EFFICIENT RARE-EVENT ESTIMATION CAN DISTORT CONDITIONAL TRAJECTORIES

Anonymous authors
Paper under double-blind review

ABSTRACT

Rare-event proposals can supply both probability estimates and trajectories used to explain how an event occurs. On a controlled diffusion, learning raises estimation efficiency, reduces estimated global conditional fidelity, and improves route-conditioned hitting-time accuracy on the same evaluation paths. This combination occurs in all five original models and all five new fits trained through the deployed control cap, with evaluation on fresh paths. RareFlow characterizes this separation through the exact factorization $S = HF$, where H is event discovery and F is the reciprocal of one plus the forward χ^2 conditional discrepancy. A mechanism decomposition separates route allocation from within-route distortion. An independent guide-family comparison shows that nearly matched route shares can conceal timing bias, while an exact conditional-law ceiling limits what route reallocation alone can repair. A finite-state control exposes severe overestimation of this ceiling when an important submode receives sparse proposal coverage. A second planar landscape reproduces the qualitative discovery–fidelity separation. Together, these results show that improved probability estimation does not by itself certify faithful unweighted event trajectories, and that full conditional fidelity and observable-specific accuracy should be evaluated separately.

1 INTRODUCTION

An importance sampler can estimate a rare-event probability accurately while producing misleading unweighted examples of how that event occurs. The distinction matters when sampled trajectories are inspected, grouped by failure mechanism, or reused as equally weighted training examples. Probability-estimator efficiency (lower variance of the probability estimate for a fixed number of attempted paths), fidelity of the full conditional path law, and accuracy of a chosen trajectory summary are distinct evaluation targets. When an approximate rare-event proposal becomes more efficient for probability estimation, do its unweighted successful trajectories also become more faithful to the target event-conditioned path law? Our controlled experiments show that learned rare-event proposals can improve probability-estimation efficiency while reducing global conditional fidelity, even as the accuracy of a selected trajectory observable improves.

For probability estimation, the central aim is to obtain accurate results with a limited simulation budget. Importance sampling and splitting address this aim by reallocating simulation effort (Sadowsky & Bucklew, 1990; Cérou & Guyader, 2007), while analyses of suboptimal importance sampling quantify how proposal mismatch affects estimator error (Hartmann & Richter, 2024). Using event trajectories to understand how an event occurs brings a second objective into focus: representing the event-conditioned path ensemble. Transition-path sampling and theory study this ensemble and its mechanisms (Dellago et al., 1998; Bolhuis et al., 2002; E & Vanden-Eijnden, 2006), and variational path sampling connects its construction to controlled dynamics and reweighting (Du et al., 2024; Singh et al., 2025). This attention to path distributions is also reflected in recent evaluations, which examine channel populations, transition statistics, and failure-sample density or coverage alongside sampling performance (Asghar et al., 2024; Seong et al., 2025; Delecki et al., 2025). Together, these lines of work characterize estimation efficiency and event-conditioned trajectories from complementary perspectives, but do not determine whether improving an approximate learned proposal for probability estimation also improves the conditional law of its unweighted event paths.

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.

The difficulty is that changing a proposal can alter both the event hit rate and the conditional path distribution, while the estimator variance combines their effects. More frequent hits can offset the variance penalty of conditional mismatch, improving probability estimation even as the generated event paths become less faithful. Importance weighting corrects estimates of target expectations, while unweighted event paths still follow the proposal's conditional law. Mechanism frequencies describe route allocation: matching them can conceal distortion within each mechanism (Figure 1; Section 2.4). We therefore separate the sources of efficiency gains, localize conditional mismatch across and within mechanisms, and quantify how much of that mismatch can be repaired by changing mechanism allocation alone.

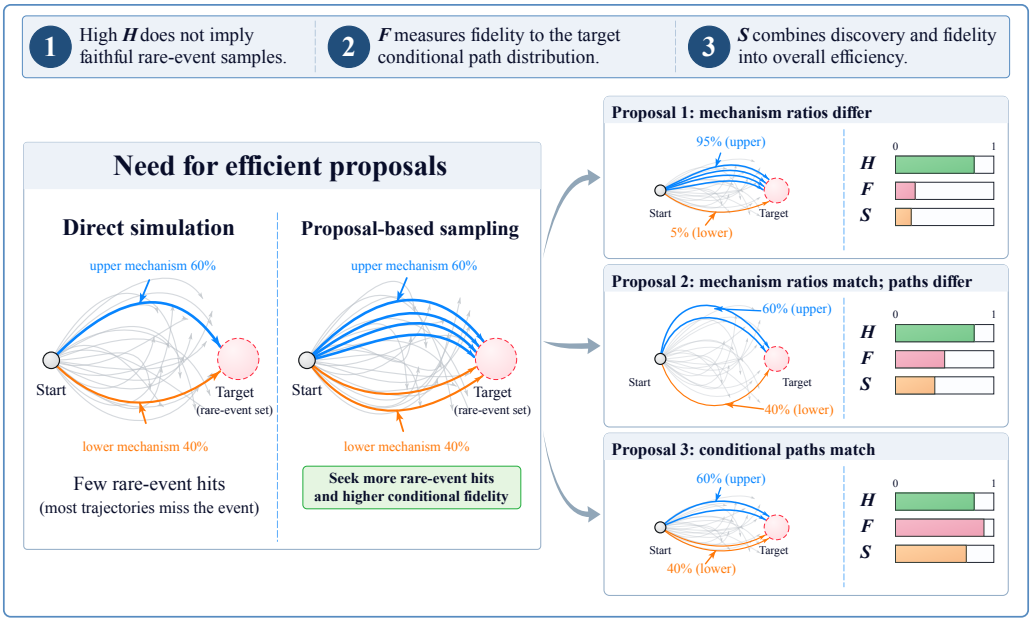

Figure 1: **Event discovery and conditional fidelity describe distinct aspects of proposal quality.** At the same event discovery H, proposals can differ in their conditional path distributions. Proposal 1 distorts mechanism allocation; Proposal 2 matches the illustrated mechanism proportions but distorts paths within mechanisms; Proposal 3 has high conditional fidelity. Percentages and bars are schematic; gray paths denote misses.

RareFlow makes this distinction quantitative using the weighted paths already produced by the proposal. The exact factorization $S = HF$ separates probability-estimation efficiency S into event discovery H and conditional path fidelity F, while two event-ESS normalizations estimate all three quantities from the same proposal sample (Agapiou et al., 2017; Hua et al., 2024). A mechanism decomposition further separates conditional mismatch into mechanism-allocation and within-mechanism components; combined with $S = HF$, it attributes relative estimator variance to event misses, mechanism misallocation, and within-mechanism distortion. We further derive the fidelity-optimal route allocation, an exact ceiling on the fidelity achievable by route reallocation alone, and an allocation-regret identity. Finally, an observable bound connects conditional mismatch to bias in trajectory summaries. Together, these results diagnose how an efficiency gain is achieved, where conditional distortion remains, and how much of it mechanism reallocation can repair.

Figure 4 previews three interventions on a modified Müller–Brown landscape (Müller & Brown, 1979): route feedback lowers fidelity at similar discovery; residual learning improves efficiency despite a fidelity loss; and guide refinement can improve both. Fresh follow-up evaluations show that efficiency and hitting-time accuracy improve together as global fidelity falls, for both the original models and new fits trained through the deployed cap. The guide-family timing study supplies a different counterexample: nearly matched route shares conceal within-route timing bias.

Gains in probability-estimation efficiency can coexist with lower global conditional fidelity, while agreement in coarse mechanism frequencies can still conceal substantial within-mechanism distortion and a selected trajectory observable can behave differently from full-law fidelity. The route-only ceiling and hidden-submode stress test further characterize what mechanism reallocation can repair and when that repairability can be reliably estimated. A second planar landscape reproduces the qualitative discovery–fidelity separation across three noise levels, showing that the central effect is not tied to a single geometry. Rare-event proposals therefore require separate evaluation as probability estimators and as generators of unweighted event trajectories, a distinction directly relevant to learned simulators and failure-search methods (Uesato et al., 2019; Delecki et al., 2025).

2 RareFlow: Auditing Conditional Path Fidelity

RareFlow separates probability-estimation efficiency from conditional path fidelity through the factorization $S = HF$, then locates the remaining conditional mismatch across and within supplied mechanisms. The same weighted paths support these diagnoses and connect them to estimator variance.

2.1 Path-Space Setting and Audit Inputs

Let $\mathcal{X} = \mathbb{R}^d$ be the state space, let $\mathcal{B} \subset \mathcal{X}$ be a measurable target set, and fix a horizon n and step size Δt. The numerical target is the law $\mathbb{P}$ of the Euler–Maruyama chain initialized at a fixed $\mathbf{x}_0 \notin \mathcal{B}$,

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{b}(\mathbf{x}_k)\Delta t + \sigma\sqrt{\Delta t}\,\boldsymbol{\xi}_k, \qquad \boldsymbol{\xi}_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_d), \qquad \sigma^2 = 2\varepsilon. \tag{1}$$

Define $\tau_{\mathcal{B}} = \inf\{k \geq 0 : \mathbf{x}_k \in \mathcal{B}\}$, $\bar{\tau} = \min\{\tau_{\mathcal{B}}, n\}$, and the rare event $\mathcal{E} = \{\tau_{\mathcal{B}} \leq n\}$. A realized stopped path is $\boldsymbol{\omega} = (\mathbf{x}_0, \ldots, \mathbf{x}_{\bar{\tau}})$, and the estimand is $p_{\mathcal{E}} = \mathbb{P}(\mathcal{E})$. All exactness statements below concern this discrete chain. Continuous-diffusion accuracy is a separate discretization question.

The audit takes a fixed proposal $\mathbb{Q}$ with $\mathbb{P} \ll \mathbb{Q}$, $p_{\mathcal{E}} > 0$, and finite event-weight second moment. Its inputs are independent stopped paths, their event labels $I_i = \mathbf{1}_{\mathcal{E}}(\boldsymbol{\omega}_i)$, and exact likelihood ratios $w_i = (\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q})(\boldsymbol{\omega}_i)$ evaluated with the same stopping rule. Define $\nu_i = I_i w_i$ and the unbiased probability estimate $\widehat{p}_N = N^{-1}\sum_i \nu_i$. Mechanism labels are optional diagnostic inputs. These requirements specify what must be recorded for any proposal under audit; the particular guided proposals used below are described in Appendix B.

2.2 Discovery, Conditional Fidelity, and Efficiency

Let $\mathbb{P}_{\mathcal{E}} = \mathbb{P}(\cdot \mid \mathcal{E})$ and $\nu = \mathbf{1}_{\mathcal{E}}\,\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}$. Define the total proposal discrepancy D_{total} by

$$D_{\text{total}} := \frac{\mathrm{Var}_{\mathbb{Q}}(\nu)}{p_{\mathcal{E}}^2} = \int \left(\frac{\mathrm{d}\mathbb{P}_{\mathcal{E}}}{\mathrm{d}\mathbb{Q}} - 1\right)^2 \mathrm{d}\mathbb{Q} = \chi^2(\mathbb{P}_{\mathcal{E}} \,\|\, \mathbb{Q})\,. \tag{2}$$

For N independent paths, D_{total}/N is the relative mean-squared error of $\widehat{p}_N$.

Define event discovery as $H = \mathbb{Q}(\mathcal{E})$ and let $\mathbb{Q}_{\mathcal{E}} = \mathbb{Q}(\cdot \mid \mathcal{E})$. Define the conditional discrepancy

$$D_{\text{shape}} := \chi^2(\mathbb{P}_{\mathcal{E}} \,\|\, \mathbb{Q}_{\mathcal{E}})\,. \tag{3}$$

If $H > 0$ and $D_{\text{total}} < \infty$, conditioning gives, on $\mathcal{E}$, $\mathrm{d}\mathbb{P}_{\mathcal{E}}/\mathrm{d}\mathbb{Q} = H^{-1}\mathrm{d}\mathbb{P}_{\mathcal{E}}/\mathrm{d}\mathbb{Q}_{\mathcal{E}}$. Consequently,

$$D_{\text{total}} = \frac{1-H}{H} + \frac{D_{\text{shape}}}{H}, \qquad 1 + D_{\text{total}} = \frac{1 + D_{\text{shape}}}{H}. \tag{4}$$

Equivalently, define conditional fidelity $F = (1 + D_{\text{shape}})^{-1}$ and overall efficiency $S = (1 + D_{\text{total}})^{-1}$. Then

$$S = HF. \tag{5}$$

Discovery H measures reaching the event; fidelity F compares the two conditional path laws. Holding $\mathbb{Q}_{\mathcal{E}}$ fixed while changing H leaves F unchanged, and $F = 1$ precisely when the conditional laws agree. Equation (5) connects these distinct notions of quality: an efficiency gain can combine more hits with worse conditional matching. For the same target, higher S means lower relative variance per attempted path; computational cost is accounted for separately.

2.3 Estimation from Weighted Proposal Paths

Importance-weight ESS is a widely used sampling diagnostic (Elvira et al., 2022), and successful-path ESS has been used to assess rare-event importance sampling (Hua et al., 2024). Here its two normalizations answer different questions. Suppose that $M > 0$ of the N paths hit the event and $\sum_i \nu_i > 0$. Let $I_i = \mathbf{1}_{\mathcal{E}}(\boldsymbol{\omega}_i)$. The empirical estimate of event discovery is $\widehat{H}_N = M/N = N^{-1}\sum_{i=1}^N I_i$. Normalizing event ESS by all attempted paths estimates per-path efficiency:

$$\widehat{S}_N = \frac{\mathrm{ESS}_{\mathcal{E}}}{N} = \frac{(\sum_{i=1}^N \nu_i)^2}{N\sum_{i=1}^N \nu_i^2}. \tag{6}$$

Then $\widehat{S}_N \le \widehat{H}_N$ by Cauchy–Schwarz. Normalizing instead by the event hits estimates conditional fidelity among those paths, giving

$$\widehat{F}_N = \frac{\widehat{S}_N}{\widehat{H}_N} = \frac{(\sum_{i=1}^N \nu_i)^2}{M\sum_{i=1}^N \nu_i^2}, \qquad \widehat{D}_{\mathrm{shape},N} = \frac{\widehat{H}_N}{\widehat{S}_N} - 1. \tag{7}$$

Thus $\widehat{S}_N = \widehat{H}_N\widehat{F}_N$ is the finite-sample accounting counterpart of the population identity. Under the stated moment assumptions (Proposition 3, Appendix A), $\widehat{H}_N \to H$, $\widehat{S}_N \to S$, and $\widehat{F}_N \to F$. The discrepancy statistic converges to D_{shape}. Thus one proposal sample and its exact likelihood ratios suffice to estimate all three coordinates. Finite-sample precision depends on how well the sample represents the event-weight distribution.

