two distinct random coordinates for nonoverlapping faults. Their onset/recovery times are $\lfloor (0.18 + U[0, 0.06))160 \rfloor$ / $\lfloor (0.54 + U[0, 0.04))160 \rfloor$ and $\lfloor (0.65 + U[0, 0.06))160 \rfloor$ / $\lfloor (0.91 + U[0, 0.04))160 \rfloor$. Each event draws faulty-coordinate standard deviation $U[0.2, 0.4)$. Healthy coordinates retain 0.05. The evaluator uses labels and fault metadata to construct and score the histories. Nominal and stationary-high conditions are also evaluated. The stuck-sensor condition holds the faulty coordinate at its corrupted pre-onset observation. Bias drift adds a random-sign ramp from zero to twice the event severity. These transfer conditions use separate random streams for each scenario.

Table 1 records the configurations selected using Eq. 12 and the robust weights in Eq. 11. For exact validation ties, adaptive families favor smaller λ and then smaller c. Robust families favor larger thresholds and then smaller c. Static ties favor smaller c. The shared static fallback remains eligible even when other candidates satisfy the nominal constraint. It wins a tie against nonzero adaptation or a finite robust threshold. Huber/IMQ cross-family ties use kind-name order. Four primary adaptive scales, one selected robust scale, and five selected robust thresholds lie at their respective grid minima. These counts describe the finite search domains used in the locked comparisons.

A.3 CONFIRMATORY COMPARISONS

Table 3 gives all sixteen primary comparisons using the paired effect in Eq. 13. All eight comparisons with the static reference establish improvement. Against the robust reference, two comparisons establish improvement and six establish higher clean-signal MSE. With ten checkpoints, the smallest two-sided sign-flip p value is $2/2^{10} = 0.001953125$. Fifteen raw p values attain this resolution limit.

Table 4 gives all eight stationary-high TOSTs. Each tests the mean paired percentage change against the two boundaries -2 and $+2$ using one-sided Student-t tests. The raw equivalence p value is the larger of the two one-sided values. Holm adjusts this secondary family separately. All eight

Table 1: Locked validation selections. T tests common noise switches and C tests local variance faults. λ is the adaptive exponent and τ is the robust threshold. Dagger marks the lower scale-grid boundary. Huber thresholds use standardized residuals. IMQ thresholds use original observation units.

Backbone / task	Rule	Static c	Adapt. c	λ	Robust	c	τ
PCWM / Pend.	T	6.7272	64	0.5	IMQ	2	0.1
PCWM / Pend.	C	8	64	0.5	Huber	4	0.5
PCWM / Cart.	T	4	16	0.5	Huber	2	0.5
PCWM / Cart.	C	3.3636	4	0.25	Huber	1	0.5
RSSM / Pend.	T	0.21022	$0.0625^{\dagger}$	0.25	IMQ	0.125	0.3
RSSM / Pend.	C	0.25	$0.0625^{\dagger}$	0.25	Huber	0.125	2.0
RSSM / Cart.	T	0.17678	$0.0625^{\dagger}$	0.25	IMQ	$0.0625^{\dagger}$	0.1
RSSM / Cart.	C	0.17678	$0.0625^{\dagger}$	0.25	Huber	0.125	2.0

Table 2: **Absolute one-step errors across all primary settings.** Values are clean-signal MSE multiplied by 10^3, averaged within checkpoint and then across ten checkpoints. Adaptive denotes T for common switches and C for local faults. The robust family is selected on validation data. Bold marks the lowest mean error in each row.

Backbone	Task	Target	K0	Static	Adaptive	Robust	Family
PCWM	Pendulum	Switch (T)	4.834	3.388	3.144	**3.112**	IMQ
PCWM	Pendulum	Local (C)	4.807	3.359	**2.766**	2.864	Huber
PCWM	Cartpole	Switch (T)	6.595	5.124	**4.715**	4.800	Huber
PCWM	Cartpole	Local (C)	5.117	4.281	3.951	**3.862**	Huber
RSSM	Pendulum	Switch (T)	3.079	3.104	2.942	**2.903**	IMQ
RSSM	Pendulum	Local (C)	3.114	3.038	2.580	**2.376**	Huber
RSSM	Cartpole	Switch (T)	3.959	4.024	3.868	**3.619**	IMQ
RSSM	Cartpole	Local (C)	3.310	3.333	3.150	**3.053**	Huber

90% intervals lie below the -2% boundary. Equivalence fails because the estimated improvements exceed the equivalence margin. The upper-bound component tests all reject, with maximum raw one-sided $p = 1.58 \times 10^{-6}$. These comparisons use the static reference selected for the mixed validation objective.

A.4 Mechanism controls and transfer conditions

The same-base control sets the selected adaptive arm's λ to zero. The scale-matched static reference uses geometric average effective residual scales from the calibration split, weighting nominal and target conditions equally and then the ten checkpoints equally. These statistics come from the same calibration split used to estimate q. The condition-informed reference selects two constants from the 45-point validation grid, minimizing nominal and stationary-high K0-normalized validation clean-signal MSE separately. True fault labels switch between these constants during evaluation. This auxiliary reference applies one common scale whenever any coordinate is faulty, including local variance faults. Its performance depends on the finite grid and this global switching rule. The gap-closure percentage $100(L_{\text{static}} - L_{\text{method}})/(L_{\text{static}} - L_{\text{informed}})$ uses grand mean checkpoint losses. We report it when the denominator is at least 2% of static clean-signal MSE. The percentage can be negative or exceed 100%. Table 5 gives the complete comparison. For common switches, T closes 29.3–74.4% of the grand-mean gap. C exceeds this common-scale reference on both Pendulum local-fault settings, yielding ratios above 100%. Table 6 compares the selected T and C configurations on each suite. C has lower error in all four local-fault settings. Table 15 supplies both configurations for each suite. Under common switches, T has lower error on PCWM/Pendulum, with the other three paired intervals spanning zero. These contrasts retain each family's independently selected scale and exponent.

Aligned replay exactly reproduces the adaptive states and predictions. Circular shifts by 53 and 106 transitions break weight alignment. Circular-shift replay uses complete recorded trajectories and is noncausal. A separate causal delay applies weights generated ten transitions earlier on the

Table 3: Sixteen primary comparisons. Values are mean checkpoint-paired changes in clean-signal MSE (%), with pointwise 95% intervals. Negative values indicate lower error. T tests common noise switches and C tests local variance faults. The two-sided sign-flip p values receive one sixteen-test Holm correction.

Backbone / task	Rule	Reference	Change [95% CI]	Raw p	Holm p
PCWM / Pend.	T	Static	-7.30 [-9.12, -5.48]	0.00195	0.03125
PCWM / Pend.	T	IMQ	+1.05 [+0.64, +1.47]	0.00195	0.03125
PCWM / Pend.	C	Static	-17.81 [-20.52, -15.10]	0.00195	0.03125
PCWM / Pend.	C	Huber	-3.49 [-4.73, -2.25]	0.00195	0.03125
PCWM / Cart.	T	Static	-8.01 [-9.11, -6.91]	0.00195	0.03125
PCWM / Cart.	T	Huber	-1.77 [-2.40, -1.14]	0.00195	0.03125
PCWM / Cart.	C	Static	-7.81 [-9.13, -6.49]	0.00195	0.03125
PCWM / Cart.	C	Huber	+2.38 [+1.71, +3.05]	0.00195	0.03125
RSSM / Pend.	T	Static	-5.18 [-5.85, -4.51]	0.00195	0.03125
RSSM / Pend.	T	IMQ	+1.44 [+0.17, +2.71]	0.03320	0.03320
RSSM / Pend.	C	Static	-15.31 [-17.15, -13.46]	0.00195	0.03125
RSSM / Pend.	C	Huber	+8.88 [+7.04, +10.72]	0.00195	0.03125
RSSM / Cart.	T	Static	-3.94 [-4.80, -3.09]	0.00195	0.03125
RSSM / Cart.	T	IMQ	+7.11 [+4.85, +9.37]	0.00195	0.03125
RSSM / Cart.	C	Static	-5.56 [-6.92, -4.19]	0.00195	0.03125
RSSM / Cart.	C	Huber	+3.21 [+2.42, +3.99]	0.00195	0.03125

Table 4: Eight stationary-high equivalence tests against the static reference. Values are mean checkpoint-paired changes in clean-signal MSE (%), with pointwise 90% intervals. The margin is ± 2 percentage points. Raw TOST p is the larger one-sided p. Holm adjusts this eight-test family separately.

Backbone / task	Rule	Change [90% CI]	Raw TOST p	Holm p
PCWM / Pend.	T	-8.43 [-10.08, -6.78]	0.99997	1.00000
PCWM / Pend.	C	-5.40 [-6.73, -4.06]	0.99941	1.00000
PCWM / Cart.	T	-10.77 [-12.09, -9.45]	1.00000	1.00000
PCWM / Cart.	C	-6.61 [-7.46, -5.77]	1.00000	1.00000
RSSM / Pend.	T	-6.18 [-6.83, -5.53]	1.00000	1.00000
RSSM / Pend.	C	-2.79 [-3.12, -2.46]	0.99911	1.00000
RSSM / Cart.	T	-5.09 [-5.98, -4.20]	0.99993	1.00000
RSSM / Cart.	C	-4.61 [-5.44, -3.78]	0.99986	1.00000

delayed arm's own history, using the base during its first ten updates. All eight main configurations have lower clean-signal MSE than their delayed and scale-matched references. These descriptive contrasts measure the sensitivity to weight timing and average scale. For PCWM/Cartpole common noise switches, the same-base contrast is -0.91% with interval $[-2.81, 0.98]$. The RSSM same-base

Table 5: Condition-informed reference on each target suite. Absolute clean-signal MSE is multiplied by 10^3. $\Delta(A, I)$ is the mean checkpoint-paired percentage change of the main adaptive arm relative to the informed reference, with a pointwise 95% interval ($n = 10$). Gap closure uses grand mean losses and the prespecified 2% denominator gate. The reference switches a finite-grid common scale using condition labels.

Backbone / task	Rule	Static	Informed	Adaptive	$\Delta(A, I)$ [95% CI]	Gap (%)
PCWM / Pend.	T	3.388	3.060	3.144	$+2.76$ $[+2.51, +3.00]$	74.4
PCWM / Pend.	C	3.359	3.104	2.766	-10.82 $[-14.27, -7.38]$	232.6
PCWM / Cart.	T	5.124	4.460	4.715	$+5.74$ $[+4.83, +6.65]$	61.6
PCWM / Cart.	C	4.281	3.867	3.951	$+2.45$ $[-0.95, +5.84]$	79.7
RSSM / Pend.	T	3.104	2.740	2.942	$+7.54$ $[+5.51, +9.56]$	44.3
RSSM / Pend.	C	3.038	2.755	2.580	-6.34 $[-8.67, -4.02]$	161.8
RSSM / Cart.	T	4.024	3.495	3.868	$+10.98$ $[+7.13, +14.82]$	29.3
RSSM / Cart.	C	3.333	3.149	3.150	$+0.16$ $[-2.30, +2.62]$	99.6

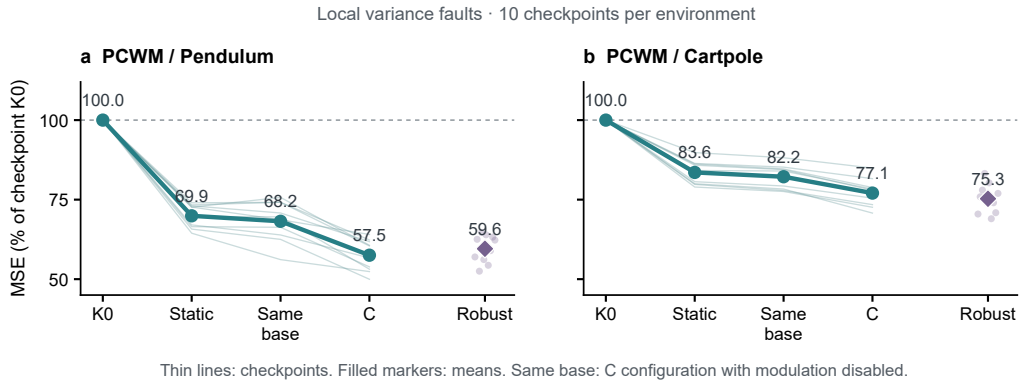

Figure 5: **Same-base and online comparisons under local faults.** PCWM configurations share each checkpoint's K0 error as the denominator. The same-base reference uses C's selected base with $\lambda = 0$. The independently selected static and robust references provide deployment comparisons. Thin lines show checkpoints, and filled markers show their means. The chain compares selected configurations on a common scale.

