WHEN DOES A NEURAL SAMPLER ACTUALLY NEED THE TARGET SCORE? ISOLATING THE ROLE OF LANGEVIN PARAMETERISATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Nearly every effective neural sampler for unnormalised targets injects the target score $\nabla \log \pi$ into its drift, a design known as *Langevin parameterisation*. He et al. (2025) showed that removing it causes severe mode collapse and concluded that an explicit score term is required, supporting this with capacity, initialisation and target-density ablations. We ask a narrower question their study leaves open: what exactly about the score is needed, and when. Varying only the score's role inside a single architecture family, we find the term is indeed causal—coverage drops from $8/8$ modes to $2/8$—but it works just as well as an ordinary input feature, and $16\times$ the capacity does not substitute for it, so what matters is the availability of the gradient rather than the residual structure around it. Most unexpectedly, within this setting the score is needed only during a specific *interval* of the sampling trajectory: restricting it to $t \in [0, 0.5)$ is numerically identical to having no score at all, $[0.5, 1)$ recovers full coverage, and the last 10% alone is worthless again. We give a partial account. The optimal control provably converges to the target score as $t \to 1$, which explains the design but deepens the puzzle; a transport bound then shows the cost of redistributing mass inside a window of width w diverges as $R^2/(2\sigma^2 w)$, providing a control-cost mechanism consistent with the failure of late windows, and varying the mode separation moves the threshold consistently with the implied $w_{\min} \propto R^2$. The failure of *early* windows remains unexplained, and we report it as open rather than attach one of the two intuitive stories our own experiments falsify.

1 INTRODUCTION

Sampling from a density known only up to normalisation, $\pi(x) \propto e^{-E(x)}$, is a core problem in Bayesian inference and statistical physics. A recent line of work learns a neural *sampler*: a diffusion or flow model trained against the energy rather than against data (Zhang & Chen, 2022; Vargas et al., 2023; Albergo & Vanden-Eijnden, 2025; Chen et al., 2025; Havens et al., 2025; Liu et al., 2025; Guo et al., 2026).

Almost all of these methods share an architectural detail that is rarely the headline of the paper. The learned drift is not a free-form network of the state; it contains the target score explicitly,

$$u_\theta(x,t) \;=\; \mathrm{NN}_1(x,t) \;+\; \mathrm{NN}_2(t) \odot \nabla \log \pi(x), \tag{1}$$

a design introduced by Zhang & Chen (2022) and inherited by most subsequent samplers. He et al. (2025) made this dependence explicit across several samplers and objectives, and supported it with ablations: they replaced the Langevin-parameterised network with a plain MLP enlarged to five layers and 256 hidden units, initialised a score-free student by distilling a converged score-using teacher, and conditioned the network on the target *density*, $\mathrm{NN}_\theta(x, \log \tilde{\pi}(x), t)$, in place of the score. All collapse. They concluded that "Langevin preconditioning cannot be replaced by alternative target information, such as $\log p_{\mathrm{target}}$", and that neural samplers "require an explicit score term". Kim et al. (2025) take the observation further, noting that this reliance on energy gradients *at inference* "erodes the primary efficiency benefit of amortized sampling".

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.

We take that conclusion as our starting point rather than our target. It settles *that* the score is needed; it leaves open *what* about the score is needed, and *when*. Two gaps motivate this paper. First, the density ablation above supplies $\log \tilde{\pi}$ as an input but never $\nabla \log \pi$ itself, so it cannot distinguish "the drift must be built around the gradient" from "the network must merely be able to read the gradient"—a distinction that decides whether the additive form of equation 1 is doing the work. Second, the score is supplied for the whole trajectory or not at all, so nothing there speaks to *when* along the trajectory it matters.

We therefore fix one architecture family—a free-form network producing the drift of a controlled SDE trained with the Girsanov reverse-KL (PIS) objective—and vary *only* the score's role, on targets with exact ground truth so that coverage can be counted rather than inferred from an aggregate distance. Our contributions:

- **A clean isolation.** Within one architecture family, removing the score term drops mode coverage from $8/8$ to $2/8$; supplying $\nabla \log \pi$ as a plain input feature is as effective as the additive form of equation 1; and a $16\times$ larger score-free network is indistinguishable from the small one (§5). The conclusion of He et al. (2025) survives, and we sharpen it: what matters is the availability of the gradient, not the residual structure around it. Their density ablation shows $\log \tilde{\pi}$ as an input is not enough; we find $\nabla \log \pi$ as an input is.
- **A schedule, not a switch.** The score is needed only inside a particular interval of trajectory time (Figure 1). Score on $[0, 0.5)$ is numerically identical to no score; on $[0.5, 1)$ it recovers full coverage; on $[0.9, 1)$ it is again worthless. In our setting every successful window contains $t \approx [0.5, 0.75]$ and spans a substantial fraction of the trajectory (§6).
- **Theory for one half of it.** The optimal control of equation 3 converges to $\sigma^2 \nabla \log \pi(x) + x$ as $t \to 1$ (Prop. 1), making the failure of terminal-only windows the more surprising; a transport bound (Prop. 2) supplies a necessary control-cost mechanism consistent with it, since carrying a fraction p of the mass a distance R inside a window of width w costs at least $pR^2/(2\sigma^2 w)$ nats. Varying the mode separation moves the threshold consistently with the implied $w_{\min} \propto R^2$ (§7).
- **An explicitly open half.** The bound has no force on early windows, which enjoy the largest transport budget and fail anyway. We formulated two intuitive explanations and falsified both with our own experiments (§8), and leave the gap visible rather than fill it with a story the data contradict.

2 Related work

Neural samplers for unnormalised targets span stochastic-control and denoising-diffusion formulations (Zhang & Chen, 2022; Vargas et al., 2023), non-equilibrium transport (Albergo & Vanden-Eijnden, 2025), adjoint and matching objectives (Havens et al., 2025; Liu et al., 2025), proximal schemes (Guo et al., 2026), simulation-free regression (Havens et al., 2026), closed-form interpolants with tempering (Castro-Macías et al., 2026), generator matching over continuous, discrete and mixed state spaces (Woo et al., 2025; Sanokowski et al., 2025), constrained mass transport (von Klitzing et al., 2026), and hybrids that correct a learned proposal with resampling or SMC (Chen et al., 2025; Wu et al., 2025). Nearly all place the target score inside the drift as in equation 1; that shared choice is what we examine. Lim (2025) studies a different under-determination of the same object—a velocity is unique only up to divergence-free fields—complementary to our question of *when* the target enters it. Remedies for mode collapse act on the prior, the temperature or the training objective (Blessing et al., 2025; Qin et al., 2025; Rissanen et al., 2025; Kim et al., 2025; Wang et al., 2025), and Blessing et al. (2024) argue the field's protocols are poorly suited to detecting collapse at all—which our Many-Well results reinforce. A separate tradition matches scores as the training *objective* rather than as an architectural input (Modi et al., 2025; Cai et al., 2025; Nguyen et al., 2025; Kim, 2026), and an adjacent literature builds diffusion and flow posteriors for latent-variable models and inverse problems (Piriyakulkij et al., 2025; Xu et al., 2026; Eijkelboom et al., 2024; 2025; Carrasco-Pollo et al., 2026; Zhai et al., 2025; Ko & Domke, 2025; Li et al., 2025; Daudel et al., 2026; Moufad et al., 2025). Appendix E expands on each of these.