Exact weights must use the full stopped-path density of the proposal being sampled. For the experimental mixture $q_{\mathrm{mix}} = \sum_j \pi_j q_j$, we evaluate $w = p/q_{\mathrm{mix}}$ by summing every component density with its mixture mass. The unguided component has $q_0 = p$ and mass $\pi_0 > 0$, so $q_{\mathrm{mix}} \ge \pi_0 p$, ensuring support and a finite event-weight second moment. Appendix B.2 gives the Gaussian path ratios and mixture calculation; Appendix A proves the bound. Figure 2 summarizes how the same records produce all three audit coordinates. For applications using estimated path ratios, the statistic also reflects ratio-estimation error. Relating it to the intended target discrepancy requires control of that error (Cranmer et al., 2015). The results below use exact ratios throughout.

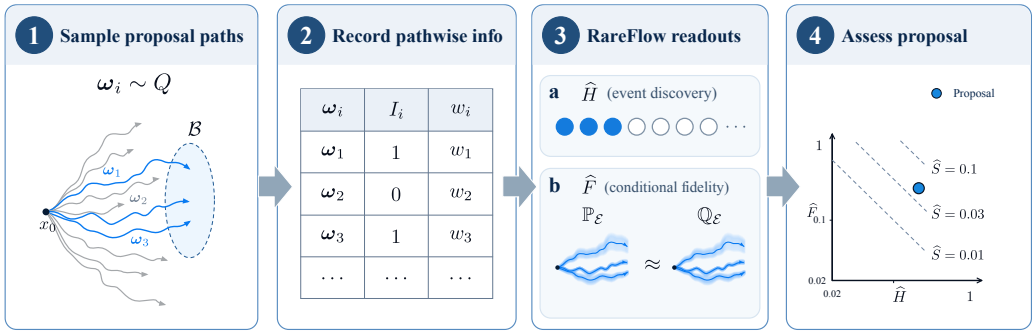

Figure 2: **RareFlow audit workflow.** Event labels and exact stopped-path likelihood ratios from the same proposal sample yield empirical readouts of event discovery, conditional fidelity, and overall efficiency. The right panel uses log–log discovery–fidelity axes; dashed lines hold $\widehat{S} = \widehat{H}\widehat{F}$ fixed. Path sketches and contour levels are schematic.

2.4 Mechanism Decomposition and Trajectory Summaries

Once conditional mismatch is detected, the next question is where it lies. Suppose a supplied, interpretable label Z partitions event paths into a finite set of mechanisms. Write $p_z = \mathbb{P}_{\mathcal{E}}(Z = z)$, $r_z = \mathbb{Q}_{\mathcal{E}}(Z = z)$, $\mathbf{p} = (p_z)_z$, $\mathbf{r} = (r_z)_z$, and $D_z = \chi^2(\mathbb{P}_{\mathcal{E}}(\cdot \mid Z = z)\|\mathbb{Q}_{\mathcal{E}}(\cdot \mid Z = z))$. With the support conditions in Proposition 5 (Appendix A.1), conditioning within each mechanism gives

$$D_{\mathrm{shape}} = \chi^2(\mathbf{p}\|\mathbf{r}) + \sum_{z:p_z>0} \frac{p_z^2}{r_z} D_z. \tag{8}$$

The first term, D_{between}, measures mechanism-allocation discrepancy; the weighted sum, D_{within}, measures discrepancy within mechanisms. Hit fractions and normalized event-weight sums estimate r_z and p_z; $\widehat{D}_z = M_z/\,\mathrm{ESS}_z - 1$ uses mechanism hit count M_z and within-mechanism ESS for positive mechanism weight sums. Substitution into Equation (8) gives the same-sample empirical decomposition (Appendix D.7). Combining Equations (4) and (8) gives

$$\frac{\mathrm{Var}_{\mathbb{Q}}(\nu)}{p_{\mathcal{E}}^2} = \frac{1-H}{H} + \frac{D_{\text{between}}}{H} + \frac{D_{\text{within}}}{H}.$$

The three terms correspond to event misses, mechanism allocation, and within-mechanism distortion. Matching mechanism frequencies sets the allocation term to zero, leaving the within-mechanism term. The supplied labels determine the diagnostic partition.

The decomposition also limits route-only corrections. For fixed within-route proposal laws and finite D_z on target-positive routes, the optimal route masses satisfy $a_z^* \propto p_z\sqrt{1+D_z}$, so target-share matching is optimal precisely when the active divergences are equal. Proposition 6 in Appendix A.1 gives the exact ceiling, allocation-regret identity, and empirical estimator. This counterfactual varies route masses at fixed within-route laws. Dynamic interventions can also change discovery, cost, and those laws. The plug-in ceiling estimates the population quantity from observed route moments.

To relate this mismatch to an interpretable trajectory property, apply Cauchy–Schwarz to a summary $g \in L^2(\mathbb{Q}_{\mathcal{E}})$. For finite D_{shape}, Proposition 5 gives

$$\left|\mathbb{E}_{\mathbb{P}_{\mathcal{E}}} g - \mathbb{E}_{\mathbb{Q}_{\mathcal{E}}} g\right|^2 \le D_{\text{shape}}\,\mathrm{Var}_{\mathbb{Q}_{\mathcal{E}}}(g). \tag{9}$$

For example, g can indicate a mechanism or record a bounded severity attribute. Under the same conditions, the argument applies within any mechanism with $p_z, r_z > 0$:

$$\left|\mathbb{E}_{\mathbb{P}_{\mathcal{E}}(\,\cdot\,|Z=z)} g - \mathbb{E}_{\mathbb{Q}_{\mathcal{E}}(\,\cdot\,|Z=z)} g\right|^2 \le D_z\,\mathrm{Var}_{\mathbb{Q}_{\mathcal{E}}(\,\cdot\,|Z=z)}(g).$$

Thus D_z controls observable bias within mechanism z, while D_{shape} controls it over the full event-conditioned population. High fidelity limits this bias relative to the corresponding proposal variance; low fidelity leaves room for distortion whose magnitude depends on the observable. Section 3.5 applies this mechanism-conditioned view to normalized first-hitting time, $g = \tau_{\mathcal{B}}/n$, compared separately within each route.

3 Controlled Evaluation

3.1 Testbed and Evaluation Design

The primary evaluation uses a rare-event system with two transition routes to examine event discovery, route allocation, and path evolution within each route separately. The task is a modified Müller–Brown potential (Müller & Brown, 1979). Paths start at $(-0.050, 0.467)$ and seek the radius-0.10 target centered at $(0.623, 0.028)$ within time 3.0, with $\Delta t = 0.005$. The lower and upper channels define the mechanism label Z used for both the decomposition and the later within-route timing comparison. Figure 3 illustrates how proposals can concentrate on one route or match route shares; matching those shares alone does not ensure fidelity within each route.

The base guiding drift is computed from a numerical approximation to the finite-horizon committor, the probability of reaching the target from the current state within the remaining time (Appendix B.1). We make three controlled comparisons. *Route feedback* switches a route-specific drift term on or off while holding the base guide and other proposal settings fixed. *Residual learning* fits a drift correction to a frozen 11×11 coarse guide with route feedback disabled, using independent development paths and exact weights. *Guide refinement* varies the base guide's grid resolution under a fixed route-feedback proposal configuration. Appendix B gives proposal construction and training details.

Within each comparison, the target chain, exact likelihood calculation, and defensive mass are fixed. Development and evaluation data are disjoint, and fitted proposals are frozen before evaluation. We report $\widehat{H}$, $\widehat{F}$, $\widehat{S}$, and path-variance ratios, suppressing the subscript N. Each interval specifies its contrast and resampling units, so its interpretation follows the corresponding evaluation design. Appendix C gives configurations, budgets, and seeds.

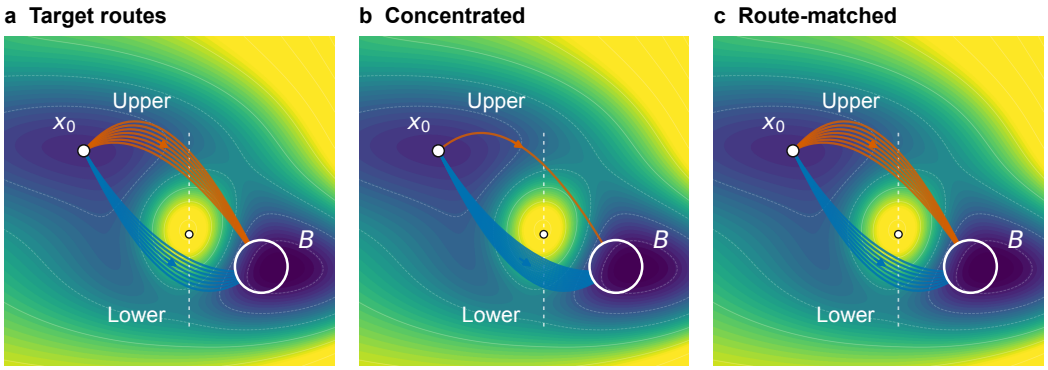

Figure 3: Two-route geometry and schematic route allocation: (a) an illustrative target allocation, (b) concentration on the lower route, and (c) matching the illustrated target shares. The filled white point marks $\mathbf{x}_0$; the white ring marks target B (radius 0.10). Blue and orange curves follow the lower and upper routes. Curves and counts are schematic, not sampled paths or measured route probabilities. The dashed line marks $x = 0.35$ and its split point $y = 0.15$. For each successful path, the linearly interpolated last left-to-right crossing before the target hit assigns Z: upper for $y \geq 0.15$, lower otherwise. Paths with no recorded crossing use the hit height.

3.2 Learning Improves Efficiency and Timing While Reducing Fidelity

Residual learning raises efficiency and improves hitting-time accuracy while reducing global conditional fidelity (Table 1). We first examine the efficiency–fidelity trade-off in the original experiment (Figure 4(b)), then combine all three measurements in fresh evaluations. Route feedback and guide refinement provide complementary controls. In the discovery–fidelity plane, rightward motion raises discovery, upward motion raises fidelity, and higher $\widehat{S}$ contours indicate greater per-path estimation efficiency.

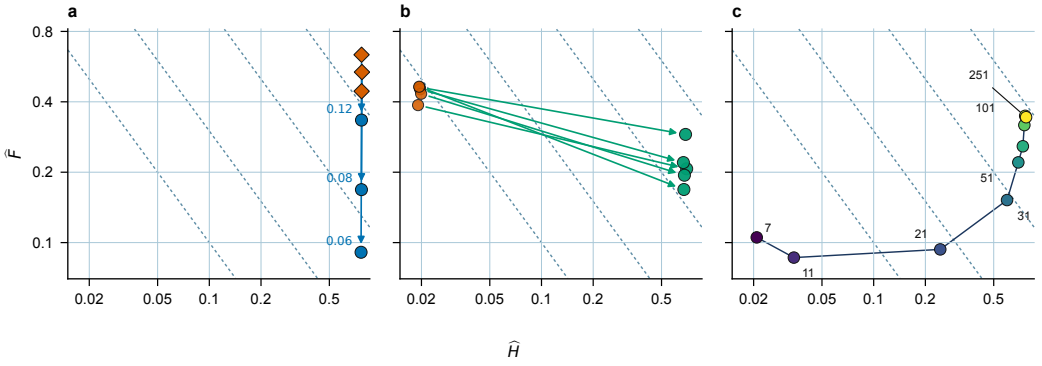

Figure 4: Three proposal interventions in the $(\widehat{H}, \widehat{F})$ plane with log–log axes. Dashed blue lines mark constant $\widehat{S} = \widehat{H}\widehat{F}$. **(a)** Arrows connect neutral proposals (orange diamonds) to route feedback (blue circles); labels give ε. **(b)** Arrows connect each frozen coarse guide (orange) to its learned residual proposal (green), across five independent runs. **(c)** Labels give guide grid width; connected points progress from 7×7 (purple) to 251×251 (yellow). Panels (b) and (c) use $\varepsilon = 0.12$.

Route feedback: similar discovery, lower fidelity. The neutral arm removes only the route-specific drift term. At $\varepsilon = 0.12$, feedback changes discovery from 0.770 to 0.769, but reduces fidelity from 0.636 to 0.334. Its path variance is 2.77 times the neutral variance (95% independent-arm batch-bootstrap interval $[2.39, 3.23]$). The same pattern holds at $\varepsilon = 0.08$ and 0.06 (Table 4). The nearly unchanged hit rate therefore conceals markedly different conditional path matching. We next test whether higher overall estimation efficiency reliably indicates higher fidelity.

Residual learning: higher overall efficiency, lower fidelity. Five networks trained on independent development samples are evaluated against their frozen coarse guides using independent evaluation paths. All five raise discovery and overall estimation efficiency while reducing fidelity (Figure 4(b)). The hierarchical-bootstrap median coarse-to-learned path-variance ratio is 18.85 (95% interval $[14.33, 24.54]$). The mean learned-minus-coarse fidelity change is -0.224 (post hoc 95% hierarchical bootstrap interval $[-0.268, -0.179]$; Appendix D.2). The large discovery gain outweighs the fidelity loss: probability estimation improves even as unweighted event paths become less faithful. Charging the 10^5 development paths to the learned arm reduces the pooled variance gain from 18.64 to 9.32 under equal total path budgets. Including recorded sampling and optimization time gives a descriptive gain of 2.57; this normalization uses the recorded stage times and assumes linear sampling throughput (Appendices D.2 and E).

Control-cap consistency: the trade-off persists. If the trade-off arose solely from fitting an auxiliary uncapped control, training through the deployed cap should remove it. A separately frozen follow-up tests five new pairs with shared development data, initialization, and training budgets, changing only this objective (Appendix D.9). All five consistent-objective fits raise efficiency, lower global fidelity, and reduce route-time RMSE relative to their coarse baselines (Table 1). The efficiency–fidelity trade-off and the timing improvement thus coexist under this implementation control.

Guide refinement: joint improvement in discovery and fidelity. Guide refinement improves both coordinates overall. With the other proposal settings fixed, the nine-resolution sweep raises discovery, conditional fidelity, and estimation efficiency (Figure 4(c)). From 7×7 to 251×251, $\widehat{H}$ rises from 0.021 to 0.772, while $\widehat{F}$ rises from 0.105 to 0.344. The trend includes local reversals: fidelity falls at the first refinement step and changes little between the two finest grids. Appendix D.3 reports the RMS control-error analysis and shared validation checks.

Second landscape: similar discovery still conceals lower fidelity. A post hoc reanalysis of the frozen three-hole transfer records gives the same qualitative audit outcome. In the higher-power execution, feedback changes $\widehat{H}$ only from 0.594 to 0.602 at $\varepsilon = 0.22$ and from 0.537 to 0.524 at $\varepsilon = 0.15$, while $\widehat{F}$ falls from 0.176 to 0.034 and from 0.086 to 0.016, respectively. The direction holds at all three noise levels in the original, higher-power, and raised-cap records. Appendix D.6 reports all nine comparisons and the original scaling analysis.