Table 6: Temporal (T) and coordinatewise (C) adaptation on common noise switches (Switch) and local variance faults (Local). All entries are mean checkpoint-paired changes in clean-signal MSE (%, $n = 10$). The first two contrasts report means. C versus T includes pointwise 95% intervals. Each family retains its own validation-selected c and λ. These within-suite comparisons are descriptive.

Backbone / task	Suite	T vs static	C vs static	C vs T [95% CI]
PCWM / Pend.	Switch	-7.30	-6.19	$+1.22\ [+0.27, +2.18]$
PCWM / Pend.	Local	-3.42	-17.81	$-14.89\ [-17.58, -12.20]$
PCWM / Cart.	Switch	-8.01	-8.66	$-0.71\ [-1.53, +0.11]$
PCWM / Cart.	Local	-2.36	-7.81	$-5.58\ [-6.73, -4.42]$
RSSM / Pend.	Switch	-5.18	-5.28	$-0.11\ [-0.32, +0.10]$
RSSM / Pend.	Local	-1.93	-15.31	$-13.65\ [-15.32, -11.97]$
RSSM / Cart.	Switch	-3.94	-3.96	$-0.01\ [-0.09, +0.07]$
RSSM / Cart.	Local	-1.74	-5.56	$-3.89\ [-4.85, -2.93]$

controls use $c = 1/16$, where the online statistic is bounded below by the unchanged base. Their larger same-base improvements are relative to this low-scale constant, while the primary comparisons use the separately calibrated static scale.

Figure 7 shows the temporal scale rising after noise increases and falling after recovery in the fixed examples. Its error panels use predictions at the existing ten-transition anchors.

Table 7 gives nominal test effects. Table 8 compares C and the selected robust reference under stuck sensors and bias drift. The 5% selection constraint applies to validation means. All four bias-drift mean changes relative to static are positive, with both PCWM intervals excluding zero. The controller averages squared residuals, whose expectation is $\mathrm{Var}(r) + (\mathbb{E}r)^2$. Mean shifts can therefore activate variance-based downweighting without correcting their direction. All transfer intervals are pointwise descriptive intervals.

A.5 Depth, cost, and reproducibility

The depth study keeps K1-selected settings and $\eta = 0.5$ fixed while using $K \in \{1, 2, 3, 5\}$ at every transition for the static reference, T, and C. Table 9 gives the main adaptive configurations. Table 10 supplies the corresponding static values. The source data include all three families and checkpoint-level values. PCWM/Pendulum clean-signal MSE under common noise switches rises sharply at larger depths. PCWM/Cartpole clean-signal MSE under local variance faults decreases through K5. The primary comparisons retain the prespecified K1 budget.

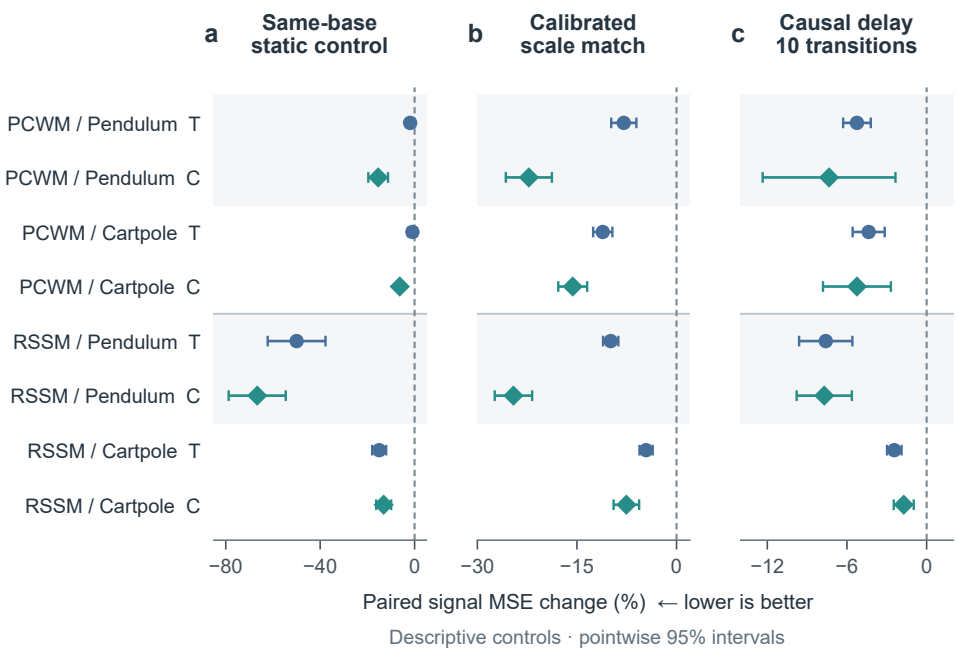

Figure 6: **Complete constant and timing controls for the one-step study.** All eight prescribed adaptive configurations are compared with their own unchanged base, calibration-derived scale-matched constants, and causal delays. The delayed arm applies ten-transition-old weights from its own history. Points and pointwise 95% t intervals summarize ten checkpoint-paired clean-signal MSE changes. Negative changes favor the adaptive rule. Each panel uses its own horizontal range.

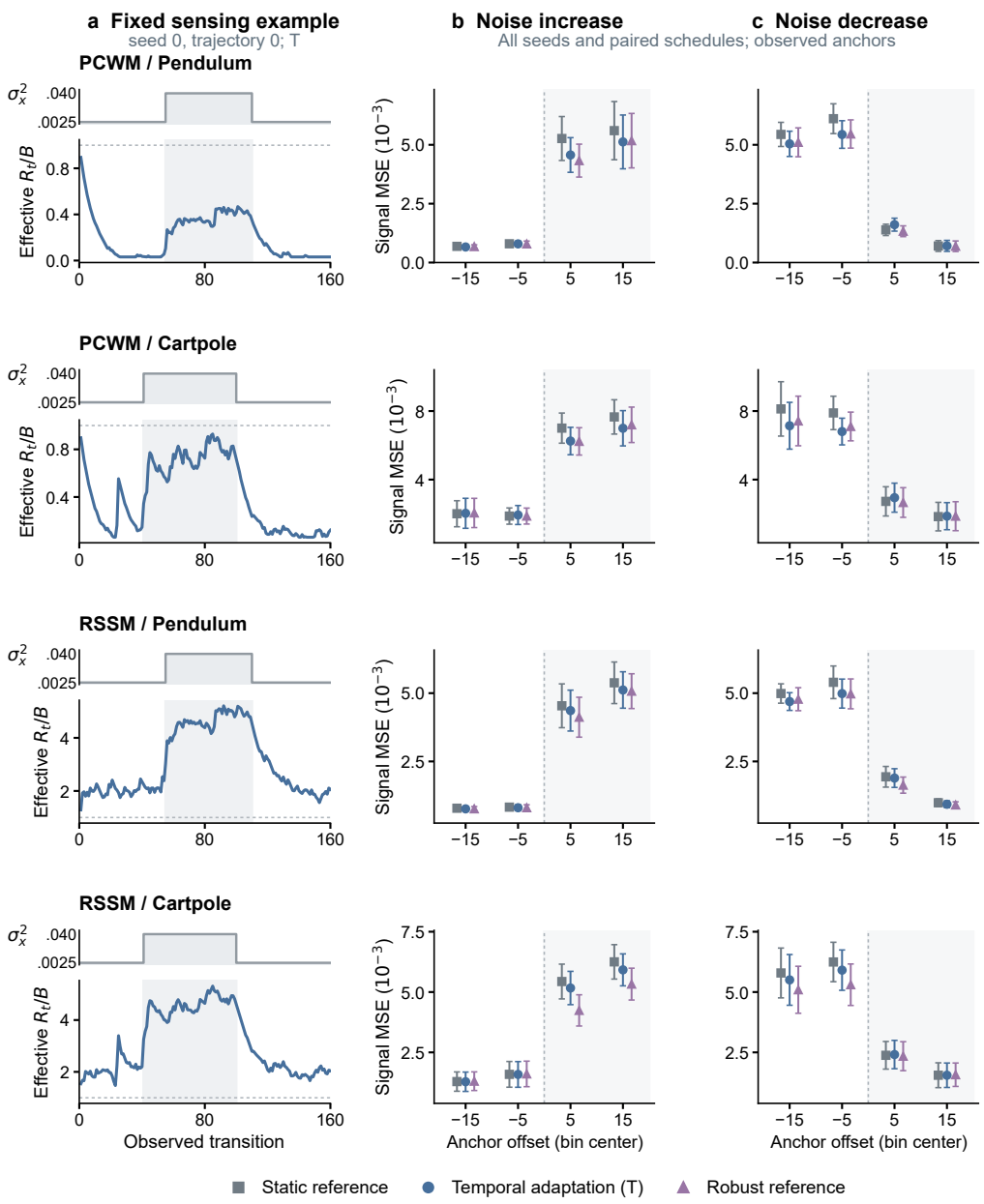

Figure 7: **Effective residual scales and errors around noise switches.** Left: seed 0, trajectory 0 from the low/high/low schedule, fixed by index. Each pair of axes shows the true observation-noise variance and T's geometric mean effective scale ratio R_t/B. Variance has squared observation units, and R_t/B is dimensionless. Middle and right: prediction errors at existing anchors around noise increases and decreases across ten checkpoints, each with twenty paired physical trajectories. Offsets use the last assimilated observation time minus the event time. Bins are $[-20, -10)$, $[-10, 0)$, $[0, 10)$, and $[10, 20]$. Errors are averaged within physical trajectory, then checkpoint. Points and pointwise 95% t intervals are descriptive summaries.

Table 7: Nominal clean-signal MSE relative to the static reference. Values are mean paired percentage changes with pointwise 95% intervals. T and C retain their respective suite's locked settings.

Backbone / task	Condition / rule	Change [95% CI]
PCWM / Pend.	Nominal / T	-0.52 [-3.40, +2.35]
PCWM / Pend.	Nominal / C	-2.08 [-5.40, +1.24]
PCWM / Cart.	Nominal / T	+1.87 [-1.90, +5.64]
PCWM / Cart.	Nominal / C	+0.89 [-1.04, +2.82]
RSSM / Pend.	Nominal / T	-3.22 [-5.08, -1.36]
RSSM / Pend.	Nominal / C	-5.15 [-7.69, -2.61]
RSSM / Cart.	Nominal / T	-0.37 [-1.03, +0.28]
RSSM / Cart.	Nominal / C	+0.45 [-0.02, +0.91]

Table 8: Transfer under stuck sensors and bias drift with settings selected for local variance faults. All four selected robust references are Huber weights. Entries are mean checkpoint-paired changes in clean-signal MSE (%, $n = 10$). C/static and robust/static report means. C/robust includes pointwise 95% intervals. These transfer comparisons are descriptive.

Backbone / task	Fault	C vs static	Robust vs static	C vs robust [95% CI]
PCWM / Pend.	Stuck	-1.12	-1.75	$+0.64$ $[+0.21, +1.06]$
PCWM / Pend.	Bias drift	+13.39	+11.18	$+1.96$ $[+1.07, +2.84]$
PCWM / Cart.	Stuck	+0.75	+0.89	-0.10 $[-1.10, +0.91]$
PCWM / Cart.	Bias drift	+5.21	+8.70	-3.19 $[-3.90, -2.49]$
RSSM / Pend.	Stuck	-3.79	-7.18	$+3.73$ $[+2.06, +5.40]$
RSSM / Pend.	Bias drift	+1.54	+13.89	-10.36 $[-14.50, -6.22]$
RSSM / Cart.	Stuck	-2.88	-3.69	$+0.91$ $[-0.42, +2.23]$
RSSM / Cart.	Bias drift	+1.25	+4.37	-2.85 $[-4.71, -0.99]$

Matched K1 filtering is timed on one 160-transition history per checkpoint using an RTX 5090, CUDA synchronization, alternating arm order, one warm-up round, and three measured rounds. The main adaptive/static latency ratios range from 1.60 to 1.67 (Table 9). Validation evaluated 237 candidate specifications per cell, or 18,960 across the eighty cells.