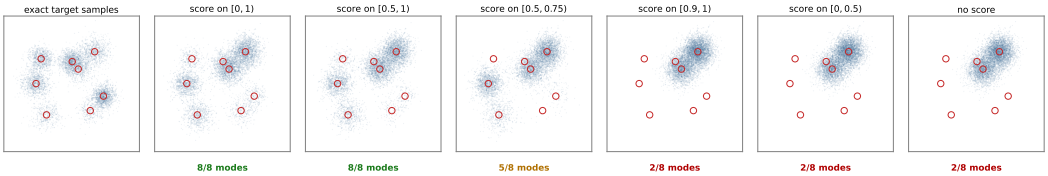

Figure 1: **The score is needed on a window of trajectory time, and the window's position decides everything.** Samples projected onto the two leading directions of the mode means; red circles mark the eight modes. All panels but the first share architecture, objective, optimiser and budget, differing only in when the drift may see $\nabla \log \pi$. Note that $[0, 0.5)$ and $[0.5, 1)$ have equal width and opposite outcomes.

3 Setup

3.1 Sampler

We use a controlled-SDE sampler in the style of the Path Integral Sampler (Zhang & Chen, 2022). Trajectories start at $x_0 = 0$ and evolve on $t \in [0, 1]$ as

$$\mathrm{d}x_t \;=\; u_\theta(x_t, t)\,\mathrm{d}t \;+\; \sigma\,\mathrm{d}W_t, \tag{2}$$

discretised with $T = 50$ Euler–Maruyama steps and $\sigma = 2$. Writing p_{ref} for the terminal law of the uncontrolled process ($\mathcal{N}(0, \sigma^2 I)$), the Girsanov form of the reverse KL on path measures gives the training loss

$$\mathcal{L}(\theta) \;=\; \mathbb{E}\left[\int_0^1 \frac{\|u_\theta(x_t, t)\|^2}{2\sigma^2}\,\mathrm{d}t \;-\; \log \tilde{\pi}(x_1) \;+\; \log p_{\mathrm{ref}}(x_1)\right]. \tag{3}$$

All arms share equation 3, the discretisation, the optimiser (Adam, 5×10^{-4}, cosine decay), the batch size (512) and the number of updates (1500). Only the parameterisation of u_θ changes.

3.2 Arms

With $\mathrm{te}(t)$ a sinusoidal time embedding, the arms are **a-additive**, $u = \mathrm{NN}_1([x, \mathrm{te}]) + \mathrm{NN}_2(\mathrm{te}) \odot \nabla \log \pi(x)$, the standard form; **b-noscore**, $u = \mathrm{NN}_1([x, \mathrm{te}])$, the same family with the score removed; **c-feature**, $u = \mathrm{NN}_1([x, \nabla \log \pi(x), \mathrm{te}])$, the score as an input feature; **d-noscore-big**, **b** with a $16\times$ larger network (capacity control); and **e-covinit**, **b** started from $x_0 \sim \mathcal{N}(0, 3^2 I)$ (mode-covering start). NN_2 is zero-initialised, so **a** and **b** coincide at initialisation. For **e** the reference terminal law is corrected to $\mathcal{N}(0, (s_0^2 + \sigma^2) I)$ so equation 3 remains the reverse KL for that process. Later sections gate the score to $t \in [\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}})$.

3.3 Targets and metrics

Gaussian mixture (GMM). $d = 16$, $K = 8$ components with unequal weights, means on a sphere of radius 6, isotropic scales in $[0.6, 1.0]$. Exact samples and exact mode weights are available for evaluation only.

Many-Well (MW). $E(x) = \sum_{i<m}(x_i^4 - 6x_i^2 - 0.5x_i) + \sum_{i \geq m} \frac{1}{2} x_i^2$ with $d = 16$, $m = 5$, giving $2^5 = 32$ modes. The geometry is deliberately unlike the GMM: modes sit on coordinate axes at $|x_i| \approx 1.7$, close to the trajectory's starting point, rather than on a distant sphere. Ground truth stays exact because the energy separates: mode weights follow from a one-dimensional quadrature and exact samples from a one-dimensional inverse CDF.

We report three quantities. **Mode coverage** counts modes receiving more than 1% of the sampler's mass. **Mode-weight error** is $\|\hat{w} - w^\star\|_1$ over modes. **Energy distance** is computed against exact reference samples. Appendix F shows these can disagree sharply, and that mode coverage is the one that tracks collapse.

4 WHY THE SCORE ENTERS THIS WAY, AND WHAT THAT PREDICTS

Before the experiments it is worth asking why equation 1 is the form everyone uses. For the objective equation 3 the answer is exact.

Proposition 1 (The optimal control is the score, at terminal time)**.** *Let* $\Phi(x,t) = \mathbb{E}[(\tilde{\pi}/p_{\mathrm{ref}})(x_1) \,|\, x_t = x]$ *under the uncontrolled process. The minimiser of equation 3 is* $u^\star(x,t) = \sigma^2 \nabla_x \log \Phi(x,t)$*, and*

$$\lim_{t \to 1^-} u^\star(x,t) \;=\; \sigma^2 \, \nabla \log \pi(x) \;+\; x. \tag{4}$$

Hence the additive parameterisation equation 1 represents the optimal control exactly *in the terminal limit, with* $\mathrm{NN}_2(t) \to \sigma^2 \mathbf{1}$ *and* $\mathrm{NN}_1(x,t) \to x$*.*

Proposition 1 explains the community's design choice: the score is not a heuristic preconditioner but the exact optimal drift in the limit $t \to 1$. It also sharpens the question this paper asks. If the parameterisation is provably exact near $t = 1$, one would expect the score to be most valuable there. Section 6 shows the opposite: supplying it *only* near $t = 1$ is worthless.

The resolution is that being able to represent the optimal drift is not the same as being able to afford it. The objective charges for control, and the charge for moving mass grows as the available window shrinks.

Proposition 2 (Transport budget)**.** *Suppose the score-informed part of the drift acts only on a window of width* $w = \tau_{\mathrm{hi}} - \tau_{\mathrm{lo}}$*, let* $C = \mathbb{E} \int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} \|u\|^2/(2\sigma^2)\, \mathrm{d}t$ *be the control cost it incurs in equation 3, and write* $\Delta = \int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} u(x_t,t)\, \mathrm{d}t$ *for the displacement the control contributes along a trajectory. Then*

$$\mathbb{E}\big[\|\Delta\|^2\big] \;\le\; 2\sigma^2 C w. \tag{5}$$

Consequently, if a fraction p *of the probability mass must be carried a distance at least* R by the control*, then*

$$C \;\ge\; \frac{p\,R^2}{2\sigma^2 w}. \tag{6}$$

The bound is stated per trajectory rather than for the mean displacement on purpose. A bound on $\|\mathbb{E}[\Delta]\|$ would be vacuous here: redistributing mass among modes can move packets in opposing directions and leave the global mean almost unchanged. Equation 5 instead controls the second moment, which is what a statement about mass transport requires. Both proofs are elementary and deferred to Appendix D. On our GMM the modes lie on a sphere of radius 6 in $d = 16$, so a typical pair is $R \approx 6\sqrt{2} \approx 8.5$ apart; with $\sigma = 2$, equation 6 charges at least $R^2/(2\sigma^2 w) \approx 9.0/w$ nats *per unit of mass* that has to cross—18 nats at $w = 0.5$, 36 at $w = 0.25$, 90 at $w = 0.1$ and 451 at $w = 0.02$. The terminal reward differences between modes are only a few nats, the weights spanning less than a factor of ten, so at small w the transport can no longer pay for itself.

