CONSTRAINT-DRIVEN MEASURE TRANSFORMATION OF PRETRAINED DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Adapting a pretrained diffusion model to a target known only through expectation constraints requires solving for the tilt that reaches it. We solve for it by minimum relative entropy against the learned prior, giving the weighted Monte Carlo measure, and carry it into the frozen sampler through the Doob h-transform. Both are prior work. Because the reference is learned, it can be varied, and the construction becomes an instrument for a question central to derivative pricing: what does a finite set of market constraints determine about a path-dependent price? On Heston paths, 53 constraints span 97% of an Asian option's payoff variance but only 76% of a barrier's, and that unspanned variance sets, to within 2.4%, how fast the admissible price range opens with relative-entropy budget. On a real AAPL option surface, every prior we test (learned, parametric and Black-Scholes alike) misprices around a scheduled earnings date in the same direction, and the pipeline recovers the known earnings-jump signature from the quotes. Carrying the tilt into the sampler, which the identification results do not require, works only at mild tilts. The learned correction is accurate as a function, but the sampler does not realise the measure it implies.

1 INTRODUCTION

Pretrained diffusion models are increasingly adapted at inference time rather than retrained: a correction added to the learned score steers the sampler toward a new target. In the settings studied so far the target arrives with its own tilt, such as a reward, a classifier, a likelihood or a prescribed endpoint law. This paper studies a target that arrives with none: a measure specified only through finitely many expectation constraints, which admit a whole family of solutions.

Derivative pricing is the canonical instance. A diffusion trained on historical returns learns the physical measure P, but no-arbitrage pricing (Harrison & Kreps, 1979) requires a risk-neutral measure Q that is never observed as samples. What a market supplies instead is a finite set of restrictions. Discounted prices must be martingales, and quoted options must price back. For path-dependent payoffs these leave the price open, and models calibrated equally well can disagree sharply on a barrier or a lookback (Schoutens et al., 2004; Cont, 2006). Robust-hedging results bound such prices over every model consistent with the quotes (Brown et al., 2001; Davis & Hobson, 2007); how much of a given price the quotes determine, and how much a chosen prior supplies, has not been measured.

We select one admissible measure by minimum relative entropy against the learned path law and carry it into the frozen sampler. The selected measure is the weighted Monte Carlo measure of Avellaneda et al. (2001), and the vehicle is the Doob h-transform; we claim neither. What the combination adds is a reference that is learned and can therefore be retrained and varied. That turns the construction into an instrument: by varying the reference and measuring how far prices can move under the same constraints, we separate what the constraints determine from what the prior supplies. The identification results run on weighted samples and never pass through the sampler; amortization is a separate capability, which we find limited. On simulated paths a learned prior can at best match the simulator it was trained on, and does; it earns its place only where no correctly specified model exists, which is why we also test on real quotes.

Contributions.

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.

- **Finite constraints span most of an Asian's payoff but leave a quarter of a barrier's unspanned.** Fifty-three constraints span 97% of an Asian's payoff variance and 76% of a barrier's, and the unspanned variance predicts, to within 2.4%, how fast the admissible price range opens with entropy budget. This confirms the local expansion of Lam (2016) numerically and makes the known weakness of vanilla information for barriers quantitative (Section 5.3).
- **Agreement across priors is weak evidence of identification.** Priors differing in one Heston parameter move exotic prices by at most about 0.13, and each is reachable within about 0.008 nats of the entropy floor; at 0.10 nats, a budget comparable to the divergence between those priors, the constraints still admit ranges 6-45 times wider than the shifts (Section 5.4).
- **On real quotes, every time-homogeneous prior misprices a scheduled earnings date the same way.** A learned diffusion, a fitted Heston, an i.i.d. Student-t and Black-Scholes all overprice expiries before the event and underprice those containing it, and the decomposition recovers a 4.3-4.6% earnings jump consistent with the established signature (Section 5.6).
- **Amortizing the tilt works only at mild tilt, and fails through the gradient.** The learned correction reproduces the tilt as a function (correlation 0.984 on Heston, ≥ 0.998 on a mixture with exact target) while the sampler under-realises the measure it implies; retraining and weighting both do better (Section 5.5).

2 Background

2.1 Score-Based Diffusion Models

A DDPM (Ho et al., 2020) corrupts data x_0 by the forward process

$$x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, I), \tag{1}$$

with $\bar{\alpha}_t = \prod_{s \leq t}(1-\beta_s)$, and a network $\varepsilon_\theta(x_t, t)$ trained to predict ε reverses it by

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\Big(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\, \varepsilon_\theta(x_t, t)\Big) + \sqrt{\tilde{\beta}_t}\, \xi_t, \quad \xi_t \sim \mathcal{N}(0, I). \tag{2}$$

By Anderson's theorem (Anderson, 1982) the reverse dynamics are driven by the score, which the noise predictor approximates (Ho et al., 2020; Song et al., 2021b):

$$\nabla_{x_t} \log p_t(x_t) \approx -\frac{1}{\sqrt{1-\bar{\alpha}_t}}\, \varepsilon_\theta(x_t, t). \tag{3}$$

Changing the law the reverse chain targets is therefore a change to ε_θ.

2.2 What a Market Supplies

Prices are expectations under a risk-neutral measure Q: a payoff Φ on the path is worth $\mathbb{E}^Q[e^{-rT}\Phi]$, while recorded returns are draws from the physical measure P. Under geometric Brownian motion the two differ only in drift and Q is unique. Under stochastic volatility, as in every experiment here, Q is not unique, and a market identifies it only through finitely many restrictions: martingale conditions on discounted prices and the prices of quoted vanillas. Those restrictions fix marginals at quoted maturities but not the joint law of the path, which is what a barrier or lookback depends on.

3 Methodology

3.1 A Measure Change Acts Additively on the Score

For measures $Q \ll P$ with positive C^1 densities,

$$\nabla \log q(x) = \nabla \log p(x) + \nabla \log \frac{dQ}{dP}(x) \tag{4}$$

(Proposition 2, Appendix A.2). The forward process (1) applies the same linear map and noise to both measures, so $Q_0 \ll P_0$ implies $q_t \ll p_t$ for $t > 0$, and substituting (4) into Anderson's reverse drift and (3) gives:

Proposition 1 (Score Measure Transformation). *Let $\{p_t\}$ and $\{q_t\}$ be diffused marginals of P_0 and Q_0 under the same DDPM forward process, with $Q_0 \ll P_0$. The reverse-time SDE targeting Q_0 is obtained from that targeting P_0 by the additive correction*

$$\varepsilon_\theta^Q(x_t, t) = \varepsilon_\theta^P(x_t, t) - \underbrace{\sqrt{1-\bar{\alpha}_t}\,\nabla \log \frac{q_t(x_t)}{p_t(x_t)}}_{\delta_t(x_t)}, \tag{5}$$

applied independently at each timestep. Proof in Appendix A.3.

This is a continuous-time statement under Assumption 1; the discrete sampler (2) adds a discretization error we measure rather than bound. The correction is minimal only in a restricted sense (Remark 1). We call the construction Score Measure Transformation (SMT) and claim nothing for the mechanism, which is the Doob h-transform.

3.2 Constraint-Driven Targets: Entropy Projection

Let $X = (Y_1, \ldots, Y_H)$ be a return path and P_θ the law of a path DDPM trained on physical paths. Given constraint functions $g_1, \ldots, g_m$ and targets $c \in \mathbb{R}^m$, we select

$$Q^* = \arg\min_{Q \ll P_\theta} \mathrm{KL}(Q \,\|\, P_\theta) \quad \text{s.t.} \quad \mathbb{E}_Q[g_j(X)] = c_j, \;\; j = 1, \ldots, m, \tag{6}$$

an exponential tilt of the prior,

$$L^*(X) = \frac{dQ^*}{dP_\theta}(X) = \exp\big(\beta^{*\top} g(X) - \psi(\beta^*)\big), \qquad \psi(\beta) = \log \mathbb{E}_{P_\theta}\big[e^{\beta^\top g(X)}\big], \tag{7}$$

where β^* minimizes the convex dual $\psi(\beta) - \beta^\top c$, solved in seconds by L-BFGS at $m \approx 50$. Since ψ is an expectation under P_θ, evaluating the objective requires samples from the generative model.

Constraint families. All are computable from the return path: discounted increments $\Delta\tilde{S}_{s+1}$ at each step; dynamic test functions $\varphi(H_s)\,\Delta\tilde{S}_{s+1}$ with $\varphi \in \{1, S_s/S_0, (S_s/S_0)^2, \text{rolling realized variance}\}$; and quoted vanillas $\mathbb{E}_Q[e^{-rT} G_j(X)] = C_j$. The first two give finitely many *unconditional* moment conditions, not the conditional martingale property, so Q^* is constraint-consistent rather than risk-neutral.

3.3 Amortization into the Sampler

L^* is a function of the clean path; the reverse chain needs a correction at every noise level. The bridge is

$$h_\tau(z) = \mathbb{E}_{P_\theta}\big[L^*(X) \,\big|\, Z_\tau = z\big], \tag{8}$$

the Doob h-transform (De Bortoli et al., 2021; Zhu et al., 2026; Denker et al., 2024), with $h_0 = L^*$ and $h_\tau \to 1$ at maximal noise; substituting $q_\tau / p_\tau = h_\tau$ into (5) gives $\delta_\tau = \sqrt{1-\bar{\alpha}_\tau}\,\nabla \log h_\tau$. We regress h on L^* directly with a positive head

$$h_\psi(z, \tau) = h_{\min} + (1 - h_{\min}) \exp\big(\bar{\alpha}_\tau\, s_\psi(z, \tau)\big), \tag{9}$$

which makes the noise-end boundary condition structural, and differentiate by autodiff (Algorithm 1). Q^* is the weighted Monte Carlo measure (Avellaneda et al., 2001), which represents it as weights on a fixed sample; amortization instead yields unweighted draws from it.

3.4 Measuring What the Constraints Identify

Four measurements, all computed on weighted samples from the prior and none passing through the sampler, separate what the constraints determine from what the prior supplies.

Span. Regressing a payoff f on $(1, g_1, \ldots, g_m)$ under Q^* splits it as $f = a + b^\top g + f_\perp$. The share R^2 of its variance in the span is fixed by the constraints; $f_\perp$ is not. R^2 is a linear and local notion, the share of Q^*-variance in the span of the chosen functions rather than a global identification percentage, and its operational meaning comes through (10).

Worst-case bounds. Optimizing $\mathbb{E}_Q[f]$ over all weights on the sample that satisfy the constraints, a sample-restricted analogue of martingale optimal transport (Beiglböck et al., 2013; Galichon et al., 2014), gives the range the constraints cannot exclude. Because the martingale conditions are unconditional and the weights live on a sample, these are ranges over constraint-consistent measures on the draw, not no-arbitrage bounds (Appendix D.11).

Budgeted intervals. Restricting to constraint-consistent measures within relative entropy Δ of the floor $\mathrm{KL}(Q^*\|P_\theta)$ grades the space in between. By the Pythagorean identity (Csiszár, 1975), $\mathrm{KL}(Q\|P_\theta) = \mathrm{KL}(Q\|Q^*) + \mathrm{KL}(Q^*\|P_\theta)$ for every constraint-consistent Q, so this budget is a relative-entropy ball around Q^*, and the local expansion of Lam (2016), restricted to perturbations that preserve the constraints, gives the interval width

$$\max_Q \mathbb{E}_Q[f] - \min_Q \mathbb{E}_Q[f] = C\sqrt{\Delta} + o(\sqrt{\Delta}), \qquad C = 2\sqrt{2\,\mathrm{Var}_{Q^*}(f_\perp)}. \tag{10}$$

What the span leaves out is therefore what sets how fast the admissible price range opens (Appendix D.20).

Varying the reference. Re-running the projection on priors trained under different assumptions, and asking what budget reaches each one's price, locates those priors inside the admissible range.

4 Related Work

Doob h-transforms for pretrained diffusions. Adding a score term to a frozen diffusion is the Doob h-transform, used for training-free adaptation (Zhu et al., 2026), inference-time alignment with gradient-regularized regression of h (Chang et al., 2026) and fine-tuning through a learned generalised h (Denker et al., 2024); diffusion Schrödinger bridges (De Bortoli et al., 2021) share its foundations, twisted SMC (Wu et al., 2023) approaches exactness through particles, and meta flow maps (Potaptchik et al., 2026) learn a reusable map for reward alignment. Guidance (Dhariwal & Nichol, 2021; Ho & Salimans, 2022; Chung et al., 2023; Fan et al., 2023) conditions or biases generation rather than imposing prescribed expectations (Appendix B). In all of these the tilt is supplied; here it is solved for, against a prior that must be sampled to evaluate the objective.

Entropic calibration and model uncertainty. Minimum relative entropy under market constraints is weighted Monte Carlo (Avellaneda et al., 2001), and our Q^* is that measure. Relatives include the minimal entropy martingale measure (Frittelli, 2000), entropic jump calibration (Cont & Tankov, 2004b), martingale Schrödinger bridges (Nutz et al., 2023) and Guyon (2024); all take a tractable reference as given and keep the result as weights. That calibrated models disagree on exotics is documented (Schoutens et al., 2004; Cont, 2006), as is the failure of time-homogeneous models around scheduled earnings announcements, which option prices treat as a dated variance increment (Patell & Wolfson, 1979; Dubinsky et al., 2019); Glasserman & Xu (2014) bound expectations over entropy balls and Lam (2016) derives their $\sqrt{\mathrm{KL}}$ scaling, which (10) specializes to a ball around Q^*. Closest on the generative side, Gierjatowicz et al. (2022) learn neural SDEs consistent with market data and bound prices over the calibrated class, and Moment Guided Diffusion (Lempereur et al., 2026) generates maximum-entropy distributions under moment constraints. Maximum entropy is relative-entropy projection against a flat reference; against a learned one the solution also inherits the prior's structure outside the span of the constraints, in the component $f_\perp$ that Section 5.3 shows sets path-dependent prices (further work: Appendix B.2).

What this paper adds. A learned, variable reference for entropic calibration; its amortization into a frozen sampler, with where that works and where it fails; and a measurement, on simulated and real quotes, of what finite constraints identify about path-dependent prices.

5 Experiments

Section 5.2 checks that the projection works, on a simulator and on a learned prior. Sections 5.3 and 5.4 measure what the constraints identify; like all of Section 3.4, they run on weighted samples and never touch the sampler. Section 5.5 tests amortization, and Section 5.6 repeats the projection on real quotes.

Table 1: Robustness across the physical-to-risk-neutral gap: projection at C3 against 10^5 simulator paths. Held-out vanilla RMSE is in units of 10^{-3}; "improvement" is against the unprojected measure. The last column is the fraction of the barrier's physical-to-risk-neutral price span the projection closes; the Asian and lookback spans cross zero across gaps, so their fractions are not meaningful (Appendix D.4). The $+2.57$ row is not the headline C3 run: the sweep fixes each gap by inverting a target volatility, giving $\theta_Q = 0.10005$ against 0.10 exactly, so the two agree on volatility to two decimals but differ in the fourth decimal of every figure.

Gap	σ_Q (21d)	ESS (%)	Held-out van.	Q floor	Improvement	Barrier closed
$+0.25$	20.25%	99.38	1.23	6.06	196×	89%
$+1.00$	21.00%	98.78	1.11	7.70	192×	81%
$+2.57$	22.57%	93.55	2.83	12.19	64×	72%
$+4.00$	24.00%	81.76	4.63	16.54	43×	68%

5.1 Setup

Paths follow Heston dynamics (Heston, 1993) over $H = 21$ daily steps, with a physical-to-risk-neutral gap of $+2.57$ volatility points (20.00% to 22.57%) set by a chosen variance risk premium; four gaps are swept (Appendix D.2). Constraint levels nest from **C0** (41 martingale test functions) to **C3** (53, adding vanillas at two maturities); 15 test functions and 9 vanillas are held out. The exotics are an Asian call, an up-and-out barrier call (K=100, B=110) and a floating-strike lookback, compared with a Heston-Q Monte Carlo benchmark treated as a reference rather than a target. The path prior is a 580k-parameter time-conditioned MLP DDPM on 21-dimensional return vectors, frozen before any projection and accepted against thresholds fixed in advance (Appendix D); a Gaussian mixture with closed-form target checks the algebra (Appendix C). The effective sample size $\mathrm{ESS} = (\sum_i w_i)^2 / \sum_i w_i^2$ is the primary health metric.