Saved manifests bind the protocol, code, checkpoints, split identities, selections, and predictions. Independent rescoring checked all 2,320 prediction arrays and 10,440 condition/checkpoint means, with zero discrepancy in clean-signal MSE. Paired targets, actions, numerical encoder variances, and prediction random streams also agree. The record contains 3,132,000 per-anchor entries from the two custom low-dimensional tasks.

Table 9: Depth diagnostics and K1 filtering cost for the main adaptive configurations. Risk is mean within-checkpoint clean-signal MSE/K0 on the target condition. The last column gives adaptive/static latency at K1. Selected settings and $\eta = 0.5$ are fixed at all depths.

Backbone / task	Rule	$K = 1$	$K = 2$	$K = 3$	$K = 5$	K1 cost
PCWM / Pend.	T	0.6498	0.7069	1.8220	7.3939	1.604×
PCWM / Pend.	C	0.5753	0.4928	0.7561	2.6518	1.598×
PCWM / Cart.	T	0.7144	0.6481	0.6743	0.8037	1.642×
PCWM / Cart.	C	0.7706	0.6887	0.6535	0.6279	1.643×
RSSM / Pend.	T	0.9569	0.9463	0.9441	0.9423	1.612×
RSSM / Pend.	C	0.8265	0.8504	0.8171	0.8269	1.609×
RSSM / Cart.	T	0.9770	0.9745	0.9736	0.9732	1.672×
RSSM / Cart.	C	0.9516	0.9420	0.9394	0.9392	1.663×

Learned residual scales. Table 11 reports the learned decoder variances in observation units. The PCWM coordinate means range from 0.00332 to 0.00630. The RSSM means range from 0.000794 to 0.00180. Nominal sensor variance is 0.0025. These learned residual scales enter both

Table 10: Static-reference depth results. Entries average the within-checkpoint signal-MSE/K0 ratio over ten checkpoints. Settings selected at $K = 1$ and $\eta = 0.5$ remain fixed across depths.

Backbone / task	Suite	$K = 1$	$K = 2$	$K = 3$	$K = 5$
PCWM / Pend.	Switch	0.7006	0.6075	0.5822	0.5764
PCWM / Pend.	Local	0.6991	0.6064	0.5831	0.5810
PCWM / Cart.	Switch	0.7765	0.7005	0.6748	0.6633
PCWM / Cart.	Local	0.8355	0.7835	0.7671	0.7594
RSSM / Pend.	Switch	1.0091	0.9867	0.9868	0.9816
RSSM / Pend.	Local	0.9758	0.9642	0.9614	0.9580
RSSM / Cart.	Switch	1.0172	1.0105	1.0121	1.0111
RSSM / Cart.	Local	1.0078	1.0019	1.0038	1.0029

the prior/observation balance and the calibration base. The selected multiplier therefore reflects the complete refinement configuration, including latent and model error.

Table 11: Learned decoder residual variance R_0 in normalized observation coordinates, multiplied by 10^3. Each entry is the mean [minimum, maximum] over ten checkpoints. Nominal and stationary-high sensor variances in these units are 2.5 and 40, respectively. Values are extracted from saved K0 residual-scale traces.

Task	Coordinate	PCWM	RSSM
Pendulum	$\cos\theta$	4.803 [4.136, 5.881]	1.582 [1.526, 1.647]
Pendulum	$\sin\theta$	4.533 [4.070, 5.195]	1.705 [1.548, 1.851]
Pendulum	$\dot{\theta}/8$	3.321 [3.019, 3.504]	1.071 [0.970, 1.236]
Cartpole	$p/2.4$	5.209 [4.620, 6.414]	1.799 [1.662, 2.030]
Cartpole	$\dot{p}/5$	6.296 [4.477, 8.794]	0.794 [0.719, 0.946]
Cartpole	$\cos\theta$	4.865 [4.355, 5.290]	1.528 [1.399, 1.624]
Cartpole	$\sin\theta$	5.146 [4.213, 5.800]	1.718 [1.518, 1.822]
Cartpole	$\dot{\theta}/10$	3.397 [3.086, 3.956]	1.115 [1.006, 1.203]

A.6 DIRECT COORDINATE COUPLING AND OBSERVED WEIGHT RESPONSES

Controlled weight-map intervention. Fix the calibrated base B and the positive, clipped excess statistic e. Let $u_j = e_j/B_j$ and $w_i = B_i/R_i$ denote normalized excess and relative precision. The two rules give

$$w_i^{\mathrm{T}} = \left(D_x^{-1} \sum_j u_j \right)^{-\lambda}, \qquad w_i^{\mathrm{C}} = u_i^{-\lambda}. \tag{15}$$

Their log sensitivities are

$$\frac{\partial \log w_i^{\mathrm{T}}}{\partial \log e_k} = -\lambda \frac{u_k}{\sum_j u_j}, \qquad \frac{\partial \log w_i^{\mathrm{C}}}{\partial \log e_k} = -\lambda \, \mathbf{1}\{i = k\}. \tag{16}$$

The T sensitivity has identical rows, so one changing statistic affects every weight. The C sensitivity is diagonal, confining the direct response to the perturbed coordinate. For pre-clipping statistics, saturation contributes a zero derivative away from clipping boundaries.

We check these identities against the implemented controller in float64. Finite differences perturb each positive excess by multiplicative factors $\exp(\pm 10^{-5})$. Sixteen checks cover three and five coordinates, both rules, clipping saturation, and the $\lambda = 0$ static identity. All analytic and numerical sensitivities agree within 10^{-8}. The illustrative sweep uses $B = 4Q$, $\lambda = 0.5$, and changes one e_j/B_j from one to sixteen. T reduces healthy-coordinate precision to 0.408 of its base for $D_x = 3$ and 0.500 for $D_x = 5$. C retains healthy-coordinate precision at one and reduces the affected coordinate to 0.250. Figure 8 displays this controlled intervention.

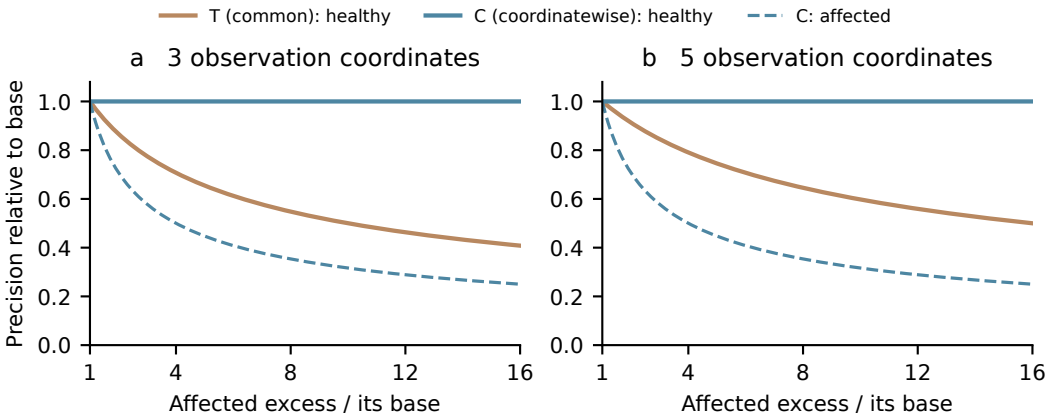

Figure 8: **Single-coordinate changes in the precision map.** The base, exponent, and unperturbed excess statistics are identical for both rules. One excess ratio varies from one to sixteen at $\lambda = 0.5$. Curves agree with the implemented controller. The experiment isolates the direct weight response at fixed input statistics.

Responses in saved filtering trajectories. We align the first local variance fault in each of twenty paired histories per checkpoint. Every setting includes ten checkpoints, with corresponding physical histories shared across the two backbones. For each rule, the fault and nominal histories use the same selected parameters and random draws. Externally assigned fault labels identify healthy and affected coordinates. The log precision ratio is $d_{s,e,t,j} = \log R^{\mathrm{nominal}}_{s,e,t,j} - \log R^{\mathrm{fault}}_{s,e,t,j}$. The time index refers to the observation assimilated by that update. Before fault onset, the paired weight sequences agree exactly.

For the summary, we average d over healthy coordinates, twenty histories, and offsets zero through twenty within each checkpoint. We then average the ten checkpoint values and report $G = \exp(\bar{d})$. Thus G is a geometric precision ratio relative to the same rule's nominal history. Table 12 reports $100(G - 1)$ for both coordinate groups. Figure 9 shows log ratios around fault onset.

Table 12: **Precision responses during the first local fault.** Entries are geometric percentage changes from each rule's paired nominal histories over offsets 0–20, with ten checkpoints per setting. T and C retain their independently selected configurations.

Setting	T healthy	C healthy	T affected	C affected
PCWM/Pendulum	−58.41	−13.78	−58.41	−90.58
PCWM/Cartpole	−47.60	−3.22	−47.60	−67.67
RSSM/Pendulum	−45.32	−6.92	−45.32	−59.51
RSSM/Cartpole	−40.22	−2.98	−40.22	−64.66

PCWM/Pendulum selects $(c, \lambda) = (16, 0.25)$ for T and $(64, 0.5)$ for C. The other three settings select the same c, λ, and ρ for both rules. Each rule still produces its own recursive latent history. Changes in shared latent states, decoder predictions, and projected prior variance can alter healthy-coordinate residual statistics under C. The trajectory summaries measure the resulting weight response after these coupled updates.

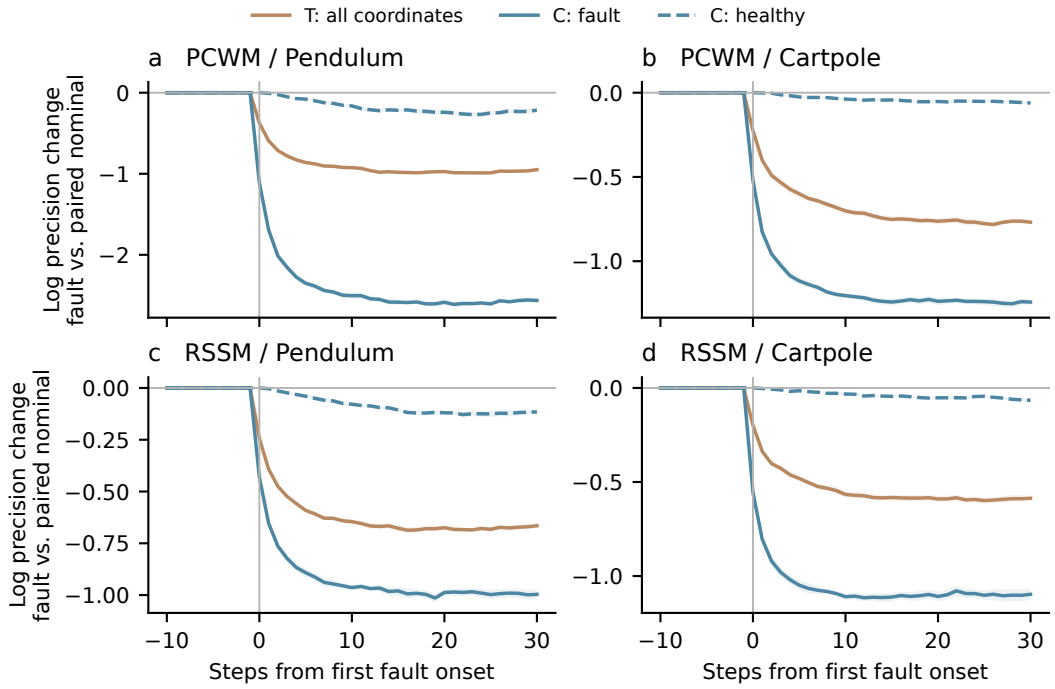

Figure 9: **Healthy and affected coordinate precision around fault onset.** Values are mean log precision ratios relative to each rule's paired nominal histories. The first fault begins at offset zero. Curves average twenty histories within checkpoint and ten checkpoints within setting. Shading shows one standard error of checkpoint means. All forty existing checkpoints contribute, with independently selected T and C configurations.