We are careful about what this does and does not deliver. It gives a *necessary* cost that diverges as the window closes, and the direction and rate of that divergence match the observed collapse of late windows: $w = 0.5$ recovers all eight modes, $w = 0.25$ recovers three, and $w = 0.10$ and $w = 0.02$ fall back to the score-free baseline (Table 3). It does not by itself locate the threshold, which depends on how much terminal reward is actually on offer and on what the optimiser finds; and it is a statement about the control, not about the diffusion, which supplies an undirected displacement of comparable size (Remark 2).

One step is worth naming rather than glossing. The bound is conditional: *if* a fraction p of the mass has to be carried a distance R by the control, the cost is at least $pR^2/(2\sigma^2 w)$. We do not prove that recovering the correct mode allocation requires exactly that of the control—the Brownian part contributes displacement of comparable magnitude (Remark 2), even if undirected. Proposition 2 should therefore be read as a necessary control-cost mechanism consistent with the late-window failure, not as a derivation of it.

Proposition 2 bears only on *late* windows. It has no force on the early side: a window $[0, \tau)$ has the largest remaining time and hence the most generous transport budget of all, yet those arms fail just as completely. What the score is doing that requires a mid-trajectory window remains open (§8).

Remark 1 (The prediction this makes, and its test)*.* The bound makes the minimum usable window width scale as $w_{\min} \propto R^2/\sigma^2$. Section 7 tests this by varying the mode separation at fixed σ.

Table 1: Isolation within one architecture family (GMM, $d = 16$, $K = 8$, three seeds). Only the role of the score changes across rows. Mode coverage is out of 8.

Arm	Modes covered	Mode-weight L_1	Params
a-additive	$8, 8, 8$	$0.803,\ 0.796,\ 0.799$	156k
b-noscore	$2, 2, 2$	$1.562,\ 1.562,\ 1.562$	148k
c-feature	$8, 8, 8$	$0.797,\ 0.790,\ 0.794$	152k
d-noscore-big	$2, 2, 2$	$1.562,\ 1.562,\ 1.562$	2,412k
e-covinit	$3, 3, 3$	$1.232,\ 1.233,\ 1.235$	148k

Table 2: Replication on Many-Well ($d = 16$, 32 modes, two seeds). Mode coverage reproduces the GMM pattern exactly. Note that energy distance *inverts* the ranking and that mode-weight error does not separate the arms at all (§5).

Arm	Modes covered (of 32)	Mode-weight L_1	Energy distance
a-additive	$32, 32$	$1.119,\ 1.122$	$0.719,\ 0.793$
b-noscore	$1, 1$	$1.142,\ 1.142$	$0.267,\ 0.254$
c-feature	$32, 32$	$1.147,\ 1.151$	$0.763,\ 0.844$
d-noscore-big	$1, 1$	$1.142,\ 1.142$	$0.265,\ 0.253$
e-covinit	$1, 1$	$1.142,\ 1.142$	$0.204,\ 0.196$

5 THE SCORE TERM IS CAUSAL, BUT NOT FOR THE REASON USUALLY GIVEN

Table 1 reports the five arms on the GMM over three seeds.

Four readings follow, and all four hold on every seed. Deleting the score term takes coverage from $8/8$ to $2/8$ with the architecture family, objective, optimiser and budget all fixed, which confirms the conclusion of He et al. (2025) at matched parameter counts and on a target whose modes can be counted exactly. Arm c, which simply concatenates $\nabla \log \pi(x)$ to the network input, then matches a on every metric: whatever the score provides survives being handed over as an ordinary feature and need not appear as an additive Langevin drift, which refines the claim that the score "cannot be replaced by alternative target information, such as $\log p_{\text{target}}$"—the experiment behind that sentence supplied $\log \pi$, the value, and supplying the *gradient* as a feature is sufficient. Arm d, with $16\times$ the parameters of b, is indistinguishable from it to four decimals. And arm e, started from a distribution that already covers every mode, still ends at $3/8$, agreeing with their distillation experiment: a favourable initialisation does not rescue a score-free sampler.

It is worth being precise about what the capacity control does and does not establish. For a fixed target the optimal drift is a function of (x, t) alone and a large enough network can represent it; the score-free arm is not blind to the target either, since its loss still contains $-\log \tilde{\pi}(x_1)$. So arm d—which plays the role their enlarged five-layer MLP plays, here at a fixed budget against a matched small baseline—shows that *increasing capacity by $16\times$ does not compensate for the absence of explicit score access under our training recipe and budget*—not that the obstruction is information-theoretic. Optimisation difficulty, sample complexity, a training horizon that is too short, credit assignment through 50 SDE steps and the inductive bias of a privileged input feature all remain live; §6.4 rules out the training-horizon explanation, and we leave the rest open.

A natural worry is that these numbers describe the GMM's particular geometry, in which the modes sit far from where trajectories begin. Table 2 repeats the study on Many-Well, whose modes instead lie close to the origin along coordinate axes.

Mode coverage reproduces the GMM pattern arm for arm (Table 2): a-additive and c-feature place more than 1% of their empirical mass in all 32 basins on both seeds, while b-noscore, d-noscore-big and e-covinit do so in one. The isolation is therefore not an artefact of the GMM's geometry.

Two caveats belong here rather than in a footnote. First, the thresholded count in Table 2 is a *basin-discovery* statistic—a mode-collapse diagnostic—not target-calibrated recall. Many-Well's weights

are extremely skewed ($w_{\text{pos}} = 0.844$, so the dominant mode carries $0.844^5 \approx 0.43$ and the smallest $\approx 10^{-4}$), and only 16 of the 32 true modes carry over 1% of the target's mass—holding 97.05% of it—so exact target samples themselves score $16/32$ under the fixed rule. Reaching every basin thus does not mean matching the target; the arms that do populate them close to uniformly. Re-running with the target-aware criterion $\hat{w}_k \geq c\, w_k^*$ leaves the ordering unchanged on both seeds: across $c = 0.1/0.25/0.5$, a-additive covers $32/31/31$ basins, c-feature $32/31/30$–31 and win$[0.50, 1)$ $32/31/31$, whereas every score-free arm covers exactly one—so what replicates is *basin discovery* under either criterion, not posterior accuracy; the GMM's mode-weight numbers should not be extrapolated. Second, energy distance ranks e-covinit ($1/32$ modes) best and a-additive ($32/32$) worst, and mode-weight error does not separate them either: both aggregate metrics are blind to exactly the failure they are usually deployed to detect, a divergence we saw three times. This supports the concern of Blessing et al. (2024) about standard evaluation protocols and argues for reporting explicit mode counts wherever the target permits; Appendix F gives the mechanism.

6 When is the score needed?

The isolation says the score is necessary. It does not say *when* along the trajectory. Because the score enters through a time-dependent gate, we can switch it off outside a chosen interval and leave everything else untouched. We denote by win$[\tau_{\text{lo}}, \tau_{\text{hi}})$ the arm in which the score term is active only for t in that interval.

We vary the window rather than the diffusion coefficient σ on purpose. Lowering σ would also change whether the uncontrolled reference can reach the modes at all, confounding any conclusion with a reachability effect; gating in time leaves the reference process untouched.

6.1 Half-trajectory gating

The first split is the obvious one. The two halves of the trajectory are not interchangeable at all: with the score on $[0, 0.5)$ the sampler covers $2/8$ modes with mode-weight error $1.5618,\ 1.5616$, against $2/8$ and $1.5617,\ 1.5615$ for having no score whatsoever, while $[0.5, 1)$ recovers all eight (Table 3).

Score on the first half is not merely less useful: it agrees with the score-free arm to four decimal places on both seeds. Note also that this arm spends exactly as many target-gradient evaluations over the first half as a-additive does, and gains nothing from them—so the advantage of the score-bearing arms is not simply a matter of querying the oracle more often.