5.2 Projection Reprices Held-Out Quotes, on Simulator and Learned Prior Alike

The projection runs first against raw simulator paths, with no network in the pipeline, then against the learned prior (Table 15, Appendix D.17). The weights stay healthy, ESS above 93% at every level with constraint error at machine precision. At C3 held-out vanillas are repriced below the benchmark's own Monte Carlo floor (0.0028 against 0.0122): the constraints tie down neighbouring strikes, so the projection acts as a control variate. The learned prior tracks the simulator, with ESS within 0.3 points and held-out error within 0.001 from C1 upward.

The result holds across the gap (Table 1). At every one of four gaps C3 reprices held-out vanillas below the benchmark's own Monte Carlo floor, improving on the unprojected measure 43- to 196-fold, while ESS falls only to 82% at the widest; the barrier closes 68-89% of its span.

Two findings qualify these results. Calibration and martingality trade off: held-out martingale residuals rise from 0.00085 to 0.01322 as vanillas enter, a property of minimum-entropy selection on a finite sample rather than of the constraint set, which Heston-Q satisfies in full (Appendix D.18). Martingale constraints alone are also not a partial step toward Q: at the $+4$-point gap they price held-out vanillas *worse* than no projection (0.250 against 0.200), removing the drift without supplying the variance. The learned-prior column required fixing a path-specific training failure, in which sign-aligned optimizer noise sums into a terminal-price shift invisible to per-coordinate checks (Appendix D.5).

The martingale block does double duty on a learned prior. On P_θ, $\|\beta^*\|$ at C0 triples on the same columns and targets, from 5.8 to 17.8: the martingale constraints are also absorbing the prior's residual level bias. A prior with its drift corrupted to $2.5\times$ bounds the effect: the block absorbs 92-98% of the corruption, at a cost in ESS (Appendix D.9).

5.3 The Constraints Span an Asian but Not a Barrier

Table 2 and Figure 1 apply the measurements of Section 3.4 at C3.

Table 2: What the 53 constraints identify at C3, on 10^6 learned-prior draws. Span R^2 is the share of payoff variance in the span of the constraints; $\mathrm{sd}(f_\perp)$ is what remains. The width constant is predicted by (10) and measured from the budgeted intervals as the budget shrinks. Intervals are at total relative entropy 0.10 and 1.00 nats (floor 0.029); "sat." marks the barrier's upper branch at its worst-case bound. The last two rows anticipate Sections 5.4 and 5.6; the AAPL exotics are an at-the-money Asian, an up-and-out barrier at 110% of spot, and a floating-strike lookback.

	Asian	UO barrier	Lookback
Q^* price	1.625	1.830	4.425
Span R^2, C0 $\to$ C3	$0.80 \to 0.97$	$0.47 \to \mathbf{0.76}$	$0.82 \to 0.93$
Residual $\mathrm{sd}(f_\perp)$	0.405	1.306	0.976
Width constant, predicted / measured	1.144 / 1.137	3.694 / 3.782	2.761 / 2.745
Interval at 0.10 nats	$[1.47, 1.77]$	$[1.36, 2.27]$	$[4.08, 4.80]$
Interval at 1.00 nat	$[0.99, 2.15]$	$[0.47,$ sat.$]$	$[3.34, 5.99]$
Worst case, no budget	$[0.24, 3.11]$	$[0.00, 2.35]$	$[2.80, 9.62]$
Largest misspecified-prior shift	$6\times$ band	$78\times$ band	$29\times$ band
AAPL: price spread across 4 priors	1.0%	$\mathbf{20.4}\%$	5.8%

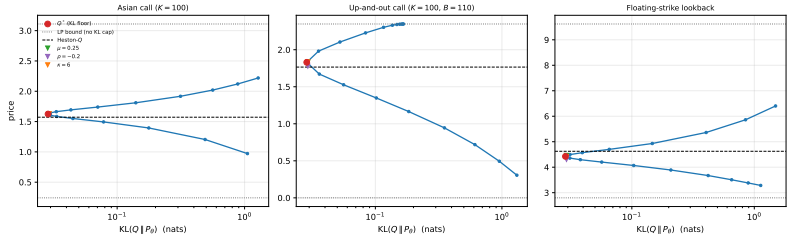

Figure 1: Admissible price range against relative-entropy budget, over measures meeting the same 53 constraints. The range opens from Q^* at the floor, 0.029 nats, as $\sqrt{\mathrm{KL} - \mathrm{floor}}$. Dotted lines mark the unbudgeted worst case, dashed lines the Heston-Q benchmark, and markers the prices of the misspecified priors of Section 5.4, all within about 0.008 nats of the floor.

The constraints span an Asian but not a barrier. 97% of the Asian's payoff variance lies in the span of the constraints, against 76% of the barrier's, and the ordering holds at every level. The barrier gains most at C2, when the strike at the barrier level enters; its one non-monotone step, C0 to C1, is a change of measure rather than lost span (Appendix D.19).

Worst-case bounds and entropy budgets. Over all constraint-consistent weights the ranges are 1.3-1.8 times the prices, with every benchmark inside, attained by degenerate 54-atom measures (Appendix D.11). A budget grades the space in between: one nat buys intervals 1.16, 1.9 and 2.65 wide, far inside the worst case.

The unspanned variance sets how fast prices open. The width constant predicted by (10) from $\mathrm{sd}(f_\perp)$ matches the measured one to 0.6%, 2.4% and 0.6%, and an independent route through $d\mathbb{E}_Q[f]/d\gamma$ agrees to within 3% (Appendix D.20). The barrier, least spanned, has the largest constant and reaches its worst-case upper bound by 0.17 nats, since the payoff is bounded above.

Responsiveness and span measure different things. Against the Heston-Q benchmark the barrier is also the exotic the tilt moves most, closing 72% of its physical-to-risk-neutral span, while the lookback barely moves (2%) and the Asian moves away (-31%). Those three figures are from the headline projection (Appendix D.13), not from the 10^6 linear-programming draw whose prices Table 2 reports; the two differ in the third decimal and give 72% against 78% on the barrier. The readings do not conflict: span closed measures how far a price moves toward one benchmark, R^2 how much of the payoff the constraints pin. The barrier responds most and keeps the largest unidentified residual, and the constraint that moves it is identifiable, namely the strike at the barrier level.

Table 3: Amortization against simpler routes to the same Q^*. *Top*: Heston at C3, each method against the weighted Q^* on the same 10^6-path pool, in standard errors (worst of three exotics). The mixture results, where Q^* is exact and both tilt and dimension are swept, are in Appendix D.22.

Heston, C3	Calibrated	Held-out	Worst exotic	10^5 draws	Adaptation
SMT: frozen prior $+ h_\psi$	1.46	2.08	2.53	19.7 s	0.9 min
Retrain on L^*-resampled paths	0.94	1.18	0.95	9.4 s	7.9 min
Weighting, 10^6 pool	0.90	1.11	1.59	-	-
Resampling, 10^5 pool	1.64	1.83	3.71	10.2 s	-

5.4 Agreement Across Priors Is Not Identification

Six priors were trained and projected: three of the baseline recipe differing only in seed, and three misspecified in one parameter each, in mean reversion, drift or leverage correlation (Appendix D.25). Seed variation moves C3 exotic prices by 0.3-2.7% of the physical-to-risk-neutral span; misspecification moves them far more, up to 78 times that band for the barrier under the wrong leverage correlation (Table 2; Figure 4, Appendix D.25).

The separation concentrates where the coupling of the path through time matters. The wrong leverage correlation moves the barrier by $78\times$ and the lookback by $29\times$ the seed band, the direction finitely many vanillas cannot reach. We had predicted that a misspecified drift would separate less than misspecified mean reversion; the observed ordering reverses that, and since the three perturbations differ in size, it does not rank the parameters by influence (Appendix D.14).

The minimum budget reaching each misspecified prior's price nonetheless lies within about 0.008 nats of the floor (Figure 1). At 0.10 nats, a budget comparable to the measured divergence between the priors, the constraints admit ranges of 0.30, 0.91 and 0.72 (Table 2): 45, 13 and 6 times the largest prior shift on each exotic. This is not a distance between priors, since matching one price is cheaper than matching a path law. Two of the three are also near as laws, within 0.052 nats of the base on the full path, but `rho02` changes the diffusion coefficient, so in continuous time it is mutually singular with the base law and its divergence is infinite; on the 21-step discrete path it is finite but we do not estimate it (Appendix D.10). Since `rho02` is also the prior with the largest price effect, the "nearby laws" reading does not cover the case that matters most. Their agreement is tight because they left the expensive directions unexplored, not because the constraints bind, and it is weak evidence of identification.

The learned prior also matches the correctly specified parametric model, the Heston-P simulator itself: span R^2 agrees within 0.003 at all twelve level-exotic cells, and budgeted interval widths within about 1% (Appendix D.12). On Heston data it cannot beat that ceiling; a learned reference is needed only where no correctly specified model exists, the setting of Section 5.6.

5.5 Amortization Reproduces Q^* Only at Mild Tilt

Everything above runs on weighted samples. Amortization is a separate capability, producing unweighted draws from Q^* out of the frozen sampler, and Table 3 compares it with simpler routes to the same measure.

At the Heston operating point. At KL ≈ 0.03 the corrected sampler sits at 1.46 standard errors of the weighted reference, rms over the calibrated columns, and up to 2.5 off on two exotics. A diffusion retrained on paths resampled by L^* is more accurate, every exotic within one standard error, and samples twice as fast; the correction's advantage is adaptation, about one minute of training against eight. It does beat resampling from a matched 10^5 pool (1.46 against 1.64 rms calibrated, worst exotic 2.53 against 3.71), where only 62% of draws are unique. Against weighted resampling of its own 10^5-path draw it agrees on all 77 constraint columns within two standard errors and is flat in reverse-step count from 999 to 100 (Appendix D.15); the exotic gaps above appear only against the more precise 10^6-pool reference.

On a mixture with exact Q^*. SMT matches weighting only at KL ≈ 0.05. As KL rises to 0.95 its error grows ninefold, and at fixed KL it grows $7\times$ from $d = 2$ to 128, against $1.67\times$ for the

Table 4: AAPL, one frozen surface as of 2026-04-20: four priors projected onto identical bid-ask interval constraints at the 25-day expiry and evaluated on the four shorter expiries. Errors are median model minus mid, in bid-ask spreads at a 0.10 spread floor (55 of 74 contracts); positive is overpricing, and median absolute errors equal the magnitudes shown. The 7- and 9-day expiries mature before the 2026-04-30 earnings date; the 11- and 18-day expiries contain it. Unfloored and per-expiry errors in Appendix D.23.

		Held-out error (spreads)		Pooled	Max daily
Prior	ESS (%)	event-free	event-bearing	abs. error	return
Path DDPM (learned)	92.2	$+1.13$	$\mathbf{-0.43}$	0.70	**0.146**
Heston (fitted, MoM)	93.7	$+0.96$	-0.71	0.87	0.221
Student-t (i.i.d., MLE)	87.7	$+1.09$	-0.49	**0.68**	2.412
Black-Scholes (hist. vol)	77.8	$\mathbf{+0.74}$	-1.12	1.11	0.091
AAPL returns (2,839 days)					0.143

base model. Weighting from a 5×10^5 pool stays near 0.01 throughout, even at 1.7% ESS. We had predicted the opposite ordering; the prediction is falsified. Weighting even beats an oracle diffusion trained on exact Q^* samples at every $d \geq 32$, since it reuses the prior's samples and applies the tilt exactly rather than fitting a second model.

Locating the failure in the gradient. In every working mixture cell h_ψ fits L^* as a function to correlation ≥ 0.998, yet evaluated on the sampler's own trajectories the sampler under-realises the measure h_ψ implies, by a margin that tracks its error, falling from 0.978 to 0.926 across dimension (Appendix D.22). The regression supervises h_ψ's values while the sampler consumes $\nabla \log h_\psi$, and h_ψ is fit on forward-noised prior samples rather than on the sampler's trajectories; the diagnostic cannot separate these two causes. Estimators that supervise the gradient directly would test the first. These include importance-weighted denoising score matching against the frozen model, close to DEFT (Denker et al., 2024), and gradient-regularized Doob's matching (Chang et al., 2026). We have not run them. A replication on an independently trained prior agrees (Appendix D.24). We claim for amortization only that it approximates Q^* at mild tilt.

5.6 Every Prior Misprices the Same Earnings Date

We repeat the projection on one frozen AAPL option surface as of 2026-04-20 (spot 272.57, five expiries of 7-25 days, 230 quoted contracts), pricing an at-the-money Asian, an up-and-out barrier at 110% of spot and a floating-strike lookback, with a path DDPM trained on 2,839 daily returns and three references fitted to the same returns: a Heston by method of moments, an i.i.d. Student-t, and Black-Scholes at historical volatility. Two setup findings shape the design (Appendix D.23): vanilla constraints must be bid-ask intervals, since equality to the mid is infeasible (ESS 0.001%), and at these maturities only the carry is identified, so the discount rate is a stated assumption.

Every prior misprices the earnings date in the same direction. AAPL reported earnings on 2026-04-30, between the 9- and 11-day expiries. All four priors overprice the expiries before it and underprice those containing it (Table 4), on 85-100% of individual contracts in every group. Each prior is time-homogeneous, so calibrating to a 25-day window that contains one earnings jump smears that variance uniformly across calendar time. The held-out set therefore tests extrapolation across a scheduled event rather than across maturity, a design limitation we found after the fact.

Evidence for a scheduled jump. That options price an anticipated earnings release as a dated variance increment is long established (Patell & Wolfson, 1979; Dubinsky et al., 2019), and the $\sigma^2 + J/T$ decomposition is theirs; recovering it here checks the pipeline against a result that is already known. Three pieces of evidence. Fitting implied variance as $\sigma^2 + J/T$ to the event-straddling expiries gives a diffusive volatility of 22.5-23.1% and an event jump of 4.3-4.6%; the fit predicts the untouched 18-day at-the-money volatility to within 0.65 points, and its diffusive level is consistent with the two pre-event expiries (22.4% and 24.2%), which it never saw. On out-of-the-money strikes the signed error swings from $+52$ to $+63\%$ before the event to -40 to -48% after it. At the 11-day

K=290 call, where the market's implied volatility is 33.7% against 28.2% at 25 days, the three parametric arms underprice by 38-46% while pricing the same strikes at the calibrated expiry to within 1.4%. The Student-t, with two orders of magnitude more tail weight than Black-Scholes, closes one percentage point of that gap, so the shortfall comes from the absence of a dated jump rather than from tail weight.

The learned prior is most accurate where the event falls. On the event-bearing expiries it errs by 0.43 of a spread, against 0.49 for the Student-t, 0.71 for the fitted Heston and 1.12 for Black-Scholes. The ordering inverts on the event-free expiries, where it is worst; pooled, it ties the Student-t, the two trading places across spread floors; and its event-bearing lead holds at the 0.10 floor but not without one. Both separate from Black-Scholes at every floor and from the fitted Heston at all but one. A prior trained on eleven years of overlapping blocks has seen many earnings windows, which would plausibly help where the event dominates and hurt where it is absent. That reading is consistent with these numbers, but they do not establish it. We report no confidence intervals: with 21 and 34 contracts per group, the DDPM-Student-t difference should not be read as significant.

Tails and cross-prior spreads. The largest daily return the learned prior generates is 0.146 against the data's 0.143; the fitted Heston reaches 0.221, and the i.i.d. Student-t, which reads volatility clustering as tail weight, 2.41. Across the four priors the barrier's price spreads 20% against the Asian's 1% (Table 2), the ordering the span analysis of Section 5.3 found on simulated paths. We did not compute span R^2 on the AAPL constraint set, so we can compare the orderings but not the underlying quantities.

6 Limitations

Unconditional restrictions. The test functions impose finitely many unconditional moment conditions, not the conditional martingale property, so Q^* is constraint-consistent rather than risk-neutral and can admit arbitrage in path-dependent claims; Section 5.3 measures how much room this leaves.

The prior's support bounds the projection. Across three seeds the learned prior's right tail is thin by the same margin (the 21-day K=115 call at 0.82-0.92 of reference), a bias shared by every prior and therefore invisible in any between-prior band; beyond the last calibrated strike the projection cannot repair it.

Where weights bind. ESS stays above 93% at 53 constraints but falls to about 47% on both laws at 86, the regime where amortization might pay; we did not run it there, nor gradient-supervised estimators of the correction (Section 5.5).