Higher efficiency, lower global fidelity, and improved hitting-time accuracy occur together in both learned comparisons. Table 1 combines the fresh original-model evaluation and the consistent-cap comparison. Each arm uses the same 10^5 paths for S, F, and route-conditioned mean times; timing errors use a common independent 8×10^5-attempt crude reference. Each of the five original-model pairs and five consistent-cap pairs has $\Delta S > 0$, $\Delta F < 0$, and $\Delta R < 0$.

Table 1: Learning improves efficiency and hitting-time accuracy while reducing global fidelity. Entries are mean learned-minus-coarse changes over five fits, with marginal 95% hierarchical bootstrap intervals below. R is the equally weighted RMSE of the two unweighted route-conditioned mean normalized hitting times. Appendix D.9 gives every fit and the declared comparisons.

learned comparison	ΔS	ΔF	ΔR
Original models	$+0.1353$	-0.246	-0.143
	$[0.1139, 0.1615]$	$[-0.278, -0.210]$	$[-0.160, -0.130]$
Consistent cap	$+0.1730$	-0.197	-0.138
	$[0.1434, 0.1988]$	$[-0.243, -0.155]$	$[-0.157, -0.120]$

The next sections locate the path-law discrepancy and examine the observable-specific interpretation. Appendix E gives the common-scale comparison and computational costs.

3.3 Near-Matched Route Shares Conceal Within-Route Distortion

We apply Equation (8) to the guide-resolution sweep. The lower-route hit fraction $\widehat{r}_{\text{lower}}$ and its importance-weighted counterpart $\widehat{p}_{\text{lower}}$ estimate proposal and target route shares. Within-route mismatch dominates all nine resolutions; Table 2 shows three and Table 9 gives the full sweep.

Table 2: Route allocation can nearly match the target while within-route mismatch remains large. The decomposition in Equation (8) uses same-sample importance-weighted target shares, with 2×10^4 paths per resolution (Appendix D.7).

grid	$\widehat{r}_{\text{lower}}$	$\widehat{p}_{\text{lower}}$	$\widehat{D}_{\text{between}}$	$\widehat{D}_{\text{within}}$	$\widehat{F}$
7	0.705	0.616	3.81×10^{-2}	8.462	0.105
31	0.560	0.557	2.66×10^{-5}	5.578	0.152
251	0.667	0.595	2.37×10^{-2}	1.885	0.344

At 31×31, the unweighted route shares nearly match the same-sample weighted target-share estimates, yet the within-route term remains 5.578. Refinement to 251×251 reduces it to 1.885, even though the allocation discrepancy increases. Route allocation and within-route evolution thus contribute separately to conditional path fidelity.

3.4 Route Reallocation Has Little Estimated Headroom

The route-only ceiling predicts at most 2.71% relative fidelity gain across the sweep, and 0.11% at the finest grid (Appendix D.4), using pooled route moments for this post hoc estimate. A finite-state control verifies the allocation prediction and tests its sensitivity to sampling coverage. The hidden-submode case gives a median estimated ceiling of 1.000 against a true value of 0.108 at 4,000 attempts (Appendix D.5). The diagnostic therefore directs attention to within-route laws and to the sampling coverage needed to estimate their discrepancies.

3.5 Within-Route Distortion Changes Hitting-Time Conclusions

Guide-family timing. The original timing audit concerns feedback-on guides, distinct from the feedback-off learned proposals. It compares normalized first-hitting time $g = \tau/T = \tau_{\mathcal{B}}/n$ separately within each route, where $\tau = \Delta t\, \tau_{\mathcal{B}}$ and $T = n\Delta t$. This isolates within-route timing from changes in the route proportions. An independent crude target reference supplies $2{,}264$ event paths. The reference and each frozen 7×7, 31×31, and 251×251 proposal are evaluated with four independent outer seeds. Reference, proposal-evaluation, and inherited guide-sweep samples are disjoint; Appendix D.8 gives sample counts and aggregation.

For route z, let Δm_z denote the proposal's mean normalized hitting time minus the corresponding target-reference mean. The 31×31 guide has a lower-route share closer to the target reference than the 251×251 guide (absolute errors 0.040 versus 0.063, computed before rounding; Appendix D.8), yet larger route-conditioned mean-time errors. Its Δm_z values are $+0.153$ in the lower route and $+0.276$ in the upper route, compared with -0.009 and $+0.073$ for 251×251 (Figure 5; Table 10).

Across the three guides, fidelity and route-time accuracy give different orderings: 7×7 has lower inherited $\widehat{F}$ than 31×31 (0.105 versus 0.152) but smaller route-time RMSE (0.077 versus 0.223); the 251×251 guide is best on both (0.344 and 0.052). This is consistent with Equation (9), which is a global upper bound on observable bias relative to the proposal variance.

The prespecified equally weighted two-route RMSE increases by 0.171 for 31×31 relative to 251×251 (90% outer-seed bootstrap interval $[0.167, 0.176]$; Table 11). The lower endpoint exceeds the practical threshold 0.02; the 7×7 contrast has a lower endpoint below that threshold. These intervals summarize variation among four outer seeds (Appendix D.8).

Self-normalized weighting brings all three guides' route-time estimates within the declared maximum absolute point-error tolerance of 0.05 (Appendix D.8).

Residual learning improves hitting-time accuracy. Table 1 shows the timing improvement alongside higher efficiency and lower global fidelity in both learned comparisons. The original-model

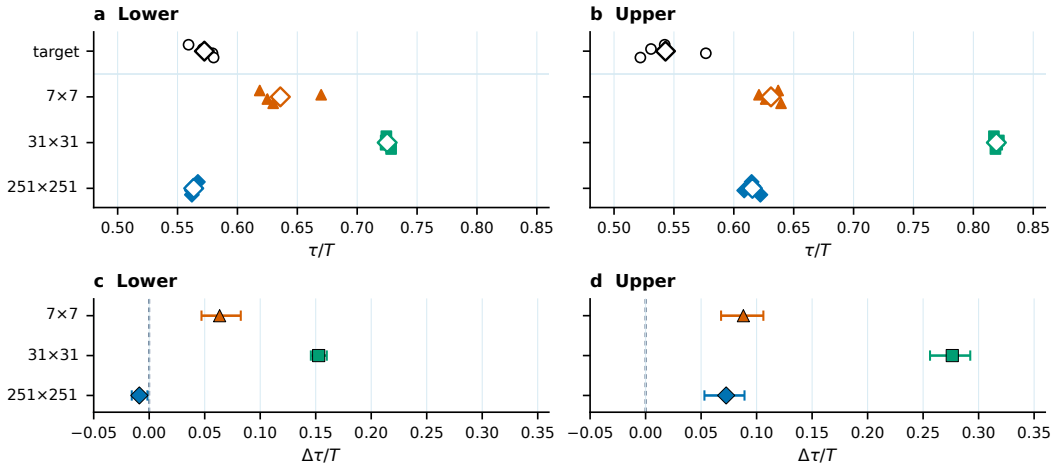

Figure 5: Independent validation of a conditional trajectory observable. A fresh crude simulation supplies unweighted draws from the target law conditioned on the rare event; fresh, disjoint balance-mixture evaluations supply each guide's displayed rare-event trajectory collection. Top panels show every outer-seed route-time mean; open diamonds mark their averages. Bottom panels form an estimation plot: points are route-conditioned mean errors and bars are the fixed 90% outer-seed bootstrap intervals. The 31×31 guide has a lower-route frequency closer to the target reference than the 251×251 guide (target: 0.602; 31×31: 0.562; 251×251: 0.664), yet has larger route-time errors in both routes. The target, guide resolutions, observable, path budgets, seeds, and practical threshold were frozen before execution.

RMSE-change interval lies entirely below the declared $+0.02$ timing-increase threshold. All 25 weighted arms pass the 0.05 point check (Appendix D.9). The joint measurements distinguish fidelity of the full conditional law from accuracy of the chosen observable: a proposal can improve this trajectory summary while moving further from the target conditional law in forward χ^2.

4 Discussion and Conclusion

The learned proposals exhibit three concurrent changes: higher estimation efficiency, lower global conditional fidelity, and more accurate route-conditioned hitting times. The same combination appears in the five original models and five new fits trained through the deployed control cap. Reporting all three quantities separates probability estimation, full-path matching, and the accuracy of the observable used to interpret trajectories. The mechanism decomposition locates the conditional discrepancy, and exact weighting supplies estimates of target expectations. The appropriate proposal therefore depends on whether the downstream goal is probability estimation, faithful trajectory generation, or accuracy of a specific trajectory observable.

The route-only ceiling sharpens the mechanism diagnosis. When within-route divergences differ, the fidelity-optimal allocation gives extra mass to the more distorted routes. Its estimated headroom is small in the guide sweep, suggesting attention to within-route laws.

The continuous-state experiments use two planar synthetic landscapes, supplied mechanism labels, and exact path ratios. Across both geometries, discovery and fidelity distinguish changes that estimation variance combines. The finite-state control quantifies the sensitivity of empirical fidelity to coverage of rare high-weight paths. Extending these measurements to implicit simulators calls for analysis of estimated likelihood ratios and mechanism definitions. Together, these results show that a rare-event proposal can be a better probability estimator without being a more faithful generator of event trajectories, while observable-specific accuracy remains a distinct third criterion.

AI USE STATEMENT

Generative AI tools assisted with polishing the paper only. The authors take responsibility for the paper's text, claims, and code.

REPRODUCIBILITY STATEMENT

The `accompanying supplementary archive` includes frozen records, source code, configurations, editable figure sources, and a common entry point for numerical, table, and figure reproduction. Appendices A–C give proofs, exact likelihoods, implementation details, and experimental settings; Appendices D–E give complete results, uncertainty procedures, and cost accounting. Development and evaluation seeds are disjoint. The primary diffusion and finite-state studies provide self-contained execution entries. The historical three-hole study provides deterministic statistical replay from its frozen records, configurations, and target code.

REFERENCES

Sergios Agapiou, Omiros Papaspiliopoulos, Daniel Sanz-Alonso, and Andrew M. Stuart. Importance sampling: Intrinsic dimension and computational cost. *Statistical Science*, 32(3):405–431, 2017. doi: 10.1214/17-STS611.

Solomon Asghar, Qing-Xiang Pei, Giorgio Volpe, and Ran Ni. Efficient rare event sampling with unsupervised normalizing flows. *Nature Machine Intelligence*, 6(11):1370–1381, 2024. doi: 10.1038/s42256-024-00918-3.

Peter G. Bolhuis, David Chandler, Christoph Dellago, and Phillip L. Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annual Review of Physical Chemistry*, 53:291–318, 2002. doi: 10.1146/annurev.physchem.53.082301.113146.

Frédéric Cérou and Arnaud Guyader. Adaptive multilevel splitting for rare event analysis. *Stochastic Analysis and Applications*, 25(2):417–443, 2007. doi: 10.1080/07362990601139628.

Kyle Cranmer, Juan Pavez, and Gilles Louppe. Approximating likelihood ratios with calibrated discriminative classifiers. *arXiv preprint arXiv:1506.02169*, 2015. doi: 10.48550/arXiv.1506.02169.

Harrison Delecki, Marc R. Schlichting, Mansur Arief, Anthony Corso, Marcell Vazquez-Chanlatte, and Mykel J. Kochenderfer. Diffusion-based failure sampling for evaluating safety-critical autonomous systems. In *2025 IEEE Engineering Reliable Autonomous Systems (ERAS)*, pp. 1–8, 2025. doi: 10.1109/ERAS63351.2025.11135908. URL https://arxiv.org/abs/2406.14761v2.

Christoph Dellago, Peter G. Bolhuis, Félix S. Csajka, and David Chandler. Transition path sampling and the calculation of rate constants. *The Journal of Chemical Physics*, 108(5):1964–1977, 1998. doi: 10.1063/1.475562.

Yuanqi Du, Michael Plainer, Rob Brekelmans, Chenru Duan, Frank Noé, Carla P. Gomes, Alán Aspuru-Guzik, and Kirill Neklyudov. Doob's lagrangian: A sample-efficient variational approach to transition path sampling. In *Advances in Neural Information Processing Systems*, volume 37, pp. 65791–65822, 2024. doi: 10.52202/079017-2103.

Paul Dupuis and Hui Wang. Importance sampling, large deviations, and differential games. *Stochastics and Stochastic Reports*, 76(6):481–508, 2004. doi: 10.1080/10451120410001733845.

Weinan E and Eric Vanden-Eijnden. Towards a theory of transition paths. *Journal of Statistical Physics*, 123(3):503–523, 2006. doi: 10.1007/s10955-005-9003-9.

Víctor Elvira, Luca Martino, and Christian P. Robert. Rethinking the effective sample size. *International Statistical Review*, 90(3):525–550, 2022. doi: 10.1111/insr.12500.

Carsten Hartmann and Lorenz Richter. Nonasymptotic bounds for suboptimal importance sampling. *SIAM/ASA Journal on Uncertainty Quantification*, 12(2):309–346, 2024. doi: 10.1137/21M1427760.

Hera Y. He and Art B. Owen. Optimal mixture weights in multiple importance sampling. *arXiv preprint arXiv:1411.3954*, 2014. doi: 10.48550/arXiv.1411.3954. URL https://arxiv.org/abs/1411.3954.

Tim Hesterberg. Weighted average importance sampling and defensive mixture distributions. *Technometrics*, 37(2):185–194, 1995. doi: 10.1080/00401706.1995.10484303.

Xinru Hua, Rasool Ahmad, Jose Blanchet, and Wei Cai. Accelerated sampling of rare events using a neural network bias potential. *arXiv preprint arXiv:2401.06936*, 2024. doi: 10.48550/arXiv.2401.06936.

Yuehaw Khoo, Jianfeng Lu, and Lexing Ying. Solving for high-dimensional committor functions using artificial neural networks. *Research in the Mathematical Sciences*, 6(1):1, 2019. doi: 10.1007/s40687-018-0160-2.

Philipp Metzner, Christof Schütte, and Eric Vanden-Eijnden. Illustration of transition path theory on a collection of simple examples. *The Journal of Chemical Physics*, 125(8):084110, 2006. doi: 10.1063/1.2335447.

Laurence I. Midgley, Vincent Stimper, Gregor N. C. Simm, Bernhard Schölkopf, and José Miguel Hernández-Lobato. Flow annealed importance sampling bootstrap. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=XCTVFJwS9LJ.

Klaus Müller and Leo D. Brown. Location of saddle points and minimum energy paths by a constrained simplex optimization procedure. *Theoretica Chimica Acta*, 53(1):75–93, 1979. doi: 10.1007/BF00547608.

Grant M. Rotskoff, Andrew R. Mitchell, and Eric Vanden-Eijnden. Active importance sampling for variational objectives dominated by rare events: Consequences for optimization and generalization. In *Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference*, volume 145 of *Proceedings of Machine Learning Research*, pp. 757–780, 2022. URL https://proceedings.mlr.press/v145/rotskoff22a.html.