6.2 Sweeping the gate

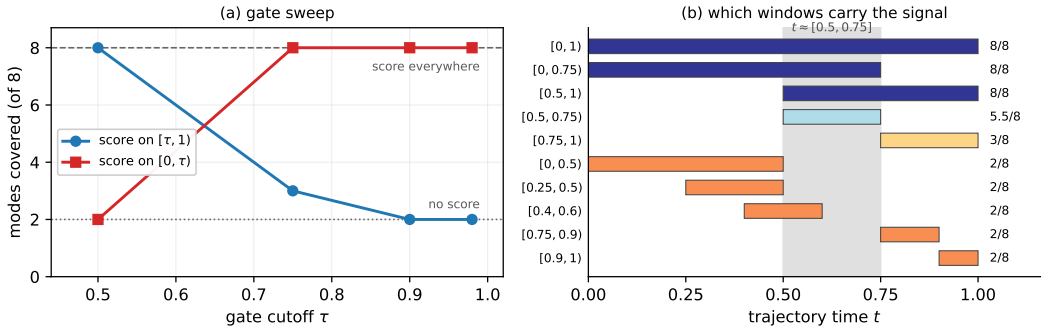

Figure 2: **The score is required on a window, not at a point in time.** (a) Coverage as the gate cutoff τ moves, for the score on $[\tau, 1)$ (blue) and $[0, \tau)$ (red); dashed and dotted lines mark the score-everywhere and score-free arms. The curves cross, so neither "early" nor "late" is the right description. (b) Every window we ran, shaded by coverage: those overlapping $t \approx [0.5, 0.75]$ (grey band) and long enough succeed. Means over two seeds.

Table 3 and Figure 2a sweep the gate position in both directions. The pattern is not "later is better".

Table 3: Gate sweep (GMM, two seeds). Left: score active on $[\tau, 1)$. Right: score active on $[0, \tau)$. Score confined to the end of the trajectory is as useless as no score at all.

Arm	Modes	L_1	Arm	Modes	L_1
$\mathsf{win}[0.50, 1)$	$8, 8$	$0.939,\ 0.944$	$\mathsf{win}[0, 0.50)$	$2, 2$	$1.562,\ 1.562$
$\mathsf{win}[0.75, 1)$	$3, 3$	$1.228,\ 1.244$	$\mathsf{win}[0, 0.75)$	$8, 8$	$\mathbf{0.741},\ \mathbf{0.753}$
$\mathsf{win}[0.90, 1)$	$2, 2$	$1.561,\ 1.561$	$\mathsf{win}[0, 0.90)$	$8, 8$	$0.802,\ 0.792$
$\mathsf{win}[0.98, 1)$	$2, 2$	$1.562,\ 1.562$	$\mathsf{win}[0, 0.98)$	$8, 8$	$0.803,\ 0.794$

Table 4: Short-window isolation (GMM, two seeds). Only the window around $[0.5, 0.75]$ carries any signal, and even it does not reach full coverage. Every other short window is bit-for-bit the score-free baseline.

Score active on	Width	Modes covered	Mode-weight L_1
$[0.50, 0.75)$	0.25	$5, 6$	$1.076,\ 1.074$
$[0.55, 0.70)$	0.15	$2, 2$	$1.562,\ 1.562$
$[0.60, 0.75)$	0.15	$2, 2$	$1.561,\ 1.561$
$[0.40, 0.60)$	0.20	$2, 2$	$1.562,\ 1.562$
$[0.25, 0.50)$	0.25	$2, 2$	$1.562,\ 1.562$
$[0.75, 0.90)$	0.15	$2, 2$	$1.561,\ 1.562$
$[0, 1)$ (full)	1.00	$8, 8$	$0.803,\ 0.796$
none	—	$2, 2$	$1.562,\ 1.562$

Restricting the score to the final 10% or 2% of the trajectory returns performance to the score-free baseline. So the score is not simply needed "at the end". Reading the two columns together, among the windows tested every one that succeeds contains $t \approx [0.5, 0.75]$ and every one that fails excludes it. The earlier half-trajectory result is explained: $[0.5, 1)$ works because it *contains* that region, not because it is late.

Incidentally, $\mathsf{win}[0, 0.75)$ attains the best mode-weight error of any arm we ran (0.741 versus 0.803 for the score everywhere): leaving the score on near terminal time appears mildly harmful.

6.3 Isolating the interval

If $[0.5, 0.75]$ were the whole story, that window alone should suffice. Table 4 shows it does not.

Two conditions are jointly necessary. The window must *cover* the region around $t \approx [0.5, 0.75]$—no other short window anywhere on the trajectory has any measurable effect—and it must be *long enough*: that region on its own reaches 5–6 of 8 modes, not 8. The sharpest contrast is between $[0, 0.5)$ and $[0.5, 1)$: identical width, opposite outcome. Duration alone does not explain the phenomenon, and neither does position.

The empirical statement we are prepared to defend is therefore narrow: *in this sampler, on these targets and across the windows tested here, the score must be available over a contiguous stretch of trajectory time that includes $t \approx [0.5, 0.75]$ and spans a substantial fraction of the trajectory.*

6.4 Two controls

Two objections are worth closing off. First, all gating so far multiplied the *additive* score term, so the window structure might be an artefact of that parameterisation. Gating the score as an input *feature* instead reproduces it (Table 5), with one exception we flag: the short window $[0.5, 0.75)$, worth five to six modes additively, is worth two as a feature. The coarse pattern is a property of score availability; the partial rescue by a short well-placed window is not. Second, retraining with $5\times$ the updates leaves the arms deep in the failing regime exactly where they were, all still at $2/8$;

Table 5: Time-gating the score as an *input feature* rather than as an additive term (GMM, two seeds). The full-window arm reduces exactly to c-feature, as it must.

Feature-score available on	Modes covered	Mode-weight L_1
$[0, 1)$ (reduces to c-feature)	$8, 8$	$0.797,\ 0.790$
$[0.5, 1)$	$8, 8$	$1.023,\ 0.998$
$[0, 0.5)$	$2, 2$	$1.562,\ 1.562$
$[0.9, 1)$	$2, 2$	$1.561,\ 1.562$
$[0.5, 0.75)$	$2, 2$	$1.562,\ 1.562$
none (b-noscore)	$2, 2$	$1.562,\ 1.562$

Table 6: Mode coverage by mode separation and window width (GMM, $d = 16$, $\sigma = 2$), per seed. $w_{\min}$ is bracketed using a refined width grid with four seeds (Table 7) under a stated criterion, mean coverage $\geq 7.5/8$. Radius 3 is degenerate—the score-free arm already solves it—and is excluded from the scaling read-out.

		score on $[1 - w, 1)$					w=1	none	$w_{\min}$
Radius	R	0.02	0.10	0.25	0.50	0.75			
3 *(degen.)*	4.2	$8, 8$	$8, 8$	$8, 8$	$8, 8$	$8, 8$	$8, 8$	$8, 8$	—
4.5 *(partial)*	6.4	$4, 4$	$4, 5$	$8, 8$	$8, 8$	$8, 8$	$8, 8$	$3, 3$	$(0.15, 0.20]$
6	8.5	$2, 2$	$2, 2$	$3, 3$	$8, 8$	$8, 8$	$8, 8$	$2, 2$	$(0.40, 0.50]$
7.5	10.6	$1, 1$	$1, 1$	$2, 2$	$4, 5$	$8, 8$	$8, 8$	$1, 1$	$(0.65, 0.75]$
9	12.7	$1, 1$	$1, 1$	$1, 1$	$2, 2$	$4, 4$	$5, 5$	$1, 1$	> 1

only win$[0.75, 1)$, the arm on the threshold of §7, improves, from $3/8$ to $5/8$ (Appendix G). The obstruction is not the training horizon.