Process. One acceptance criterion was amended after it rejected a prior; the replacement is stricter on 20 of 21 columns and widens the reference band, working against our own claims (Appendix D.8).

Scope. Simulated results use one Heston specification at $H = 21$. The market experiment is one surface, one underlying and one date, and tests extrapolation across an earnings event rather than across maturity; its held-out comparisons carry no confidence intervals.

7 Conclusion

A target specified only by constraints leaves its tilt to be solved for. Solving it by entropy projection against a learned path prior makes the reference variable, and so makes identification measurable. Finite market constraints span most of an Asian's payoff variance but leave a quarter of a barrier's unspanned, and that residual sets how fast admissible prices diverge; agreement among nearby priors says little about it. On a real surface every time-homogeneous prior misprices a scheduled earnings date in the same direction. Carrying the tilt into a frozen sampler works at mild tilt, and beyond it fails through the gradient of the correction rather than its value.

AI USE STATEMENT

We used Claude, a large language model developed by Anthropic, in preparing this submission. Classified against the ICLR 2027 policy:

Required-disclosure tasks for which it was used. Assisting in the writing of proofs and formulating mathematical claims: it redrafted Proposition 1 to separate the continuous-time result from the finite-step sampler, restated minimality in the restricted form of Remark 1 with the supporting argument in Appendix A.3, and corrected the claim in Appendix B relating Bayes-optimal classifier log-odds to density-ratio scores. Interpreting results: it drafted the interpretive prose of Sections 5 and 6 from result summaries provided by the authors. Feedback on methodology: it flagged numerical inconsistencies between tables for the authors to check. Across the project as a whole, we used generative AI tools for drafting and editing the manuscript, for implementing and running the experimental code, and for preparing tables and figures. The research question and all final decisions are our own. The AI proposed several experimental designs and analyses, which we reviewed, modified and approved, and it identified related work that we verified before citing. We take responsibility for all claims.

Recommended-disclosure tasks. It drafted substantial parts of the text from author-provided outlines and results, covering the abstract, introduction, Sections 3.2, 4, 5, 6 and 7, and Appendices B and D. It converted the manuscript to the ICLR template, restructured it to meet the page limit, formatted the references, and checked reference author lists, years and venues against publisher and arXiv records, which we then re-checked ourselves and corrected where the AI had erred.

We reviewed all AI-assisted work, reproduced reported results from saved artifacts, and verified numerical claims against the code that produced them. We take responsibility for the final content of this work, including text, claims and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work involves no human subjects and no personally identifying data. The simulated experiments use synthetic dynamics; the market experiment uses one frozen, publicly quoted equity option surface and a daily return history, with no personal or proprietary data. The method is a research prototype for derivative pricing and is not calibrated or validated for deployment in production trading or risk systems.

REPRODUCIBILITY STATEMENT

Experimental settings are given in Section 5.1, with full architectural and optimization details in Appendices F and D. All theoretical claims are stated with their regularity conditions in Assumption 1 and proved in Appendix A. Simulation details for the Heston experiments are in Appendix D, together with the acceptance criteria for the frozen path prior, the learned-correction diagnostics, and the solve-draw noise characterization. Experiments were run on Apple Silicon (MPS) and on an A100 with TF32 disabled: the frozen checkpoint failed the acceptance criteria on the A100 under default TF32 and passed under full fp32, over 999 reverse steps. Anonymized code covering all experimental settings is provided as supplementary material.

REFERENCES

Anderson, B. D. O. (1982). Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313-326.

Avellaneda, M., Buff, R., Friedman, C., Grandechamp, N., Kruk, L., and Newman, J. (2001). Weighted Monte Carlo: A new technique for calibrating asset-pricing models. *International Journal of Theoretical and Applied Finance*, 4(1):91-119.

Beiglböck, M., Henry-Labordère, P., and Penkner, F. (2013). Model-independent bounds for option prices - a mass transport approach. *Finance and Stochastics*, 17(3):477-501.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637-654.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621-651.

Brown, H., Hobson, D., and Rogers, L. C. G. (2001). Robust hedging of barrier options. *Mathematical Finance*, 11(3):285-314.

Buehler, H., Gonon, L., Teichmann, J., and Wood, B. (2019). Deep hedging. *Quantitative Finance*, 19(8):1271-1291.

Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61-73.

Chang, J., Duan, C., Jiao, Y., Xu, Y., and Yang, J. Z. (2026). Inference-time alignment for diffusion models via variationally stable Doob's matching. *arXiv preprint arXiv:2601.06514*.

Chung, H., Kim, J., McCann, M. T., Klasky, M. L., and Ye, J. C. (2023). Diffusion posterior sampling for general noisy inverse problems. In *International Conference on Learning Representations*.

Clark, K., Vicol, P., Swersky, K., and Fleet, D. J. (2023). Directly fine-tuning diffusion models on differentiable rewards. *arXiv preprint arXiv:2309.17400*.

Cont, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519-547. DOI 10.1111/j.1467-9965.2006.00281.x.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC, Boca Raton.

Cont, R. and Tankov, P. (2004b). Non-parametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, 7(3):1-49. DOI 10.21314/JCF.2004.123.

Csiszár, I. (1975). I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1):146-158.

Davis, M. H. A. and Hobson, D. (2007). The range of traded option prices. *Mathematical Finance*, 17(1):1-14.

De Bortoli, V., Thornton, J., Heng, J., and Doucet, A. (2021). Diffusion Schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems*, volume 34, pages 17695-17709.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(3):463-520.

Denker, A., Vargas, F., Padhy, S., Didi, K., Mathis, S., Dutordoir, V., Barbano, R., Mathieu, E., Komorowska, U. J., and Liò, P. (2024). DEFT: Efficient fine-tuning of diffusion models by learning the generalised h-transform. In *Advances in Neural Information Processing Systems*, volume 37, pages 19636-19682.

Dhariwal, P. and Nichol, A. (2021). Diffusion models beat GANs on image synthesis. In *Advances in Neural Information Processing Systems*, volume 34, pages 8780-8794.

Dolinsky, Y. and Soner, H. M. (2014). Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields*, 160(1-2):391-427.

Dubinsky, A., Johannes, M., Kaeck, A., and Seeger, N. J. (2019). Option pricing of earnings announcement risks. *The Review of Financial Studies*, 32(2):646-687. DOI 10.1093/rfs/hhy060.

Fan, Y., Watkins, O., Du, Y., Liu, H., Ryu, M., Boutilier, C., Abbeel, P., Ghavamzadeh, M., Lee, K., and Lee, K. (2023). DPOK: Reinforcement learning for fine-tuning text-to-image diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 79858-79885.

Frittelli, M. (2000). The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, 10(1):39-52.

Galichon, A., Henry-Labordère, P., and Touzi, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312-336.

Gierjatowicz, P., Sabate-Vidales, M., Šiška, D., Szpruch, L., and Žurič, Ž. (2022). Robust pricing and hedging via neural stochastic differential equations. *Journal of Computational Finance*, 26(3):1-32. DOI 10.21314/JCF.2022.025.

Glasserman, P. and Xu, X. (2014). Robust risk measurement and model risk. *Quantitative Finance*, 14(1):29-58.

Guo, G. and Obłój, J. (2019). Computational methods for martingale optimal transport problems. *The Annals of Applied Probability*, 29(6):3311-3347.

Guyon, J. (2024). Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *Finance and Stochastics*, 28(1):27-79.

Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381-408.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327-343.

Ho, J. and Salimans, T. (2022). Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*.

Ho, J., Jain, A., and Abbeel, P. (2020). Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pages 6840-6851.

Kawar, B., Elad, M., Ermon, S., and Song, J. (2022). Denoising diffusion restoration models. In *Advances in Neural Information Processing Systems*, volume 35, pages 23593-23606.

Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations*.

Kunsági-Máté, S., Fáth, G., Csabai, I., and Molnár-Sáska, G. (2022). Deep weighted Monte Carlo: A hybrid option pricing framework using neural networks. *arXiv preprint arXiv:2208.14038*.

Lam, H. (2016). Robust sensitivity analysis for stochastic systems. *Mathematics of Operations Research*, 41(4):1248-1275. DOI 10.1287/moor.2015.0776.

Lempereur, E., Cuvelle-Magar, N., Coeurdoux, F., Mallat, S., and Vanden-Eijnden, E. (2026). MGD: Moment guided diffusion for maximum entropy generation. *arXiv preprint arXiv:2602.17211*.

Ni, H., Szpruch, L., Sabate-Vidales, M., Xiao, B., Wiese, M., and Liao, S. (2021). Sig-Wasserstein GANs for time series generation. *arXiv preprint arXiv:2111.01207*.

Nutz, M., Wiesel, J., and Zhao, L. (2023). Martingale Schrödinger bridges and optimal semistatic portfolios. *Finance and Stochastics*, 27:233-254.

Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*, 6th edition. Springer, Berlin.

Patell, J. M. and Wolfson, M. A. (1979). Anticipated information releases reflected in call option prices. *Journal of Accounting and Economics*, 1(2):117-140. DOI 10.1016/0165-4101(79)90003-X.

Potaptchik, P., Saravanan, A., Mammadov, A., Prat, A., Albergo, M. S., and Teh, Y. W. (2026). Meta flow maps enable scalable reward alignment. *arXiv preprint arXiv:2601.14430*.

Schoutens, W., Simons, E., and Tistaert, J. (2004). A perfect calibration! Now what? *Wilmott Magazine*, March 2004:66-78. DOI 10.1002/wilm.42820040216.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, New York.

Sirignano, J. and Spiliopoulos, K. (2018). DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, 375:1339-1364.

Song, J., Meng, C., and Ermon, S. (2021a). Denoising diffusion implicit models. In *International Conference on Learning Representations*.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. (2021b). Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*.

Song, J., Vahdat, A., Mardani, M., and Kautz, J. (2023). Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations*.

Sugiyama, M., Suzuki, T., and Kanamori, T. (2012). *Density Ratio Estimation in Machine Learning*. Cambridge University Press, Cambridge.

Wiese, M., Knobloch, R., Korn, R., and Kretschmer, P. (2020). Quant GANs: Deep generation of financial time series. *Quantitative Finance*, 20(9):1419-1440.

Wu, L., Trippe, B. L., Naesseth, C. A., Blei, D. M., and Cunningham, J. P. (2023). Practical and asymptotically exact conditional sampling in diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pages 31372-31403.

Zhu, Q., Ye, Z., Liu, H., Wang, Z., and Chen, M. (2026). Training-free adaptation of diffusion models via Doob's h-transform. *arXiv preprint arXiv:2602.16198*.

A PROOFS

We provide complete proofs of all theoretical results. Assumption 1 is in force throughout.

Assumption 1. *1. The forward SDE admits a unique weak solution on* $[0, 1]$.

2. *For each* $t \in (0, 1]$*, the forward marginals* p_t *and* q_t *admit strictly positive* C^1 *densities on* $\mathbb{R}^d$.

3. *The score fields* $\nabla \log p_t$ *and* $\nabla \log q_t$ *are measurable with at most linear growth in* x*, locally uniformly in* t.

4. *The reverse-time SDEs associated with* p_t *and* q_t *are well defined in the sense of Anderson time reversal (Anderson, 1982).*

5. *The score-noise relation* (3) *holds exactly in the population limit.*

Conditions 1-4 are standard in reverse-time diffusion theory (Song et al., 2021b; Anderson, 1982) and are satisfied for all distributions considered in this paper. Condition 5 holds exactly when the noise predictor is optimal; in practice it holds approximately, with quality governed by training.

A.1 THE MIXTURE SETTING IN DETAIL

Mixture setting. Between two equal-variance Gaussians the score difference is constant; between mixtures it is not: the score difference becomes a nonlinear function of the current state, and no constant shift can implement the correct measure change.

Let $P_0 = \sum_k \pi_k^P \mathcal{N}(\mu_k, \sigma_k^2)$ and $Q_0 = \sum_k \pi_k^Q \mathcal{N}(\mu_k, \sigma_k^2)$ with *identical* component parameters but different weights $\pi^P \neq \pi^Q$. The forward noising process preserves mixture structure: the diffused marginal of a mixture is a mixture of diffused components, with means $\sqrt{\bar{\alpha}_t}\mu_k$ and variances $s_{k,t}^2 = \bar{\alpha}_t \sigma_k^2 + (1 - \bar{\alpha}_t)$. The score of a Gaussian mixture is a responsibility-weighted average of per-component scores:

$$\partial_y \log p_t(y) = \sum_k r_k^P(y, t) \cdot \frac{-(y - \sqrt{\bar{\alpha}_t}\mu_k)}{s_{k,t}^2}, \tag{11}$$

where $r_k^P(y, t) = \pi_k^P \, \mathcal{N}(y;\, \sqrt{\bar{\alpha}_t}\mu_k,\, s_{k,t}^2) \,/\, p_t(y)$ are the posterior responsibilities. The SMT correction is:

$$\delta_t(y) = \sqrt{1 - \bar{\alpha}_t}\,\big(\partial_y \log q_t(y) - \partial_y \log p_t(y)\big), \tag{12}$$

a responsibility-weighted difference of per-component scores by (11). This is nonlinear in y through the responsibility weights. The correction is zero where $r_k^Q = r_k^P$, and nonzero precisely where the two measures assign different regime probabilities to the same observation. No constant shift can approximate this sign-changing, state-dependent function. The proof that no constant shift suffices is in Appendix A.4.

A.2 PROOF OF PROPOSITION 2

Proposition 2 (Score factorization under measure change)**.** *Let* P *and* Q *be probability measures on* $\mathbb{R}^d$ *with strictly positive* C^1 *densities* p *and* q*, and suppose* $Q \ll P$ *with Radon-Nikodym derivative* $\Lambda = dQ/dP$*. Then*

$$\nabla \log q(x) = \nabla \log p(x) + \nabla \log \Lambda(x). \tag{13}$$

Proof. Since $Q \ll P$, the Radon-Nikodym theorem gives $\Lambda = dQ/dP \geq 0$ such that $q(x) = p(x)\Lambda(x)$ Lebesgue-almost everywhere. Since p and q are both strictly positive by assumption, $\Lambda(x) > 0$ everywhere. Taking logarithms: $\log q(x) = \log p(x) + \log \Lambda(x)$. Since p, q, and $\Lambda = q/p$ are C^1 and strictly positive, differentiating with respect to x gives

$$\nabla \log q(x) = \nabla \log p(x) + \nabla \log \Lambda(x). \qquad \square$$

A.3 Proof of Proposition 1

Remark 1 (Minimality)**.** *Within the class of corrections that act additively on ε_θ^P and depend only on (x, τ), any correction reproducing the marginals $\{q_\tau\}$ under Anderson's formula must agree with δ_τ pointwise, since the reverse drift determines the score field and the score field determines the correction. We state this as a remark rather than a theorem: it presumes that the target is specified by the full family of diffused marginals, and it says nothing about corrections that match only the terminal law Q_0, of which there may be many. We make no uniqueness claim over that larger class.*

Proof. Step 1: Absolute continuity of diffused marginals. Since $Q_0 \ll P_0$ and the forward process (1) applies the same linear map plus independent Gaussian noise to both, for any $t > 0$ and measurable set A: if $p_t(A) = 0$ then $\mathcal{N}(x_t; \sqrt{\bar\alpha_t}x_0, (1-\bar\alpha_t)I) = 0$ on A for all x_0, which implies $q_t(A) = 0$. Hence $q_t \ll p_t$ for all $t > 0$ and the density ratio $\Lambda_t = q_t/p_t$ is well defined and strictly positive.

Step 2: Score correction in continuous time. Anderson's reverse-time formula (Anderson, 1982) gives the reverse SDEs:

$$dX_t = \big(f(X_t, t) - g(t)^2 \nabla \log p_t(X_t)\big)dt + g(t)\, d\bar W_t, \tag{14}$$

$$dX_t = \big(f(X_t, t) - g(t)^2 \nabla \log q_t(X_t)\big)dt + g(t)\, d\widetilde{W}_t, \tag{15}$$

where f, g are the forward SDE drift and diffusion. Both reverse SDEs share the same diffusion coefficient $g(t)$ and differ only in drift. Applying Proposition 2 to the diffused marginals:

$$\nabla \log q_t(x) = \nabla \log p_t(x) + \nabla \log \frac{q_t(x)}{p_t(x)},$$

so the target reverse drift is the source reverse drift plus the additive correction $-g(t)^2 \nabla \log(q_t/p_t)$.