John S. Sadowsky and James A. Bucklew. On large deviations theory and asymptotically efficient Monte Carlo estimation. *IEEE Transactions on Information Theory*, 36(3):579–588, 1990. doi: 10.1109/18.54903.

Kiyoung Seong, Seonghyun Park, Seonghwan Kim, Woo Youn Kim, and Sungsoo Ahn. Transition path sampling with improved off-policy training of diffusion path samplers. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/e897b4d0914f213896f3e5c25732b1ed-Abstract-Conference.html.

Aditya N. Singh, Avishek Das, and David T. Limmer. Variational path sampling of rare dynamical events. *Annual Review of Physical Chemistry*, 76:639–662, 2025. doi: 10.1146/annurev-physchem-083122-115001.

Jonathan Uesato, Ananya Kumar, Csaba Szepesvári, Tom Erez, Avraham Ruderman, Keith Anderson, Krishnamurthy Dvijotham, Nicolas Heess, and Pushmeet Kohli. Rigorous agent evaluation: An adversarial approach to uncover catastrophic failures. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=B1xhQhRcK7.

Eric Vanden-Eijnden and Jonathan Weare. Rare event simulation of small noise diffusions. *Communications on Pure and Applied Mathematics*, 65(12):1770–1803, 2012. doi: 10.1002/cpa.21428.

Eric Veach and Leonidas J. Guibas. Optimally combining sampling techniques for Monte Carlo rendering. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pp. 419–428, 1995. doi: 10.1145/218380.218498.

Aki Vehtari, Daniel Simpson, Andrew Gelman, Yuling Yao, and Jonah Gabry. Pareto smoothed importance sampling. *Journal of Machine Learning Research*, 25(72):1–58, 2024. URL https://jmlr.org/papers/v25/19-556.html.

Wei Zhang, Han Wang, Carsten Hartmann, Marcus Weber, and Christof Schütte. Applications of the cross-entropy method to importance sampling and optimal control of diffusions. *SIAM Journal on Scientific Computing*, 36(6):A2654–A2672, 2014. doi: 10.1137/14096493X.

A Theoretical Results and Proofs

The stopped chain in Section 2 induces all measures on a common path space. Because $\mathbb{Q}_0 = \mathbb{P}$ and $\pi_0 > 0$, $\mathbb{P} \ll \mathbb{Q}_{\text{mix}}$ holds independently of the guided controls. The likelihood in Equation (24) is exact because each proposal retains the target covariance and changes only its conditional mean.

Proposition 1 (Unbiasedness). *If the densities in Equation (25) are evaluated exactly, then* $\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\widehat{p}_m] = p_{\mathcal{E}}$ *for every* $m \geq 1$.

Proof. By linearity, it suffices to compute the expectation of one summand in Equation (25). A change of measure yields

$$\mathbb{E}_{\mathbb{Q}_{\text{mix}}}\left[\mathbf{1}_{\mathcal{E}} \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_{\text{mix}}}\right] = \int_{\mathcal{E}} \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}_{\text{mix}}} \, \mathrm{d}\mathbb{Q}_{\text{mix}} = \mathbb{P}(\mathcal{E}) = p_{\mathcal{E}}.$$

Consequently, averaging the m summands preserves this expectation, and hence $\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\widehat{p}_m] = p_{\mathcal{E}}$. □

Thus admissible controls preserve the probability estimand, while guide quality changes variance. The controlled experiments vary the proposal while holding this estimand fixed.

Proposition 2 (Total-discrepancy variance identity). *Let* $p_{\mathcal{E}} = \mathbb{P}(\mathcal{E}) > 0$, $\mathbb{P}_{\mathcal{E}} = \mathbb{P}(\cdot \mid \mathcal{E})$, *and* $\nu = \mathbf{1}_{\mathcal{E}} \, \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_{\text{mix}}$. *Then*

$$\frac{\mathrm{Var}_{\mathbb{Q}_{\text{mix}}}(\nu)}{p_{\mathcal{E}}^2} = \chi^2(\mathbb{P}_{\mathcal{E}} \| \mathbb{Q}_{\text{mix}}). \tag{10}$$

Consequently, $\mathrm{Var}_{\mathbb{Q}_{\text{mix}}}(\widehat{p}_m)/p_{\mathcal{E}}^2 = \chi^2(\mathbb{P}_{\mathcal{E}} \| \mathbb{Q}_{\text{mix}})/m$ *for* m *independent paths.*

Proof. We first observe that $\mathrm{d}\mathbb{P}_{\mathcal{E}}/\mathrm{d}\mathbb{Q}_{\text{mix}} = p_{\mathcal{E}}^{-1} \mathbf{1}_{\mathcal{E}} \, \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_{\text{mix}}$. Consequently,

$$\frac{\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\nu^2]}{p_{\mathcal{E}}^2} - 1 = \int \left(\frac{\mathrm{d}\mathbb{P}_{\mathcal{E}}}{\mathrm{d}\mathbb{Q}_{\text{mix}}}\right)^2 \mathrm{d}\mathbb{Q}_{\text{mix}} - 1 = \chi^2(\mathbb{P}_{\mathcal{E}} \| \mathbb{Q}_{\text{mix}}).$$

Taking $m = 1$ in Proposition 1 gives $\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\nu] = p_{\mathcal{E}}$, so the left-hand side equals $\mathrm{Var}_{\mathbb{Q}_{\text{mix}}}(\nu)/p_{\mathcal{E}}^2$. Finally, independence of the m paths yields the asserted factor $1/m$ for $\widehat{p}_m$. □

Proposition 3 (Consistency of the empirical conditional-discrepancy statistic). *Let* $(\boldsymbol{\omega}_i)_{i \geq 1}$ *be i.i.d. under* $\mathbb{Q}_{\text{mix}}$. *Assume* $p_{\mathcal{E}} > 0$, $H > 0$, *and* $\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\nu^2] < \infty$. *With probability one,* $\widehat{S}_N$ *is well-defined for all sufficiently large* N. *For those* N, *let* $\widehat{H}_N$ *and* $\widehat{S}_N$ *be the statistics in Equations (6)–(7). Then*

$$\widehat{S}_N \longrightarrow S = \frac{H}{1 + D_{\text{shape}}}, \qquad \frac{\widehat{H}_N}{\widehat{S}_N} - 1 \longrightarrow D_{\text{shape}} \quad \textit{almost surely}. \tag{11}$$

Proof. The strong law gives $\widehat{H}_N \to H$, $N^{-1} \sum_i \nu_i \to p_{\mathcal{E}}$, and $N^{-1} \sum_i \nu_i^2 \to \mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\nu^2]$. Therefore

$$\widehat{S}_N \longrightarrow \frac{p_{\mathcal{E}}^2}{\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[\nu^2]} = \frac{1}{1 + D_{\text{total}}} = \frac{H}{1 + D_{\text{shape}}},$$

where the last equality is the second form of Equation (4). Combining this limit with $\widehat{H}_N \to H$ proves the second statement. Consequently,

$$\widehat{F}_N = \frac{\widehat{S}_N}{\widehat{H}_N} \longrightarrow F \quad \text{almost surely}.$$

□

Proposition 4 (Mixture domination). *For any component* j *with* $\pi_j > 0$ *and* $\mathbb{P} \ll \mathbb{Q}_j$, $\mathbb{E}_{\mathbb{Q}_{\text{mix}}}[(\mathbf{1}_{\mathcal{E}} \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_{\text{mix}})^2] \leq \pi_j^{-1} \, \mathbb{E}_{\mathbb{Q}_j}[(\mathbf{1}_{\mathcal{E}} \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_j)^2]$, *provided the right-hand side is finite.*

Proof. Since $q_{\mathrm{mix}} \geq \pi_j q_j$ pointwise, we have $q_{\mathrm{mix}}^{-1} \leq \pi_j^{-1} q_j^{-1}$ wherever $q_j > 0$. Consequently,

$$\int_{\mathcal{E}} \frac{p^2}{q_{\mathrm{mix}}} \leq \pi_j^{-1} \int_{\mathcal{E}} \frac{p^2}{q_j},$$

which is the stated second-moment inequality. □

The Gaussian components here satisfy this support condition. Taking $j = 0$ bounds the mixture's second moment by π_0^{-1} times the unguided one, independently of the guided components. This direct density-domination argument is the path-space counterpart of the protection afforded by defensive mixture importance sampling (Hesterberg, 1995). For $0 < p_{\mathcal{E}} < 1$, subtracting $p_{\mathcal{E}}^2$ and dividing by the crude path variance gives

$$\frac{\mathrm{Var}_{\mathbb{Q}_{\mathrm{mix}}}(\nu)}{\mathrm{Var}_{\mathbb{P}}(\mathbf{1}_{\mathcal{E}})} \leq \frac{\pi_0^{-1} - p_{\mathcal{E}}}{1 - p_{\mathcal{E}}} = \frac{1}{\pi_0} + \frac{p_{\mathcal{E}}}{1 - p_{\mathcal{E}}} \left(\frac{1}{\pi_0} - 1 \right). \tag{12}$$

For a rare event, the mixture-to-crude variance ratio is therefore at most approximately $1/\pi_0$. Equivalently, the reciprocal variance-ratio gain is bounded below by approximately π_0.

A.1 Mechanism interpretation

The next identity is a conditional decomposition of the same second moment. It makes mechanism-level interpretation explicit.

Proposition 5 (Mechanism-stratified discrepancy)**.** *Let Z be a finite measurable partition of $\mathcal{E}$. For each z with $p_z > 0$, assume $r_z > 0$ and $\mathbb{P}_{\mathcal{E}}(\cdot \mid Z = z) \ll \mathbb{Q}_{\mathcal{E}}(\cdot \mid Z = z)$. Then Equation (8) holds, allowing infinite values. For any $g \in L^2(\mathbb{Q}_{\mathcal{E}})$ and finite D_{shape}, Equation (9) holds.*

Proof. On $\{Z = z\}$ with $p_z > 0$, the conditional density ratio equals p_z/r_z times the within-mechanism density ratio. Splitting its squared integral over the partition gives

$$1 + D_{\mathrm{shape}} = \sum_{z : p_z > 0} \frac{p_z^2}{r_z} (1 + D_z). \tag{13}$$

Subtracting one and using $\chi^2(\mathbf{p} \| \mathbf{r}) = \sum_{z:p_z>0} p_z^2/r_z - 1$ proves the decomposition, with zero-mass target terms omitted. For the second assertion, write $R = \mathrm{d}\mathbb{P}_{\mathcal{E}}/\mathrm{d}\mathbb{Q}_{\mathcal{E}}$. Since $\mathbb{E}_{\mathbb{Q}_{\mathcal{E}}}(R - 1) = 0$, the difference of expectations is $\mathbb{E}_{\mathbb{Q}_{\mathcal{E}}}[(R - 1)(g - \mathbb{E}_{\mathbb{Q}_{\mathcal{E}}} g)]$. Cauchy–Schwarz gives Equation (9). Applying the same argument to $\mathbb{P}_{\mathcal{E}}(\cdot \mid Z = z)$ and $\mathbb{Q}_{\mathcal{E}}(\cdot \mid Z = z)$ gives the within-mechanism bound with D_{shape} replaced by D_z. □

The decomposition also predicts what a route-only intervention can achieve. Hold the conditional proposal law within each mechanism fixed at $\mathbb{Q}_z$ and vary only its conditional route masses $\mathbf{a} = (a_z)_z$, so that $\mathbb{Q}_{\mathcal{E}}^{(\mathbf{a})} = \sum_z a_z \mathbb{Q}_z$.

Proposition 6 (Route-reallocation ceiling)**.** *Let $(\mathcal{E}_z)_z$ be a finite measurable partition of $\mathcal{E}$ and let $p_z = \mathbb{P}_{\mathcal{E}}(\mathcal{E}_z)$. Let $\mathbb{Q}_z$ be a probability law supported on $\mathcal{E}_z$. For each z with $p_z > 0$, let $\mathbb{P}_z = \mathbb{P}_{\mathcal{E}}(\cdot \mid \mathcal{E}_z)$, assume $\mathbb{P}_z \ll \mathbb{Q}_z$, and let $D_z = \chi^2(\mathbb{P}_z \| \mathbb{Q}_z) < \infty$. Let $\mathbf{a}$ be a probability vector with $a_z > 0$ whenever $p_z > 0$. The fidelity of $\mathbb{Q}_{\mathcal{E}}^{(\mathbf{a})} = \sum_z a_z \mathbb{Q}_z$ is*

$$F(\mathbf{a}) = \left\{ \sum_{z : p_z > 0} \frac{p_z^2 (1 + D_z)}{a_z} \right\}^{-1}. \tag{14}$$

Its maximum and an optimizing allocation are

$$F_{\mathrm{route}}^{\max} = \left\{ \sum_{z : p_z > 0} p_z \sqrt{1 + D_z} \right\}^{-2}, \qquad a_z^* = \frac{p_z \sqrt{1 + D_z}}{\sum_{j : p_j > 0} p_j \sqrt{1 + D_j}} \quad (p_z > 0), \tag{15}$$

with $a_z^* = 0$ for $p_z = 0$. For every admissible $\mathbf{a}$,

$$\frac{F_{\mathrm{route}}^{\max}}{F(\mathbf{a})} - 1 = \chi^2(\mathbf{a}^* \| \mathbf{a}). \tag{16}$$

Thus target-share matching, $\mathbf{a} = \mathbf{p}$, is optimal precisely when the active D_z are equal; when they differ, the fidelity-optimal proposal allocates extra mass to routes with larger within-route divergence. If every $D_z = 0$, the positive-control prediction is $\mathbf{a}^* = \mathbf{p}$ and $F^{\max}_{\text{route}} = 1$. For empirical route count M_z, event-weight sum A_z, and squared-weight sum B_z, substitution yields the plug-in ceiling

$$\widehat{F}^{\max}_{\text{route}} = \frac{(\sum_z A_z)^2}{(\sum_z \sqrt{M_z B_z})^2}. \tag{17}$$

Equation (16) identifies relative improvement from route reallocation exactly with allocation discrepancy from $\mathbf{a}^*$. This is a disjoint-support specialization of second-moment mixture optimization (He & Owen, 2014). The counterfactual concerns conditional laws: a dynamic intervention may also change discovery, cost, or within-route laws. Its empirical ceiling is a point estimate based on the observed route moments.