7 Testing the Scaling Prediction

Proposition 2 makes a prediction that does not depend on any fitted quantity: the minimum usable window width should grow as $w_{\min} \propto R^2$ at fixed σ. We test it by moving the modes. Keeping everything else fixed, we place the eight components on spheres of several radii and sweep the window $[1 - w, 1)$.

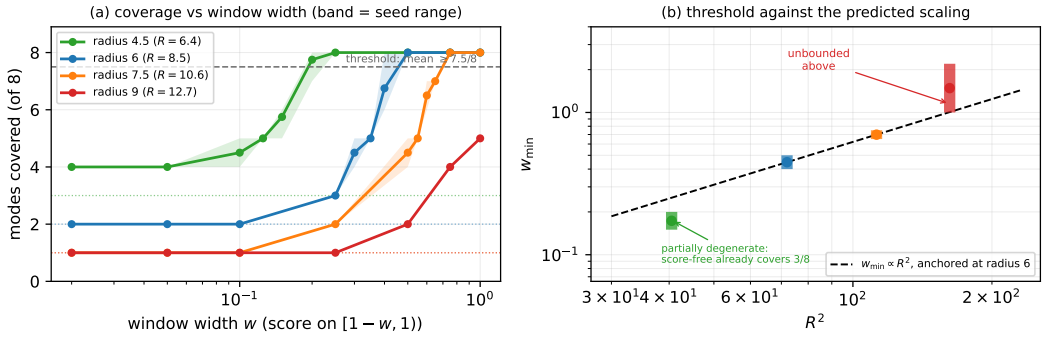

Figure 3: **Testing the prediction of Proposition 2.** (a) Coverage against window width for four separations; bands give the seed range, dotted lines each setting's score-free baseline. The transition moves right as the modes separate and at radius 9 has left the trajectory. (b) The bracketed threshold against R^2, anchored at radius 6. Radius 9 is open above, since even $w = 1$ is insufficient; radius 4.5 falls below, and is the one partially degenerate setting.

Figure 3 shows the movement; per-seed counts are in Table 6. Note first that Proposition 2 is a *necessary* condition: redistribution costs at least $pR^2/(2\sigma^2 w)$, so success becomes impossible once that exceeds the available terminal reward, but not necessarily possible when it does not. The prediction is one-sided—a constraint on where the transition can sit, not a formula for it. At radius 6 the transition sits between $w = 0.25$ (three modes) and $w = 0.5$ (all eight). Increasing the radius to 9 multiplies R^2 by 2.25, so the bound predicts $w_{\min}$ should move into the range 0.56–1.13; that is, it should approach or exceed the entire trajectory. This is what happens: at radius 7.5 the transition has moved to $(0.50, 0.75]$, and at radius 9 the window $w = 0.5$ recovers two modes, $w = 0.75$ recovers four, and even the score-everywhere arm ($w = 1$) reaches only five of eight. This is the qualitative consequence the bound does predict: a $1.5\times$ increase in separation is enough to exhaust the whole trajectory.

Refining the grid and raising the seed count to four inside each transition sharpens this. Fixing the threshold by a stated criterion—the smallest w at which mean coverage reaches 7.5 of 8—brackets it at $(0.40, 0.50]$, $(0.65, 0.75]$ and > 1 for radii 6, 7.5 and 9. Using the *midpoints* of those brackets gives $w_{\min}/R^2 \approx 0.0062$ at radii 6 and 7.5, while radius 9 yields the corresponding lower bound; the normalised intervals overlap. Radius 4.5 sits at 0.0043 and is the one partially degenerate setting here, its score-free arm already recovering three of eight modes against two and one at larger radii—an ordering that follows from the baselines, not the fit. We read this as *consistent with* R^2 scaling, not a measurement of the exponent: four radii cannot separate R^2 from $R^{1.8}$, and the constancy rests on three points, one of them one-sided.

Two caveats. Radius 3 is degenerate—there the score-free arm already covers all eight modes, and attains a *better* mode-weight error (0.23) than the score-everywhere arm (0.37)—so there is nothing to repair and no threshold to locate; we report it as the boundary of the regime in which the question is meaningful rather than as a confirmation. And the separations we ran establish the direction and magnitude of the scaling, not its exponent.

8 On the mechanism: one half explained, one half open

Proposition 2 is consistent with the late-window failures; the early window remains unexplained. We record two dead ends, since they constrain a full account.

Hypothesis 1 (distance). Early on x_t is near the origin, far from the modes, where $\nabla \log \pi$ says little about *which* mode a sample will end in—it points at the nearest or heaviest mass regardless. This predicts that a target whose modes lie close to the start should benefit from an early score. Many-Well is that target, with modes at $|x_i| \approx 1.7$ rather than on a radius-6 sphere, and the prediction fails: its first-half arm still covers 1 of 32 modes, exactly as with no score.

Hypothesis 2 (dilution). Early control is washed out by the diffusion, so only control near $t = 1$ survives into the terminal distribution the loss sees. This predicts a terminal score should be the *most* valuable; it is the least (Table 3). Proposition 2 shows the intuition is backwards: the binding constraint at the end is not that control is forgotten but that there is no time left to pay for it.

Appendix B records the closest thing to a candidate explanation we know of.

The asymmetry is real. Late windows fail for a reason we can write down; early windows have everything the bound cares about—the longest remaining time, the largest transport budget, the same oracle calls—and fail anyway, matching the score-free baseline to four decimals on two geometries. An account of the early side must also explain the minimum-duration requirement that persists when the window is well placed (Table 4). We do not have one.

9 Conclusion

The claim that neural samplers need an explicit target-score term survives a properly controlled test; the reason usually attached to it does not. The score is causal within a fixed family, yet works as well as a plain input feature, and neither $16\times$ the parameters, a mode-covering start nor $5\times$ the budget substitutes for it. In our controlled setting the requirement is a schedule, not a switch: the score must be present over a window containing a mid-trajectory interval, and is worthless elsewhere—including at terminal time, where the optimal control provably equals it.

Reproducibility statement

All targets admit exact ground truth: the Gaussian mixture by construction, and Many-Well through a one-dimensional quadrature for the normalising constant and mode weights and a one-dimensional inverse CDF for exact samples. Every arm shares the objective, discretisation, optimiser, batch size and number of updates; the arms differ only in the parameterisation of the drift. Each table reports per-seed values rather than means, so the reader can see that the seed-to-seed spread is far smaller than the effects discussed. Architecture, optimiser and quadrature details are given in Appendix H.

AI use statement

Generative AI was used for language editing, code implementation and debugging of the experiments, and exploratory analysis of the resulting data. The authors take full responsibility for the final content of this paper, including the design of the experiments, the statements of the propositions and their proofs, and any text produced with the assistance of generative AI.

References

Michael S. Albergo and Eric Vanden-Eijnden. NETS: A non-equilibrium transport sampler. In *International Conference on Machine Learning*, 2025.

Denis Blessing, Xiaogang Jia, Johannes Esslinger, Francisco Vargas, and Gerhard Neumann. Beyond ELBOs: A large-scale evaluation of variational methods for sampling. In *International Conference on Machine Learning*, 2024.

Denis Blessing, Xiaogang Jia, and Gerhard Neumann. End-to-end learning of gaussian mixture priors for diffusion sampler. In *International Conference on Learning Representations*, 2025. arXiv:2503.00524.

