
Bellman Defect Correction

Inference Time Algorithms
for SciML and GenAI

Yiping Lu

Beijing International Center for Mathematical Research
Peking University

More thought can improve the answer

• Flash

Quick response

YOU

A bat and a ball cost \$1.10 together.
The bat costs \$1.00 more than the ball.
What does the ball cost?

AI

The ball costs **\$0.10**.
 $1.10 - 1.00 = 0.10$

✗ Fails the check: $\$0.10 + \$1.10 = \$1.20$

• Pro

Reason and verify

YOU

A bat and a ball cost \$1.10 together.
The bat costs \$1.00 more than the ball.
What does the ball cost?

AI

Let the ball cost x . The bat then costs $x + 1$.
 $x + (x + 1) = 1.10 \implies x = 0.05$

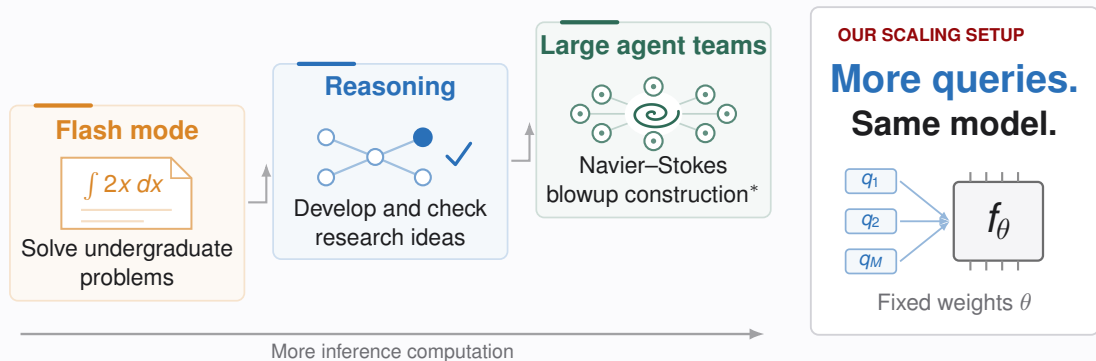
The ball costs **\$0.05**.

✓ Check: $\$0.05 + \$1.05 = \$1.10$

Same question. **More computation for reasoning and checks.**

Illustrative AI dialogues

From fast answers to scientific discovery



Which algorithm turns more queries into better answers?

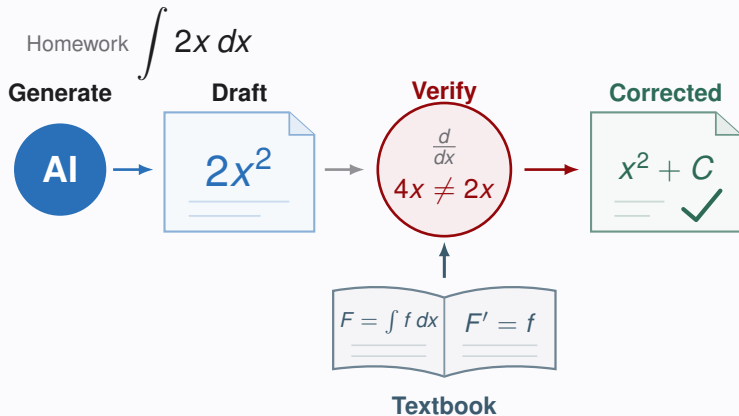
Illustrative task progression. * Reported NS case: $\sim 10,000$ agents, with model upgrades. OpenAI (2026).

Textbook knowledge checks an AI answer

MY AI PHILOSOPHY

**Generate first.
Verify with
known rules.**

The error tells us
what to correct.



Defect correction has exactly two steps

Write the Law of Defect

Derive the equation obeyed by $\check{u} := u - \hat{u}$.

1

known equation + autodiff

Simulate the Law of Defect

Spend new compute on stochastic paths for $\check{u}$.