Proof of Proposition 6. For the reallocated conditional law, the within-route density ratios and divergences remain unchanged, while r_z in the partition identity becomes a_z. Therefore

$$1 + D_{\text{shape}}(\mathbf{a}) = \sum_{z:p_z>0} \frac{p_z^2}{a_z}(1 + D_z),$$

which proves Equation (14). Cauchy–Schwarz and $\sum_z a_z = 1$ give

$$\left(\sum_{z:p_z>0} p_z\sqrt{1+D_z}\right)^2 \le \left(\sum_{z:p_z>0} \frac{p_z^2(1+D_z)}{a_z}\right)\left(\sum_{z:p_z>0} a_z\right) \le \sum_{z:p_z>0} \frac{p_z^2(1+D_z)}{a_z}.$$

Equality holds when all proposal mass lies on target-positive routes and a_z there is proportional to $p_z\sqrt{1+D_z}$, proving Equation (15). Let $C = \sum_{z:p_z>0} p_z\sqrt{1+D_z}$. Since $p_z^2(1+D_z) = C^2(a_z^*)^2$,

$$\frac{F^{\max}_{\text{route}}}{F(\mathbf{a})} = \sum_{z:a_z^*>0} \frac{(a_z^*)^2}{a_z} = 1 + \chi^2(\mathbf{a}^* \| \mathbf{a}),$$

proving Equation (16). Finiteness of each active D_z ensures that C and every stated allocation are well-defined. Substituting $\widehat{p}_z = A_z / \sum_j A_j$ and $1 + \widehat{D}_z = M_z B_z / A_z^2$ gives Equation (17). □

The allocation result concerns independent draws from a random mixture. With a fixed number of draws per route, the estimator variance instead depends on within-route variances and gives a separate stratified-sampling allocation problem.

B Guided Proposals and Exact Path Likelihoods

The constructions below instantiate proposals for the controlled evaluation. RareFlow's audit in Section 2 applies once their event labels and exact path ratios are available, independently of how a guide or residual was obtained. Control-based changes of measure are well established in rare-event importance sampling (Dupuis & Wang, 2004; Vanden-Eijnden & Weare, 2012).

B.1 Guided proposals for controlled comparisons

Related neural approaches approximate committor functions using variational formulations (Khoo et al., 2019) and active importance sampling (Rotskoff et al., 2022). For r remaining steps, the finite-horizon committor is $h_r(\mathbf{x}) = \mathbb{P}_{\mathbf{x}}(\tau_{\mathcal{B}} \le r)$. With $p(\mathbf{x}' \mid \mathbf{x})$ denoting the transition density in Equation (1), it obeys the backward recursion

$$h_0(\mathbf{x}) = \mathbf{1}_{\mathcal{B}}(\mathbf{x}), \qquad h_r(\mathbf{x}) = \mathbf{1}_{\mathcal{B}}(\mathbf{x}) + (1 - \mathbf{1}_{\mathcal{B}}(\mathbf{x})) \int_{\mathcal{X}} h_{r-1}(\mathbf{x}')p(\mathbf{x}' \mid \mathbf{x})\,\mathrm{d}\mathbf{x}'. \tag{18}$$

We approximate this recursion on a bounded spatial grid by tensor Gauss–Hermite quadrature and bilinear interpolation. Quadrature points outside the grid use committor value zero (constant-zero extension in the backward recursion). During proposal simulation, only the base-guide control is zero

outside the grid; route feedback or a residual can remain active. The chain itself evolves on the full state space. If $\widetilde{h}_r$ is the resulting table, the stored base control is

$$\mathbf{u}_r^{\mathrm{com}}(\mathbf{x}) = 2\varepsilon\, \nabla_{\mathbf{x}} \mathbf{S}_\eta \Big[\log\big(\max\{\widetilde{h}_r(\mathbf{x}), h_{\min}\}\big)\Big], \tag{19}$$

where $\mathbf{S}_\eta$ is the fixed grid-smoothing operator. Equation (19) uses the continuous-time Doob control as a guide. Exactness for the declared discrete chain comes from the stopped-path likelihood. After smoothing and differentiation, the builder sets the stored control to zero at target grid nodes. Proposal evaluation interpolates this stored table.

The proposal family contains an unguided component $j = 0$ and $j_{\max}$ guided components. For the multichannel planar experiments, let $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^2$ be the coordinate basis, y_j the transverse location assigned to channel j, x_{rel} the release location, $w > 0$ the release width, and $\gamma \geq 0$ the feedback gain. The route-specific perturbation is

$$\mathbf{g}_j(\mathbf{x}) = \gamma\big(y_j - \mathbf{e}_2^\top \mathbf{x}\big) \left[1 + \exp\left(\mathrm{clip}_{[-40,40]}\left(\frac{\mathbf{e}_1^\top \mathbf{x} - x_{\mathrm{rel}}}{w}\right)\right)\right]^{-1} \mathbf{e}_2. \tag{20}$$

Let $s_k = (n-k)/n$ be normalized remaining time, $\alpha > 0$ a scale, $f_\phi : \mathcal{X} \times [0,1] \to \mathbb{R}^d$ an optional learned residual, and $\mathbf{C}_c(\mathbf{v}) = \mathbf{v} \min\{1, c/\|\mathbf{v}\|_2\}$ for $\mathbf{v} \neq \mathbf{0}$ (and $\mathbf{C}_c(\mathbf{0}) = \mathbf{0}$) the norm cap at $c > 0$. With $\mathbf{u}_{0,k} \equiv \mathbf{0}$ for the unguided component,

$$\mathbf{u}_{j,k}(\mathbf{x}) = \mathbf{C}_c\big(\alpha \mathbf{u}_{n-k}^{\mathrm{com}}(\mathbf{x}) + \mathbf{g}_j(\mathbf{x}) + f_\phi(\mathbf{x}, s_k)\big), \qquad j = 1, \ldots, j_{\max}. \tag{21}$$

The feedback ablation sets $f_\phi \equiv \mathbf{0}$ and compares $\gamma > 0$ with $\gamma = 0$. The learned-residual experiment sets $\gamma = 0$ and uses a single guided component. Under component j, the chain evolves as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \big[\mathbf{b}(\mathbf{x}_k) + \mathbf{u}_{j,k}(\mathbf{x}_k)\big]\Delta t + \sigma\sqrt{\Delta t}\, \boldsymbol{\xi}_k. \tag{22}$$

Thus every proposal changes only the drift and retains the target covariance.

B.2 Exact stopped-path weights and defensive mixture

Let $\mathbb{Q}_j$ and $q_j(\boldsymbol{\omega})$ denote the stopped-path law and density induced by Equation (22); $\mathbb{Q}_0 = \mathbb{P}$ and $q_0 = p$. For the target innovation

$$\Delta \mathbf{w}_k = \mathbf{x}_{k+1} - \mathbf{x}_k - \mathbf{b}(\mathbf{x}_k)\Delta t, \tag{23}$$

the Gaussian transition densities give the exact component log-ratio

$$\ell_j(\boldsymbol{\omega}) := \log \frac{p(\boldsymbol{\omega})}{q_j(\boldsymbol{\omega})} = \sum_{k=0}^{\bar{\tau}-1} \left[-\frac{\mathbf{u}_{j,k}(\mathbf{x}_k)^\top \Delta \mathbf{w}_k}{\sigma^2} + \frac{\|\mathbf{u}_{j,k}(\mathbf{x}_k)\|_2^2 \Delta t}{2\sigma^2}\right]. \tag{24}$$

We draw paths from the defensive mixture $\mathbb{Q}_{\mathrm{mix}} = \sum_{j=0}^{j_{\max}} \pi_j \mathbb{Q}_j$, where $\pi_j > 0$, $\sum_{j=0}^{j_{\max}} \pi_j = 1$, and π_0 is the unguided mass. Every sampled path is reweighted by the full mixture density, which accounts for each component capable of generating it (Veach & Guibas, 1995). Using $\ell_0 \equiv 0$, the exact path weight and probability estimator are

$$\begin{aligned} \log w(\boldsymbol{\omega}) &= \log \frac{p(\boldsymbol{\omega})}{q_{\mathrm{mix}}(\boldsymbol{\omega})} = -\log\left[\sum_{j=0}^{j_{\max}} \pi_j \exp\{-\ell_j(\boldsymbol{\omega})\}\right], \\ \widehat{p}_m &= \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{\mathcal{E}}(\boldsymbol{\omega}_i) w(\boldsymbol{\omega}_i), \qquad \boldsymbol{\omega}_i \overset{\text{i.i.d.}}{\sim} \mathbb{Q}_{\mathrm{mix}}. \end{aligned} \tag{25}$$

Because p/q_{mix} is evaluated exactly, $\mathbb{E}_{\mathbb{Q}_{\mathrm{mix}}}[\widehat{p}_m] = p_{\mathcal{E}}$ for every admissible collection of controls. Guide quality changes variance while unbiasedness remains fixed. Writing $\nu = \mathbf{1}_{\mathcal{E}}\, \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_{\mathrm{mix}}$ for the single-path contribution, the inequality $q_{\mathrm{mix}} \geq \pi_0 p$ yields the defensive second-moment bound

$$\mathbb{E}_{\mathbb{Q}_{\mathrm{mix}}}\left[\left(\mathbf{1}_{\mathcal{E}} \frac{p}{q_{\mathrm{mix}}}\right)^2\right] \leq \frac{p_{\mathcal{E}}}{\pi_0}, \qquad \mathrm{Var}_{\mathbb{Q}_{\mathrm{mix}}}(\nu) \leq \frac{p_{\mathcal{E}}}{\pi_0} - p_{\mathcal{E}}^2. \tag{26}$$

Appendix A gives the general component-wise bound and proofs.

B.3 Learned residual intervention

The residual is learned from the frozen, capped coarse-guide component $\mathbb{Q}_{\mathrm{dev}}$, which supplies every development path. For development path i, define its exact event contribution and realized effective control by

$$v_i = \mathbf{1}_{\mathcal{E}}(\boldsymbol{\omega}_i)\frac{p(\boldsymbol{\omega}_i)}{q_{\mathrm{dev}}(\boldsymbol{\omega}_i)}, \qquad \mathbf{a}_{ik} = \frac{\Delta \mathbf{w}_{ik}}{\Delta t}. \tag{27}$$

Let $\overline{\mathbf{u}}_{ik} = \mathbf{C}_c(\alpha \mathbf{u}^{\mathrm{com}}_{n-k}(\mathbf{x}_{ik}))$ be the frozen coarse-guide control, and let $\mathcal{S}$ be every fifth entry of the stored transition stream within each development batch, ordered by time and then path. Each retained transition inherits its path's event weight v_i; the denominator below sums these retained transition weights directly. The implemented full-batch objective is

$$\ell_{\mathrm{res}}(\phi) = \frac{\sum_{(i,k)\in\mathcal{S}} v_i \left\|\mathbf{a}_{ik} - \overline{\mathbf{u}}_{ik} - f_\phi(\mathbf{x}_{ik}, s_k)\right\|_2^2}{\sum_{(i,k)\in\mathcal{S}} v_i} + \frac{\lambda}{d|\mathcal{S}|}\sum_{(i,k)\in\mathcal{S}} \left\|f_\phi(\mathbf{x}_{ik}, s_k)\right\|_2^2 . \tag{28}$$

Up to a positive scale factor, the unregularized population objective is the Gaussian cross-entropy, or forward-KL projection, of $\mathbb{P}_{\mathcal{E}}$ onto an auxiliary *uncapped* family with control $\overline{\mathbf{u}} + f_\phi$. This interpretation connects the residual fit to cross-entropy methods for importance sampling and controlled diffusions (Zhang et al., 2014). Equation (28) is its self-normalized, strided empirical approximation. The baseline $\overline{\mathbf{u}}$ is capped before regression; the evaluated proposal instead applies $\mathbf{C}_c$ after adding the residual to $\alpha \mathbf{u}^{\mathrm{com}}$. Forward-KL projection and forward-χ^2 fidelity order proposals differently in general. The held-out discovery and fidelity metrics therefore remain the evaluation criteria for the fitted residual. In flow-based sampling, Flow AIS Bootstrap targets importance-weight variance through an $\alpha = 2$ divergence objective (Midgley et al., 2023).

For the planar experiment, f_ϕ is a shallow tanh MLP whose output layer is initialized to zero. Optimization therefore starts at the coarse guide. The architecture, full-batch Adam schedule, and regularization are frozen before training. Table 3 records them in full. Each fitted residual is frozen before its evaluation seeds are opened.

C Experimental Setup and Reproducibility

Table 3 consolidates the experimental configuration; coordinates are rounded for display, while the supplied machine-readable configurations retain full precision and seed arrays. The Gaussian obstacle creates the route ambiguity targeted by the channel-specific architecture. Channel assignment uses an interpolated crossing, avoiding the discretization noise of a nearest-step label. The stored committor control is a continuous-time-motivated approximation. The exact discrete likelihood corrects its proposal. The neutral feedback arm changes only the route term, retaining a controlled component-level attribution.

Primary target dynamics. For the primary task, the unclipped potential describing the landscape is

$$V(x,y) = s\sum_{m=1}^{4} A_m e^{\psi_m(x,y)} + A_{\mathrm{obs}} \exp\left[-\frac{1}{2}\left\{\left(\frac{x-x_{\mathrm{obs}}}{\sigma_x}\right)^2 + \left(\frac{y-y_{\mathrm{obs}}}{\sigma_y}\right)^2\right\}\right],$$
$$\psi_m(x,y) = a_m(x-X_m)^2 + b_m(x-X_m)(y-Y_m) + c_m(y-Y_m)^2,$$

with coefficient tuples

$$(A_m)_m = (-200, -100, -170, 15), \quad (a_m)_m = (-1, -1, -6.5, 0.7),$$
$$(b_m)_m = (0, 0, 11, 0.6), \quad (c_m)_m = (-10, -10, -6.5, 0.7),$$
$$(X_m)_m = (1, 0, -0.5, -1), \quad (Y_m)_m = (0, 0.5, 1.5, 1).$$

Here $s = 0.02$, $A_{\mathrm{obs}} = 3.0$, $(x_{\mathrm{obs}}, y_{\mathrm{obs}}) = (0.35, 0.15)$, and the Gaussian standard deviations are $(\sigma_x, \sigma_y) = (0.12, 0.14)$. Writing the Gaussian obstacle term as V_{obs}, the actual numerical drift in

Equation (1) is

$$\mathbf{b}(\mathbf{x}) = -s \sum_{m=1}^{4} A_m e^{\min\{\psi_m(\mathbf{x}),40\}} \nabla\psi_m(\mathbf{x}) - \nabla V_{\mathrm{obs}}(\mathbf{x}).$$

This definition includes the implementation's exponent-overflow safeguard and retains the uncapped polynomial derivative factors. Clipping acts on the exponent during drift evaluation. The target and proposal transitions and their likelihood ratios use this same numerical drift. The frozen primary-task evaluation records report zero safeguard activations.