Step 3: Translation to DDPM parameterization. Under Assumption 1.5:

$$\nabla_{x_t} \log p_t(x_t) = -\frac{1}{\sqrt{1-\bar\alpha_t}} \varepsilon_\theta^P(x_t, t), \quad \nabla_{x_t} \log q_t(x_t) = -\frac{1}{\sqrt{1-\bar\alpha_t}} \varepsilon_\theta^Q(x_t, t).$$

Subtracting and multiplying by $-\sqrt{1-\bar\alpha_t}$:

$$\varepsilon_\theta^Q(x_t, t) = \varepsilon_\theta^P(x_t, t) - \sqrt{1-\bar\alpha_t}\, \nabla_{x_t} \log \frac{q_t(x_t)}{p_t(x_t)}.$$

Terminal law, in continuous time. The corrected reverse SDE (15) has terminal law Q_0. This is a statement about the continuous-time process with exact scores; it does not by itself establish the terminal law of the finite-step chain that Algorithm 1 runs, which differs by a discretization error we do not bound.

Minimality (Remark 1). Suppose $\varepsilon_\theta^P + c_\tau$ also reproduces the diffused marginals $\{q_\tau\}$ under Anderson's formula. Since the reverse drift is determined by those marginals and the map from drift to additive noise-predictor correction is injective at each (x, τ), $c_\tau = \delta_\tau$ pointwise. Note what this does *not* show: the hypothesis is that the whole family $\{q_\tau\}$ is reproduced, which is strictly stronger than reproducing Q_0 alone. A correction matching only the terminal law need not agree with δ_τ, and we do not claim uniqueness over that class. This is why the statement appears as a remark and not as part of Proposition 1.

Decomposability. $\delta_t(x_t)$ at timestep t depends only on q_t, p_t, and the current state x_t, not on corrections at other timesteps. This follows from the pointwise application of (13) at each t independently. $\square$

A.4 Proof of the Mixture Score Formula

Lemma 1. *Under the DDPM forward process* (1)*, the diffused marginal of* $P_0 = \sum_k \pi_k^P \mathcal{N}(\mu_k^P, \sigma_k^2)$ *is:*

$$p_t(y) = \sum_k \pi_k^P \, \mathcal{N}(y;\, \sqrt{\bar{\alpha}_t}\mu_k^P,\, s_{k,t}^2), \quad s_{k,t}^2 = \bar{\alpha}_t\sigma_k^2 + (1-\bar{\alpha}_t),$$

and analogously for Q_0*.*

Proof. By the law of total probability, the diffused marginal of a mixture is a mixture of diffused marginals. The forward process maps each component $\mathcal{N}(\mu_k, \sigma_k^2)$ to $\mathcal{N}(\sqrt{\bar{\alpha}_t}\mu_k,\, \bar{\alpha}_t\sigma_k^2 + (1-\bar{\alpha}_t))$, since the forward process maps a Gaussian to a Gaussian with mean scaled by $\sqrt{\bar{\alpha}_t}$ and variance $\bar{\alpha}_t\sigma^2 + (1-\bar{\alpha}_t)$. Mixture weights π_k are preserved since the forward process does not depend on the component index. □

Proposition 3. *The score of* p_t *is:*

$$\partial_y \log p_t(y) = \sum_k r_k^P(y,t) \cdot \frac{-(y - \sqrt{\bar{\alpha}_t}\mu_k^P)}{s_{k,t}^2},$$

where $r_k^P(y,t) = \pi_k^P \mathcal{N}(y; \sqrt{\bar{\alpha}_t}\mu_k^P, s_{k,t}^2)/p_t(y)$*.*

Proof. Differentiating $\log p_t(y) = \log \sum_k \pi_k^P \mathcal{N}(y; \sqrt{\bar{\alpha}_t}\mu_k^P, s_{k,t}^2)$:

$$\partial_y \log p_t(y) = \frac{\sum_k \pi_k^P \mathcal{N}(y; \sqrt{\bar{\alpha}_t}\mu_k^P, s_{k,t}^2) \cdot (-(y - \sqrt{\bar{\alpha}_t}\mu_k^P)/s_{k,t}^2)}{p_t(y)} = \sum_k r_k^P(y,t) \cdot \frac{-(y - \sqrt{\bar{\alpha}_t}\mu_k^P)}{s_{k,t}^2}.$$

□

The exact SMT correction (12) follows from Proposition 1 and Proposition 3. A constant shift $\delta_t = c$ would require the score difference $\partial_y \log q_t - \partial_y \log p_t$ to be independent of y. By Proposition 3, this holds only when the responsibility-weighted per-component scores of q_t and p_t differ by a constant, which fails in the intermodal region where responsibilities transition between components. No constant shift suffices.

A.5 Proof Chain Summary

For clarity, the logical dependencies among all results:

1. Radon-Nikodym theorem ⇒ Proposition 2.
2. Proposition 2 + Anderson reverse-time formula + score-noise relation ⇒ Proposition 1.
3. Proposition 1 + Lemma 1 + Proposition 3 ⇒ equation (12).

Every result follows from the single elementary observation in Proposition 2: a change of measure acts additively on the score.

B Connection to Classifier Guidance

B.1 Formal Comparison

Classifier guidance (Dhariwal & Nichol, 2021) modifies the noise predictor as:

$$\varepsilon_\theta^{\text{guided}}(x_t, t) = \varepsilon_\theta(x_t, t) - s \cdot \sqrt{1-\bar{\alpha}_t}\, \nabla_{x_t} \log p_\phi(c \mid x_t), \tag{16}$$

aiming at $p_\theta(x_0) \cdot p(c|x_0)^s$. As in Section 4, the comparison with SMT is a difference in what each method is given, not a hierarchy between them.

What the correction is. The SMT correction $\nabla \log(q_t/p_t)(x_t)$ is the score of the density ratio between the diffused target and source marginals. The classifier-guidance correction $s \nabla \log p_\phi(c|x_t)$ is the gradient of a trained discriminative model. By Bayes' rule, a Bayes-optimal classifier gives $\nabla \log p(c|x_t) = \nabla \log p_t(x_t|c) - \nabla \log p_t(x_t)$, which is itself a density-ratio score, for the class-conditional law against the marginal. At $s = 1$ the two are therefore the same construction applied to different pairs of distributions. In practice the classifier is estimated, and $s \neq 1$ moves the correction away from any density-ratio score.

What the sampler targets. By Proposition 1, the SMT-corrected reverse-time SDE has terminal law Q_0, in continuous time and with exact scores; a finite-step sampler carries discretization, base-model and correction-estimation error in addition. Classifier guidance with a perfect classifier and $s = 1$ targets $p(x_0|c)$ in the same sense. For $s \neq 1$ the guided scores are not the scores of any noised version of $p_\theta(x_0)\, p(c|x_0)^s$, so even a perfect classifier does not make the sampler target that tilted law.

What each requires. Classifier guidance requires labeled data to train $p_\phi(c|x_t)$ across noise levels, which amounts to samples from the target class. SMT in the mixture setting requires the analytical form of q_t/p_t. In the path setting it requires constraint functions and their targets, from which L^* is solved and h_ψ is regressed on samples from P_θ (Section 3.2). No SMT variant uses labels or samples from Q.

Martingale restrictions. In the path setting the martingale property enters as finitely many unconditional restrictions imposed on Q^* (Section 3.2), and Section 6 states what that leaves open. Guidance methods have no built-in analogue, because their target is defined by a condition rather than by constraints on expectations.

The classifier route. For any P and Q, and any class priors, the log-odds gradient of a Bayes-optimal domain classifier equals $\nabla \log(q_t/p_t)$, since the priors contribute only a constant (Appendix B.3). The practical difference lies in how the ratio is obtained. A domain classifier needs samples from both P and Q. In the mixture setting the ratio is available in closed form instead, and in the pricing setting samples from Q are not available at all, which is why the ratio is obtained from constraints (Section 3.2). In the mixture setting the closed form removes the need to learn a nonlinear decision boundary; it is exact as a formula, while the sampler that uses it still carries the base-model and discretization error described after Proposition 1.

B.2 Adjacent Work

Deep weighted Monte Carlo (Kunsági-Máté et al., 2022) learns WMC weights without amortizing into a generator. Deep hedging (Buehler et al., 2019) learns hedging policies from simulated paths; deep PDE solvers such as DGM (Sirignano & Spiliopoulos, 2018) price by solving the pricing equation of a model already specified under Q, and so presuppose the measure this paper constructs; and generative path models (Wiese et al., 2020; Ni et al., 2021) generate under P without imposing risk-neutral restrictions.

B.3 Connection to Density Ratio Estimation

The estimated SMT correction is related to density ratio estimation (Sugiyama et al., 2012). A binary classifier $f_\psi(y_t, t)$ trained to distinguish p_t (label 0) from q_t (label 1) with equal class priors satisfies at the Bayes optimum:

$$f^*_\psi(y_t, t) = \frac{q_t(y_t)}{p_t(y_t) + q_t(y_t)},$$

from which the log-density ratio and correction are:

$$\log \frac{q_t}{p_t} = \log \frac{f^*_\psi}{1 - f^*_\psi}, \quad \eta_t(y_t) = \nabla_{y_t} \log \frac{f^*_\psi(y_t, t)}{1 - f^*_\psi(y_t, t)}.$$

With unequal priors the log-odds shift by a constant, which leaves η_t unchanged.

In the mixture setting, the regression estimator of Appendix E.1 is trained against analytical targets η_t^* and needs no Q samples; the classifier route needs samples from both P and Q. That requirement is what rules the classifier route out for pricing, where Q is not observed as a dataset. Where dQ/dP is not available analytically, this paper obtains the ratio from constraints by entropy projection (Section 3.2) rather than by classification. A classifier trained on samples from an option-calibrated simulator would be an alternative, but it would inherit whatever that simulator assumes, and we do not pursue it.

B.4 Summary Table

Table 5 collects the comparison above in one place.

Table 5: SMT variants and guidance methods. ✓ holds; ≈ holds in the limit of perfect estimation and, for the guidance columns, only at guidance scale 1; × does not hold in general. †Finitely many unconditional restrictions (Section 3.2). The closed-form and estimated variants are shown on a mixture target, which carries no martingale restriction. Rows about terminal laws refer to the continuous-time dynamics; every finite-step sampler carries additional discretization error.

	SMT			Clf.	Clf.-
Property	closed form	est.	projected	guidance	free
Correction is a density-ratio score	✓	≈	≈	≈	≈
Target law given in advance	✓	✓	×	✓	✓
Continuous-time terminal law equals target	✓	≈	≈	≈	≈
Requires samples from the target	×	×	×	✓	✓
Target satisfies martingale restrictions	×	×	✓†	×	×
Closed form available	✓	-	×	×	×

C Extended Mixture Results

C.1 Complete Results Table

Table 6: Complete mixture results including the constant shift baseline. The constant shift corrects only the mean, reducing W_1 from 1.174 to 0.812 but leaving standard deviation unchanged at 1.574 versus the target 2.200. SMT (closed form) recovers standard deviation to 2.087 (95% of target) and reduces W_1 $6.8\times$ over the unshifted baseline and $4.7\times$ over the constant shift. The constant shift leaves the crash-regime tail essentially untouched ($P(Y<-3) = 0.098$ against 0.095 unshifted); the state-dependent correction recovers 0.332, against 0.341 under the true target.

Method	$W_1 \downarrow$	KS $\downarrow$	$P(Y<-3)$	Mean	Std
DDPM P	1.174	0.269	0.095	-0.449	1.574
Constant shift	0.812	0.192	0.098	-1.049	1.574
SMT (closed form)	0.172	0.050	0.332	-1.655	2.087
SMT (estimated)	0.171	0.047	0.333	-1.653	2.084
Retrained Q	0.075	0.021	0.329	-1.564	2.182
True Q	0.000	0.000	0.341	-1.621	2.200

The mixture setting is the controlled demonstration that the correction must be state-dependent. We transport $P = 0.90\,\mathcal{N}(0,1) + 0.10\,\mathcal{N}(-4,1)$ to $Q = 0.60\,\mathcal{N}(0,1) + 0.40\,\mathcal{N}(-4,1)$: identical component means and variances, differing only in weights. This is a pure regime reweighting with no location shift, so no constant additive correction can recover Q from P - which is exactly what a drift-style adjustment would supply.

The constant shift leaves the standard deviation at 1.574 against a target of 2.200 and moves the tail probability by 0.003; the state-dependent correction reduces W_1 from 1.174 to 0.172, recovers 97% of the target tail mass, and is tracked by the learned estimator ($W_1 = 0.171$; Table 6). Figure 2 shows

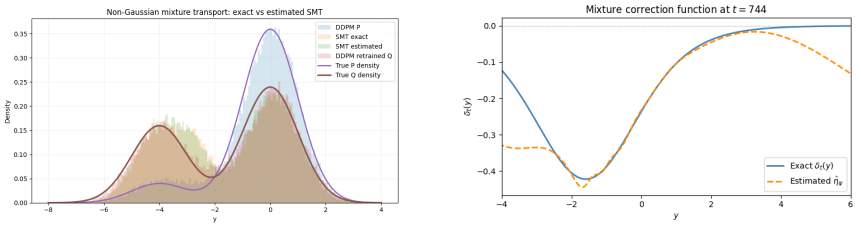

Figure 2: **Mixture setting.** *Left*: the corrected sampler recovers the heavier crash-regime mass of Q; the unshifted DDPM mirrors P. *Right*: the correction $\delta_t(y)$ is strongly negative near the crash component ($y \approx -2$ in standardised coordinates).

why: $\delta_t(y)$ is strongly negative near the crash component, suppressing normal-regime generation and amplifying crash-regime mass, and no constant approximates that shape.

The constant shift applies the closed-form correction for a pure mean shift to a non-Gaussian target, correcting only the mean while leaving the standard deviation unchanged at 1.574. Its performance ($W_1 = 0.812$) versus exact SMT ($W_1 = 0.172$) directly quantifies the value added by the state-dependent correction of Appendix A.1: a $4.7\times$ improvement in Wasserstein distance, and a tail probability of 0.332 against 0.098 for the constant shift and 0.341 under the true target. The constant shift cannot recover the crash-regime mass of Q because it applies the same additive displacement to every sample regardless of which regime the sample is near. The measure change here is a pure weight rebalancing ($\pi^P_{\text{crash}} = 0.10$) $\to$ ($\pi^Q_{\text{crash}} = 0.40$); only a state-dependent correction that treats normal-regime and crash-regime samples differently can implement it.

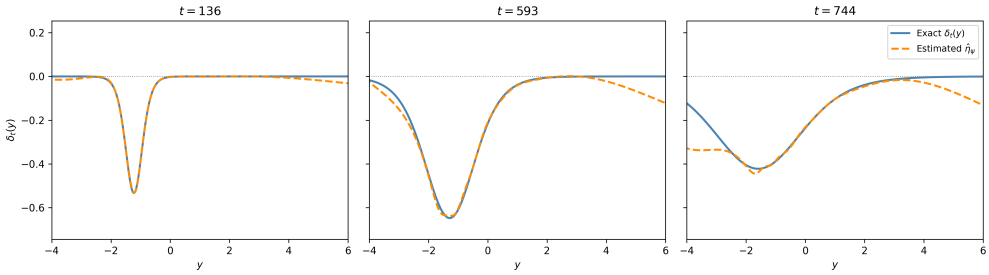

Figure 3: State-dependent correction $\delta_t(y)$ at three diffusion timesteps ($t = 136, 593, 744$), corresponding to $\bar{\alpha}_t \approx 0.95, 0.35, 0.15$: $t = 136$ is near-clean and $t = 744$ the noisiest of the three. At $t = 136$ the correction is sharply concentrated near the crash component mean ($y \approx -2$ in standardised coordinates), because near the data the two regimes are cleanly separated. At $t = 593$ added noise blurs the components and the correction broadens; it also deepens, as the growing prefactor $\sqrt{1-\bar{\alpha}_t}$ outweighs the flattening score difference. At $t = 744$ it is widest in support but smaller in peak magnitude: the two diffused marginals now overlap heavily, and their score difference flattens faster than the prefactor grows. The correction is negative throughout: Q assigns more weight to the crash regime than P, so the reverse diffusion must be suppressed in the normal-regime region at every noise level. The learned estimator (orange dashed) tracks the exact correction (blue) closely at all three timesteps, with the largest deviation at $t = 744$ near the tails.