A.7 Sensitivity to search domain and memory rate

Search-domain sensitivity. An exploratory extension added three smaller scales to each adaptive family and two smaller thresholds to each robust family. The added scales were $c \in \{2^{-6}, 2^{-8}, 2^{-10}\}$. Huber added thresholds $\{0.25, 0.125\}$, and IMQ added $\{0.05, 0.03\}$. The complete pool grew from 237 to 309 specifications. Selection retained the original calibration, validation objective, nominal constraint, and test histories. Both pools appear in Table 13. The original-pool results retain their primary-test status.

The adaptive rule had lower point-estimate error than the selected robust reference in three settings after expansion, compared with two originally. RSSM/Pendulum local-fault C changed from

$+8.88\%$ to -2.95% relative to the selected robust reference. The selected RSSM adaptive scales included 2^{-6} and 2^{-8}. Because clipping remains anchored at Q, these scales keep adaptation in the region $e_{t,j}/B_j \geq 1/(16c) \geq 1$. The added search therefore explores smaller bases within the existing downweighting regime.

Table 13: **Search-domain sensitivity and matched filtering cost.** MSE columns give mean checkpoint-paired changes relative to the separately selected robust reference (%), with pointwise 95% t intervals ($n = 10$). Each pool selects its own adaptive and robust configurations on validation data. The expanded pool is an exploratory re-selection evaluated on the original test histories. Prediction uses the indicated rule. Timing uses C on both suites. The final column divides mean C time by mean robust time for original-pool configurations, over ten checkpoints and three measured rounds each.

Setting	Rule	Original pool	Expanded pool	C/robust time
PCWM/Pend. Switch	T	$+1.05$ $[+0.64, +1.47]$	$+1.05$ $[+0.64, +1.47]$	1.521
PCWM/Pend. Local	C	-3.49 $[-4.73, -2.25]$	-3.49 $[-4.73, -2.25]$	1.173
PCWM/Cart. Switch	T	-1.77 $[-2.40, -1.14]$	-1.22 $[-1.88, -0.55]$	1.146
PCWM/Cart. Local	C	$+2.38$ $[+1.71, +3.05]$	$+2.38$ $[+1.71, +3.05]$	1.141
RSSM/Pend. Switch	T	$+1.44$ $[+0.17, +2.71]$	$+0.22$ $[-0.72, +1.17]$	1.532
RSSM/Pend. Local	C	$+8.88$ $[+7.04, +10.72]$	-2.95 $[-4.28, -1.62]$	1.165
RSSM/Cart. Switch	T	$+7.11$ $[+4.85, +9.37]$	$+2.76$ $[+1.67, +3.85]$	1.602
RSSM/Cart. Local	C	$+3.21$ $[+2.42, +3.99]$	$+1.04$ $[+0.19, +1.90]$	1.136

The timing column uses the original configurations and matched single-history workloads. Each workload contains 160 transitions, with device synchronization, alternating arm order, one warm-up round, and three measured rounds per checkpoint. The ratio divides mean C time by mean robust time across ten checkpoints. These measurements cover filtering on an RTX 5090.

Common-grid memory comparison. We reselected each rate independently on saved validation losses using twenty configurations per family and rate. The common grid was $c \in \{0.25, 1, 4, 16, 64\}$ and $\lambda \in \{0, 0.25, 0.5, 1\}$. The shared fallback was selected from the original 45-point static grid. The nominal constraint and tie-breaking rules remained unchanged. Table 14 compares $\rho = 0.2$ with instantaneous statistics, $\rho = 1$, using paired percentage changes in the validation objective J. C had lower selected validation risk with accumulated statistics in all four PCWM settings, with paired changes of 0.42–1.51% from $\rho = 0.2$ to $\rho = 1$. T favored instantaneous statistics in three settings and accumulated statistics in one.

Table 14: **Memory-rate sensitivity on a common validation grid.** Entries are mean paired changes in J from $\rho = 0.2$ to $\rho = 1$ (%), conditional on independent within-rate selection, with descriptive 95% t intervals ($n = 10$). Positive values favor accumulated statistics. The intervals describe selected validation configurations.

PCWM setting	T	C
Pend. Switch	-0.34 $[-1.33, +0.66]$	$+0.64$ $[-0.08, +1.36]$
Pend. Local	-2.01 $[-3.06, -0.97]$	$+0.63$ $[-0.71, +1.97]$
Cart. Switch	$+0.88$ $[-0.61, +2.36]$	$+1.51$ $[+0.16, +2.86]$
Cart. Local	-0.50 $[-1.34, +0.34]$	$+0.42$ $[-0.54, +1.38]$

All eight RSSM family/suite comparisons selected identical static weights across rates. The Pendulum local-fault families selected $c = 0.25$, $\lambda = 0$. The other six comparisons selected their common static fallback. The common grid excludes $c < 0.25$, which contains the originally selected RSSM adaptive scales. These validation summaries are conditional on the selected configurations and shared calibration shape q. Their descriptive intervals represent between-checkpoint variation after selection.

Parameters for comparisons within each suite. Table 15 records the original T and C settings for every suite. Each family used the validation objective of the suite on which it was scored. The parameter differences accompany the prediction contrasts in Table 6.

Table 15: **Independently selected T and C configurations for each suite.** These original-pool settings accompany the within-suite comparisons in Table 6. Both families use $\rho = 0.2$. Each scale is selected for the suite in its row.

Setting	T (c, λ)	C (c, λ)
PCWM/Pend. Switch	$(2^6, 0.5)$	$(2^4, 0.25)$
PCWM/Pend. Local	$(2^4, 0.25)$	$(2^6, 0.5)$
PCWM/Cart. Switch	$(2^4, 0.5)$	$(2^5, 0.5)$
PCWM/Cart. Local	$(2^2, 0.25)$	$(2^2, 0.25)$
RSSM/Pend. Switch	$(2^{-4}, 0.25)$	$(2^{-4}, 0.25)$
RSSM/Pend. Local	$(2^{-4}, 0.25)$	$(2^{-4}, 0.25)$
RSSM/Cart. Switch	$(2^{-4}, 0.25)$	$(2^{-4}, 0.25)$
RSSM/Cart. Local	$(2^{-4}, 0.25)$	$(2^{-4}, 0.25)$

A.8 STANDARDIZED-INNOVATION IMQ

This follow-up standardizes innovations before applying the global IMQ weight. For innovation r_t, decoder-projected prior variance V_t, and calibrated base B, the factor is

$$w_t = \max\left(10^{-6}, \frac{\tau^2}{\tau^2 + \sum_j r_{tj}^2/(V_{tj} + B_j)}\right), \qquad R_{tj} = B_j/w_t.$$

The variance normalization matches the Huber reference. The factor applies uniformly across coordinates and uses the existing one-step state update. Its threshold is dimensionless.

The balanced72 pool combines the original 12 scales with thresholds $\{0.03, 0.1, 0.3, 1, 3, 10\}$. We also evaluate the original48 threshold pool $\{0.1, 0.3, 1, 3\}$ and lower72 pool $\{0.03, 0.05, 0.1, 0.3, 1, 3\}$. The 84-candidate union is computed once. These pools distinguish matched candidate counts from sensitivity to the dimensionless threshold range. The 72-candidate pools provide more choices than the original 48-candidate family. Every pool retains the original nominal constraint and static fallback.

Calibration shapes, checkpoints, forecasting randomness, and the original T/C references remain fixed. Selections use validation trajectories and are locked before the new held-out evaluation. The test trajectories were evaluated previously, so this follow-up provides exploratory evidence. Each setting contains ten checkpoints and twenty test trajectories per checkpoint. Switching comparisons average the two switching schedules. Local comparisons use coordinate variance switches. The original locked tests retain their original definitions. Tables 16 and 17 report all three pools. Original48 and lower72 yield identical selections and predictions, so their table rows are shared. Their 24 target comparisons share one Holm correction.

Table 16: Balanced standardized-IMQ comparison. Parameters c and τ belong to standardized IMQ. Entries report checkpoint-paired relative MSE changes in percent, with 95% confidence intervals. Negative values favor the locked T or C rule. Holm correction includes all 24 comparisons across the three pools.

Backbone	Task	Target rule	c	τ	Change [95% CI]	Holm p
PCWM	Pendulum	Switch (T)	2	1	+1.34 [+0.75, +1.92]	0.047
PCWM	Pendulum	Local (C)	2	1	-8.16 [-10.59, -5.73]	0.047
PCWM	Cartpole	Switch (T)	$\frac{1}{2}$	1	+2.09 [+1.37, +2.81]	0.047
PCWM	Cartpole	Local (C)	2	3	-1.13 [-2.84, +0.57]	1.000
RSSM	Pendulum	Switch (T)	$\frac{1}{8}$*	10^*	-0.01 [-0.95, +0.93]	1.000
RSSM	Pendulum	Local (C)	$\frac{1}{16}$*	3	+1.12 [-2.21, +4.45]	1.000
RSSM	Cartpole	Switch (T)	$\frac{1}{16}$*	3	+5.89 [+3.90, +7.87]	0.047
RSSM	Cartpole	Local (C)	$\frac{1}{16}$*	3	+4.24 [+2.88, +5.60]	0.047

An asterisk marks a selected grid boundary. Intervals are conditional on validation selection.

Table 17: Original48 and lower72 sensitivities, reported together. Each predeclared pool was selected separately and yielded identical parameters and predictions in all eight settings. Parameters c and τ belong to standardized IMQ. Endpoints and Holm correction match Table 16.

Backbone	Task	Target rule	c	τ	Change [95% CI]	Holm p
PCWM	Pendulum	Switch (T)	2	1	+1.34 [+0.75, +1.92]	0.047
PCWM	Pendulum	Local (C)	2	1	-8.16 [-10.59, -5.73]	0.047
PCWM	Cartpole	Switch (T)	$\frac{1}{2}$	1	+2.09 [+1.37, +2.81]	0.047
PCWM	Cartpole	Local (C)	2	3^*	-1.13 [-2.84, +0.57]	1.000
RSSM	Pendulum	Switch (T)	$\frac{1}{16}{}^*$	3^*	+3.39 [-0.33, +7.12]	0.668
RSSM	Pendulum	Local (C)	$\frac{1}{16}{}^*$	3^*	+1.12 [-2.21, +4.45]	1.000
RSSM	Cartpole	Switch (T)	$\frac{1}{16}{}^*$	3^*	+5.89 [+3.90, +7.87]	0.047
RSSM	Cartpole	Local (C)	$\frac{1}{16}{}^*$	3^*	+4.24 [+2.88, +5.60]	0.047

An asterisk marks a selected grid boundary. Intervals are conditional on validation selection.

B Latent optimization under a common filtering budget

B.1 Fixed interfaces and optimizer grids

We evaluated alternative latent optimizers on the forty existing PCWM/RSSM checkpoints across Pendulum and Cartpole. The local-suite calibration vector q from Appendix A.1 and previously selected c, λ, and ρ remained fixed. Table 19 lists the scalar parameters and selected optimizers. The comparison isolates optimizer choices at the original calibrated precision scales.

All candidates started from the encoder proposal, used Eq. 2 as their objective, and retained the latent clamp $[-30, 30]$. Each candidate maintained its own recursive state and causal precision trajectory. Within each transition, precision, prior, and RSSM context remained fixed throughout optimization and line search. Precision statistics were updated once per observation, without inner-loop updates.

Table 18: Optimizer grids applied separately to static and coordinatewise precision. Each precision receives 48 candidates. Armijo rates initialize each line search.

Family	Steps K	Step sizes η	Count
Fixed diagonal GN	$1, 2, 4$	$0.125, 0.25, 0.5, 1$	12
Prior-scaled GD	$1, 2, 4$	$0.03, 0.1, 0.3, 1$	12
Adam	$1, 2, 4$	$0.01, 0.03, 0.1, 0.3$	12
GN with Armijo	$1, 2, 4$	$0.5, 1, 2, 4$	12

Both GN families froze diagonal curvature at the encoder proposal. Prior-scaled GD multiplied the gradient by the diagonal prior variance. Adam used $(\beta_1, \beta_2, \epsilon) = (0.9, 0.999, 10^{-8})$, with moments reset at each transition. Armijo used contraction 0.5, sufficient-decrease constant 10^{-4}, and at most six contractions. Acceptance was evaluated separately for each history using the projected displacement. A failed search retained the preceding state and stopped that history's inner optimization. The original fixed GN configuration, $K = 1$ and $\eta = 0.5$, remained an unretuned anchor within its twelve-candidate family.