Table 3: Fixed task, proposal, learning, and evaluation configuration.

category	field	fixed value
landscape	potential and obstacle	Müller–Brown scale 0.02; obstacle amplitude 3.0, center $(0.35, 0.15)$, standard deviations $(0.12, 0.14)$
landscape	source and target	source $(-0.050, 0.467)$; target $(0.623, 0.028)$; target radius 0.10
landscape	channel rule	last interpolated crossing of $x = 0.35$, split at $y = 0.15$
dynamics	numerical chain	$\Delta t = 0.005$; horizon 3.0; noise level reported per experiment
guide	domain and default grid	$[-0.8, 1.2] \times [-0.65, 1.3]$; 251×251
guide	quadrature and stabilization	Gauss–Hermite order 3; $h_{\min} = 10^{-8}$; Gaussian smoothing s.d. 0.5 cells
guide	storage and interpolation	every fifth remaining-time slice plus remaining steps 0–20; linear in time, bilinear in space; zero base-guide control outside grid
matched feedback	control parameters	$\alpha = 0.90$, $\gamma = 1.50$, $\mathbf{y} = (y_1, y_2) = (-0.10, 0.50)$, $x_{\mathrm{rel}} = 0.48$, $w = 0.08$, $c = 4.0$
matched feedback	mixture and neutral arm	$\boldsymbol{\pi} = (\pi_0, \pi_1, \pi_2) = (0.20, 0.40, 0.40)$; neutral arm sets $\gamma = 0$ only; arms use distinct frozen seeds
learned residual	guide and mixture	11×11 guide; $\gamma = 0$; one guided component; $\boldsymbol{\pi} = (\pi_0, \pi_1) = (0.20, 0.80)$
learned residual	architecture	two hidden tanh layers; 64 units per layer; zero-initialized output
learned residual	optimizer	full-batch Adam, 400 epochs; $\eta = 10^{-3}$; PyTorch defaults $\boldsymbol{\beta} = (0.9, 0.999)$ and $\delta = 10^{-8}$
learned residual	regularization and development	$\lambda = 10^{-3}$; five independent models; 100,000 development paths per model; disjoint seed families
evaluation	batching and intervals	feedback: 20 batches of 5,000 paths per arm and 10,000-resample batch bootstrap; learned residual: the same path budget per arm and model, with a 1,999-resample hierarchical bootstrap
external observable	independent reference and evaluation	crude target: four fresh seeds of 200,000 paths; each frozen 7×7, 31×31, and 251×251 guide: four fresh seeds of 25,000 paths; route-conditioned normalized first-hitting time with a 2,000-replicate outer-seed bootstrap

Evaluation. The feedback, learned-residual, and timing studies retain their declared batch or outer-seed summaries. The guide sweep retains route-specific sufficient statistics pooled across batches. Development and evaluation seeds are disjoint where learning occurs, and all seeds are recorded in the supplied configurations or results.

Independent conditional-observable audit. The target, three inherited guide resolutions, route classifier, normalized first-hitting-time observable, path budgets, seeds, bootstrap rule, and practical contrast threshold were prespecified. Four fresh crude target simulations supply accepted paths from the target law conditioned on the rare event. Four disjoint fresh balance-mixture evaluations are performed for each guide. The inherited guide-sweep paths are disjoint by design. The record at `experiments/04_conditional_observable/results/formal_`

`results.json` contains source and guide hashes, support checks, seed information, and observable summaries. The figure displays unweighted rare-event trajectory collections; self-normalized weighting is retained as a separate calibration diagnostic for target expectations.

Code and data. Raw result files and regeneration scripts for every table and figure are included in the `accompanying supplementary archive`.

D Additional Empirical Results and Validation

D.1 Complete Route-Feedback Comparison

The feedback and neutral arms use independent simulation seeds. Independent-arm batch bootstraps quantify the path-variance ratio; all three source noise levels are retained.

Table 4: Route feedback changes conditional fidelity at nearly fixed event discovery. $\widehat{H}$ and $\widehat{F} = \widehat{S}/\widehat{H}$ are the two audit coordinates. The path-variance ratio $\rho_{\mathrm{mix}} = V_{\mathrm{fb}}/V_0$ uses independent simulation arms and batch-bootstrap intervals. Values above one favor the neutral arm.

ε	$\widehat{H}_{\mathrm{fb}}$	$\widehat{H}_0$	$\widehat{F}_{\mathrm{fb}}$	$\widehat{F}_0$	ρ_{mix}	95% CI
0.12	0.769	0.770	0.334	0.636	2.77	$[2.39, 3.23]$
0.08	0.769	0.774	0.168	0.536	4.67	$[4.21, 5.18]$
0.06	0.765	0.771	0.091	0.444	6.92	$[5.65, 8.30]$

At population level, both arms estimate the same $p_{\mathcal{E}}$, so

$$\rho_{\mathrm{mix}} = \frac{S_{\mathrm{fb}}^{-1} - 1}{S_0^{-1} - 1}. \tag{29}$$

For the recorded empirical variances, which use denominator N, the corresponding identity also contains $(\widehat{p}_{\mathrm{fb}}/\widehat{p}_0)^2$. The bootstrap therefore resamples batch summaries of path contributions, retaining variation in both probability and second-moment estimates.

D.2 Learned-Residual Replication

Across all five independent runs, residual learning increases discovery and overall efficiency while decreasing conditional fidelity.

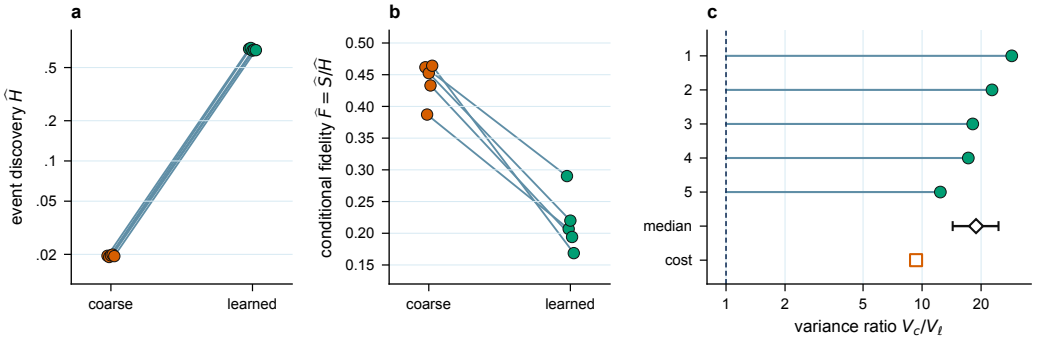

Figure 6: The learned residual reallocates overall efficiency toward event discovery. In **(a)** and **(b)**, orange marks the coarse guide, green marks the learned proposal, and each line joins run-index-matched independent arms. Event discovery rises in every run while conditional fidelity $\widehat{F}$ shifts downward; panel (a) reports $\widehat{H}$. **(c)** Green circles give per-run coarse-to-learned variance ratios. The diamond gives the hierarchical-bootstrap median and 95% interval; the square applies equal-total-path-budget accounting for the development sample.

Design and budget. The committor is deliberately degraded to an 11×11 resolution with substantial headroom relative to the fine-grid diagnostic. The 11×11 arm in the separate resolution sweep uses a route-feedback configuration. The learned-residual study uses a single-guide, feedback-removed defensive mixture. Here "coarse" denotes resolution relative to the fine guide; the two coarse arms use different proposals. The zero-initialized residual begins exactly at the learned study's coarse guide.

The study fits five models. Each receives an independent $100{,}000$-path development sample, Torch initialization, $100{,}000$-path coarse-mixture evaluation group, and $100{,}000$-path learned-mixture evaluation group. These seed families are mutually disjoint. Architecture, optimizer, mixture, control cap, and budgets are fixed across runs. The quality checks cover convergence, finite exact weights, potential clipping, agreement with the independent probability reference, and minimum event support. Model checkpoints and the source committor are hash verified.

Table 5: Frozen values for the five learned-residual runs. Subscripts c and ℓ denote the coarse and learned defensive mixtures; hats denote empirical estimates. $R_s = V_{c,s}/V_{\ell,s}$ is the held-out per-path variance ratio. The main text visualizes run-index-matched independent-arm changes and reports the replication interval.

run	$\widehat{H}_c$	$\widehat{H}_\ell$	$\widehat{F}_c$	$\widehat{F}_\ell$	$\widehat{S}_c$	$\widehat{S}_\ell$	R_s
1	0.0195	0.6900	0.462	0.290	0.0090	0.2002	28.67
2	0.0191	0.7005	0.387	0.207	0.0074	0.1449	22.74
3	0.0197	0.6703	0.453	0.220	0.0089	0.1474	18.13
4	0.0199	0.6794	0.433	0.194	0.0086	0.1319	17.21
5	0.0194	0.6748	0.464	0.169	0.0090	0.1138	12.39
pooled variance ratio							18.64
hierarchical median (95% interval)							18.85[14.33, 24.54]
equal-total-path-budget pooled ratio							9.32

Uncertainty and result. Each model s contributes a per-model ratio $R_s = V_{c,s}/V_{\ell,s}$. Each hierarchical-bootstrap replicate draws five model indices with replacement. For each selected occurrence, it independently draws 20 evaluation batches with replacement from each of that model's coarse and learned arms. Within each arm, it pools all selected batches using their path counts n_b, means m_b, and empirical path variances v_b (with denominator n_b):

$$N_* = \sum_b n_b, \qquad A_* = \sum_b n_b m_b, \qquad B_* = \sum_b n_b(v_b + m_b^2),$$
$$V_* = \max\{0, B_*/N_* - (A_*/N_*)^2\}.$$

The replicate statistic is the pooled variance ratio $V_{c,*}/V_{\ell,*}$. The reported 18.85 is the median of these $1{,}999$ pooled ratios (bootstrap seed 810000). The five R_s values describe individual fits. The 2.5th and 97.5th percentiles, using linear quantile interpolation, give $[14.33, 24.54]$. This interval quantifies variation across the recorded training and evaluation runs. Every model improves the held-out path variance relative to its coarse guide. The secondary path-budget calculation assigns twice as many evaluation paths to the coarse arm because the learned arm also consumes an equally large development sample; it divides the pooled ratio by two. This counts paths, while model optimization and proposal-evaluation speeds enter the separate accounting in Appendix E.

Post hoc uncertainty of the fidelity change. To assess whether the observed fidelity decrease is stable to recorded training and evaluation variation, we additionally analyze $\Delta_F = \frac{1}{5}\sum_{s=1}^{5}(\widehat{F}_{\ell,s} - \widehat{F}_{c,s})$. Each of 1,999 hierarchical-bootstrap replicates (seed 810000) draws five fit indices with replacement, then independently resamples the 20 batches in each arm for each selected occurrence. For each occurrence, batch path counts, hit counts, weight sums, and squared sums are pooled before computing $\widehat{F}$. We then average the five learned-minus-coarse differences; linear 2.5th and 97.5th percentiles give $[-0.268, -0.179]$ around the observed change -0.224. This statistic averages per-fit contrasts and assigns each fitted proposal equal weight. It is a post hoc analysis of the retained record, separate from the original variance-ratio protocol. All three coordinate contrasts are retained in the

derived record. With five fits, this bootstrap describes variation among the observed fits and batches. The hidden-submode control separately examines sensitivity to rare high-weight paths.

D.3 Guide-Resolution Sweep and Shared Validation Checks

Across the nine guide resolutions, discovery, fidelity, and efficiency improve overall, with local reversals in fidelity.

The RMS probe reuses the same 500 points sampled uniformly from $[-0.8, 1.2] \times [-0.65, 1.3]$ at remaining steps $600, 500, 400, 300, 200, 100, 50, 10$, with equal weighting over the resulting $4{,}000$ point–time comparisons.

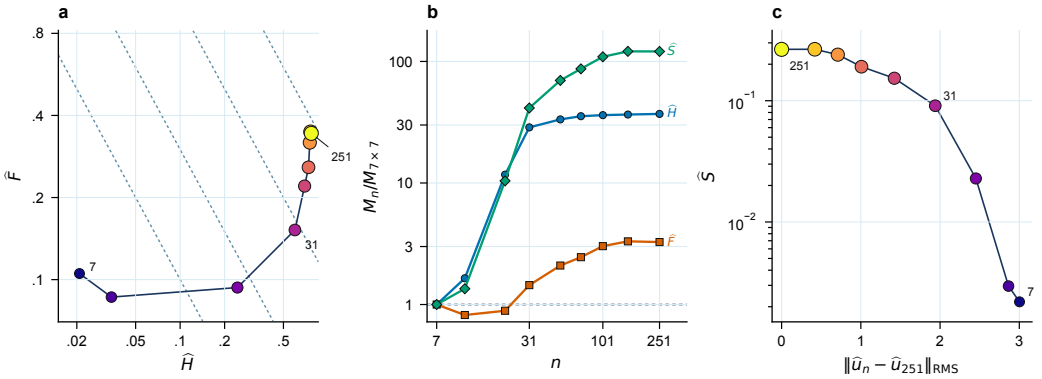

Figure 7: Guide-resolution sweep, where n is grid width and $M \in \{\widehat{H}, \widehat{F}, \widehat{S}\}$ in panel (b). Marker color, size, and connecting order progress from coarse to fine; labels identify the initial, intermediate, and final grids. **(a)** The connected points trace the effect of guide resolution in the discovery–fidelity plane; dashed curves mark constant overall efficiency. **(b)** Each coordinate is divided by its 7×7 value, revealing the coupled gains in event discovery, conditional fidelity, and overall efficiency. **(c)** The untransformed efficiency coordinate increases as RMS control error falls. RMS is the root mean squared Euclidean control-vector difference from the stored 251×251 guide over 500 uniformly sampled spatial points and eight remaining-time slices, with equal weight for each point–time pair. Across all nine resolutions, the descriptive Spearman rank correlation between error and $\widehat{S}$ is $\rho_{\mathrm{Spearman}} = -0.98$.

Configurations specify the tasks, proposals, seed families, path budgets, and analysis rules. Feedback and learned-residual records retain per-batch estimates, variances, hit counts, event ESS, likelihood diagnostics, and wall times. The guide sweep retains pooled route moments and validation checks; these pooled moments support a point estimate of the route-only ceiling.

Likelihood and reference checks. The learned-residual record contains a direct Gaussian likelihood fixture, a core-implementation regression, and the recorded validity checks for every training run. These checks cover finite likelihood ratios, optimization convergence, minimum event support, and compatibility with the independently seeded crude probability reference. The guide sweep stores its reused crude reference and checks finite weights, minimum event counts, and potential-exponent clipping at every resolution.

Uncertainty and discretization. The matched feedback intervals resample independent evaluation batches, and the learned-residual intervals also resample training replicates. The target is a fixed Euler chain; the exact path weights establish unbiasedness for that chain. Accuracy relative to the continuous diffusion is governed separately by discretization error.

Reference precision. Variance-reduction figures normalized against low-hit-count crude references carry corresponding uncertainty. The reference-free ratio ρ_{mix} used in the matched feedback experiment removes that crude-reference uncertainty.

D.4 Route-Only Counterfactual on the Frozen Guide Sweep

If route allocation explains most of a population fidelity loss, optimizing the route masses in Equation (15) must recover most of that loss. We examine the corresponding plug-in prediction on the frozen sweep. This post hoc diagnostic uses the pooled route moments retained by the sweep and reports the resulting point estimates.