C.2 Why the Correction is Negative Everywhere

Write $\Lambda_t(y) = q_t(y)/p_t(y)$. With shared components and weights $\pi^P = (0.90, 0.10)$ and $\pi^Q = (0.60, 0.40)$ on the normal and crash components,

$$\Lambda_t(y) = \frac{\pi^Q_{\text{c}}\,\varrho(y) + \pi^Q_{\text{n}}}{\pi^P_{\text{c}}\,\varrho(y) + \pi^P_{\text{n}}}, \qquad \varrho(y) = \frac{\mathcal{N}(y; \sqrt{\bar{\alpha}_t}\mu_{\text{c}}, s^2_{\text{c},t})}{\mathcal{N}(y; \sqrt{\bar{\alpha}_t}\mu_{\text{n}}, s^2_{\text{n},t})}.$$

As a function of ϱ this is increasing, since $\pi^Q_{\text{c}}\pi^P_{\text{n}} - \pi^Q_{\text{n}}\pi^P_{\text{c}} = 0.30 > 0$, and ϱ is decreasing in y because the crash component sits to the left. So Λ_t is decreasing in y, falling monotonically from

$\pi_{\mathrm{c}}^{Q}/\pi_{\mathrm{c}}^{P} = 4$ as $y \to -\infty$ to $\pi_{\mathrm{n}}^{Q}/\pi_{\mathrm{n}}^{P} = 2/3$ as $y \to +\infty$. Hence $\partial_y \log \Lambda_t < 0$ at every y, and the correction $\delta_t(y) = \sqrt{1-\bar{\alpha}_t}\,\partial_y \log \Lambda_t(y)$ is negative everywhere, vanishing in both tails where Λ_t flattens to a constant and largest in magnitude in the intermodal region where the responsibilities shift fastest.

The sign of the correction is the sign of the *derivative* of the log ratio, not of the ratio: $\Lambda_t > 1$ near the crash component does not make δ_t positive there. Since $\varepsilon^Q = \varepsilon^P - \delta_t$, a negative correction raises the predicted noise and moves each reverse step toward smaller y, that is, toward the crash regime. The net effect is the transfer of mass from the normal regime to the crash regime that the weight rebalancing $\pi^P \to \pi^Q$ requires, applied dynamically at every reverse step rather than post-hoc by importance reweighting.

C.3 Terminal Distribution Validation

Table 7: Terminal distribution statistics for the mixture setting. SMT recovers 95% of the target standard deviation and 97% of the true tail probability. The shortfall is not base-model capacity alone: a DDPM of the same architecture retrained on Q reaches 2.182.

Method	Mean	Std	W_1	KS
SMT (closed form)	-1.655	2.087	0.172	0.050
SMT (est.)	-1.653	2.084	0.171	0.047
Retrained Q	-1.564	2.182	0.075	0.021
True Q	-1.621	2.200	0.000	0.000

The residual variance shortfall (2.087 against 2.200) is not a capacity limit of the base architecture alone, since a DDPM of the same architecture retrained on Q reaches 2.182. Part of it therefore arises in the corrected sampler, consistent with the under-realisation diagnosed in Section 5.5.

D Path Prior, Projection, and Amortization Details

D.1 Sampling Algorithm

Algorithm 1 is the sampling loop of Section 3, with the correction supplied either by a target known in closed form or by the projection of Section 3.2.

Algorithm 1 Score Measure Transformation (SMT)

Require: Pretrained DDPM ε_θ, schedule $\{\beta_t\}_{t=1}^{T}$, and a correction source. Either ESTIMATED, a known target Q_0 in closed form (MIXTURE) or distilled into $\hat{\eta}_\psi$ (Appendix E.1); or PROJECTED, constraints g with targets c, giving h_ψ by Section 3.2, with no known Q_0

Ensure: Sample $z_0 \sim Q_0$ (ESTIMATED) or $z_0 \sim Q^*$ (PROJECTED)

Sample $z_T \sim \mathcal{N}(0, I)$
for $t = T, T-1, \ldots, 1$ **do**
  if ESTIMATED **then** ▷ target known in advance
    $\delta_t(z_t) \leftarrow$ Eq. (12), or $\sqrt{1-\bar{\alpha}_t}\,\hat{\eta}_\psi(z_t, t)$
  else ▷ PROJECTED: target defined by g, c
    $\delta_t(z_t) \leftarrow \sqrt{1-\bar{\alpha}_t}\,\nabla \log h_\psi(z_t, t)$, h_ψ from (9)
  end if
  $\varepsilon^Q \leftarrow \varepsilon_\theta(z_t, t) - \delta_t(z_t)$
  $z_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}}\left(z_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\varepsilon^Q\right) + \sqrt{\tilde{\beta}_t}\,\xi_t$
end for
return z_0

D.2 Path Settings

Heston dynamics (Heston, 1993) with $\kappa = 3.0$, $\theta = 0.04$, $\xi = 0.35$, $\rho = -0.7$, $\mu = 0.10$, $r = 0.05$, $S_0 = 100$, $H = 21$, $\Delta t = 1/252$. The benchmark risk-neutral world is fixed by a chosen

volatility risk premium ($\kappa_Q = 5.0$, $\theta_Q = 0.10$), putting the headline gap at $+2.57$ volatility points ($20.00\% \to 22.57\%$), where one-month index implied minus realized typically sits; gaps of $+0.25$, $+1.0$, $+2.57$ and $+4.0$ are swept (Appendix D.4). Constraint levels are nested: **C0** 41 martingale test functions only; **C1** + the ATM vanilla; **C2** + several strikes; **C3** + multiple maturities and strikes (53 constraints). Held-out sets of 15 test functions and 9 vanillas are never calibrated on. The path prior is a time-conditioned MLP over the 21-dimensional return vector; training and gate details follow in this appendix.

D.3 Heston Simulation

The Heston model specifies joint dynamics for asset price S_t and instantaneous variance v_t:

$$dS_t = \mu S_t\, dt + \sqrt{v_t}\, S_t\, dW_t^S, \tag{17}$$

$$dv_t = \kappa(\theta - v_t)\, dt + \xi\sqrt{v_t}\, dW_t^v, \tag{18}$$

with $d\langle W^S, W^v\rangle_t = \rho\, dt$ and the parameters of Appendix D.2. Paths are discretized via Euler-Maruyama:

$$S_{t+\Delta t} = S_t \exp\big((\mu - \tfrac{1}{2}v_t)\Delta t + \sqrt{v_t \Delta t}\, Z_1\big), \tag{19}$$

$$v_{t+\Delta t} = \max\big(v_t + \kappa(\theta - v_t)\Delta t + \xi\sqrt{v_t \Delta t}\, Z_2,\ 0\big), \tag{20}$$

where (Z_1, Z_2) are correlated standard normals with $\mathrm{Corr}(Z_1, Z_2) = \rho$, generated as $Z_1 = W_1$, $Z_2 = \rho W_1 + \sqrt{1-\rho^2}W_2$ with $W_1, W_2 \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. Vanilla benchmark prices are computed via the Carr-Madan FFT method (Carr & Madan, 1999) using the Heston characteristic function. The Feller condition $2\kappa\theta > \xi^2$ is satisfied under both measures (Appendix D.4), so in continuous time the variance process stays strictly positive almost surely; the truncated Euler scheme does not inherit that guarantee, which is why it floors v at zero.

D.4 The Heston-Q Benchmark

The evaluation benchmark is a Heston world under Q specified by a *chosen* volatility risk premium rather than a fitted one: we set $\kappa_Q = 5.0$ and $\theta_Q = 0.10$, holding v_0, ξ and ρ at their physical values, since v_0 is a state variable and not a parameter of the measure change. The volatility gap is an output of that choice, not an input to it. Every "span closed" figure in Section 5.3, and the widths of Table 11 relative to the physical-to-risk-neutral span, are measured against this benchmark.

The consequence for pricing at this horizon is summarised by the mean variance,

$$\frac{1}{T}\int_0^T \mathbb{E}_Q[v_t]\, dt \;=\; \theta_Q + (v_0 - \theta_Q)\,\frac{1 - e^{-\kappa_Q T}}{\kappa_Q T}, \tag{21}$$

which gives $\sigma_Q(21\text{d}) = 22.57\%$ against $\sigma_P(21\text{d}) = 20.00\%$, a gap of 2.57 volatility points. Because $v_0 = \theta_P$, $\sigma_P(T)$ equals 20% for every horizon, so the gap is exact rather than horizon-dependent. The same identity inverted at fixed κ_Q,

$$\theta_Q = \frac{\sigma^2 - v_0 f}{1 - f}, \qquad f = \frac{1 - e^{-\kappa_Q T}}{\kappa_Q T}, \tag{22}$$

is used in three of the four arms of the robustness sweep. The Feller margin $2\kappa\theta - \xi^2$ is $+0.1175$ under P and $+0.8775$ under Q.

The sweep comprises four arms:

Arm	Construction	κ_Q	θ_Q	$\sigma_Q(21\text{d})$	Gap
`lam_v` $= -1$	single-parameter convention	2.650	0.04991	20.25%	+0.25
21.0%	target-vol inversion	5.000	0.06251	21.00%	+1.00
22.57%	target-vol inversion	5.000	0.10005	22.57%	+2.57
24.0%	target-vol inversion	5.000	0.13661	24.00%	+4.00

The first arm is not a target-vol inversion. It applies the single-parameter convention $\kappa_Q = \kappa_P + \xi\lambda_v$ and $\kappa_Q\theta_Q = \kappa_P\theta_P - \xi\rho(\mu - r)$, which at $\lambda_v = -1$ gives $\kappa_Q = 2.650$ and $\theta_Q = 0.13225/2.650 = 0.049906$. Note that $\kappa\theta$ is *not* preserved. The correlation term $-\xi\rho(\mu - r)$ is what moves the variance level: with $\rho < 0$ and $\mu > r$ it is positive, $+0.01225$ here, and raises it. Preserving $\kappa\theta$ would instead give $\theta_Q = 0.04528$ and a gap of $+0.14$ points, against the $+0.25$ this arm reaches.[1]

Results across the four arms. Projection quality at $N = 10^5$, with held-out vanilla RMSE in units of 10^{-3}:

	ESS (%)				Held-out vanilla RMSE ($\times 10^{-3}$)					
Gap	C0	C1	C2	C3	C0	C1	C2	C3	Q floor	Improvement
$+0.25$	99.41	99.41	99.41	99.38	2.88	1.59	1.53	1.23	6.06	$196\times$
$+1.00$	99.41	99.04	98.82	98.78	47.44	14.26	1.37	1.11	7.70	$192\times$
$+2.57$	99.41	96.03	93.91	93.55	151.62	39.15	3.12	2.83	12.19	$64\times$
$+4.00$	99.41	90.83	83.91	81.76	249.69	63.66	5.55	4.63	16.54	$43\times$

ESS at C0 is identical across arms by construction, since C0 contains no vanilla constraints. "Improvement" is the C3 RMSE against the unweighted Heston-P measure for that arm (0.2408, 0.2135, 0.1811, 0.2001), and C3 RMSE sits below the Q-sample Monte Carlo floor in all four.

Exotic prices at C3, against the Heston-Q benchmark for each arm. The physical values are fixed at 1.61319 (Asian), 2.05449 (barrier) and 4.43604 (lookback); "span" is physical minus benchmark for that arm.

	Asian			UO barrier				Lookback	
Gap	C3	Heston-Q	span	C3	Heston-Q	span	closed	C3	Heston-Q
$+0.25$	1.4961	1.4912	$+0.1220$	1.9325	1.9181	$+0.1364$	89%	4.2226	4.2224
$+1.00$	1.5383	1.5171	$+0.0961$	1.8973	1.8612	$+0.1933$	81%	4.2900	4.3501
$+2.57$	1.6261	1.5721	$+0.0410$	1.8460	1.7653	$+0.2892$	72%	4.4317	4.6254
$+4.00$	1.7070	1.6238	-0.0106	1.7937	1.6682	$+0.3863$	68%	4.5631	4.8772

Lookback spans are $+0.2136$, $+0.0859$, -0.1894 and -0.4412. The barrier is the stable result: its span grows monotonically with the gap and the projection closes 89, 81, 72 and 68% of it, a fraction that decreases monotonically as the gap widens. The Asian and lookback fractions are not interpretable, because their physical-to-risk-neutral spans shrink toward zero and change sign as the gap varies. The Asian benchmark crosses the physical value between $+2.57$ and $+4.00$, and the lookback between $+1.00$ and $+2.57$. The resulting fractions reach 884% for the Asian at $+4.00$ and 170% for the lookback at $+1.00$, artifacts of a near-zero denominator rather than large effects, which is why both are reported here as price levels with the span shown beside them.

The 22.57% arm is a near-duplicate of the headline configuration rather than the headline itself: it inverts a target volatility and lands on $\theta_Q = 0.10005401$, against $\theta_Q = 0.10$ exactly for every main-text result, a difference of 0.0022 volatility points. The two agree on volatility to two decimals but are different parameter sets, so this row's prices differ from Table 13 in the fourth decimal and should not be read as the same numbers: this arm's benchmarks are 1.57214, 1.76527 and 4.62539, against 1.57207, 1.76547 and 4.62501 for the headline.

Standard errors are available for the widest arm only, and apply to $+4.00$ alone: 0.01237 (Asian), 0.00885 (barrier) and 0.02116 (lookback) at C3.

D.5 Path DDPM Training

The prior is a time-conditioned MLP mapping the 21-dimensional return vector to itself, which already represents joint dependence across steps without a sequence architecture. Returns are globally standardized. Sampling uses a frozen ancestral DDPM sampler started at $t_{\text{start}} = 998$: the clipped

[1] The 22.57% arm uses the inverted $\theta_Q = 0.10005$, while the headline runs use $\theta_Q = 0.10$ exactly. The two differ in the fifth decimal and ESS agrees to 0.01 percentage points, but they are not the same parameters: the sweep reproduces the headline gap in volatility, not bitwise.

final cosine step ($t = 999$, $\beta = 0.999$) derails 21-dimensional chains, producing 57.7% rejection against 0% from $t = 998$, with results identical for any start in $[950, 998]$. A rejection cap at $|z| > 8$ binds on nothing (training maximum 6.16; 0 of 4×10^5 generated paths exceed 7). Three disjoint draws are used throughout, with the dual solved on one only.

A level bias from optimizer noise. Acceptance thresholds were recorded before training, together with an escalation ladder of remedies. The ladder was falsified: longer training (450 epochs) and Min-SNR-γ weighting both made the criteria *worse*, and the specific prediction that SNR weighting would reduce high-noise $\hat{\varepsilon}$ bias was wrong (0.033 to 0.036). The diagnosis came from outside the ladder. The bias is the last-iterate noise floor of constant-learning-rate AdamW without decay or averaging: it is smooth in t (adjacent-step correlation 0.8-0.9, so not the time embedding), the same magnitude across epoch counts and loss weightings, differently signed on each run, and drops $2.5\times$ under five epochs at lr $= 10^{-4}$. Cosine decay $10^{-3} \to 10^{-5}$ with EMA 0.999 fixes it.

Table 8: The escalation ladder and the fix. Root-mean-square per-coordinate bias barely moves across the three rungs, but its *sum* does, and the exotics track the sum.

	Rung 0 (150 ep)	Rung 1 (450 ep)	Rung 2 (Min-SNR-γ)	lr decay + EMA
Tail / mean criterion	0.041 / 0.197	0.148 / 0.137	0.173 / 0.150	**0.0075 / 0.0106**
Asian (SE)	5.7	25.6	35.2	**1.36**
Column bias rms / sum	0.077 / $+0.13$	0.069 / -0.56	0.074 / -0.73	0.074 / $\mathbf{-0.022}$
$\hat{\varepsilon}$ bias rms at t=998	0.032	0.033	0.036	**0.0046**

The substantive point is in the third row. Per-coordinate bias is essentially unchanged across rungs, but a per-coordinate high-noise $\hat{\varepsilon}$ bias accumulates into a *level shift* in S_{21} whose severity is governed by sign alignment across coordinates rather than by magnitude. This is specific to path-valued diffusion, where the coordinates are summed into a terminal price, and it is not visible in any per-coordinate summary.

D.6 Frozen-Prior Acceptance Criteria

Table 9 lists the criteria the frozen prior had to meet, the value it achieved on each, and the bound; null values are simulator-against-simulator baselines.