Diana Cai, Robert M. Gower, David M. Blei, and Lawrence K. Saul. Fisher meets feynman: Score-based variational inference with a product of experts. In *Advances in Neural Information Processing Systems*, 2025. arXiv:2510.21598.

Jorge Carrasco-Pollo, Floor Eijkelboom, and Jan-Willem van de Meent. Pawsterior: Variational flow matching for structured simulation-based inference. In *Proceedings of GRaM: the Second Edition of the Workshop on Geometry-grounded Representation Learning and Generative Modeling*, volume 326 of *Proceedings of Machine Learning Research*, pp. 75–87, 2026. arXiv:2602.13813.

Francisco M. Castro-Macías, Pablo Morales-Álvarez, Saifuddin Syed, Daniel Hernández-Lobato, Rafael Molina, and José Miguel Hernández-Lobato. Conditional diffusion sampling. 2026. arXiv:2605.04013.

Junhua Chen, Lorenz Richter, Julius Berner, Denis Blessing, Gerhard Neumann, and Anima Anandkumar. Sequential controlled langevin diffusions. In *International Conference on Learning Representations*, 2025. arXiv:2412.07081.

Kamélia Daudel, Minh-Ngoc Tran, and Cheng Zhang. Importance weighted variational inference without the reparameterization trick. 2026. arXiv:2602.01412.

Andrew Dennehy, Ramchandran Muthukumar, Rebecca Willett, and Nisha Chandramoorthy. Diffusion models recover accurate mixture weights despite score function insensitivity. *arXiv preprint arXiv:2607.15485*, 2026.

Floor Eijkelboom, Grigory Bartosh, Christian A. Naesseth, Max Welling, and Jan-Willem van de Meent. Variational flow matching for graph generation. In *Advances in Neural Information Processing Systems*, 2024.

Floor Eijkelboom, Heiko Zimmermann, Sharvaree Vadgama, Erik J Bekkers, Max Welling, Christian A. Naesseth, and Jan-Willem van de Meent. Controlled generation with equivariant variational flow matching. In *International Conference on Machine Learning*, 2025.

Wei Guo, Jaemoo Choi, Yuchen Zhu, Molei Tao, and Yongxin Chen. Proximal diffusion neural sampler. In *International Conference on Learning Representations*, 2026.

Aaron Havens, Benjamin Kurt Miller, Bing Yan, Carles Domingo-Enrich, Anuroop Sriram, Daniel Levine, Brandon Wood, Bin Hu, Brandon Amos, Brian Karrer, Xiang Fu, Guan-Horng Liu, and Ricky T. Q. Chen. Adjoint sampling: Highly scalable diffusion samplers via adjoint matching. In *International Conference on Machine Learning*, 2025.

Aaron Havens, Brian Karrer, and Neta Shaul. Flow sampling: Learning to sample from unnormalized densities via denoising conditional processes. *arXiv preprint arXiv:2605.03984*, 2026.

Jiajun He, Yuanqi Du, Francisco Vargas, Dinghuai Zhang, Shreyas Padhy, RuiKang OuYang, Carla Gomes, and José Miguel Hernández-Lobato. No trick, no treat: Pursuits and challenges towards simulation-free training of neural samplers. In *ICLR Workshop on Frontiers in Probabilistic Inference*, 2025. arXiv:2502.06685.

Minkyu Kim, Kiyoung Seong, Dongyeop Woo, Sungsoo Ahn, and Minsu Kim. On scalable and efficient training of diffusion samplers. In *Advances in Neural Information Processing Systems*, 2025.

Minyoung Kim. Large-scale score-based variational posterior inference for bayesian deep neural networks. 2026. arXiv:2602.05873.

Joohwan Ko and Justin Domke. Model-informed flows for bayesian inference. In *Advances in Neural Information Processing Systems*, 2025.

Chengkun Li, Bobby Huggins, Petrus Mikkola, and Luigi Acerbi. Normalizing flow regression for bayesian inference with offline likelihood evaluations. In *Symposium on Advances in Approximate Bayesian Inference*, 2025.

Soon Hoe Lim. On the hidden biases of flow matching samplers. *arXiv preprint arXiv:2512.16768*, 2025.

Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin Kurt Miller, and Ricky T. Q. Chen. Adjoint schrödinger bridge sampler. In *Advances in Neural Information Processing Systems*, 2025.

Chirag Modi, Diana Cai, and Lawrence K. Saul. Batch, match, and patch: Low-rank approximations for score-based variational inference. In *International Conference on Artificial Intelligence and Statistics*, 2025.

Badr Moufad, Yazid Janati, Lisa Bedin, Alain Durmus, Randal Douc, Eric Moulines, and Jimmy Olsson. Variational diffusion posterior sampling with midpoint guidance. In *International Conference on Learning Representations*, 2025.

Dai Hai Nguyen, Tetsuya Sakurai, and Hiroshi Mamitsuka. Wasserstein gradient flow over variational parameter space for variational inference. In *International Conference on Artificial Intelligence and Statistics*, 2025.

Wasu Top Piriyakulkij, Yingheng Wang, and Volodymyr Kuleshov. Denoising diffusion variational inference: Diffusion models as expressive variational posteriors. In *AAAI Conference on Artificial Intelligence*, 2025.

Juehang Qin, Shixiao Liang, and Christopher Tunnell. FlowVAT: Normalizing flow variational inference with affine-invariant tempering. 2025. arXiv:2505.10466.

Severi Rissanen, RuiKang OuYang, Jiajun He, Wenlin Chen, Markus Heinonen, Arno Solin, and José Miguel Hernández-Lobato. Progressive tempering sampler with diffusion. In *International Conference on Machine Learning*, 2025.

Sebastian Sanokowski, Wilhelm Berghammer, Haoyu Wang, Martin Ennemoser, Sepp Hochreiter, and Sebastian Lehner. Scalable discrete diffusion samplers: Combinatorial optimization and statistical physics. In *International Conference on Learning Representations*, 2025.

Francisco Vargas, Will Grathwohl, and Arnaud Doucet. Denoising diffusion samplers. In *International Conference on Learning Representations*, 2023.

Christopher von Klitzing, Denis Blessing, Henrik Schopmans, Pascal Friederich, and Gerhard Neumann. Learning boltzmann generators via constrained mass transport. In *International Conference on Learning Representations*, 2026. arXiv:2510.18460.

Chenguang Wang, Xiaoyu Zhang, Kaiyuan Cui, Weichen Zhao, Yongtao Guan, and Tianshu Yu. Importance weighted score matching for diffusion samplers with enhanced mode coverage. *arXiv preprint arXiv:2505.19431*, 2025.

Dongyeop Woo, Minsu Kim, Minkyu Kim, Kiyoung Seong, and Sungsoo Ahn. Energy-based generator matching: A neural sampler for general state space. In *Advances in Neural Information Processing Systems*, 2025. arXiv:2505.19646.

Luhuan Wu, Yi Han, Christian A. Naesseth, and John P. Cunningham. Reverse diffusion sequential monte carlo samplers. In *Advances in Neural Information Processing Systems*, 2025.

Jian Xu, Delu Zeng, Qibin Zhao, and John Paisley. Diffusion bridge variational inference for deep gaussian processes. In *International Conference on Learning Representations*, 2026. arXiv:2509.19078.

Percy S. Zhai, So Won Jeong, and Veronika Ročková. Conditional flow matching for bayesian posterior inference. 2025. arXiv:2510.09534.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: A stochastic control approach for sampling. In *International Conference on Learning Representations*, 2022.