2

linear: Feynman Kac + Monte Carlo
semilinear: multilevel Picard

$$u_{\text{corr}} = \hat{u} + \hat{\check{u}}_N$$

Write analytically. Spend inference compute stochastically.

No retraining: additional compute targets only what the model missed.

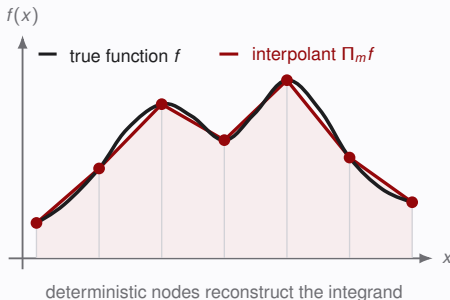
Source: SCaSML paper, structure preserving Law of Defect and stochastic correction.

Integration from random samples

Rare events, minimax rates, and optimal correction.

Linear case: function approximation

For intuition, take $q = 1$ and $I(f) = \int_{\Omega} f d\mu$.



APPROXIMATE, THEN INTEGRATE

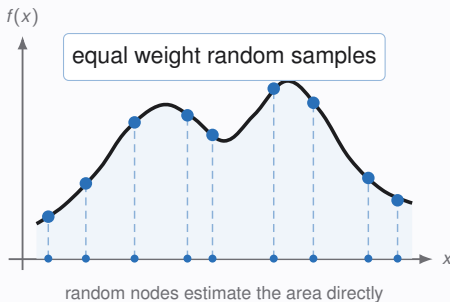
$$f \approx \Pi_m f = \sum_{j=1}^m f(x_j) \ell_j,$$

$$Q_m(f) = \int_{\Omega} \Pi_m f d\mu = \sum_{j=1}^m w_j f(x_j).$$

APPROXIMATION RATE

$$\begin{aligned} |I(f) - Q_m(f)| &\leq \|f - \Pi_m f\|_{L^1(\mu)} \\ &\lesssim m^{-s/d}. \end{aligned}$$

Linear case: Monte Carlo



SAMPLE, THEN AVERAGE

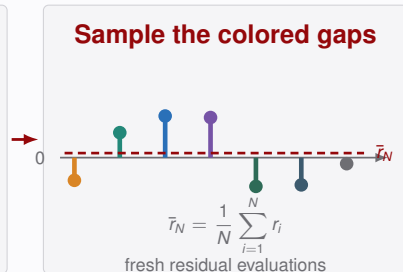
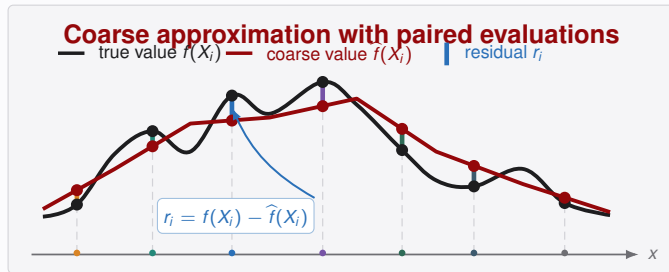
$$X_i \stackrel{\text{iid}}{\sim} \mu, \quad \hat{I}_N^{\text{MC}} = \frac{1}{N} \sum_{i=1}^N f(X_i).$$

$$\mathbb{E}[\hat{I}_N^{\text{MC}}] = I(f).$$

Sampling rate

$$\text{RMSE} = \frac{\sqrt{\text{Var}_{\mu}(f)}}{\sqrt{N}}.$$

Learn the bulk, simulate only the defect



$$I(f) = I(\hat{f}) + \mathbb{E}[r(X)], \quad \hat{I}_{\text{CV}} = I(\hat{f}) + \bar{r}_N.$$

Training learns the bulk; inference averages only what the surrogate missed.