Table 9: Acceptance criteria for the frozen prior, evaluated at 10^5 samples. Null values are simulator-against-simulator baselines. The split, fixed in advance, of the first criterion was never exercised: 0 of 77 columns exceeded the uniform bound.

Check	Value	Bound
Tail means (21 vanilla)	0.0075	≤ 0.02 (null 0.0099)
Martingale means (56 columns)	0.0106	≤ 0.05 (null 0.0092)
Martingale sd ratio	0.0245	≤ 0.05
Vanilla sd ratio (K=115)	0.0576	≤ 0.10 (null 0.052)
Asian / barrier / lookback	1.36 / 0.89 / 1.49 SE	≤ 5
$\mathrm{CV}(\mathrm{RV}_{21})$	0.405 (Heston 0.409)	$[0.385, 0.435]$
$\mathrm{ACF}_1(Y^2)$	0.037 (0.041)	$[0.030, 0.050]$
Leverage (5-day)	-0.104 (-0.103)	$[-0.129, -0.077]$

D.7 Dual Re-Solved on the Learned Prior

Table 10 compares the dual solved on 10^6 draws from P_θ against the raw-simulator solution of Table 16.

Table 10: Dual re-solved on P_θ samples (10^6 solve draw), against the raw-simulator values of Table 16. The feasibility screen first binds at the 10-day $K{=}108$ column, at essentially the margin seen on the simulator (0.147).

Level	Screen margin	$\|\beta^*\|$ (P_θ / sim.)	ESS % (P_θ / sim.)	Held-out van. (P_θ / sim.)
C0	4.28	17.8 / 5.8	99.36 / 99.41	0.158 / 0.152
C1	0.81	19.5 / 8.4	95.76 / 96.04	0.0397 / 0.0391
C2	0.29	37.4 / 39.2	93.78 / 93.92	0.0039 / 0.0031
C3	0.149	34.9 / 33.3	93.40 / 93.56	0.0031 / 0.0028

D.8 Amendment to the Frozen-Prior Criteria

We amended one acceptance criterion partway through the project, after it had already rejected a prior. We record the amendment here in full, because it readmits that prior and because it widens the reference band against which our own prior-dependence claims are measured.

The original criterion (G2b) bounded the per-column ratio of vanilla standard deviations, and `retrain1` failed it on one column of twenty-one: the 21-day $K{=}115$ strike, at 0.870 against a 0.90 bound. Before deciding what to do about the failure, we specified a check: how often does the criterion reject when nothing is wrong? Two independent simulations of the *same* Heston model, at $n = 10^5$ paths, fail G2b 27% of the time. The failures are driven almost entirely by that one column. The $K{=}115$ call pays off on 0.3% of paths, so roughly 300 of the 10^5 carry its estimate; with a sample kurtosis of about 1590, the null standard deviation of a ratio of two such estimates is $\sqrt{(\kappa-1)/2n} \approx 9\%$, and the empirical null matches that to within 10%. A 0.10 bound on that column is therefore tighter than the noise, and rejects sound priors at a rate the criterion never intended.

The check was specified before its outcome was known; the amendment followed from it. G2b is replaced by a studentised maximum: each column's ratio is standardised by its own null standard deviation, and the maximum over columns is compared against a threshold calibrated so that two independent simulations of the same model fail 5% of the time, against 27% before. The replacement is *stricter* on twenty of the twenty-one columns and looser only in the far wing, where the original bound was below the noise floor. `retrain1` passes under it, and its binding column moves off $K{=}115$.

The honest cost is that this widens the band of same-recipe priors in Section 5.4, which is the reference against which we measure prior dependence. A wider reference band makes the separation of the misspecified priors harder to establish, not easier, so the amendment works against the claims it sits beside. It does not affect any constraint, projection or sampling result.

D.9 Corrupted-Drift Prior

Under a drift corrupted to $2.5\times$ its true value, the martingale block works visibly harder: $\|\beta_{\text{raw}}\|$ at C0 rises $1.35\times$ and ESS falls from 99.32% to 91.76%. It absorbs 92-98% of the corruption; the residual at C3 sits 6-13$\times$ outside the retrain band of Section 5.4.

D.10 Divergence Between the Trained Priors

Relative entropy from the base Heston-P law to the laws the misspecified priors were trained on, on the full (S, v) path by exact CIR sampling, is at most 0.0523 nats for `mu25` and 0.0334 for `kappa6`, which are upper bounds for return paths alone, since v is latent. `rho02` changes the diffusion coefficient, making the laws mutually singular in continuous time, so its divergence is infinite there; on the 21-step discrete path it is finite, but we do not estimate it. Since `rho02` carries the largest price effect of the three, the divergence scale below does not cover it. At 0.10 nats of total budget, comparable to these divergences, the constraints already admit intervals 0.30-0.91 wide. Where the residual prior dependence does show is in exotics depending on the coupling of the path through time rather than on terminal marginals, the direction finitely many vanillas cannot reach.

Table 11: Constraint-set price bounds at C3, over weights on the baseline prior's 10^6-path draw satisfying the same 53 constraints. Prices here are from that 10^6-path draw, and the bounds are for the $+2.57$ arm only. Strikes: Asian $K{=}100$; up-and-out barrier $K{=}100$, $B{=}110$; lookback floating-strike. "Position" locates a price within its own interval. Absolute prices, as Q^* / Heston-Q / Heston-P: Asian 1.6247 / 1.5721 / 1.6132; barrier 1.8297 / 1.7655 / 2.0545; lookback 4.4251 / 4.6250 / 4.4360.

				Width relative to		Position in interval (%)		
Exotic	Lower	Upper	Width	price	span	Q^*	Heston-Q	Heston-P
Asian	0.2420	3.1096	2.8676	1.8×	70×	48.2	46.4	47.8
UO barrier	0.0000	2.3473	2.3473	1.3×	8×	77.9	75.2	87.5
Lookback	2.7976	9.6204	6.8228	1.5×	36×	23.9	26.8	24.0

D.11 Constraint-Set Price Bounds

Table 11 gives the bounds behind Section 5.3, with each benchmark's position inside its own interval.

Two approximations separate these intervals from true martingale optimal transport bounds, and they pull in opposite directions. Restricting to a finite sample narrows them: enlarging the draw from a 50k prefix to 10^6 widened the Asian 16% and the lookback 22%. Imposing only finitely many *unconditional* martingale restrictions, rather than the conditional property that MOT imposes, widens them, since it admits measures true MOT excludes. We do not know which effect dominates, so the intervals below are not bounds on the MOT interval in either direction. They are exact for the question actually posed, the range over measures on the draw satisfying the same 53 constraints, with 54-atom support and non-negative reduced costs at termination. Bounding over measures rather than sample weights is its own numerical literature (Guo & Obłój, 2019).

Table 12: Price intervals by relative-entropy budget, over measures on the baseline prior's 10^6-path draw meeting the same 53 constraints at C3. The budget is total $\mathrm{KL}(\cdot \,\|\, P_\theta)$; the floor 0.02869 nats is $\mathrm{KL}(Q^* \| P_\theta)$, below which no constraint-consistent measure lies. "sat." marks the barrier's upper branch attaining its linear-programming bound of 2.3473 at $\mathrm{KL} \approx 0.17$, after which further budget buys nothing upward because the payoff is bounded above. The last row is the unbudgeted linear program of Table 11.

KL (nats)	Asian		UO barrier		Lookback	
	lower	upper	lower	upper	lower	upper
floor 0.02869	1.6247		1.8297		4.4250	
0.05	1.5422	1.7045	1.5707	2.0857	4.2349	4.6235
0.10	1.4724	1.7693	1.3554	2.2651	4.0778	4.7976
0.25	1.3503	1.8777	1.0779	sat.	3.8470	5.1031
0.50	1.2001	1.9939	0.8156	sat.	3.6154	5.4697
1.00	0.9938	2.1494	0.4699	sat.	3.3372	5.9889
∞ (LP)	0.2420	3.1096	0.0000	2.3473	2.7976	9.6204

The intervals widen as $\sqrt{\mathrm{KL} - \text{floor}}$, the local quadratic behaviour. Writing the full interval width as $C\sqrt{\Delta\mathrm{KL}}$ with $\Delta\mathrm{KL}$ measured above the floor, C is 1.112 for the Asian at both 0.05 and 0.10 nats and 2.68 for the lookback throughout. These are values read at finite budget, which is what Table 12 tabulates. The barrier's upper branch is the exception, attaining its linear-programming bound at $\mathrm{KL} \approx 0.17$ because the payoff is bounded above.

D.12 Budgeted Intervals Under a Parametric Reference

Repeating the budgeted-interval construction of Table 12 with the Heston-P simulator in place of the learned prior tests whether the identification diagnostics depend on which reference is used. Both arms impose the same 53 constraints at C3; the floors differ slightly, 0.02700 under the simulator against 0.02869 under the learned prior.

Table 13: Exotic prices under the projected measure, from the headline configuration ($\theta_Q = 0.10$ exactly). "Span closed" is the fraction of the Heston-P to Heston-Q gap closed at C3. The barrier responds to the vanilla constraints; the lookback does not move; the Asian moves away from the benchmark. Stable across seeds (sd $\leq 1.0\times$ the within-run standard error).

Exotic	Heston-P	C0	C1	C2	C3	Heston-Q	Span closed
Asian K=100	1.6132	1.4970	1.6162	1.6256	1.6260	1.5721	-31%
UO barrier K=100, B=110	2.0545	1.9304	2.0199	1.8374	1.8461	1.7655	$+72\%$
Lookback (floating)	4.4360	4.2268	4.4147	4.4334	4.4315	4.6250	-2%

Table 14: Exotic-price spread at C3 across priors. The retrain band is the spread over three same-recipe priors differing only in seed; the misspecified priors are reported as multiples of that band, per exotic. The final column averages the per-exotic spread as a fraction of the physical-to-risk-neutral span.

Prior	Asian	Barrier	Lookback	Mean (% of span)
Retrain band (absolute)	0.0011	0.0009	0.0043	0.3-2.7
`retrain1`		within band		-
`retrain2`		within band		-
`kappa6`	0×	10×	2×	2.7
`mu25`	6×	13×	8×	13.3
`rho02`	4×	78×	29×	33.0

	Asian		UO barrier		Lookback	
KL (nats)	simulator	learned	simulator	learned	simulator	learned
0.05	[1.5357, 1.7096]	[1.5399, 1.7069]	[1.5565, 2.1046]	[1.5617, 2.0897]	[4.2254, 4.6448]	[4.2296, 4.6321]
0.10	[1.4648, 1.7752]	[1.4676, 1.7737]	[1.3527, 2.2776]	[1.3538, 2.2700]	[4.0748, 4.8206]	[4.0757, 4.8117]
0.25	[1.3396, 1.8845]	[1.3409, 1.8832]	[1.0626, sat.]	[1.0641, sat.]	[3.8368, 5.1379]	[3.8367, 5.1296]
0.50	[1.2003, 1.9989]	[1.1992, 1.9971]	[0.8027, sat.]	[0.8067, sat.]	[3.6080, 5.4935]	[3.6077, 5.4849]
1.00	[0.9985, 2.1547]	[0.9910, 2.1515]	[0.4616, sat.]	[0.4679, sat.]	[3.3322, 6.0123]	[3.3338, 6.0018]
∞ (LP)	[0.2699, 3.1302]	[0.2420, 3.1096]	[0.0000, 2.3473]	[0.0000, 2.3473]	[2.7976, 9.6769]	[2.7976, 9.6204]

"sat." marks the barrier's upper branch at its linear-programming bound of 2.3473, as in Table 12. Interval widths agree within about 1% at every budget from 0.10 nats upward. At 0.05 nats they differ by 3.8-4.2% across the three exotics, which is what the two floors imply rather than a difference between the references: by the $\sqrt{\text{KL} - \text{floor}}$ law of Appendix D.20, the excess budgets 0.0230 and 0.02131 predict a width ratio of $\sqrt{0.0230/0.02131} = 1.039$, a 3.9% gap.

The one substantive difference is the Asian's unbudgeted lower bound, 0.2699 against 0.2420, an 11% gap. That bound is set by a handful of extreme atoms, so it is sample-sensitive rather than law-sensitive; the corresponding widths still agree to 0.25%. The identification diagnostics are therefore the same whether the reference is the correctly specified parametric model or the learned prior.

D.13 Exotic Prices Against the Heston-Q Benchmark

Table 13 gives the C3 exotic prices at each constraint level against the Heston-Q benchmark, and the fraction of the physical-to-risk-neutral span closed. Section 5.3 explains why this is a secondary reading.

D.14 Prior Sweep: Per-Exotic Spread

Table 14 gives the C3 spread behind Section 5.4: the retrain band in absolute price units, and each misspecified prior as a multiple of that band, per exotic.

Table 15: Entropy projection at the $+2.57$-point gap, against 10^5 Heston-P paths (sim.) and 10^6 draws from the learned prior P_θ. Held-out columns are never calibrated on; the Q floor is Monte Carlo noise on the benchmark itself. Full diagnostics in Tables 16 and 10.

		ESS (%)		Held-out	Held-out vanilla RMSE	
Level	m	sim.	P_θ	martingale (sim.)	sim.	P_θ
Unprojected	-	100	100	0.01793	0.1811	-
C0	41	99.41	99.36	0.00085	0.1515	0.158
C1	42	96.04	95.76	0.01238	0.0391	0.0397
C2	46	93.92	93.78	0.01319	0.0031	0.0039
C3	53	93.56	93.40	0.01322	**0.0028**	0.0031
Q floor	-	-	-	0.00243	0.0122	-

Averaged over the three exotics as a fraction of the span, the observed ordering is κ 2.7% $<$ μ 13.3% $<$ ρ 33.0%, against an ordering $\mu < \kappa < \rho$ fixed in advance: we record that prediction as falsified as stated, with only the claim that ρ dominates surviving. The perturbation sizes were not normalized. Each parameter was moved by a different amount, so the ordering reflects how far each was moved as much as how strongly it shapes the path law, and the divergences of Appendix D.10 differ between the perturbed laws (0.0523 nats for `mu25`, 0.0334 for `kappa6`, infinite for `rho02`). Comparing parameters on equal footing would require perturbations normalized to a common divergence, which a change in the diffusion coefficient does not admit.

D.15 Amortization Unit Test

The claim that a frozen physical diffusion plus a learned score correction generates Q^* is testable directly, by comparing SMT-generated paths against weighted resampling of P_θ under L^*. On 10^5 corrected paths with zero rejections: all 53 calibrated constraint columns agree within 2 standard errors, as do all 24 held-out columns; exotics agree within 1 standard error on that draw. This does not hold in general, since against the weighted Q^* on the same pool (Appendix D.21) two of three exotics sit 2.4-2.5 standard errors away and held-out columns 2.08 rms; martingale deviation is 5.7×10^{-4} against 5.8×10^{-4}. Sliced Wasserstein distance between the two is 0.00866 against a null of 0.00843, where the signal being tested - unweighted P_θ against the corrected sampler - is 0.02795. The agreement is also flat in reverse-step count from $K = 999$ down to $K = 100$, so neither discretization nor regression error is visible above sample noise at this scale.

D.16 Exotics on the Corrected Sampler

At C3 on a fresh 10^5-path draw from the corrected sampler, against Table 13 and the Heston-Q benchmark: Asian 1.628 ± 0.008 / 1.626 / 1.572; barrier 1.844 ± 0.009 / 1.846 / 1.766, closing the same 72% of span; lookback 4.433 ± 0.015 / 4.432 / 4.625. The learned prior behaves as the simulator does, which is what the gate existed to protect.

D.17 Projection Quality by Constraint Level

Table 16 gives the projection against raw Heston-P paths at each constraint level summarized in Section 5.2.

D.18 The Calibration-Martingality Trade-Off

On a tier-2 held-out set used in neither arm, adding 15 further martingale constraints moved tier-2 martingale RMSE from $1.98\times$ to $1.70\times$ the noise floor without reaching it, while degrading vanilla generalization $1.5\times$ on tier 1 and $6.3\times$ on tier 2. It is therefore not an artifact of a thin test-function family. Nor is it a property of the constraint set, which Heston-Q satisfies in full, calibrated and held-out columns alike: it arises from selecting one measure by minimum relative entropy under finitely many constraints on a finite sample.

Table 16: Projection against raw Heston-P paths at the $+2.57$-point gap. Held-out columns were never calibrated on. The Q floor row is Monte Carlo noise on the benchmark itself. Seed-to-seed standard deviation of ESS is at most 0.23 percentage points over three seeds.