A LIMITATIONS

Scope. One sampler (controlled-SDE/PIS with a free-form drift) and two target families, both $d = 16$ and both with modes gradient search can locate; rugged landscapes are untested. We did not reproduce the original DDS-versus-NF-DDS comparison but built an analogous setting, so our claim is that the score term is causal *in this family* and that the published evidence did not establish it in general. The Many-Well replication holds for mode *coverage* only, and the scaling study varies separation but not σ, which would also change whether the reference process reaches the modes.

Budgets. The horizon control tests $5\times$ the updates, not asymptotically long training, and the score-bearing arms evaluate $\nabla \log \pi$ every rollout step while the score-free arms do not—we ask why the score helps, not whether it is efficient. Two internal controls limit that objection: win$[0, 0.5)$ spends half the score-bearing arm's oracle calls and gains nothing, and d-noscore-big spends $16\times$ the parameters and gains nothing. The mechanism remains half open.

B A LEAD WE HAVE NOT FOLLOWED

A lead we have not followed. Dennehy et al. (2026) trace a related puzzle: the target score is insensitive to mixture weights when components are well separated, yet diffusion models recover the weights anyway, which they trace to scores at *intermediate* noise levels being informative when the noiseless score is not. Proposition 1 exhibits our value function as a Gaussian convolution at bandwidth $\sigma^2(1 - t)$, so intermediate time corresponds to intermediate smoothing and the two statements rhyme. They are not the same—our arms receive the raw $\nabla \log \pi(x)$, and what varies with t is where it is evaluated rather than how much it is smoothed—but an account built on which scales carry mode-weight information seems the most promising direction available.

C REFINED SCALING GRID

Widths inside each measured transition, four seeds each. These fix the $w_{\min}$ brackets quoted in Table 6.

Table 7: Refined window grid, per seed.

Radius	w	Modes covered
4.5	0.125	5, 5, 5, 5
4.5	0.15	5, 6, 6, 6
4.5	0.2	8, 8, 7, 8
6	0.3	4, 5, 4, 5
6	0.35	5, 5, 5, 5
6	0.4	6, 6, 8, 7
7.5	0.55	5, 5, 5, 5
7.5	0.6	7, 7, 6, 6
7.5	0.65	7, 7, 7, 7

D PROOFS

Throughout, $\tilde{\pi}$ denotes the unnormalised target, π its normalisation, and $p_{\mathrm{ref}} = \mathcal{N}(0, \sigma^2 I)$ the terminal law of the uncontrolled process $\mathrm{d}x_t = \sigma\,\mathrm{d}W_t$ started at $x_0 = 0$. Since $\tilde{\pi}$ and π differ by a constant factor, $\nabla \log \tilde{\pi} = \nabla \log \pi$ everywhere below.

PROOF OF PROPOSITION 1

The identification $u^\star(x,t) = \sigma^2 \nabla_x \log \Phi(x,t)$ with $\Phi(x,t) = \mathbb{E}[(\tilde{\pi}/p_{\mathrm{ref}})(x_1) \mid x_t = x]$ is the standard verification result for the stochastic-control (Schrödinger half-bridge) formulation of the objective equation 3; see Zhang & Chen (2022). We prove the terminal limit.

Under the uncontrolled process, conditioning on $x_t = x$ gives $x_1 \sim \mathcal{N}\big(x, \sigma^2(1-t)I\big)$, so

$$\Phi(x,t) \;=\; \int \mathcal{N}\big(y; x, \sigma^2(1-t)I\big)\, \frac{\tilde{\pi}(y)}{p_{\mathrm{ref}}(y)}\, \mathrm{d}y. \tag{7}$$

This exhibits $\Phi(\cdot, t)$ as the Gaussian convolution of $\tilde{\pi}/p_{\mathrm{ref}}$ at bandwidth $\sigma^2(1-t)$. For the targets considered here $\tilde{\pi}/p_{\mathrm{ref}}$ is continuous and of at most Gaussian growth, so the convolution and its gradient converge locally uniformly as the bandwidth vanishes:

$$\Phi(x,t) \;\longrightarrow\; \frac{\tilde{\pi}(x)}{p_{\mathrm{ref}}(x)}, \qquad \nabla_x \log \Phi(x,t) \;\longrightarrow\; \nabla \log \tilde{\pi}(x) - \nabla \log p_{\mathrm{ref}}(x) \qquad (t \to 1^-). \tag{8}$$

Since $\nabla \log p_{\mathrm{ref}}(x) = -x/\sigma^2$, the right-hand side equals $\nabla \log \pi(x) + x/\sigma^2$. Multiplying by σ^2,

$$\lim_{t \to 1^-} u^\star(x,t) \;=\; \sigma^2 \nabla \log \pi(x) + x. \tag{9}$$

Comparing with the parameterisation $u_\theta = \mathrm{NN}_1(x,t) + \mathrm{NN}_2(t) \odot \nabla \log \pi(x)$ of equation 1, the limit is attained exactly by $\mathrm{NN}_2(t) \to \sigma^2 \mathbf{1}$ and $\mathrm{NN}_1(x,t) \to x$. □

PROOF OF PROPOSITION 2

Write $[\tau_{\mathrm{lo}}, \tau_{\mathrm{hi}})$ for the window, w for its width, $\Delta = \int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} u(x_t, t)\,\mathrm{d}t$ for the displacement the control contributes along a trajectory, and $C = \mathbb{E}\int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} \|u\|^2/(2\sigma^2)\,\mathrm{d}t$.

Fix a trajectory. Applying the Cauchy–Schwarz inequality in t to each coordinate of the vector-valued integral and summing,

$$\|\Delta\|^2 \;=\; \Big\| \int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} u(x_t,t)\,\mathrm{d}t \Big\|^2 \;\le\; (\tau_{\mathrm{hi}} - \tau_{\mathrm{lo}}) \int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} \|u(x_t,t)\|^2\,\mathrm{d}t \;=\; w \int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} \|u(x_t,t)\|^2\,\mathrm{d}t. \tag{10}$$

Equation 10 holds pathwise, so taking expectations preserves it:

$$\mathbb{E}\big[\|\Delta\|^2\big] \;\le\; w\,\mathbb{E}\int_{\tau_{\mathrm{lo}}}^{\tau_{\mathrm{hi}}} \|u\|^2\,\mathrm{d}t \;=\; w \cdot 2\sigma^2 C, \tag{11}$$

which proves Equation 5. Note that no property of the Brownian increment is used: the bound constrains the control contribution Δ alone, and holds conditionally on the noise realisation.

For the second claim, suppose a set of trajectories of probability at least p must have $\|\Delta\| \geq R$. Then $\mathbb{E}[\|\Delta\|^2] \geq pR^2$, and combining with the display above gives $pR^2 \leq 2\sigma^2 Cw$, i.e. $C \geq pR^2/(2\sigma^2 w)$. □

Remark 2 (What the diffusion contributes). The bound constrains the displacement produced by the *control*. Over a window of width w the Brownian part contributes $\mathbb{E}\|\sigma(W_{\tau_{\text{hi}}} - W_{\tau_{\text{lo}}})\|^2 = \sigma^2 d\, w$, i.e. an RMS norm of $\sigma\sqrt{dw}$. For our $\sigma = 2$ and $d = 16$ that is 5.66 at $w = 0.5$ and 2.53 at $w = 0.1$, which is *comparable* to the inter-mode distance $R \approx 8.5$, not negligible against it. (An earlier version of this remark compared the per-coordinate standard deviation $\sigma\sqrt{w}$ against a Euclidean distance in $d = 16$; the correct comparison carries the factor $\sqrt{d}$.)

The diffusion therefore cannot be dismissed on magnitude. What it cannot do is *direct* the mass: it is isotropic and carries no information about where the modes are, so it produces undirected spread rather than the targeted reallocation that correct mode weights require—and the score-free arm, which experiences exactly the same diffusion, does not find the modes. The bound should be read as constraining the directed component of transport, which is the component that has to do the work.