B.2 Timing and validation selection

Timing used checkpoint zero's first local-fault validation history, with batch size one and 160 transitions. After warmup, three synchronized rounds used a fixed randomized candidate order. Each setting included 27 interleaved original C/GN1 timings to define its reference median. Both precision rules received the common absolute cap of 1.05 times that median. The original C/GN1 anchor satisfied the timing constraint by definition. All final selections also satisfied the observed cap using their measured medians. This common maximum allowed different realized times, reported in Table 19.

All families retained the original host energy trace with $K + 1$ scalar values. Timing included this logging, extra line-search evaluations, and associated synchronizations. Timing included fil-

Table 19: Locked precision parameters and validation-selected optimizers under the common filtering-time cap. All coordinatewise rules retain $\rho = 0.2$.

Setting	Rule	c	λ	Optimizer	(K, η)	Time ratio
PCWM/Pend.	Static	8	0	Prior-GD	$(2, 0.3)$	0.559
PCWM/Pend.	C	64	0.5	Prior-GD	$(2, 0.3)$	0.926
PCWM/Cart.	Static	$2^{7/4}$	0	Prior-GD	$(4, 0.3)$	0.759
PCWM/Cart.	C	4	0.25	Prior-GD	$(2, 0.3)$	0.817
RSSM/Pend.	Static	2^{-2}	0	Fixed GN	$(2, 0.25)$	0.785
RSSM/Pend.	C	2^{-4}	0.25	Fixed GN	$(1, 0.5)$	1.042
RSSM/Cart.	Static	$2^{-5/2}$	0	Fixed GN	$(1, 0.25)$	0.655
RSSM/Cart.	C	2^{-4}	0.25	Fixed GN	$(1, 0.5)$	0.974

Time ratios divide candidate medians by the corresponding original C/GN1 anchor median. Both rules share the same upper bound of 1.05. Static uses $\lambda = 0$.

tering but excluded the shared K0 pass and future forecasts. Measurements used an RTX 5090 with Python 3.13.5 and PyTorch 2.9.1+cu128. The cap therefore describes this implementation and timing workload, including diagnostic overhead.

Selection reused eight original validation trajectories per checkpoint and preserved their forecast random streams. For checkpoint s and condition d, let $L_{s,d}$ average clean-signal MSE over coordinates, anchors, and trajectories. The validation objective was

$$J = \frac{1}{20} \sum_{s=1}^{10} \left(\frac{L_{s,\mathrm{nom}}}{L_{s,\mathrm{nom}}^{K0}} + \frac{L_{s,\mathrm{fault}}}{L_{s,\mathrm{fault}}^{K0}} \right), \tag{17}$$

using the original stored validation K0 denominators. Mean normalized nominal error could not exceed 1.05 times the original static/GN1 value. We minimized J among feasible candidates within each family and across all 48 candidates per precision rule. Ties were resolved by time, step count, step size, and method name. The shared fallback was original static/GN1, including when a C family lacked feasible candidates. All four C/Armijo families required this fallback under the joint nominal and timing constraints. These fallback outcomes reflect eligibility under the nominal and timing constraints. The archived logs retain every family's descriptive results, raw timings, and choices without the time constraint.

B.3 Prospective forecasts and paired analysis

Selections and source/model hashes were locked before fresh test trajectories were generated. The test namespace was `latent-optimizer-prospective-20260919-v1`. Each environment and checkpoint seed received twenty new 160-transition physical trajectories. Backbones shared corresponding trajectories, giving 400 distinct physical trajectories across both environments. One-step forecasts used anchors $10, 20, \ldots, 150$ and $P = 150$ particles. Arms shared clean targets, future actions, K0 forecast variance, and forecast random streams. This generated new test trajectories while reusing the existing training checkpoints.

Within each checkpoint, we computed $100(L_C/L_{\mathrm{static}} - 1)$ after averaging losses over coordinates, anchors, and trajectories. The statistical unit was the checkpoint, with ten paired effects per setting. We report pointwise 95% t intervals and exact two-sided sign-flip tests over all 2^{10} assignments. Sign-flip inference assumes paired-null sign symmetry, conditional on the locked selection and existing checkpoints. One Holm correction covered the eight planned comparisons in Table 20. Both RSSM selections recovered original C, giving one shared effect per setting across their two planned comparisons.

Original C/GN1 improved local-fault error in three settings after Holm correction. Its PCWM/Cartpole comparison had a -0.47% effect with an interval spanning zero. Selected C improved all four settings against selected static optimization under the common cap. Both PCWM selections used two prior-scaled GD steps, whereas both RSSM selections retained original GN1. Seven raw sign-flip p values were $2/1024$, while original PCWM/Cartpole yielded $536/1024$.

Table 20: All eight primary local-fault comparisons against validation-selected static optimization. Absolute MSE values are multiplied by 10^3. Intervals are pointwise 95% paired t intervals.

Setting	C choice	C MSE	Static MSE	Δ [95% CI] (%)	Holm p
PCWM/Pend.	Original	2.540	3.081	$-17.80\ [-21.19, -14.41]$	0.0156
PCWM/Pend.	Selected	2.319	3.081	$-24.90\ [-28.89, -20.92]$	0.0156
PCWM/Cart.	Original	3.759	3.777	$-0.47\ [-2.11, +1.18]$	0.5234
PCWM/Cart.	Selected	3.570	3.777	$-5.60\ [-7.35, -3.86]$	0.0156
RSSM/Pend.	Original	2.499	2.953	$-15.56\ [-17.78, -13.33]$	0.0156
RSSM/Pend.	Selected†	2.499	2.953	$-15.56\ [-17.78, -13.33]$	0.0156
RSSM/Cart.	Original	2.946	3.137	$-6.16\ [-8.27, -4.05]$	0.0156
RSSM/Cart.	Selected†	2.946	3.137	$-6.16\ [-8.27, -4.05]$	0.0156

† Selected C equals original C in both RSSM settings. Both planned tests remain in the eight-test Holm family. Effects average percentage changes computed within checkpoints before aggregation.

Table 21: Descriptive percentage changes against validation-selected static optimization under nominal and stationary-high observations. Values show checkpoint means and pointwise 95% t intervals.

Setting	C choice	Nominal Δ [95% CI]	Stationary-high Δ [95% CI]
PCWM/Pend.	Original	$-2.48\ [-7.07, +2.12]$	$-3.46\ [-6.19, -0.73]$
PCWM/Pend.	Selected	$-1.40\ [-6.26, +3.46]$	$-4.71\ [-7.69, -1.74]$
PCWM/Cart.	Original	$+0.44\ [-2.00, +2.87]$	$+7.00\ [+5.34, +8.67]$
PCWM/Cart.	Selected	$-0.10\ [-2.59, +2.38]$	$+3.15\ [+1.55, +4.75]$
RSSM/Pend.	Original	$-3.08\ [-6.37, +0.21]$	$-3.09\ [-3.87, -2.31]$
RSSM/Pend.	Selected†	$-3.08\ [-6.37, +0.21]$	$-3.09\ [-3.87, -2.31]$
RSSM/Cart.	Original	$-0.55\ [-1.63, +0.53]$	$-3.17\ [-4.07, -2.27]$
RSSM/Cart.	Selected†	$-0.55\ [-1.63, +0.53]$	$-3.17\ [-4.07, -2.27]$

Negative changes favor C. These outcomes are outside the primary Holm family. † Selected and original C coincide.

Nominal and stationary-high outcomes were descriptive and excluded from the primary Holm family. The selected configurations' nominal-change intervals included zero in every setting. Under stationary-high noise, selected C increased PCWM/Cartpole error by 3.15% [95% CI $1.55, 4.75$]. An independent CPU audit reconstructed all 216,000 saved forecast losses and reproduced the validation selections and eight primary tests.

C MULTISTEP PREDICTION AND SELECTIVE PREDICTION

C.1 DATA, FORECASTS, AND SELECTION

The multistep evaluation uses ten existing checkpoints for each PCWM/RSSM and Pendulum/Cartpole combination, giving forty frozen checkpoints. Calibration, validation, and development holdout each contain eight physical trajectories per checkpoint, with distinct phase seeds. Every trajectory has 160 transitions. Nominal conditions and local variance faults share physical states, actions, and standard Gaussian measurement draws. Both backbones receive the same external inputs. The nominal measurement standard deviation is 0.05. Faults affect two distinct coordinates in successive intervals. Onset fractions are sampled uniformly from $[0.18, 0.24)$ and $[0.65, 0.71)$. Recovery fractions lie in $[0.54, 0.58)$ and $[0.91, 0.95)$. Each event time is the integer floor of 160 times its sampled fraction. Each interval uses a measurement standard deviation sampled uniformly from $[0.2, 0.4)$ on its affected coordinate.

Filtering uses observations through anchor $a \in \{10, 20, \ldots, 140\}$. The same fourteen anchors support horizons $h \in \{1, 5, 10, 20\}$. Each forecast consumes actions $a, \ldots, a + 19$ and is scored against the clean signal at $a + h$. All arms initialize forecast particles with the corresponding K0 latent variance, while retaining their own filtered state means. Forecasts use $P = 60$ particles and fixed batches of 64 anchors. Within a backbone, arms share forecast random streams and batch boundaries. The shared refinement settings are $K = 1$ and $\eta = 0.5$.

For checkpoint s, condition d, and horizon h, let $L_{s,d,h}$ denote clean-signal MSE, averaged over coordinates, anchors, and trajectories. The selection objective is

$$J_s = \frac{1}{6} \sum_{d \in \{\text{nominal,fault}\}} \sum_{h \in \{5,10,20\}} \frac{L_{s,d,h}}{\max(10^{-8}, L^{\text{cal},K0}_{s,d,h})}. \tag{18}$$

The denominator comes from independent calibration trajectories. The diagonal shape q also uses calibration K0 innovations and decoder-projected prior variances, through the shape normalization in the one-step protocol. The controller retains the calibrated-range clipping interval $[Q_j/16, 128Q_j]$, with $Q_j = R_{0,j} q_j$ and base $B_j = cQ_j$.

Each family minimizes mean validation J_s over the ten checkpoints. Its mean nominal risk must remain within 5% of the selected static reference, with that reference available as a fallback. Static weighting searches 69 scales $c = 2^{k/4}$, $k = -40, \ldots, 28$. C and T each search 72 configurations. Their scales are $\{1/4, 1, 4, 16, 64, 128\}$ for PCWM and $\{2^{-10}, 2^{-8}, 2^{-6}, 2^{-4}, 1/4, 1\}$ for RSSM. Each scale combines $\lambda \in \{1/4, 1/2, 1\}$ with $\rho \in \{0.05, 0.2, 0.5, 1\}$. Huber and IMQ each search 84 configurations. Their common scale grid is $\{2^{-10}, 2^{-8}, 2^{-6}, 2^{-4}, 1/8, 1/4, 1, 2, 4, 16, 64, 128\}$. Huber thresholds are $\{1/8, 1/4, 1/2, 1, 2, 4, 8\}$. IMQ thresholds are $\{0.03, 0.05, 0.1, 0.3, 1, 3, 10\}$. Thus, family budgets differ while the validation data, objective, and nominal constraint are shared.

A predetermined rule allows one boundary extension, capped at 96 candidates per family. An active scale boundary adds an outward fourfold scale, retaining its other parameter combinations. The static grid extends by eight quarter-log steps at an active boundary. A robust threshold boundary adds a halved or doubled threshold when scale extension is unnecessary. The selected configurations triggered no extensions in any of the four settings. Table 22 lists all selected families. The robust reference is chosen between Huber and IMQ on validation data. Development holdout trajectories are generated after all selections are locked.

C.2 AGGREGATE FORECASTS AND COMPUTATION

Table 24 gives the complete aggregate comparison. The aggregate J_s is dimensionless because each MSE term is divided by its calibration counterpart. Table 23 reports mean paired changes

Table 22: Multistep validation-selected configurations. Static scales are rounded for display and lie on the stated quarter-log grid.