Level	m	ESS %	max w	$\|\beta^*\|$	KL (nats)	held-out mart.	held-out van.
Heston-P	-	100	1.0	-	-	0.01793	0.1811
C0	41	99.41	1.6	5.8	0.0030	0.00085	0.1515
C1	42	96.04	4.3	8.4	0.0185	0.01238	0.0391
C2	46	93.92	11.7	39.2	0.0262	0.01319	0.0031
C3	53	93.56	14.6	33.3	0.0265	0.01322	0.0028
Q floor	-	-	-	-	-	0.00243	0.0122

Table 17: Share of payoff variance spanned by the constraints, by level. Each row projects the payoff onto that level's columns under that level's own Q^*, on the baseline prior's 10^6 draw; $\mathrm{sd}(f_\perp)$ is the residual standard deviation. Column sets are nested, $41 \subset 42 \subset 46 \subset 53$.

		Asian		UO barrier		Lookback	
Level	Columns	R^2	$\mathrm{sd}(f_\perp)$	R^2	$\mathrm{sd}(f_\perp)$	R^2	$\mathrm{sd}(f_\perp)$
C0	41	0.804	0.8891	0.468	1.9102	0.817	1.4494
C1	42	0.915	0.6212	0.453	2.0183	0.918	1.0319
C2	46	0.925	0.6108	0.737	1.3543	0.929	1.0068
C3	53	0.967	0.4046	0.757	1.3061	0.933	0.9762

D.19 Span by Constraint Level, at Fixed Measure

Table 17 gives the span at every constraint level, behind the C3 column of Table 2.

Each row of Table 17 changes two things at once: the column set and the measure the projection is taken under. Nesting ($41 \subset 42 \subset 46 \subset 53$, verified) forces weak monotonicity of R^2 only at a fixed measure, so the barrier's $0.468 \to 0.453$ step at C0 to C1 is not evidence that adding a constraint lost span. Re-running the same nested column sets under one held-fixed Q^* gives monotone R^2 everywhere, the barrier included: 0.468, 0.511, 0.776, 0.791 under Q^*(C0), and 0.375, 0.386, 0.733, 0.757 under Q^*(C3). The span therefore grows at every level for all three payoffs, and the dip is a measure effect, since the first vanilla moves Q^* enough (barrier price $1.916 \to 2.005$, total standard deviation $2.618 \to 2.730$) to outrun the single column it adds.

D.20 The Widening Constant

By the Pythagorean identity (Csiszár, 1975), $\mathrm{KL}(Q\|P_\theta) = \mathrm{KL}(Q\|Q^*) + \mathrm{KL}(Q^*\|P_\theta)$ for every constraint-consistent Q, so the measures within Δ of the floor are exactly those within Δ of Q^* and the local expansion of Lam (2016) applies around Q^*. It predicts $C = 2\sqrt{2\,\mathrm{Var}(f_\perp)}$, where $f_\perp$ is the payoff after projecting out the constraints; what follows is a numerical check of that prediction on our draw, not an independent result. Projecting each payoff under Q^* at C3 on the 10^6 draw onto the 53 *calibrated* constraint columns by weighted least squares, using the same columns the dual solves rather than the full 77, gives:

Payoff	$\mathrm{sd}(f_\perp)$	Projection R^2	Predicted C	Measured C	Agreement
Asian	0.4046	0.967	1.1444	1.1370	0.6%
UO barrier	1.3061	0.757	3.6941	3.7818	2.4%
Lookback	0.9762	0.933	2.7611	2.7446	0.6%

Measured C is read from the curves and extrapolated to $\Delta\mathrm{KL} \to 0$, which is the only limit in which the constant is defined. An independent route through the same artifact gives ratios of 1.000, 0.974 and 1.003. It uses the derivative $d\mathbb{E}_Q[f]/d\gamma \to \mathrm{Var}(f_\perp)$ at $\gamma = 0$, reading price against γ rather than width against $\sqrt{\Delta\mathrm{KL}}$. Sampling error is negligible: enlarging the draw from 10^5 to 10^6 moves $\mathrm{sd}(f_\perp)$ by under 0.2%.

C is not constant along the curve, and the two branches differ away from the limit: for the Asian the upper branch runs $1.143 \to 1.064$ across the budget range while the lower runs $1.134 \to 1.290$. The 1.112 quoted above is the full-width coefficient at the finite budgets of Table 12, and happens to coincide with the Asian upper branch near $\Delta\mathrm{KL} \approx 0.111$; comparisons against theory use the extrapolated values in the table here, not that figure.

The barrier is the weak case on every count: the lowest projection R^2 at 0.757, the worst constant match at 2.4%, and an upper branch that saturates at its linear-programming bound of 2.347 by $\Delta\mathrm{KL} \approx 0.14$ because the payoff is bounded above. For that instrument the $\sqrt{\Delta\mathrm{KL}}$ law describes the interval only very near the floor.

D.21 Amortization Against Simpler Alternatives

At C3 in the Heston world, each method is compared against the weighted Q^* on the same pool, which is the reference an amortized sampler should reproduce. Standard errors are those of that reference.

	Constraint columns (rms, SE)			
Method	calibrated	held-out	Exotics (SE)	Cost
SMT (this paper)	1.46	2.08	+2.41 / +0.69 / +2.53	19.7 s per 10^5; h in ≈1 min
Diffusion retrained on L^*-resampled paths	0.94	1.18	all within 1	9.4 s per 10^5; ≈8 min
Resampling from a 10^6 pool	0.90	1.11	within 1.6	-
Resampling from a matched 10^5 pool	1.64	1.83	up to 3.7	10.2 s; 62% unique

Exotics are Asian / barrier / lookback. Retraining reproduces Q^* more accurately than SMT at lower per-sample cost; SMT adapts about eight times faster, since only h_ψ is trained. Resampling from a large pool is accurate, but from a matched 10^5 pool only 62% of draws are unique and its errors exceed SMT's. An oracle diffusion trained on Heston-Q paths comes closest to Heston-Q, as it must; it requires target samples, which the setting does not provide.

D.22 Dimension Scaling on a Gaussian Mixture

A Gaussian mixture has an exact Q^*, so each method's sliced Wasserstein distance to it can be measured directly, and we rank methods on that distance. Two sweeps run from $d = 2$ to 128. In (a) the tilt is per-coordinate, so KL rises from 0.05 to 2.7 and ESS falls from 90% to 1.65%; in (b) KL is fixed at 0.5 and ESS stays near 40%. Weighting uses a 5×10^5 pool with resampling.

Sweep	Row	$d = 2$	$d = 8$	$d = 32$	$d = 128$
(a) per-coordinate tilt	ESS	90.15%	64.91%	21.50%	1.65%
	SMT	0.01117	0.02856	0.09735	failed[†]
	weighting + SIR	0.01177	0.01047	0.00969	0.01238
	oracle	0.01652	0.00907	0.01330	0.02471
	resampling-retrain	0.01050	0.01073	0.02879	failed[‡]
	β err., L_∞	0.0047	0.0060	0.0140	0.0461
	β err., rel. L_2	0.0252	0.0187	0.0449	0.1018
(b) fixed KL $= 0.5$	ESS	≈ 40% throughout (39.75% at $d = 128$)			
	SMT	0.02101	0.04460	0.03872	0.14672
	weighting + SIR	0.00901	0.00929	0.01196	0.01385
	oracle	0.01560	0.00985	0.01216	0.03011
	resampling-retrain	0.00900	0.01361	0.02227	0.03845
	β err., L_∞	0.0076	0.0085	0.0133	0.0178
	β err., rel. L_2	0.0121	0.0189	0.0536	0.1214
both	base model P_θ to exact P	0.01120	0.00938	0.01514	0.01874

Sliced Wasserstein distances to the exact Q^*, except the last row, which is the base model's distance to the exact P. [†]A failed run, not evidence of a limit: the adaptive floor rule, half the minimum

L^*, set h_ψ's clamp so low that training stalled and h_ψ collapsed to its floor. The failure is one of optimization, caused by the floor rather than by a capacity limit. The degradation claim therefore rests on the remaining cells: in (a) SMT's error rises $0.011 \to 0.029 \to 0.097$ as KL rises from 0.05 to 0.95, and in (b) it rises $7\times$ across dimension at fixed KL, against $1.67\times$ for the base model. [‡]The arm that retrains a diffusion on resampled paths fails at $d = 128$ in sweep (a) for a separate reason: about 8,250 distinct atoms are replicated into 10^5 training rows.

Weighting also beat the oracle diffusion, trained on samples from the exact Q^*, on sliced Wasserstein in every cell at $d \geq 32$. It reuses P_θ's samples and applies the tilt exactly, so it is not bounded by the accuracy of fitting a second diffusion. The prediction, fixed in advance, that SMT would overtake weighting as dimension grew is falsified; SMT is reliable only at mild tilts, $\mathrm{KL} \lesssim 0.05$, which includes the Heston operating point at 0.029. β recovery degrades with dimension in both sweeps, and the two error measures must be read separately. The L_∞ column is the maximum component error and the relative column is $\|\hat{\beta} - \beta\|_2 / \|\beta\|_2$; read from L_∞ alone, (a) degrades $9.8\times$ against (b)'s $2.3\times$, which would suggest that falling ESS drives the error. The relative figures reverse that: (b) degrades $10.0\times$ with ESS held near 40% throughout, against $4.0\times$ in (a). They differ because $\|\beta\|_2$ moves in opposite directions. In (a) the tilt is fixed per coordinate, $\beta = 0.15 \cdot \mathbf{1}_d$, so $\|\beta\|_2 = 0.15\sqrt{d}$ grows from 0.212 to 1.697 and absorbs most of the absolute growth; in (b) β is rescaled to hold KL fixed and $\|\beta\|_2$ slightly shrinks, 0.676 to 0.533, so the relative error carries the absolute growth and adds to it. The cross-sweep comparison then isolates the cause: at $d = 128$ relative recovery is worse in (b) than in (a), 0.1214 against 0.1018, despite $24\times$ the effective sample size (39.75% against 1.65%). The degradation tracks the dimension of β rather than weight collapse.

Why SMT degrades is open, but the failure is now localised. A third diagnostic evaluates h_ψ on the sampler's own reverse trajectories, where the correct target is not 1: since $q = p\,L^*$, the forward kernel gives $q_\tau(z) = p_\tau(z)\,h_\tau(z)$, so on trajectories whose marginal should be q_τ the expectation is $\mathbb{E}_{q_\tau}[h_\tau] = \mathbb{E}_{p_\tau}[h_\tau^2] \geq 1$, by Jensen and the tower property. We therefore report the measured mean divided by $\mathbb{E}_{p_\tau}[h_\tau^2]$, which is 1 when the sampler realises the tilt.

Across all seven working cells the sampler under-realises it, and the shortfall broadly tracks SMT's error. At $\tau = 300$ in the growing-tilt sweep the ratio is 1.001, 0.990 and 0.924 as KL rises $0.05 \to 0.24 \to 0.95$, against sliced Wasserstein of 0.011, 0.029 and 0.097; at fixed KL it falls $0.978 \to 0.980 \to 0.959 \to 0.926$ from $d = 2$ to 128. Over the same cells h_ψ's value accuracy stays near-perfect, with correlation between h_0 and L^* at least 0.998. The correction also shrinks against the base score it corrects: $\|\nabla \log h_\psi\| / \|\nabla \log p\|$ at $\tau = 300$ falls from 0.38 at $d = 2$ to 0.042 at $d = 128$.

So h_ψ is accurate as a function, and the sampler does not realise the measure that h_ψ implies. That points at the gradient and trajectory channel rather than at an error in h_ψ itself, but the diagnostic cannot separate the two candidates: that $\nabla \log h_\psi$ is never supervised by a regression on values, and that h_ψ was fit on forward-noised prior samples rather than on the sampler's own reverse trajectories. Both predict this pattern, and we leave the question open.

We rank on sliced Wasserstein rather than on errors normalized by standard error, which can mask heavy-tailed failure because a failing model's inflated variance sits in their denominator; nor do we compare arms on calibrated-mean error, which is zero for weighting by construction.

D.23 The AAPL Surface

The surface is committed as data rather than refetched, because the vendor's chain endpoint returns live expiries and cannot reproduce a past as-of date. The return history is likewise pinned: a fresh fetch now differs from the committed file by 1.78×10^{-3} after a vendor revision, so the frozen file is the reproducible source.

Interval constraints, and what is identified. Forcing equality to the quoted mid is infeasible: it gives ESS 0.001% at $\mathrm{KL} = 1.6 \times 10^{26}$, against 93.6% for the martingale columns alone. Reweighting a sample cannot hit 21 smiled strikes exactly, and a mid asserts a precision the spread denies. We therefore impose bid $\leq \mathbb{E}_Q[\text{payoff}] \leq$ ask as an inequality-constrained I-projection, which leaves every arm well posed at ESS 78-94% with violations under a cent. Separately, only the carry is identified at this maturity: put-call parity pins the forward to within 0.1%, but the implied rate ranges

Table 18: Held-out error on the AAPL surface at the 0.10 spread floor, in bid-ask spread units, split by whether the expiry straddles the 2026-04-30 earnings date. Signed medians are model minus mid, so positive is overpricing: every arm overprices the event-free expiries and underprices the event-bearing ones, and the ordering of the four arms inverts exactly between the two groups. The median absolute errors equal the absolute signed medians to five decimals in all sixteen cells, a property of this sample rather than an identity, since few contracts carry the opposite sign and their magnitudes sit inside the bulk. Errors without the floor, which the 2-8 cent denominators of out-of-the-money quotes dominate, are in the per-expiry table below.

	Event-free ($n = 21$)		Event-bearing ($n = 34$)		Pooled ($n = 55$)
Prior	signed med.	med. abs. err.	signed med.	med. abs. err.	med. abs. err.
Path DDPM (learned)	$+1.1320$	1.1320	-0.4276	$\mathbf{0.4276}$	0.6971
Heston (fitted, MoM)	$+0.9615$	0.9615	-0.7086	0.7086	0.8674
Student-t (i.i.d., MLE)	$+1.0940$	1.0940	-0.4887	0.4887	$\mathbf{0.6781}$
Black-Scholes (hist. vol)	$+0.7386$	$\mathbf{0.7386}$	-1.1157	1.1157	1.1142

from -4.8% to $+26\%$, because at 7-25 days the discount factor differs from 1 by under 0.35%. The carry, -0.788%/yr, comes from the data; the discount rate is pinned at 4% and stated as an assumption, entering only as a common factor.

Projection quality. Every arm is projected onto the identical constraint set, with vanilla constraints imposed as bid-ask intervals.

			max	active bounds	calibrated fit	
Prior	ESS (%)	KL (nats)	violation	at bid / ask	median	max
Path DDPM (learned)	92.22	0.0373	1.12×10^{-4}	23 / 19	0.3106	0.5007
Heston (fitted, MoM)	93.71	0.0305	1.27×10^{-3}	20 / 23	0.2342	0.5246
Student-t (i.i.d., MLE)	87.67	0.0517	9.39×10^{-4}	20 / 23	0.3094	0.5080
Black-Scholes (hist. vol)	77.77	0.0650	8.17×10^{-4}	30 / 12	0.4108	0.5348

Every arm stays well posed, ESS 77.8-93.7%, with the largest constraint violation at 1.27×10^{-3}, under a cent. "Active bounds" counts the 42 calibrated inequalities resting on the bid and on the ask; calibrated fit is in spread units, so a median near 0.3 and a maximum just above 0.5 place the fitted prices inside the quoted spread throughout. Imposing equality to the mid instead is infeasible: ESS falls to 0.001% at $\mathrm{KL} = 1.6 \times 10^{26}$, against 93.6% for the martingale columns alone.