Table 6: Post hoc route-only counterfactual on the frozen guide sweep. Target-share matching $\widehat{F}(\widehat{\mathbf{p}})$ is compared with the optimum; $\widehat{F}^{\max}_{\text{route}}$ and $\widehat{a}^{*}_{\text{lower}}$ use Equation (17). Maximum gain is relative to observed $\widehat{F}$. The full nine-resolution calculation is included in the derived record.

grid	$\widehat{F}$	$\widehat{F}(\widehat{\mathbf{p}})$	$\widehat{F}^{\max}_{\text{route}}$	$\widehat{a}^{*}_{\text{lower}}$	max. gain (%)
7	0.105257	0.097119	0.108023	0.779	2.63
21	0.093527	0.094521	0.096062	0.668	2.71
31	0.152032	0.151866	0.153747	0.612	1.13
251	0.343780	0.333346	0.344169	0.683	0.11

The point estimates suggest little headroom: the largest relative gain is 2.71% across all nine resolutions and 0.11% at the finest grid (Table 6). Matching target shares can even reduce predicted fidelity at the coarsest grid because it allocates too little mass to the more distorted route. These calculations prioritize within-route changes for future evaluation. Their precision depends on how well the sample represents the within-route laws.

A separately frozen finite-state control checks the allocation prediction with known target laws. If allocation is the only mismatch, target-share matching should restore $F = 1$; if within-route divergences differ, the predicted optimum should outperform share matching but stop below one. Both predictions hold in the specified controls (Appendix D.5). This is an oracle check of the formula using known conditional laws.

Sparse submode coverage inflates the estimated ceiling. In the stress control, a cell with 2.5% of target-conditional mass has probability 2×10^{-5} per baseline attempt. At $N = 4{,}000$, it is absent with probability $(1 - 2 \times 10^{-5})^{4000} \approx 0.923$. The frozen simulation observes it in only 8.1% of replications; the median estimated ceiling is 1.000 although the true ceiling is 0.108. Appendix D.5 reports the calibration and undercoverage analysis. The population assumptions of support and finite moments hold in this example, while the rare cell's small observation probability governs finite-sample reliability. Diagnostics based on observed weight tails (Vehtari et al., 2024) describe the sampled distribution; the stress control quantifies the effect of target mass outside that sample.

D.5 Finite-State Route-Reallocation Calibration

The calibration has one miss state and four event states indexed by two routes and two subtypes. The target event probability is 0.01, proposal discovery is 0.4, and event-state likelihood ratios are exact. The frozen ladder uses $N \in \{250, 1000, 4000, 16000, 64000\}$ attempted paths and 2,000 independent replications per scenario and arm. Baseline and theorem-predicted route-optimal arms use independent deterministic streams. The allocation-only case sets the target and proposal within-route laws equal. The mismatch case changes the within-route law in one route. The stress case assigns proposal probability 0.0001 to a target-important subtype. The protocol, parameters, formal seed, and decision rules were frozen before the retained formal execution. The route-optimal arm uses the known population parameters. This oracle control isolates the allocation formula; estimating its inputs from development data defines a further empirical comparison.

For completeness, write each within-route law as a two-by-two matrix, with routes in rows and subtypes in columns. The three cases use

	$\mathbf{p}$	$\mathbf{a}_{\text{base}}$	(P_z)	(Q_z)
allocation only	$(.65, .35)$	$(.9, .1)$	$\left(\begin{smallmatrix} .7 & .3 \\ .4 & .6 \end{smallmatrix}\right)$	$\left(\begin{smallmatrix} .7 & .3 \\ .4 & .6 \end{smallmatrix}\right)$
within-route mismatch	$(.6, .4)$	$(.45, .55)$	$\left(\begin{smallmatrix} .95 & .05 \\ .25 & .75 \end{smallmatrix}\right)$	$\left(\begin{smallmatrix} .2 & .8 \\ .25 & .75 \end{smallmatrix}\right)$
hidden submode	$(.5, .5)$	$(.5, .5)$	$\left(\begin{smallmatrix} .95 & .05 \\ .5 & .5 \end{smallmatrix}\right)$	$\left(\begin{smallmatrix} .9999 & .0001 \\ .5 & .5 \end{smallmatrix}\right)$.

Sampling uses multinomial counts over the miss state and four event cells, which is equivalent to independent path draws for these statistics. Each event-cell weight is $p_{\mathcal{E}} p_z P_z(u)/(H a_z Q_z(u))$.

Table 7: **Oracle allocation reaches the predicted ceiling; sparse coverage inflates its estimate.** Population quantities are exact. Simulation columns are medians over the declared replications. The first two rows use the largest budget to check convergence; the stress row uses its prespecified $N = 4{,}000$ decision point. "Stress cell seen" is the fraction of baseline replications observing the rare high-weight cell.

case	N	F_{base}	$F^{\max}_{\text{route}}$	median $\widehat{F}_{\text{repair}}$	median $\widehat{F}^{\max}_{\text{route}}$	stress cell seen
allocation only	64,000	0.590	1.000	1.000	1.000	–
within-route mismatch	64,000	0.256	0.356	0.356	0.356	–
hidden-submode stress	4,000	0.074	0.108	0.530	1.000	8.1%

The controls recover the predicted allocation-only behavior and heterogeneous-divergence optimum, while the hidden-submode case exposes severe ceiling overestimation under sparse coverage. The last row measures the gap between the empirical estimate and the population ceiling. Samples comprising the common cells have nearly constant within-route weights and yield an estimated ceiling close to one. The frequency of rare-cell observation therefore determines the reliability of this finite-sample estimate.

The baseline stress cell has per-attempt probability $q_* = H a_1 Q_1(2) = 0.4 \times 0.5 \times 0.0001 = 2 \times 10^{-5}$, so its observation probability is $1 - (1 - q_*)^N$. A 95% chance of seeing it at least once requires

$$N \geq \left\lceil \frac{\log(0.05)}{\log(1 - q_*)} \right\rceil = 149{,}786.$$

This analytical threshold gives the budget for observing the specified cell at least once; precise estimation of F requires further weight information. At $N = 64{,}000$, 27.2% of the retained baseline replications still miss the cell. At $N = 4{,}000$, the median largest observed weight is only 0.064% of total event weight despite the severely optimistic median fidelity. A small observed maximum-weight fraction thus coexists with substantial unobserved target mass. We retain raw weights throughout; smoothing weights would change the empirical identities being audited.

D.6 Second-Geometry Comparison and Scaling Results

A post hoc analysis of nine feedback–neutral comparisons on the three-hole landscape shows similar event discovery and lower estimated conditional fidelity across three noise levels and two control caps. These records originally tested an exponential efficiency-ratio scaling hypothesis, which was not supported in any of the three executions. The three result records and their hash-bound configurations are included in the supplement. All bootstrap intervals below are regenerated from the stored batch summaries.

The second target uses $\mathbf{b} = -\nabla V$ and the three-hole potential (Metzner et al., 2006)

$$V(x, y) = 3e^{-x^2-(y-1/3)^2} - 3e^{-x^2-(y-5/3)^2} - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2} + 0.2x^4 + 0.2(y - 1/3)^4.$$

Paths start near $(-1.048055, -0.042094)$ and target a radius-0.1 disk near $(1.048055, -0.042094)$, with $\Delta t = 0.005$, horizon 3, and $\varepsilon \in \{0.22, 0.18, 0.15\}$. The hash-bound configurations retain full coordinate precision. The guide uses a 251×251 grid, scale 0.9, and mixture masses $(0.2, 0.4, 0.4)$; the paired arms differ only in feedback gain (1.5 versus zero). Each arm has 20 independent batches: 5,000 paths per batch originally and 20,000 in each later execution. All nine arm pairs use disjoint simulation-seed ranges. The three executions comprise the original paired study, a fourfold path-budget rerun, and a raised-cap rerun, using control caps 4 and 40. The distributed historical records support deterministic statistical replay; fresh simulation entries in the archive cover the primary diffusion studies.

Table 8: Post hoc RareFlow readout of every retained three-hole comparison. Feedback (fb) and neutral (0) arms use independent simulations. The path-variance ratio $\rho_{\mathrm{mix}} = V_{\mathrm{fb}}/V_0$ and its original independent-arm batch-bootstrap interval are retained; values above one favor the neutral arm. The raised-cap row at $\varepsilon = 0.15$ is directionally consistent but its interval includes one.

execution	ε	$\widehat{H}_{\mathrm{fb}}$	$\widehat{H}_0$	$\widehat{F}_{\mathrm{fb}}$	$\widehat{F}_0$	ρ_{mix}	95% CI
original	0.22	0.603	0.594	0.051	0.204	4.37	[2.54,7.11]
original	0.18	0.572	0.572	0.046	0.176	4.15	[3.09,5.59]
original	0.15	0.525	0.536	0.012	0.095	10.05	[3.72,21.03]
higher power	0.22	0.602	0.594	0.034	0.176	5.79	[3.29,9.48]
higher power	0.18	0.570	0.572	0.034	0.134	4.10	[2.79,5.95]
higher power	0.15	0.524	0.537	0.016	0.086	6.11	[3.04,11.54]
raised cap	0.22	0.788	0.794	0.074	0.367	6.66	[3.54,11.48]
raised cap	0.18	0.780	0.788	0.098	0.241	2.76	[1.05,8.69]
raised cap	0.15	0.721	0.738	0.106	0.178	1.78	[0.68,6.93]

Feedback has similar discovery and lower point-estimate fidelity in all nine rows, so the qualitative separation holds across both tested geometries and the recorded control caps. This post hoc reanalysis addresses a different question from the original exponential-scaling hypothesis. The frozen scaling decisions and the discovery–fidelity comparisons report both outcomes for the same three executions. The resulting evidence concerns these two planar systems.

D.7 Mechanism Decomposition

The frozen guide sweep retains route-specific hits and weighted moments. Let $\mathcal{I}_z = \{i : I_i = 1,\, Z_i = z\}$, $M_z = |\mathcal{I}_z|$, $A_z = \sum_{i\in\mathcal{I}_z} w_i$, $B_z = \sum_{i\in\mathcal{I}_z} w_i^2$, and $\widehat{\mu}_z = A_z/N$. The route-specific effective sample size is $\mathrm{ESS}_z = A_z^2/B_z$; $I_i = 1$ on $\mathcal{I}_z$. The same-sample quantities

$$\widehat{p}_z = \frac{\widehat{\mu}_z}{\sum_{z'} \widehat{\mu}_{z'}}, \qquad \widehat{r}_z = \frac{M_z}{M}, \qquad \widehat{D}_z = \frac{M_z}{\mathrm{ESS}_z} - 1 \tag{30}$$

are used when the total event-weight sum is positive; $\widehat{D}_z$ is defined only for mechanisms with positive event-weight sums. They give the exact empirical decomposition $\widehat{D}_{\mathrm{between}} = \chi^2(\widehat{\mathbf{p}}\|\widehat{\mathbf{r}})$ and $\widehat{D}_{\mathrm{within}} = \sum_{z:\widehat{p}_z>0}(\widehat{p}_z^2/\widehat{r}_z)\widehat{D}_z$. The within-mechanism sum runs over routes with positive estimated target share. Computing each included ESS_z and $\widehat{D}_z$ requires a positive route-weight sum. Table 9 reports all stored resolutions. These are descriptive same-sample estimates based on importance-weighted target shares; independent route-conditioned validation is reported in Appendix D.8.

Evaluation records. The 251×251 feedback entries in Tables 16 and 2 use separate frozen evaluations, with $20 \times 5{,}000$ and $5 \times 4{,}000$ paths, respectively, and disjoint seed ranges 873000–873019 and 855000–855004. Both declare the same target chain ($\varepsilon = 0.12$, $\Delta t = 0.005$, horizon 3.0), 251×251 guide construction, scale 0.90, route gain 1.50, weights $(0.20, 0.40, 0.40)$, and control cap 4.0. Both feedback arms share the saved guide at each noise level. The guide sweep retains its separate reference guide and evaluation. Feedback intervals use independent-arm batch bootstrapping at 95% confidence. Both statistics pool path-weight sums and squared sums across batches before computing $\widehat{F} = \mathrm{ESS}_{\mathcal{E}}/M$ from the pooled moments. The resulting values are 0.334155 and 0.343780.

Table 9: Within-route variation dominates the same-sample empirical discrepancy throughout the frozen guide sweep. The lower-route frequency $\widehat{r}_{\text{lower}}$ and importance-weighted target-share estimate $\widehat{p}_{\text{lower}}$ can nearly agree while the within-route term remains large. Every stored resolution is included.

grid	$\widehat{r}_{\text{lower}}$	$\widehat{p}_{\text{lower}}$	$\widehat{D}_{\text{between}}$	$\widehat{D}_{\text{within}}$	$\widehat{F}$
7	0.705	0.616	3.81×10^{-2}	8.462	0.105
11	0.754	0.640	6.95×10^{-2}	10.521	0.086
21	0.587	0.606	1.44×10^{-3}	9.691	0.094
31	0.560	0.557	2.66×10^{-5}	5.578	0.152
51	0.603	0.594	3.54×10^{-4}	3.539	0.220
71	0.614	0.588	2.93×10^{-3}	2.871	0.258
101	0.641	0.591	1.10×10^{-2}	2.133	0.318
151	0.658	0.585	2.41×10^{-2}	1.847	0.348
251	0.667	0.595	2.37×10^{-2}	1.885	0.344

D.8 Independent Route-Conditioned Timing Audit

Independent reference support and frequency aggregation. Each target seed supplies 200,000 paths. For seeds 2026090411–2026090414, the (lower, upper; total) event counts are $(339, 240; 579)$, $(345, 237; 582)$, $(325, 201; 526)$, and $(352, 225; 577)$, respectively. Pooling gives 800,000 paths, 2,264 hits, and a crude probability estimate $2{,}264/800{,}000 = 0.002830$; these are separate from the guide sweep's reference sample. The pooled lower-route fraction is $1{,}361/2{,}264 \approx 0.601148$. The frequency reported in Section 3.5 and Table 10 instead averages the four seed-level fractions equally: $(339/579+345/582+325/526+352/577)/4 \approx 0.601550$, rounded to 0.602. The proposal frequencies use the same seed averaging, giving 0.705758, 0.561777, and 0.664310 for 7×7, 31×31, and 251×251.