Setting	Static c	C (c, λ, ρ)	T (c, λ, ρ)	Huber (c, τ)	IMQ (c, τ)	Robust
PCWM/Pendulum	4.757	16, 0.5, 0.5	16, 0.5, 1	2, 0.5	1, 0.1	Huber
PCWM/Cartpole	9.514	64, 0.5, 0.2	16, 0.5, 1	2, 0.5	0.25, 0.05	Huber
RSSM/Pendulum	0.2102	2^{-8}, 0.5, 0.05	2^{-8}, 0.5, 0.5	2^{-4}, 2	2^{-6}, 0.05	Huber
RSSM/Cartpole	0.1487	2^{-8}, 0.5, 0.05	2^{-8}, 0.5, 1	2^{-4}, 2	2^{-6}, 0.1	IMQ

Table 23: **Multistep development risk and inference cost.** Risk averages independently normalized MSE at $h = 5, 10, 20$ in nominal and local-fault conditions. Changes are checkpoint-paired percentages over ten checkpoints. Nominal reports C/static changes over the same three horizons under nominal sensing. Latency is the synchronized C/static K1 ratio on the first checkpoint of each setting.

Setting	C/static (%)	C/robust (%)	Nominal (%)	Latency
PCWM/Pendulum	-3.80	+0.39	+0.62	1.59×
PCWM/Cartpole	-2.27	-0.05	-1.39	1.64×
RSSM/Pendulum	-10.04	-0.22	-4.95	1.46×
RSSM/Cartpole	-2.93	-0.82	+0.24	1.65×

$10^{-1} \sum_s 100(J_s^C / J_s^{\mathrm{ref}} - 1)$. These differ from ratios of the mean risks in Table 24. All multistep comparisons describe the development holdout with fixed validation-selected parameters. The four-setting mean change relative to robust weighting is -0.18%, below the 2% engineering improvement target in magnitude. The nominal-cost and latency targets are met.

For RSSM, C selects $c = 2^{-8}$ and $\lambda = 1/2$ in both environments. Then $B = Q/256$ and the clipping rule gives $R_t \in [Q/64, Q/\sqrt{2}]$ coordinatewise. Its same-base control disables modulation at a scale four times below the active rule's minimum. The independently selected static scales are approximately 54 and 38 times larger than this base on Pendulum and Cartpole, respectively. Same-base differences quantify disabling modulation at C's chosen base, while separately calibrated static weights provide the deployment reference.

Table 24: Mean calibration-normalized multistep risk J_s over ten checkpoints. Lower values indicate lower risk. Same base retains the selected C base and sets $\lambda = 0$. Robust retains its validation-selected family.

Setting	K0	Static	C	T	Huber	IMQ	Robust	Same base
PCWM/Pendulum	1.1617	1.0816	1.0367	1.0456	1.0350	1.0390	1.0350	1.0956
PCWM/Cartpole	1.1481	1.0404	1.0156	1.0205	1.0172	1.0181	1.0172	1.0431
RSSM/Pendulum	1.2894	1.3419	1.2196	1.2985	1.2281	1.2966	1.2281	24.0270
RSSM/Cartpole	1.2044	1.2252	1.1910	1.2082	1.1914	1.1939	1.1939	2.0786

Filtering latency is measured on checkpoint zero of each setting on an RTX 5090. Each measurement processes sixteen histories of 160 transitions. Two warmup runs precede five synchronized measurements. Ratios use median C and static filtering times under the same workload. The ratios in Table 23 measure the filtering pass before future rollout.

C.3 Horizon-specific risk ranking

The score is the coordinate-averaged clean-signal particle variance at each forecast horizon. For n forecasts ranked by ascending score, normalized AURC is $(n\bar{L})^{-1} \sum_{k=1}^{n} R_k$, where R_k is the retained mean loss. Exact score ties use expected uniformly random ordering. A constant score therefore has nAURC 1. Pooled curves contain 224 forecasts per checkpoint and horizon. Each condition-specific curve contains 112 forecasts.

Table 25 compares K0, static, and C variance scores on the same C forecast errors. The innovation score is the coordinate-averaged squared innovation at the anchor. Separate condition-wise curves show that ranking extends within each sensor condition. K0 and static variance provide similar ranking quality. All arms share initial particle variances, while their corrected means and subsequent nonlinear dynamics can alter forecast variances. The comparison separates prediction error levels from the additional ranking information supplied by refinement.

Table 25: Horizon-20 nAURC, averaged across ten checkpoints. The first four scores rank the same pooled C errors. The final columns rank C errors within nominal and fault conditions separately. Lower is better.

Setting	K0	Static	C	Innovation	C nominal	C fault
PCWM/Pendulum	0.2875	0.2832	0.2816	0.9677	0.2711	0.2961
PCWM/Cartpole	0.4665	0.4632	0.4628	0.9721	0.4590	0.4716
RSSM/Pendulum	0.2056	0.2057	0.2011	0.8952	0.1900	0.2137
RSSM/Cartpole	0.4171	0.4157	0.4146	0.8597	0.4199	0.4154

Thresholds target 50% or 75% acceptance on independent calibration scores from the selected controller. They are fixed before application to development forecasts. Exact threshold ties receive a fixed acceptance probability determined during calibration. Retained risk divides expected accepted loss by expected accepted count. Table 26 reports actual acceptance and the within-checkpoint retained-to-full error change. These quantities characterize selective prediction and its coverage tradeoff.

Table 26: Horizon-20 selective prediction with independent calibration thresholds. All entries are percentages averaged across ten checkpoints. Accept denotes actual development acceptance. $\Delta = 100(L_{\text{retained}}/L_{\text{full}} - 1)$ uses each arm's own complete forecast set as its reference.

Setting	Target	K0 accept	K0 Δ	Static accept	Static Δ	C accept	C Δ
PCWM/Pendulum	50	46.5	-82.2	46.4	-82.5	46.6	-82.4
PCWM/Pendulum	75	71.2	-62.7	70.8	-64.5	70.9	-64.5
PCWM/Cartpole	50	48.5	-60.3	49.1	-60.5	48.7	-60.4
PCWM/Cartpole	75	72.8	-52.2	73.0	-52.5	73.0	-52.7
RSSM/Pendulum	50	46.3	-89.9	46.2	-90.8	46.1	-91.2
RSSM/Pendulum	75	70.9	-77.9	70.8	-78.3	70.7	-80.0
RSSM/Cartpole	50	46.5	-65.5	46.8	-65.7	47.1	-65.2
RSSM/Cartpole	75	72.2	-54.1	72.0	-54.6	72.3	-53.3

C.4 Particle sensitivity and reproducibility

Checkpoint zero in each setting is also evaluated with $P = 120$ for K0, static, and C, retaining the selected configurations and calibration denominators. The two particle resolutions use separately sized Monte Carlo draws. Table 27 reports particle-resolution sensitivity for these four checkpoints. The C/static change shifts by $+0.61$, -0.05, $+0.30$, and $+0.02$ percentage points in table order. All four C/static comparisons retain their improvement direction at both resolutions. This check evaluates K0, static, and C on one checkpoint per setting.

The saved manifest binds source snapshots, checkpoints, candidate pools, calibration records, and selected configurations. Model tensor hashes agree before and after evaluation. An independent verifier recomputes saved prediction losses, action and target indices, risk curves, calibration thresholds, and actual acceptance. The accompanying source data include all 2,560 checkpoint-level raw MSE entries for eight arms, two conditions, and four horizons. The table exporter also retains selected configurations, paired aggregate changes, risk scores, threshold summaries, and input hashes.

Table 27: Particle-resolution sensitivity on checkpoint zero. Entries are percentage changes in J_s. $\Delta_P = 100(J_s^{P=120}/J_s^{P=60} - 1)$. The first two columns compare C with static at their respective resolutions.

Setting	C/static P=60	C/static P=120	K0 Δ_P	Static Δ_P	C Δ_P
PCWM/Pendulum	-2.79	-2.17	-5.49	-6.56	-5.97
PCWM/Cartpole	-2.71	-2.76	-1.22	-1.18	-1.23
RSSM/Pendulum	-10.69	-10.39	-0.80	-0.74	-0.40
RSSM/Cartpole	-4.34	-4.32	-1.20	-1.15	-1.13

D Environments and frozen backbones

Section 4 gives the observation mappings and collection policy. Appendix A gives additional evaluation details.

Environment dynamics and observations. Both tasks are custom stochastic simulators with one continuous action $a_t \in [-1, 1]$. Pendulum has physical state $(\theta, \dot{\theta})$. Mass and length are one, gravity is 10, torque is $2a_t$, and each transition lasts 0.05 seconds. Angular velocity is updated before the angle. Gaussian process noise of standard deviation 0.05 is added to velocity before clipping it to $[-8, 8]$. Initial angle and velocity are uniform on $[-\pi, \pi]$ and $[-1, 1]$. Writing angular velocity as ω, the transition is

$$\omega_{t+1} = \mathrm{clip}_{[-8,8]}\big[\omega_t + 0.05(15\sin\theta_t + 6a_t) + \epsilon_t\big], \qquad \theta_{t+1} = \theta_t + 0.05\omega_{t+1}, \quad \epsilon_t \sim \mathcal{N}(0, 0.05^2).$$

Cartpole has state $(p, \dot{p}, \theta, \dot{\theta})$. Cart mass, pole mass, pole half-length, and gravity are 1, 0.1, 0.5, and 9.8. Force $10a_t$ acts over two 0.02-second integration substeps, followed by angular-velocity process noise of standard deviation 0.05. Cart position is restricted to $[-2.4, 2.4]$, with velocity reset to zero on reaching a wall. Initial position is uniform on $[-0.2, 0.2]$ and pole angle on $[\pi - 0.2, \pi + 0.2]$. Both velocities start at zero. For each Cartpole substep, write $m_c = 1$, $m_p = 0.1$, $l = 0.5$, and $f = 10a_t$. The accelerations are

$$b = \frac{f + m_p l\dot{\theta}^2 \sin\theta}{m_c + m_p}, \qquad \ddot{\theta} = \frac{9.8\sin\theta - b\cos\theta}{l(4/3 - m_p\cos^2\theta/(m_c + m_p))},$$
$$\ddot{p} = b - \frac{m_p l\ddot{\theta}\cos\theta}{m_c + m_p}.$$

Each velocity advances by its acceleration times 0.02, followed by its position using the updated velocity. Wall handling follows each substep, and angular process noise follows both substeps.

Training collection lasts 200 transitions for Pendulum and 250 for Cartpole. The prospective evaluation uses 160 transitions for both. These horizons determine episode termination. Nominal measurement noise has standard deviation 0.05 independently in each normalized signal coordinate.

D.1 Checkpoint training and network architecture

The forty checkpoint/configuration pairs comprise two backbones, two environments, and seeds $0, \ldots, 9$. Each training run collected 200 episodes, reserved the last twenty, and trained on the remaining 180. The prospective calibration, validation, and test sets use separate episodes. All checkpoints use 6,000 training iterations, batches of 128 subsequences with sixteen transitions, Adam (Kingma & Ba, 2015) with learning rate 3×10^{-4}, and gradient-norm clipping at 100. The trainer skips a nonfinite loss or gradient and aborts after more than 100 skips. The controller experiment uses all forty checkpoints with their trained parameters frozen.

PCWM. The stochastic latent dimension is 32. Five separately parameterized transition networks map (z_{t-1}, a_{t-1}) to a residual latent mean and a diagonal Gaussian variance. Each has two width-256 SiLU hidden layers. The observation encoder and separate initializer have the same hidden

architecture and return diagonal Gaussian means and variances. Their inputs are (x_t, z_{t-1}, a_{t-1}), with zero previous state/action for initialization. The shared decoder has two width-256 SiLU hidden layers and a learned, state-independent observation log variance per coordinate. All Gaussian log-variance readouts use $-10 + 12\,\mathrm{sigmoid}(\ell_{\mathrm{raw}})$.

PCWM training uses sampled amortized latent states at $K = 0$ with reparameterized Gaussian sampling (Kingma & Welling, 2014). The loss averages Gaussian reconstruction NLL over seventeen observations. It also includes analytic Gaussian $\mathrm{KL}(q_t \| p_{m,t})$, averaged over sixteen transitions and included ensemble members with weight one. A squared latent norm penalty has weight 10^{-3}. For each member, 180 episode indices are drawn with replacement once. Binary membership determines each episode's fixed mask, which its subsequences inherit. The shared encoder, decoder, and auxiliary objective couple the members. The overshooting loss starts from detached inferred states and propagates a deterministic ensemble-mean chain. At offsets $\{2, 4, 8\}$, each member's final Gaussian scores the detached inferred target. Intermediate means are clipped to $[-30, 30]$ and auxiliary log variances floored at -4. Its weight is $0.5 \min(1, u/1000)$ at zero-indexed training iteration u.