Residual Monte Carlo multiplies the two rates

Condition on a surrogate trained with independent data.

$$\mathbb{E}[\hat{l}_{\text{CV}} \mid \hat{f}] = l(f), \quad \text{RMSE}(\hat{l}_{\text{CV}} \mid \hat{f}) \leq \frac{\|f - \hat{f}\|_{L^2(\mu)}}{\sqrt{N}}.$$

Two independent gains multiply

$$\underbrace{\|f - \hat{f}_m\|_2}_{O(m^{-\gamma})} \times \underbrace{N^{-1/2}}_{\text{residual samples}} = \underbrace{O(m^{-\gamma} N^{-1/2})}_{\text{corrected RMSE}}.$$

MATCHED BUDGETS: $N \asymp m$ $\text{RMSE} = O(m^{-\gamma-1/2})$

Source: Blanchet, Chen, Lu, and Ying, arXiv:2305.16527, Theorems 4 and 5 for the noisy linear functional.

Moment estimation with rare peaks

Estimate a moment from random function values

$$\mathcal{D}_n = \{(X_i, f(X_i))\}_{i=1}^n, \quad X_i \stackrel{\text{iid}}{\sim} \text{Unif}(\Omega)$$

$\Omega = [0, 1]^d$, n exact function values



$$T_q(f) = \int_{\Omega} f(x)^q dx = \mathbb{E}[f(X)^q]$$

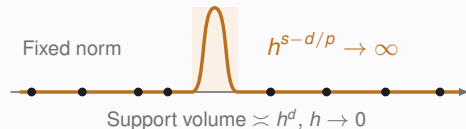
$q = 1$: integral $q \geq 2$: nonlinear moment

Sobolev regularity allows rare peaks

$$\mathcal{F} = \{f : \|f\|_{W^{s,p}(\Omega)} \leq 1\}, \quad \alpha = s/d$$

$$W^{s,p} \hookrightarrow L^{p^*}, \quad \frac{1}{p^*} = \frac{1}{p} - \alpha$$

$0 < \alpha < 1/p$: finite p^* , no uniform height bound



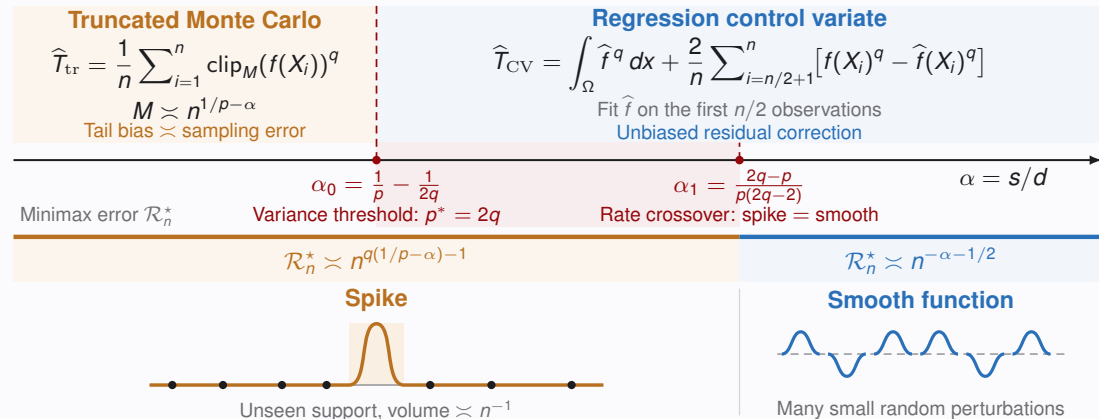
What is the optimal estimation rate?

$$\mathcal{R}_n^* = \inf_{\hat{T} = \hat{T}(\mathcal{D}_n)} \sup_{f \in \mathcal{F}} \mathbb{E} |\hat{T} - T_q(f)|$$

Moment theorem: integer $q \geq 2$, $p > 2$, $q < p < 2q$. Expected absolute error. Boundary cases omitted.