Table 10: External conditional-observable audit. The target reference is generated by fresh crude target simulation and is independent of the fresh proposal evaluations. Time errors are unweighted normalized first-hitting-time errors relative to that reference; Figure 5 gives fixed outer-seed bootstrap intervals. The 31×31 guide is closer in lower-route frequency than the 251×251 guide while showing a larger route-conditioned time error. The frozen result includes all three guide resolutions; the self-normalized route-time calibration is reported separately from this unweighted trajectory-collection comparison. The inherited $\widehat{F}$ column is copied from the 2×10^4-path-per-resolution guide sweep in Table 9. The timing columns use the fresh observable-evaluation paths.

guide	inherited $\widehat{F}$	$\widehat{r}_{\text{lower}}$	Δm_{lower}	Δm_{upper}	route-time RMSE
7×7	0.105	0.706	$+0.063$	$+0.088$	0.077
31×31	0.152	0.562	$+0.153$	$+0.276$	0.223
251×251	0.344	0.664	-0.009	$+0.073$	0.052
target reference	–	0.602	0	0	0

Observable protocol and decision rule. The prespecified protocol defines $g = \tau/T$ and the two-route error $R_a = \{\frac{1}{2}\sum_z(\widehat{m}_{a,z} - \widehat{m}_{P,z})^2\}^{1/2}$, where each $\widehat{m}$ is the arithmetic mean of four independent outer-seed route-conditioned sample means. For each arm a, define the normalized route-mean error $\Delta m_z := \widehat{m}_{a,z} - \widehat{m}_{P,z}$. Routes receive equal weight, irrespective of their frequencies. The primary contrasts are $R_7 - R_{251}$ and $R_{31} - R_{251}$. For each of 2,000 bootstrap replicates, the four complete seed summaries are resampled within the reference and independently within each arm; the same resampled reference and comparator are used in both contrasts. With the frozen bootstrap seed, the 5th and 95th percentiles give the 90% intervals. Evidence requires a lower endpoint strictly greater than 0.02, equivalent to 0.06 simulation-time units at $T = 3$. Table 11 reports both prespecified comparisons and decisions. Figure 5 displays the individual route-mean errors and their 90% intervals; the table gives the intervals for the RMSE contrasts.

Weighted calibration. Within each seed and route, the weighted mean is the self-normalized ratio $\sum_i w_i g_i / \sum_i w_i$ over event paths in that route. The analysis averages these four seed-level ratios,

Table 11: Prespecified primary contrasts and separate weighted calibration check. The primary criterion is a 90% interval lower endpoint > 0.02. Calibration requires maximum absolute weighted route-time error ≤ 0.05; it is a point-error tolerance check.

guide	$R_a - R_{251}$	90% interval	criterion met	max. weighted error
7×7	0.025053	$[0.019045, 0.032395]$	no	0.011150
31×31	0.171491	$[0.167113, 0.175502]$	yes	0.001361
251×251	–	–	comparator	0.003852

then compares each route with the corresponding mean of the four independent crude-reference summaries. The frozen configuration sets a tolerance of 0.05: each arm passes if its maximum absolute error across the two routes is at most this value. All three arms pass (Table 11). This calibration rule compares point errors with the declared tolerance.

D.9 Follow-up on Learned Trajectories and Control-Cap Consistency

The original learned-residual comparison and the guide-family timing audit use different proposals. We therefore froze a separate follow-up after inspecting those results. Its evaluation data are new; the earlier findings remain retrospective. The target chain, event, route labels, coarse guide, network architecture, optimizer, regularization, and stopping rule are inherited from the learned comparison. The accompanying `07_learned_path_audit` directory contains the protocol, configuration, source hashes, raw event records, checkpoints, and execution entry.

Same-proposal trajectory comparison. The bridge study evaluates each of the five original frozen checkpoints and its coarse baseline on 20 fresh batches of $5{,}000$ attempted paths per arm. A common independent crude reference uses 20 batches of $40{,}000$ attempts. We retain each successful path's route, normalized first-hitting time, exact full-mixture weight, and batch index. Batch sums and counts are pooled before computing $\widehat{H}$, $\widehat{F}$, $\widehat{S}$, and route-conditioned mean times. This pooled-path timing estimator gives each event path equal weight; Appendix D.8 instead gives each outer-seed mean equal weight. Each study retains its declared aggregation rule.

For model j and arm a, the timing error is

$$R_{j,a} = \left[\frac{1}{2} \sum_{z \in \{\text{lower},\text{upper}\}} (\widehat{m}_{j,a,z} - \widehat{m}_{\text{ref},z})^2 \right]^{1/2}.$$

The primary effect is the mean of $R_{j,\text{learned}} - R_{j,\text{coarse}}$ over the five fixed checkpoints. The prespecified timing-increase criterion declares a meaningful degradation only when the lower endpoint of the 95% interval exceeds $+0.02$. The observed interval instead lies entirely below zero, indicating improved timing accuracy. The early-hit fraction, $\Pr(\tau_{\mathcal{B}}/n \leq 0.5 \mid \mathcal{E}, Z = z)$, is a descriptive secondary endpoint. The decision uses the prespecified mean-time RMSE contrast. Self-normalized weighted route-time means provide a separate point check: the maximum absolute error across the two routes must be at most 0.05.

Matched control-cap comparison. The cap study trains five new pairs on independent 10^5-path development samples. Within each pair, the two networks share every retained transition, weight, initialization, optimizer setting, and 400-epoch training schedule. Each coarse, legacy-objective, and consistent-objective arm then receives 20 fresh batches of $5{,}000$ evaluation paths. Let $\mathbf{u}$ denote the raw scaled guide, $\mathbf{C}_c$ the norm cap, f_ϕ the residual, and $\mathbf{a}$ the realized control used in Appendix B.3. The legacy objective regresses $\mathbf{a} - \mathbf{C}_c(\mathbf{u})$ on f_ϕ. The consistent objective instead regresses it on $\mathbf{C}_c(\mathbf{u} + f_\phi) - \mathbf{C}_c(\mathbf{u})$, fitting the control that is deployed. Both use the same residual penalty and are evaluated with $\mathbf{C}_c(\mathbf{u} + f_\phi)$. Only the regression objective changes.

The cap study's primary comparison is consistent minus coarse: an upper 95% endpoint below zero for ΔF, together with a lower endpoint above zero for ΔS, supports persistence of the trade-off under this matched objective. Consistent-minus-legacy fidelity is also reported. This contrast estimates the effect of changing cap placement in the objective. Cap activation is measured among simulated

guided-component states. The separate potential-exponent safeguard has a validity threshold of zero activations.

Uncertainty, validity, and accounting. Each of 1,999 bootstrap replicates resamples the reference's 20 batches once, draws five model indices with replacement, and independently resamples 20 evaluation batches within each selected model and arm. The same resampled reference is used throughout the replicate. Ratios are recomputed from pooled moments, then model effects are averaged. The intervals use the 2.5th and 97.5th percentiles. For the bridge's fixed checkpoints, this hierarchical interval is a sensitivity summary over the five observed fits and their evaluation batches. Sensitivity to high-weight submodes depends on their representation in those samples.

The frozen validity checks require at least 1,500 crude events and 400 per reference route, at least 100 events per route in every complete proposal arm, finite likelihoods, zero potential-safeguard activations, and probability estimates within four combined standard errors of the fresh reference. The cap study also checks the inherited finite-schedule optimization diagnostic: final loss within 1% of the minimum over the last 50 epochs. This check measures stability over the final training window. Claim eligibility requires every validity check to pass with the recorded batches, paths, seeds, and endpoints matching the fixed schedule.

The complete follow-up uses 3.8×10^6 attempts: 0.8 million for the common reference, 1.0 million for the bridge, 0.5 million for cap-study development, and 1.5 million for cap-study evaluation. The two objectives use the same development and evaluation budgets. These comparisons use equal attempted-path budgets; the concurrent execution timings document the run. Tables 12–14 report all five fits and the declared contrasts.

Outcomes. Both studies pass every frozen validity check. The bridge repeats the fidelity decrease, $\Delta F = -0.246$ with interval $[-0.278, -0.210]$, while improving timing in every pair. Its RMSE-change interval lies below the declared $+0.02$ threshold. In the cap study, all five consistent-objective fits also combine higher efficiency, lower fidelity, and better timing. The cap study's primary comparison concerns F and S; timing is an additional reported outcome. Its mean fidelity difference from legacy training is 0.014 with interval $[-0.014, 0.056]$, which spans zero. All 25 proposal arms pass the weighted timing point check. The trade-off persists with consistent cap placement, and both learned comparisons combine lower global fidelity with improved accuracy of this observable. The raw records also retain the declared early-hit summaries as secondary outcomes.

Table 12: Fresh evaluations of the five original checkpoints and their coarse baselines. Each row uses 10^5 attempts. Time RMSE is equally weighted across the two routes; wt. error is the maximum absolute weighted mean-time error. The common 8×10^5-attempt crude reference has 2014 events and route-time means 0.573 (lower) and 0.543 (upper). Cap columns give the percentage of guided states whose raw combined control exceeds the cap.

fit	arm	$\widehat{H}$	$\widehat{F}$	$\widehat{S}$	time RMSE	wt. error	cap (%)
1	coarse	0.0198	0.471	0.0093	0.211	0.0112	0.5
1	learned	0.6885	0.269	0.1854	0.079	0.0014	0.7
2	coarse	0.0199	0.476	0.0095	0.216	0.0074	0.5
2	learned	0.6957	0.203	0.1412	0.040	0.0008	0.5
3	coarse	0.0192	0.443	0.0085	0.214	0.0216	0.5
3	learned	0.6720	0.178	0.1195	0.082	0.0039	0.7
4	coarse	0.0197	0.456	0.0090	0.213	0.0188	0.5
4	learned	0.6784	0.177	0.1198	0.078	0.0036	0.7
5	coarse	0.0198	0.444	0.0088	0.218	0.0157	0.5
5	learned	0.6764	0.230	0.1559	0.079	0.0047	0.7

Table 13: Matched cap-objective study. Each fit uses one shared 10^5-path development sample and initialization for legacy and consistent training; each displayed arm uses 10^5 independent evaluation attempts. The columns and common reference are as in Table 12.

fit	arm	$\widehat{H}$	$\widehat{F}$	$\widehat{S}$	time RMSE	wt. error	cap (%)
1	coarse	0.0195	0.471	0.0092	0.215	0.0095	0.5
1	legacy	0.7193	0.248	0.1785	0.084	0.0015	0.7
1	consistent	0.7166	0.237	0.1702	0.084	0.0036	0.7
2	coarse	0.0198	0.447	0.0088	0.212	0.0106	0.5
2	legacy	0.7146	0.187	0.1335	0.042	0.0040	0.5
2	consistent	0.7144	0.269	0.1918	0.042	0.0052	0.5
3	coarse	0.0201	0.447	0.0090	0.210	0.0104	0.5
3	legacy	0.6489	0.312	0.2021	0.088	0.0011	0.7
3	consistent	0.6488	0.299	0.1938	0.087	0.0005	0.7
4	coarse	0.0195	0.461	0.0090	0.216	0.0196	0.5
4	legacy	0.6803	0.168	0.1144	0.102	0.0101	0.6
4	consistent	0.6818	0.189	0.1290	0.101	0.0014	0.6
5	coarse	0.0211	0.469	0.0099	0.214	0.0066	0.5
5	legacy	0.7090	0.327	0.2318	0.062	0.0038	0.9
5	consistent	0.7089	0.319	0.2260	0.061	0.0049	0.9

Table 14: Follow-up contrasts, averaged over five fits, with 95% hierarchical bootstrap intervals. In the bridge, Δ denotes learned minus coarse; in the cap study it denotes consistent minus coarse. R is route-time RMSE. The two additional fidelity contrasts isolate the legacy objective. These intervals summarize variation among the five fits and their recorded evaluation batches.

study	contrast	estimate	95% interval
bridge	ΔF	-0.246	$[-0.278, -0.210]$
bridge	ΔS	0.1353	$[0.1139, 0.1615]$
bridge	ΔH	0.663	$[0.655, 0.670]$
bridge	ΔR	-0.143	$[-0.160, -0.130]$
cap	ΔF	-0.197	$[-0.243, -0.155]$
cap	ΔS	0.1730	$[0.1434, 0.1988]$
cap	ΔH	0.674	$[0.649, 0.695]$
cap	ΔR	-0.138	$[-0.157, -0.120]$
cap	$F_{\text{consistent}} - F_{\text{legacy}}$	0.014	$[-0.014, 0.056]$
cap	$F_{\text{legacy}} - F_{\text{coarse}}$	-0.211	$[-0.268, -0.155]$

E COMPUTATIONAL COST AND COMMON-SCALE ACCOUNTING

Recorded-time normalization. Table 15 separates the recorded stages. For one trained model evaluated on the frozen path budget, a descriptive comparison under linear coarse-sampling throughput is

$$R_{\text{time}} = R_{\text{pooled}} \frac{\bar{t}_c}{\bar{t}_{\text{dev}} + \bar{t}_{\text{train}} + \bar{t}_\ell}. \tag{31}$$

This is the coarse-to-learned estimator-variance ratio after scaling the coarse path count to the learned workflow's recorded time. It excludes the shared coarse-guide construction and assumes independent paths with constant per-path cost. It is a plug-in normalization of the recorded computational components; the executed comparisons use fixed path budgets. Reusing a learned proposal would amortize development differently.

Table 15: Recorded-time accounting from the five frozen learned-residual runs. Stage times are arithmetic means for the declared workloads. The range spans per-model plug-ins and is a descriptive minimum-to-maximum band across the five models.

quantity	recorded or derived value
development sampling	27.2218 s
optimization	72.0835 s
coarse evaluation	18.6714 s
learned evaluation	36.0557 s
pooled recorded-time normalization	2.57
per-model recorded-time range	$[1.71, 3.86]$

E.1 COMMON-SCALE CONFIGURATION AND PREPARATION COSTS

Table 16 places the interventions on the same $\widehat{H}, \widehat{F}, \widehat{S}$ scale. It compares route feedback at 251×251, residual learning at 11×11 with feedback off, and refinement from 11×11 to 251×251 with feedback on.

Table 16: Three controlled patterns of efficiency change at $\varepsilon = 0.12$. Across these contrasts, efficiency differences arise from discovery, fidelity, or both. Columns give discovery ($\widehat{H}$), fidelity ($\widehat{F}$), and overall efficiency ($\widehat{S}$). Ranges span five independent residual runs. Evaluations use 10^5 paths per configuration (2×10^4 for the 11×11 feedback-on guide).

guide	route feedback	$\widehat{H}$	$\widehat{F}$	$\widehat{S}$
11×11	on ($\gamma = 1.5$)	0.034	0.086	0.0030
11×11	off ($\gamma = 0$)	0.019–0.020	0.387–0.464	0.0074–0.0090
11×11 + residual	off ($\gamma = 0$)	0.670–0.701	0.169–0.290	0.114–0.200
251×251	on ($\gamma = 1.5$)	0.769	0.334	0.257
251×251	off ($\gamma = 0$)	0.770	0.636	0.490

The fine-feedback $\widehat{F} = 0.334$ here and 0.344 in Table 2 come from separate evaluations at the same declared settings (Appendix D.7). The resolution-sweep 11×11 guide took 0.074 s to build, while the 251×251 feedback guide took 7.96 s. Per learned model, development sampling averaged 27.2 s and optimization 72.1 s; the evaluation stage times are in Table 15. These are preparation costs, accounted for separately from per-path S. The coarse sweep arm uses 2×10^4 paths; the learned and matched-feedback comparisons use 10^5 per arm. With $\gamma = 0$, the feedback mixture $(0.2, 0.4, 0.4)$ is equivalent to the learned study's $(0.2, 0.8)$ mixture, and the two neutral code paths agree bitwise in the supplied check.