Gaussian RSSM. A GRU with deterministic dimension 128 receives the preceding 32-dimensional stochastic state and action. Prior and posterior MLPs each have one width-256 SiLU hidden layer. The posterior also receives the observation. A two-layer width-256 SiLU decoder maps the concatenated deterministic and stochastic states to observation means. Variance readouts use the same bounded parameterization as PCWM. Training samples stochastic posterior states and minimizes reconstruction NLL plus the per-transition KL term $0.8\,\mathrm{KL}(\mathrm{sg}(q)\|p) + 0.2\,\mathrm{KL}(q\|\,\mathrm{sg}(p))$, floored at one nat after summing latent coordinates. The operator sg stops gradients. RSSM training uses amortized inference. Ensemble masks, overshooting, and the shared configuration's `k` argument apply only to PCWM. This Gaussian RSSM implements observation prediction and latent dynamics.

D.2 Filtering and particle prediction

At every PCWM transition, moment matching across the five member priors gives $\mu_t = M^{-1} \sum_m \mu_{m,t}$ and $v_t = M^{-1} \sum_m [v_{m,t} + (\mu_{m,t} - \mu_t)^2]$ at the previous state estimate. RSSM computes the next deterministic state and its stochastic prior. The deterministic state is held fixed during the current stochastic correction. Both adapters initialize from their observation encoder and refine at every observed transition. The corrected state enters the next transition. The initial observation is encoded without refinement. For RSSM, zero states and a zero-action recurrent step precede this initial posterior. Evaluation depth is independent of checkpoint training options.

For observation dimensions three and five, diagonal GN curvature and decoder-projected prior variance V_t use all decoder Jacobian rows. These contractions use a linearization of the decoder. The diagonal preconditioner is evaluated at the encoder proposal and fixed within each transition's K updates. Corrected stochastic coordinates are clipped to $[-30, 30]$. The initial precision and multistep studies use $K = 1$ and $\eta = 0.5$. For $K > 1$, this is a preconditioned gradient iteration with frozen initial curvature. Appendix B specifies the alternative state optimizers and their selected iteration budgets. K0 uses the amortized proposal.

At each scoring anchor, the predictor samples from that arm's state mean using the shared numerical K0 encoder variance. In the one-step study, PCWM allocates thirty particles to each of its five members. Each particle keeps its member identity. Initial draws follow the Gaussian encoder distribution. Subsequent PCWM transition draws are clipped to $[-30, 30]$ before decoding. RSSM uses 150 particles and samples its stochastic state initially. Its transition draws follow the Gaussian prior without clipping. These rollout conventions are shared across arms within each backbone. The forecast mean averages decoded particles. The model's observation predictive variance adds fixed R_0 to their variance. The multistep study uses $P = 60$ and reads clean-signal variance directly from decoded particles, excluding R_0 (Appendix C).

D.3 Linear analysis for refinement design

This analysis explains how decoder curvature constrains the update's step size. For one linear measurement $g(z) = Cz$, fix the prior mean μ_t, $P = \mathrm{diag}(v_t)$, and $R = \mathrm{diag}(R_t)$, with positive variances. The objective $F_t(z) = \frac{1}{2}(x_t - Cz)^\top R^{-1}(x_t - Cz) + \frac{1}{2}(z - \mu_t)^\top P^{-1}(z - \mu_t)$ has

Hessian $\Lambda = P^{-1} + C^\top R^{-1} C \succ 0$ and minimizer $z_* = \Lambda^{-1}(P^{-1}\mu_t + C^\top R^{-1} x_t)$. Write $D = \mathrm{diag}(\mathrm{diag}\,\Lambda)$ and $\delta_k = z^{(k)} - z_*$. The unclipped error recurrence is $\delta_{k+1} = (I - \eta D^{-1}\Lambda)\delta_k$. Because $D^{-1}\Lambda$ is similar to the positive definite $D^{-1/2}\Lambda D^{-1/2}$, convergence from every initial state holds for

$$0 < \eta < \frac{2}{\lambda_{\max}(D^{-1/2}\Lambda D^{-1/2})}.$$

At $\eta = 0.5$, the largest eigenvalue must therefore be below four. The guarantee applies to an unclipped linear update with fixed P and R. It provides a design reference for the nonlinear implementation, whose curvature varies with state and whose update includes clipping. The depth study measures prediction error under the implemented finite iteration budgets.

The latent-to-observation dimension ratio constrains curvature through its observation contribution. Let $M = D^{-1/2}\Lambda D^{-1/2} = A + G$, where $A = D^{-1/2}P^{-1}D^{-1/2} = \mathrm{diag}(a_i)$ and $G = D^{-1/2}C^\top R^{-1} C D^{-1/2}$. Then $G \succeq 0$, $\mathrm{rank}(G) \le D_x$, and $\mathrm{tr}(G) = D_z - \sum_i a_i$. Consequently,

$$\lambda_{\max}(M) \ge \lambda_{\max}(G) \ge \frac{D_z - \sum_i a_i}{D_x}.$$

With $D_z = 32$, a mean prior share below 0.625 for Pendulum or 0.375 for Cartpole forces this bound above four. These thresholds describe the fixed linear model and depend on its prior precision and decoder Jacobian.

The encoder also enters the update through its initial state. If all observation precisions tend to zero, an unclipped first step satisfies $z^{(1)} \longrightarrow (1-\eta)z^{(0)} + \eta\mu_t$. Thus, one step retains part of the encoder proposal even when the observation contribution vanishes. Under coordinate-specific weighting, the remaining observation gradient and curvature enter the update together with the prior terms.

E Visual-input development pilot

E.1 Images, checkpoints, and local corruption

We extend the observation format of Pendulum to 32×32 grayscale images. Each clean image renders a unit rod from the origin to $(\sin\theta, \cos\theta)$ on a grid spanning $[-1.2, 1.2]^2$. A pixel at distance d from the rod has intensity $\max(0, 1 - d/0.14)$. Angular velocity is inferred from observation history. The flattened observation has 1,024 coordinates, while the stochastic latent state has 32. All five existing PCWM PixelPendulum checkpoints, seeds zero through four, are frozen. They use the five-member architecture and 6,000-step training configuration described in Appendix D.1, with image-sized encoder inputs and decoder outputs.

Each checkpoint receives four calibration, four validation, and eight holdout physical trajectories with 80 transitions each. Separate phase seeds generate physical states, actions, measurement draws, and fault events. The collection policy and physical dynamics follow Appendix D. Nominal and patch-fault histories share physical trajectories and standard-normal measurement draws. Nominal pixel noise has standard deviation 0.05. During two recoverable events, an 8×8 patch has standard deviation 0.30. Each patch is sampled uniformly among valid image locations. The first onset and recovery fractions are uniform in $[0.20, 0.235)$ and $[0.47, 0.505)$. The second uses $[0.62, 0.655)$ and $[0.89, 0.925)$. Event times are the integer floor of 80 times these fractions. Pixels outside the active patch retain exactly the nominal observations. Fault masks and event times are used for scoring and visualization only.

E.2 Inference, selection, and paired prediction

All refining arms use $K = 1$, $\eta = 0.5$, and the current frozen-model interface. T and C use $\rho = 0.2$ and clip excess statistics to $[Q/16, 128Q]$. The high-dimensional implementation uses eight Rademacher probes for each diagonal decoder-curvature and projected-prior-variance estimate. Random streams are paired by numerical role and transition. Requesting a projected variance therefore preserves the curvature probe sequence. The static identity at $\lambda = 0$ is checked with these randomized estimates.

Calibration pools normalized excess statistics with equal condition and checkpoint weights, then constructs one shared shape q across the five checkpoints. The initial shared scale grid is $c = 2^k$ for $k = -6, -4, \ldots, 10$. T and C cross these nine scales with $\lambda \in \{0.25, 0.5, 1\}$. Huber thresholds are $\{0.25, 1, 4\}$ and IMQ thresholds are $\{0.5, 2, 8\}$. This gives nine static and 27 candidates for each of the four other families. Each family also receives the selected static fallback. The selection objective averages six calibration-K0-normalized MSE terms, covering nominal and patch-fault conditions at $h \in \{1, 5, 10\}$. The denominator floor is 10^{-12}. After averaging checkpoints equally, each family minimizes this objective subject to nominal risk at most 1.05 times the selected static risk. The selected Huber and IMQ configurations compete for the robust reference.

The protocol permits one outward grid extension before holdout generation. A selected scale at either boundary adds both endpoints divided and multiplied by four for every family. A selected robust threshold at either boundary extends both threshold grids fourfold at both ends. Existing candidate results are reused, and the full budget is at most 187 candidates per checkpoint. Calibration, selection, and source hashes are locked before generating the holdout.

Predictions use anchors $10, 20, \ldots, 70$, horizons $h \in \{1, 5, 10\}$, and $P = 30$ particles. All arms share the K0 encoder variance, forecast random draws, and batches of eight anchors. Seven arms are evaluated: K0, static, T, C, Huber, IMQ, and static refinement at C's selected base. Whole-image clean-target MSE is the primary score. We report checkpoint-paired percentage changes and standard errors over five checkpoints. The pilot uses existing checkpoints and a small development cohort. Foreground MSE uses target intensity above 0.1, and background MSE uses its complement. Observed-frame persistence and the calibration mean image provide additional prediction references.

E.3 Selected configurations and holdout results

Only robust thresholds triggered the outward extension, giving 153 candidates per checkpoint. The final Huber thresholds were $\{0.0625, 0.25, 1, 4, 16\}$ and IMQ thresholds were $\{0.125, 0.5, 2, 8, 32\}$. All selected scales and robust thresholds were interior to these grids. Static refinement selected $c = 4$. T selected $c = 256$ and $\lambda = 1$. Huber selected $c = 4$ and $\tau_{\mathrm{H}} = 4$, while IMQ selected $c = 1$ and $\tau_{\mathrm{I}} = 2$. The C family selected the shared static fallback, equivalent to $c = 4$ and $\lambda = 0$. The best positive-λ C candidate used $c = 64$ and $\lambda = 0.5$ and satisfied the nominal constraint. Its validation risk was 1.05213, compared with 1.05185 for static refinement. The validation objective therefore selected static refinement.

Table 28 reports all seven evaluated arms. The selected C and its same-base control produced predictions identical to static refinement. Its mean checkpoint-paired aggregate risk changes were -0.0573% relative to K0, -0.0124% relative to T, and $+0.0242\%$ relative to selected IMQ weighting. Across individual condition/horizon pairs, changes from K0 ranged from -0.255% to $+0.178\%$. Figure 10 displays all horizons and both conditions. These results characterize the selected pipeline at the static operating point. Table 29 gives image-region errors and simple image predictors. Figure 11 shows a fixed history and forecast anchor.

Table 28: **Whole-image prediction in the visual pilot.** MSE is multiplied by 10^3 and averaged over five checkpoints. The C family selects the static fallback. IMQ is the validation-selected robust reference.

	Nominal			Patch faults		
Arm	$h = 1$	$h = 5$	$h = 10$	$h = 1$	$h = 5$	$h = 10$
K0	6.932	11.784	13.152	7.023	11.812	13.262
Static	6.945	11.756	13.146	7.024	11.783	13.257
T	6.948	11.758	13.145	7.029	11.779	13.254
C selection ($\lambda = 0$)	6.945	11.756	13.146	7.024	11.783	13.257
Huber	6.944	11.756	13.145	7.023	11.783	13.256
IMQ	6.941	11.754	13.145	7.021	11.781	13.256
C's same base	6.945	11.756	13.146	7.024	11.783	13.257

Verification. All 11,760 stored per-anchor forecast losses were reconstructed from saved means and targets, with maximum absolute discrepancy zero. Targets, future actions, K0 variances, and

Table 29: **Image-region and simple-predictor diagnostics.** The C selection uses $\lambda = 0$. Its foreground and background MSE use clean-target intensity above 0.1 and its complement, respectively. Persistence and calibration-mean columns report whole-image MSE. All values are multiplied by 10^3 and averaged over five checkpoints.