E Extended related work

Neural samplers for unnormalised targets. The modern line begins with stochastic-control formulations (Zhang & Chen, 2022) and denoising-diffusion samplers (Vargas et al., 2023), and has since split into several families: non-equilibrium transport with learned drift and unbiased weights (Albergo & Vanden-Eijnden, 2025); adjoint- and matching-based objectives that avoid importance weighting during training (Havens et al., 2025; Liu et al., 2025); proximal-point schemes on path measures (Guo et al., 2026); simulation-free regression onto a denoising drift (Havens et al., 2026); closed-form conditional interpolants combined with parallel tempering (Castro-Macías et al., 2026); and discrete state spaces (Sanokowski et al., 2025). A parallel line corrects a learned proposal with exact machinery, either by interleaving resampling with learned transitions (Chen et al., 2025) or by treating the diffusion sampler as an SMC proposal (Wu et al., 2025). Almost all of these place the target score inside the drift, in the form of equation 1; that shared choice is what this paper examines.

Mode collapse and what is done about it. Reverse-KL training is mode-seeking, and the standard remedies act on the initialisation or the temperature: learned Gaussian-mixture priors (Blessing et al., 2025), affine-invariant tempering (Qin et al., 2025), progressive tempering that distils across temperatures (Rissanen et al., 2025), and off-policy exploration driven by a novelty energy (Kim et al., 2025). Our e-covinit arm is the minimal version of the first and shows a mode-covering start is not sufficient. Blessing et al. (2024) argue the field's evaluation protocols are poorly suited to detecting collapse at all, which our Many-Well results reinforce.

Score-based variational inference. A separate tradition matches scores rather than maximising an ELBO, with Gaussian and low-rank families (Modi et al., 2025), products of heavy-tailed experts (Cai et al., 2025), gradient flows over variational parameters (Nguyen et al., 2025), and stochastic-gradient variants for Bayesian neural networks (Kim, 2026). That literature uses the score as the *objective*; the samplers above use it as an architectural *input*. Our c-feature arm suggests the distinction matters less than the mere availability of $\nabla \log \pi$.

Diffusion and flow models as variational posteriors. A large adjacent literature builds expressive posteriors from diffusions (Piriyakulkij et al., 2025), observation-conditioned bridges (Xu et al., 2026), flow-matching formulations of variational inference (Eijkelboom et al., 2024; 2025; Carrasco-Pollo et al., 2026), conditional flow matching (Zhai et al., 2025), model-informed architectures (Ko & Domke, 2025), offline surrogate regression (Li et al., 2025), improved importance-weighted estimators (Daudel et al., 2026), and diffusion priors for inverse problems (Moufad et al., 2025). These target latent-variable models or learned priors; our setting is the unnormalised target, where the score is available by construction and using it is a design decision rather than a necessity.

Table 8: Training-horizon control (GMM, one seed). Five times the updates leaves every arm where it was.

	1500 updates		7500 updates	
Arm	Modes	L_1	Modes	L_1
a-additive $[0, 1)$	$8, 8$	0.799	8	0.861
win$[0.50, 1)$	$8, 8$	0.941	8	0.891
win$[0.75, 1)$	$3, 3$	1.236	5	1.046
win$[0.90, 1)$	$2, 2$	1.561	2	1.562
win$[0.98, 1)$	$2, 2$	1.562	2	1.562
win$[0, 0.50)$	$2, 2$	1.562	2	1.562
win$[0, 0.75)$	$8, 8$	0.747	8	0.741
win$[0, 0.90)$	$8, 8$	0.797	8	0.848
win$[0, 0.98)$	$8, 8$	0.798	8	0.865

F WHY THE AGGREGATE METRICS FAIL ON MANY-WELL

Table 2 contains a result worth isolating. On Many-Well, energy distance against exact reference samples ranks e-covinit, which covers 1 of 32 modes, at 0.204; it ranks a-additive, which covers all 32, at 0.719. The ordering is inverted. Mode-weight error does not help either: every arm lands in $[1.117, 1.151]$ whether it covers 1 mode or 32.

The reason is structural: Many-Well's modes are close together relative to the target's spread, so a sampler on one mode still produces roughly the right pairwise distances; and because the true weights are so skewed, covering all modes with nearly uniform weights is about as wrong in L_1 as covering one. Both aggregate metrics are insensitive to exactly the failure they are usually deployed to detect. We saw this divergence three times. It supports the concern of Blessing et al. (2024) that standard protocols are ill-suited to detecting collapse, and argues for reporting explicit mode counts wherever the target permits.

All gating so far multiplied the *additive* score term. Since §5 found that the additive structure is not what matters, the window phenomenon should survive when the score is supplied as an input feature and that feature is gated in time instead. Table 5 reports the test.

The main structure replicates exactly. Gating the feature to $[0, 1)$ reproduces c-feature to four decimals, as it must; $[0.5, 1)$ recovers all eight modes; and $[0, 0.5)$ and $[0.9, 1)$ fall back to the score-free baseline. So which half of the trajectory carries the score is a property of when the score *information* is available, not of the additive parameterisation.

One result does not replicate, and we flag it rather than bury it. The short window $[0.5, 0.75)$, which recovers five to six of eight modes in the additive form (Table 4), recovers only two in the feature form—the score-free value. The partial rescue by a short, well-placed window therefore *does* depend on the additive structure, even though the coarse half-trajectory pattern does not. We do not have an account of why the two parameterisations differ in this narrow regime and consider it part of the open half of the mechanism.

G TRAINING-HORIZON CONTROL

A natural objection to §5 is that the score-free and badly-windowed arms are merely slower to train. We re-ran the gate sweep at $5\times$ the number of updates (Table 8). The arms well inside the failing regime—win$[0.90, 1)$, win$[0.98, 1)$ and win$[0, 0.50)$—do not move at all: still $2/8$, with mode-weight errors agreeing to three decimals with the short run. The one arm that does improve is win$[0.75, 1)$, from $3/8$ to $5/8$, which is the arm sitting on the threshold identified in §7; a borderline case becoming somewhat better with five times the compute is what one expects near a transition, and it is not a rescue. The obstruction is therefore not the training horizon, though we note that we have tested $5\times$, not asymptotically long training.

H ADDITIONAL EXPERIMENTAL DETAILS

All arms use a three-layer GELU network of width 256 for NN_1 (≈ 150k parameters); d-noscore-big uses five layers of width 768 (≈ 2.4M). The gate NN_2 is a two-layer network of width 64 acting on the time embedding, with a zero-initialised final layer, so all score-bearing arms coincide with b-noscore at initialisation. The time embedding is 32-dimensional sinusoidal. The target score is norm-clipped at 100 before use, in every arm that uses it.

Optimisation uses Adam at 5×10^{-4} with cosine decay to zero over 1500 updates, batch size 512, gradient-norm clipping at 10. The SDE uses $T = 50$ Euler–Maruyama steps with $\sigma = 2$. Evaluation draws $20{,}480$ samples.

For the Many-Well target, the normalising constant of the one-dimensional factor $e^{-(x^4 - 6x^2 - 0.5x)}$ and the probability w_{pos} of the positive well are obtained by trapezoidal quadrature on $[-6, 6]$ with 2×10^5 points; the exact mode weights are then products of w_{pos} and $1 - w_{\mathrm{pos}}$ over the $m = 5$ well coordinates, and exact samples are drawn by inverse-CDF sampling on the same grid for the well coordinates and standard normal draws for the remaining $d - m$ coordinates.