Held-out error by floor. Errors are $|\text{model} - \text{mid}|$ in spread units; the floor excludes contracts whose quoted spread falls below it, since those denominators nearly vanish. Median absolute error, with and without the floor:

	no floor			0.10 floor		
Prior	pre (24)	post (50)	pooled (74)	pre (21)	post (34)	pooled (55)
Path DDPM (learned)	1.3391	2.7537	2.0429	1.1320	$\mathbf{0.4276}$	0.6971
Heston (fitted, MoM)	1.3007	2.8823	2.0800	0.9615	0.7086	0.8674
Student-t (i.i.d., MLE)	1.2522	$\mathbf{2.7482}$	$\mathbf{2.0190}$	1.0940	0.4887	$\mathbf{0.6781}$
Black-Scholes (hist. vol)	$\mathbf{1.1521}$	3.0440	2.2602	$\mathbf{0.7386}$	1.1157	1.1142

Counts are contracts surviving the floor. The pre-event ordering is the same under both floors, and the sign pattern is unchanged, but the post-event ordering is not: without the floor the Student-t edges the learned prior, 2.7482 against 2.7537, and leads pooled as well. The learned prior's post-event lead therefore holds at the 0.10 floor and not without it. The raw metric is dominated by near-zero denominators. The DDPM's 11-day error falls from 3.1765 to 0.3363, a factor of 9.4, as the floor drops its count from 32 contracts to 20, which is why we report against a floor, but the dependence is a limitation of the metric and we state the ordering at both. Contract counts are identical across arms, since the filter is on the quoted spread and not on the model:

dte	0.00	0.05	0.10	0.15	0.25
7	13	12	11	10	8
9	11	10	10	9	7
11	32	27	20	19	17
18	18	16	14	11	10
pooled	74	65	55	49	42

Median absolute error in spread units, by arm, expiry and floor:

Prior	dte	0.00	0.05	0.10	0.15	0.25
Path DDPM	7	1.7196	1.4258	1.1320	1.0585	0.6720
	9	0.9869	0.9869	0.9869	0.6348	0.3318
	11	3.1765	0.8568	0.3363	0.3172	0.2265
	18	1.8440	1.1475	**0.6714**	0.3738	0.3365
	pooled	2.0429	1.1320	0.6971	**0.3738**	0.3245
Heston	7	1.6389	1.3002	0.9615	0.9144	0.5649
	9	0.9354	0.9354	0.9354	0.5700	0.2107
	11	3.1722	0.9802	0.5953	0.4806	0.3590
	18	1.9366	1.3266	0.9055	0.5023	0.4595
	pooled	2.0800	1.1039	0.8674	0.5700	0.3879
Student-t	7	1.6015	1.3477	1.0940	1.0086	0.6386
	9	0.9150	0.9150	0.9150	0.5778	0.3305
	11	3.1169	0.8575	**0.2656**	0.2608	0.1930
	18	1.9058	1.2931	0.8082	0.3980	0.3454
	pooled	**2.0190**	**1.0940**	**0.6781**	0.3980	**0.2817**
Black-Scholes	7	1.4182	1.0784	**0.7386**	0.7031	0.4138
	9	0.7820	0.7820	**0.7820**	0.4118	0.1524
	11	3.4275	1.4134	0.8786	0.6430	0.4749
	18	2.2143	1.8268	1.4221	0.6898	0.6230
	pooled	2.2602	1.4134	1.1142	0.6677	0.4433

Three readings of this table matter. The 11-day expiry is where the floor bites hardest, with the DDPM falling from 3.1765 to 0.3363 between the 0.00 and 0.10 floors as its count drops from 32 contracts to 20, so the raw column is not like-for-like across expiries. The 9-day expiry is flat from 0.00 to 0.10 for every arm, because only one of its eleven contracts has a spread under ten cents. And the pooled row is not a summary of the rows above it: pre- and post-event errors carry opposite signs and cancel, so the signed pooled medians at the 0.10 floor are only -0.044, -0.091, -0.006 and -0.152 for the DDPM, Heston, Student-t and Black-Scholes, against absolute medians three to seven times larger.

The pooled ordering is also unstable between the two leading arms. The Student-t leads at the 0.00, 0.05, 0.10 and 0.25 floors and the DDPM at 0.15, with gaps of 0.019 to 0.043 throughout: on the pooled metric the two are tied, and which one prints first is a property of the floor. What does survive every floor is the gap to Black-Scholes, and the gap to the fitted Heston at all floors but 0.05, where Heston's 1.1039 falls between the Student-t's 1.0940 and the DDPM's 1.1320.

Tail behaviour. Maximum daily return under each arm, against the data:

	max daily	count beyond			
Source	return	0.15	0.25	0.50	1.00
AAPL returns (2,839 obs.)	0.1426	0	0	0	0
Path DDPM (learned)	0.1455	0	0	0	0
Black-Scholes (hist. vol)	0.0906	0	0	0	0
Heston (fitted, MoM)	0.2209	41	0	0	0
Student-t (i.i.d., MLE)	2.4124	3,236	675	70	6

Counts are over 4.2×10^6 draws. The learned prior is closest to the data and, like Black-Scholes and the data itself, never exceeds 0.15. The Student-t admits a 241% daily move and puts 70 draws

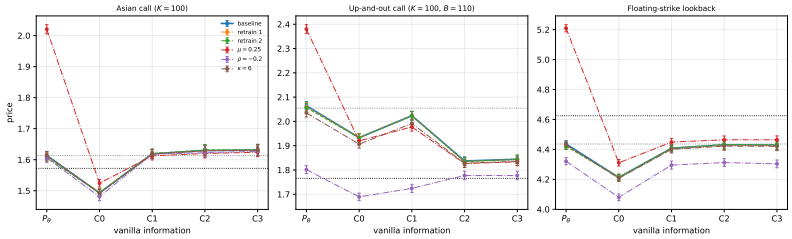

Figure 4: Exotic prices under six priors as vanilla information is added, from the unprojected prior through C0 to C3. The three seeds of the baseline recipe overlap throughout; the misspecified priors separate, most strongly on the barrier and lookback under the wrong leverage correlation.

beyond 50%, because fitting an i.i.d. heavy tail by maximum likelihood reads volatility clustering as unconditional tail weight. Black-Scholes errs the other way, its maximum draw falling short of the data's.

Exotics across arms. The identification result of Section 5.3 reproduces on a real surface. Under the same constraints the barrier price spreads 0.7481 across the four arms, 20.45% of their mean, from a minimum of 3.3085 under Black-Scholes to a maximum of 4.0566 under the fitted Heston; the Student-t lies between the extremes, so excluding it leaves 20.15% and the spread is not an artefact of that arm. The lookback spreads 0.8827 (5.76%), with the Student-t at the minimum, so excluding it leaves 0.4494 (2.90%). The Asian spreads 0.0554, 1.00%. The ordering barrier $\gg$ lookback $>$ Asian is the one the constraint-span analysis predicts.

D.24 Replication on an Independently Trained Prior

On a separately trained path prior, run on the A100 at C3 with the same 53 constraints as the primary run (ESS 92.9%), the corrected sampler and the weighted representation of the same Q^* agree on all three exotics within 1.65 standard errors and on 20 of 24 held-out constraint columns within 2 standard errors, with sliced Wasserstein distance at the null and no trend in reverse-step count from $K = 999$ down to 50. The held-out agreement is weaker than the 24 of 24 obtained on the primary prior (Appendix D.15); with the constraint set unchanged, we attribute the difference to the independent prior and training run.

D.25 Prior Retrains

Two retrains of the baseline recipe, differing only in initialization seed, bound the contribution of seed variance. Under the original criteria one of them (`retrain1`) failed a single check, the 21-day K=115 vanilla standard deviation ratio, at 0.870 against a 0.90 bound. That check was subsequently found to fail 27% of the time between two independent simulations of the same model, and was replaced (Appendix D.8); `retrain1` passes under the amended criterion and is included. All six priors in the sweep passed, with no retries, so no prior was dropped and the band of Section 5.4 is not conditioned on a discarded run. For the record, `retrain1`'s unweighted exotics differ from its own reference by -0.66, -0.48 and -0.63%.

D.26 The Learned Correction h_ψ

Training targets are $L^*(X_i)$ normalized so that $\frac{1}{N}\sum_i L_i^* = 1$, with MSE taken on h directly. In (9) we set $h_{\min} = 0.05$, which is $6.8\times$ below the smallest L^* observed at C3 (0.3405); the floor is never active across 1.96×10^7 samples. Diagnostics were declared before the run:

- **Tower property** $\mathbb{E}[h_\tau] = 1$: maximum deviation 0.0038 over the full range, at most 0.0026 for $\tau \geq 300$, and 1.0000 to four figures for $\tau \geq 950$.
- **Recovery at the clean end**: correlation between h_0 and L^* is 0.984 ($R^2 = 0.967$, RMSE 0.051).
- **Where the correction acts**: the prediction that the gradient ratio would peak at $\tau \in [100, 500]$ is *falsified as stated* - it peaks at $\tau \approx$ 40-60, because $\|\nabla \log p\|$ grows as $\tau \to 0$. The quantity

that actually enters the sampler, $\sqrt{1-\bar{\alpha}_\tau}\,\|\nabla \log h\|$, does peak in the predicted band (0.10 at $\tau \approx 400$-450). We record the prediction as falsified rather than reinterpreting it.

- **Vanishing at the noise end**: ratio 0.0017 for $\tau > 900$.

These diagnostics check h_ψ's values on forward-noised samples, not the gradient the sampler uses nor the sampler's own trajectories: in the dimension sweep of Appendix D.22 they stay near-perfect (correlation between h_0 and L^* about 0.999, tower property about 1) in cells where the sampler is already wrong. The trajectory diagnostic introduced there does detect it.

D.27 Solve-Draw Noise

Cross-sample held-out RMSE initially appeared alarming (0.033 against 0.0047 in-sample). Solving β^* on three of six draws and evaluating on all six localizes the cause: correlation between a draw's level gap and its mean held-out error is $+0.95$, with slope 0.76 price units per z-unit, errors centred at zero across 15 pairs (mean -0.0000, sd 0.021) and signs flipping with the gap. One pair was a 2.9σ outlier. The offset lives in the pair, not in either draw: β^* absorbs the solve draw's own level noise and hands it to every other draw. Moving the solve draw to 10^6 reduces in-sample C3 held-out RMSE from 0.0047 to 0.0031 (simulator 0.0028, i.e. within noise) and cross-sample from 0.033 to 0.0138 (simulator 0.012), with ESS rising from 92.8% to 93.4%. All duals reported use the 10^6 draw.

D.28 Computational Resources

Two machines were used: an Apple-silicon laptop (12-core CPU, Metal GPU via PyTorch MPS) for development and smaller diagnostics, and a single NVIDIA A100-SXM4-40GB on Google Colab for every reported result. No other CUDA hardware was used, and there is no multi-GPU or distributed training. TF32 is disabled for both matmul and cuDNN and float32 matmul precision is set to highest, because per-step relative error compounds over 999 reverse steps: under the A100's default TF32 the frozen checkpoint fails the acceptance criteria of Table 9, and passes under full fp32. The path model is a 580k-parameter time-conditioned MLP on 21-dimensional path vectors, with $T = 1000$ diffusion steps, $N = 10^5$ training paths and batch 512.

Stage	Scale	Device	Wall-clock
Path-model training (450 epochs, cosine decay, EMA)	10^5 paths	A100	7.2 min
the same run on the laptop	10^5 paths	MPS	3.5 min
Sampling, 999 reverse steps	10^5 paths	A100	0.2 min
Sampling the solve draw	10^6 paths	A100	1.6 min
Acceptance gate (77 columns, 3 exotics, 3 structural)	10^5 vs 10^5	CPU	≈ 1 s
Convex dual, C0-C3	10^6 paths	CPU	1.3 min
h_ψ training (100 epochs)	10^5 paths	A100	0.9 min
Corrected sampling ($+\nabla \log h_\psi$ per step)	10^5 paths	A100	0.3 min
Sample-restricted price bounds (6 LPs)	10^6 atoms	CPU	0.6 min
End-to-end, amortisation experiment		A100	11.9 min
End-to-end, prior overlay (six priors)		A100	≈ 70 min

Every figure above is read from run artifacts except the prior overlay, which is an estimate: six priors at roughly 12 minutes each. Training runs faster on the laptop than on the A100 because a 580k-parameter MLP leaves the A100 launch-bound rather than compute-bound, and disabling TF32 pushes the same way.

The correction's inference overhead is one extra forward-and-backward pass of h_ψ, a 256-wide MLP, per reverse step, at 0.3 min against 0.2 min per 10^5 paths, roughly 50%. The convex dual is the only iterative optimisation in the pricing pipeline; it takes seconds per constraint level and is solved once per level.

Each draw is identified by seed, sample size and device, and is bitwise reproducible on the same device but not across CPU and GPU, since the backends use different RNG streams and kernels. On a seed-matched 2×10^4 draw the two agree on every gate statistic within Monte Carlo error: the largest column-mean gap is 0.026 standard deviations against an expected maximum of 0.027 across 77 columns at that size, and all three exotics agree within 1.1 standard errors.

E ESTIMATED SMT: DETAILED ANALYSIS

E.1 ESTIMATOR DEFINITION AND SCOPE

This subsection concerns the tractable mixture setting of Appendix A.1 only. There, the target Q_0 is given in closed form, and the estimator is a distillation of a correction that is already known. It serves as a consistency check and as evidence that the correction is learnable, but it is not a way to obtain an unknown target. When the target is not known in closed form the estimator has nothing to regress against, which is the gap Section 3.2 exists to fill; the learned h_ψ used in the path setting is trained against L^* from the projection, not against an analytical formula.

In the mixture setting, we parameterize the correction as a neural network $\hat{\eta}_\psi(x_t, t)$ and train by regression against the analytical target:

$$\mathcal{L}(\psi) = \mathbb{E}_{t,\, x_t \sim p_t}\left[\left\|\hat{\eta}_\psi(x_t, t) - \eta_t^*(x_t)\right\|^2\right], \tag{23}$$

where $\eta_t^*(x_t) = \nabla \log q_t(x_t) - \nabla \log p_t(x_t)$. Training requires only samples from P and the analytical correction formula, with no samples from Q at any stage. In the mixture setting it learns the nonlinear state-dependent correction from equation (12). The estimated correction substitutes into (5) directly. Algorithm 1 summarizes the full procedure.

E.2 GAP ANALYSIS

Table 19: Gap between closed-form and estimated corrections in the mixture setting, where the correction is genuinely nonlinear. The gap is small in both relative and absolute terms.

Setting	Metric	Exact	Estimated	Gap
Mixture	W_1	0.172	0.171	0.6%
Mixture	$P(Y<-3)$	0.332	0.333	0.3%

In the mixture setting, where the correction is genuinely nonlinear and state-dependent, the estimated correction tracks the closed form to within 0.6% on W_1 and 0.3% on the tail probability, confirming that the MLP learns the smooth logistic-like transition between modes governed by the mixture responsibilities. The corresponding analysis for the learned h_ψ used in the path setting is in Appendix D.

E.3 PER-TIMESTEP ESTIMATOR GAP

We analyse how the gap between estimated and exact corrections varies across diffusion timesteps t in the mixture setting. The per-timestep gap is measured as the mean absolute error between $\hat{\eta}_\psi(y_t, t)$ and $\eta_t^*(y_t)$, averaged over $y_t \sim p_t$.

Table 20: Per-timestep estimator gap in the mixture setting. The relative gap is approximately uniform at 6-10% across all noise levels, indicating the estimator fits the correction shape uniformly rather than concentrating capacity at particular timesteps.

Timestep range	$\lvert\hat{\eta} - \eta^*\rvert$ (mean)	$\lvert\eta^*\rvert$ (mean)	Relative gap
$t \in [1, 200]$	0.042	0.631	6.7%
$t \in [201, 500]$	0.018	0.284	6.3%
$t \in [501, 800]$	0.009	0.128	7.0%
$t \in [801, 1000]$	0.003	0.031	9.7%

The uniform relative gap is consistent with the smooth sinusoidal time embedding, which lets one network share the correction's shape across noise levels; we did not test alternative embeddings, so this is an interpretation rather than a finding.

E.4 When the Estimated Correction Adds Most Value

In the mixture setting the correction is nonlinear and state-dependent, and the estimator is the only practical route when the analytical form is unavailable. The regime where estimation is not optional is the path setting of Section 5.5, where dQ^*/dP_θ is defined by a constraint set rather than a formula and no closed form exists at intermediate noise levels.

F Experimental Details for the Mixture Setting

A time-conditioned DDPM parameterized by a 3-layer MLP (hidden 256, SiLU, sinusoidal time embeddings), cosine variance schedule ($s = 0.008$), $T = 1000$, AdamW (lr $= 10^{-3}$, weight decay 10^{-4}, batch 512, 60 epochs), $N = 30{,}000$ source samples. The correction estimator is a separate 3-layer MLP (hidden 64, 32-dim time embedding) trained for 20 epochs by MSE against analytical targets on 10^5 samples; no Q samples are used at any stage.