Condition	Horizon	C foreground	C background	Persistence	Mean image
Nominal	1	106.597	1.553	14.490	13.285
Nominal	5	202.922	1.440	26.453	13.256
Nominal	10	231.137	1.396	28.477	13.405
Patch faults	1	107.751	1.573	18.090	13.285
Patch faults	5	203.187	1.453	30.064	13.256
Patch faults	10	233.112	1.407	32.097	13.405

forecast seeds agree exactly across arms. Independent indexing checks match saved targets to the clean frame at $a + h$. Model tensor hashes remain unchanged, and a second selection calculation recovers every chosen configuration.

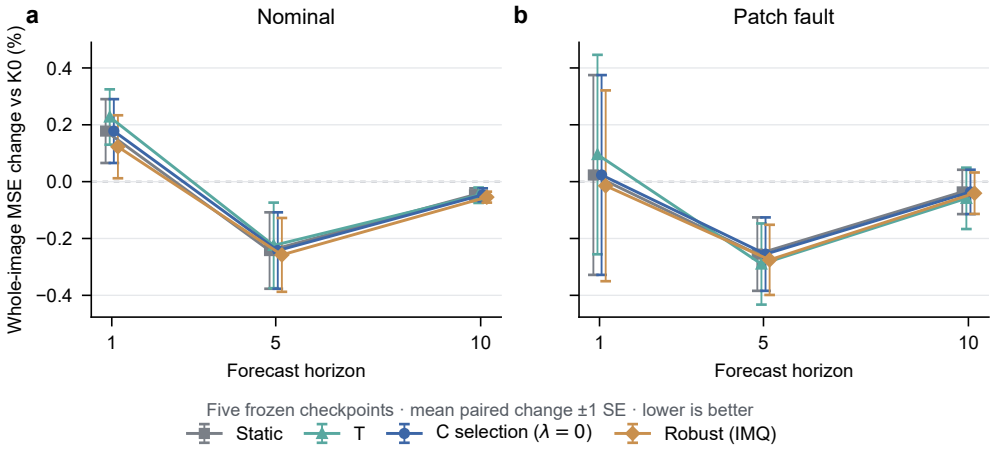

Figure 10: **Visual-pilot prediction changes relative to K0.** Points show mean checkpoint-paired whole-image MSE changes and one standard error over five checkpoints. Static, T, the C selection, and validation-selected IMQ use the same observations and forecast random draws. C selects $\lambda = 0$ and coincides with static refinement. Small horizontal offsets separate overlapping markers.

E.4 Spatial precision and numerical sensitivity

We evaluate the best positive-λ C configuration from the original validation grid, $c = 64$ and $\lambda = 0.5$, together with its same-base static control. This posthoc diagnostic reuses all five checkpoints and eight heldout physical trajectories per checkpoint. It retains the native eight-probe estimators, paired forecast draws, and the original selected static fallback. Figure 12 quantifies the active controller's spatial response. During patch faults, mean precision ratios B/R are 2.40 inside the patch and 10.85 on simultaneous healthy coordinates. The paired nominal values are 12.36 and 10.92. The controller assigns lower relative precision to the corrupted region while retaining larger precision on healthy pixels. These ratios are measured against the $c = 64$ base. The resulting aggregate risk changes are $-0.221 \pm 0.145\%$ against same-base static and $+0.266 \pm 0.050\%$ against independently selected static $c = 4$ (mean checkpoint-paired change $\pm$ one SE). Table 30 separates whole-image and foreground effects at every horizon.

Numerical sensitivity. We fix validation episode zero and its first 40 transitions from each of five checkpoints, then compare four configurations under native and exact diagonal contractions. Exact observation curvature uses 32 decoder JVPs at the encoder proposal. Exact projected prior variance

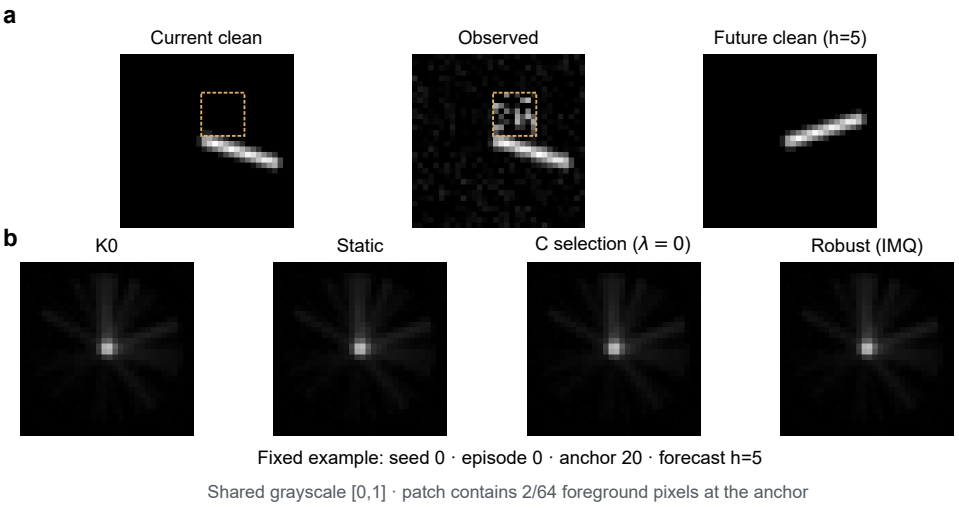

Figure 11: **Fixed visual-pilot input and five-step forecasts.** The example uses checkpoint zero, history zero, and the first scoring anchor during its first fault, $a = 20$. The patch covers two of 64 pixels with current clean intensity above 0.1. All panels share grayscale limits $[0, 1]$. Observed values are clipped only for display, and filters receive the original observations. Forecasts show particle means, with C explicitly retaining its selected static fallback.

uses another 32 at the prior mean. Both native estimators average squared projections and remain nonnegative. For C with $\lambda = 0.5$, native estimates have mean relative L_2 errors of 2.93% for D and 29.75% for V along native traces. Table 31 reports the corresponding forecast risks. A 2×2 follow-up separately replaces D and V by exact contractions for C with $c = 64$, $\lambda = 0.5$ and both static controls. It retains this fixed subset, paired forecast draws, and the original calibration denominators. Making D exact narrows the mean paired C/static difference toward zero. Replacing V alone has a smaller effect here (Table 32). The mean reduction against same-base static remains 0.70–0.98% across all four settings. The fixed subset measures numerical sensitivity of selected configurations. The complete validation grid supplies the reported deployment selection.

Table 30: **Active visual C against both static references.** Entries are checkpoint-paired MSE changes in percent (mean $\pm$ one SE, five checkpoints). Whole denotes whole-image MSE. FG uses clean-target intensity above 0.1. C has $c = 64$, $\lambda = 0.5$ and native probes. Negative values indicate lower error.

		vs. same-base $c = 64$		vs. selected $c = 4$	
Condition	h	Whole	FG	Whole	FG
Nominal	1	-0.785 ± 0.313	-1.690 ± 0.251	$+0.564 \pm 0.095$	$+0.932 \pm 0.145$
	5	$+0.087 \pm 0.093$	-0.115 ± 0.093	$+0.080 \pm 0.029$	$+0.154 \pm 0.029$
	10	$+0.041 \pm 0.077$	-0.056 ± 0.055	-0.002 ± 0.032	$+0.016 \pm 0.025$
Patch fault	1	-0.770 ± 0.265	-1.608 ± 0.177	$+0.937 \pm 0.203$	$+1.251 \pm 0.145$
	5	$+0.110 \pm 0.137$	-0.103 ± 0.118	$+0.040 \pm 0.070$	$+0.142 \pm 0.053$
	10	$+0.076 \pm 0.074$	$+0.010 \pm 0.058$	-0.051 ± 0.036	-0.020 ± 0.034

Table 31: **Exact-contraction sensitivity on a fixed validation subset.** The subset contains episode zero and transitions 1–40 from each of five checkpoints. Forecasts use anchors $10, 20, 30$ and $h \in \{1, 5, 10\}$. J retains the original calibration denominators. D and V errors are checkpoint-mean relative L_2 errors of native estimates at the same local inputs along native traces. The four configurations are fixed diagnostics. The full-grid selection remains unchanged.

Configuration	Native J	Exact J	D error (%)	V error (%)
Static $c = 4$	1.010343	1.009136	3.73	—
Static $c = 64$	1.019038	1.015928	0.24	—
C $c = 64$, $\lambda = 0.5$	1.009471	1.009231	2.93	29.75
C $c = 64$, $\lambda = 1$	1.012007	1.007446	39.75	29.70

Table 32: **Separating curvature and projected-variance contractions.** On the fixed five-checkpoint validation subset in Table 31, entries are $100(J_C/J_{\text{static}} - 1)$ in percent (mean of checkpoint-paired changes $\pm$ one SE). C uses $c = 64$, $\lambda = 0.5$. Each static reference uses the same D setting. V is unused by static controls. Negative values favor C. A mean of paired percentages differs from a ratio of mean J values and can give a different ordering from Table 31.

D	V	C vs. selected static $c = 4$	C vs. same-base static $c = 64$
Native	Native	-0.092 ± 0.061	-0.985 ± 0.234
Exact	Native	-0.004 ± 0.065	-0.697 ± 0.137
Native	Exact	-0.078 ± 0.071	-0.970 ± 0.233
Exact	Exact	-0.008 ± 0.080	-0.701 ± 0.155

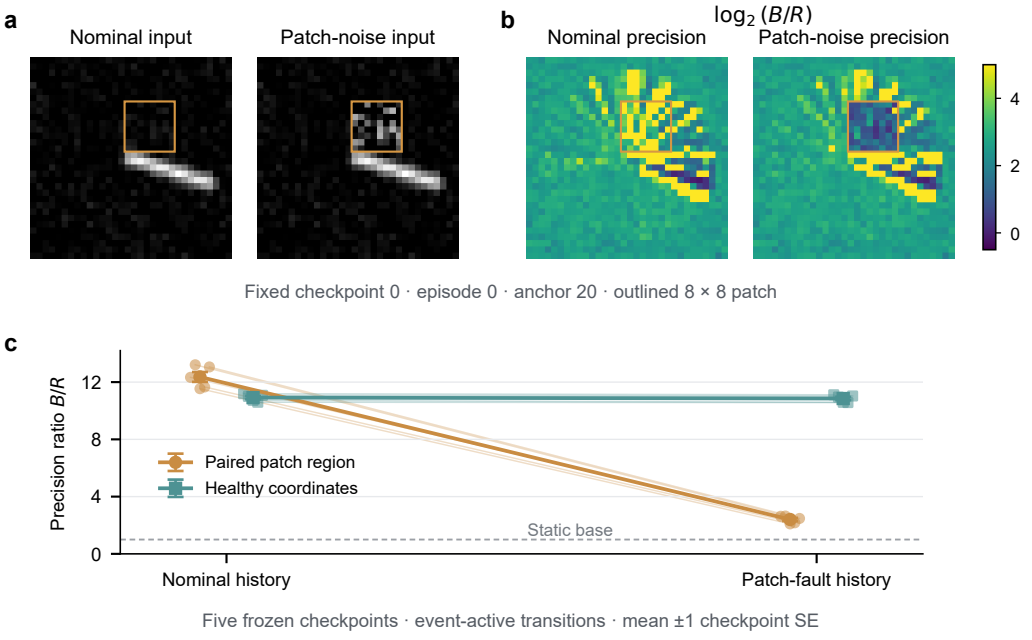

Figure 12: **Spatial precision response of active C in image space.** **a** Paired nominal and patch-noise inputs for fixed checkpoint zero, episode zero, and anchor 20. Grayscale values are clipped to $[0, 1]$ only for display. **b** Effective precision maps share the scale $\log_2(B/R)$ for $c = 64$, $\lambda = 0.5$. Outlines mark the paired patch location. **c** Precision ratios averaged over event-active heldout transitions. Small points and thin lines show the five checkpoint pairs. Large markers and bars show checkpoint means and one SE. The nominal patch region uses the paired fault schedule as an evaluation label. Masks enter scoring only. The dashed line denotes the static $c = 64$ base.