Source: Blanchet, Chen, Lu, and Ying, [arXiv:2305.16527](https://arxiv.org/abs/2305.16527), Section 2, Figure 1, Sections 3.1 and 3.2.

Optimal algorithms and minimax rates



Expected absolute error. CV: optimal regression oracle, exact surrogate integral. Strict regimes, schematic axis.

Source: Blanchet, Chen, Lu, and Ying, [arXiv:2305.16527](https://arxiv.org/abs/2305.16527), Theorems 1 to 3, Figure 2, Assumption 3.1.

Influence functions control nonlinear variance

Sample split estimation of $T_q(f) = \int_{\Omega} f^q$, $q \geq 2$

$$\hat{T}_{q,\text{CV}} = \int_{\Omega} \hat{f}^q + \frac{2}{n} \sum_{i=n/2+1}^n [f(X_i)^q - \hat{f}(X_i)^q]$$

Influence expansion

$$e = f - \hat{f}$$

$$f^q - \hat{f}^q = \underbrace{q f^{q-1}}_{\text{IF}_f(x)} e + R_q$$

$$\text{Var}(\hat{T}_{q,\text{CV}} \mid \hat{f}) \leq \frac{C_q}{n} \left(\underbrace{\int_{\Omega} |f|^{2q-2} |e|^2}_{\text{A: influence fluctuation}} + \underbrace{\int_{\Omega} |e|^{2q}}_{\text{B: nonlinear propagation}} \right)$$

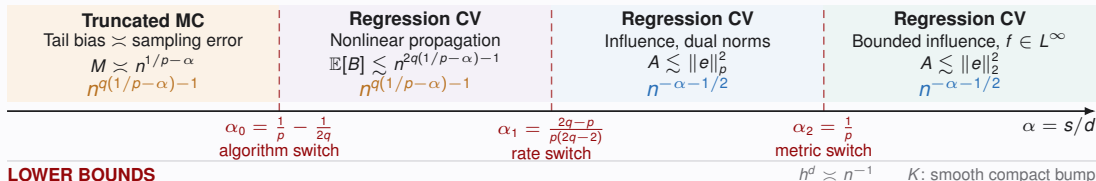
Source: Blanchet, Chen, Lu, and Ying, arXiv:2305.16527, equation 7 and Theorem 2.

Proof mechanisms behind the phase diagram

$$e = f - \hat{f}, \quad f^q - \hat{f}^q = \underbrace{q f^{q-1} e}_{\text{IF}_f} + R_q, \quad n \text{Var}(\hat{T}_{q,\text{CV}} | \hat{f}) \lesssim \underbrace{\|f^{q-1} e\|_2^2}_{\text{A: influence}} + \underbrace{\|e\|_{2q}^{2q}}_{\text{B: propagation}}.$$

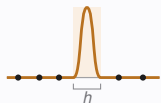
$$\mathcal{R}_n^* \asymp \max\{n^{q(1/p-\alpha)-1}, n^{-\alpha-1/2}\}, \quad \alpha = s/d, \quad p > 2, \quad q < p < 2q.$$

UPPER BOUNDS



LOWER BOUNDS

Spike

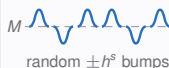


$$g(x) = h^{s-d/p} K\left(\frac{x-c}{h}\right)$$

Unseen support, $\Pr(\text{miss}) \asymp 1$

$$\int g^q \asymp n^{q(1/p-\alpha)-1}$$

Smooth function



$$f_\eta(x) = M + h^s \sum_j \eta_j K\left(\frac{x-c_j}{h}\right)$$

Sign bias $\asymp n^{-1/2}$, $\eta_j \in \{\pm 1\}$

$$\Delta T_q \asymp h^s n^{-1/2} = n^{-\alpha-1/2}$$

Expected absolute error, bounded Sobolev ball. Boundary cases omitted. CV upper bounds: regression oracle.

Source: Blanchet, Chen, Lu, and Ying, [arXiv:2305.16527](https://arxiv.org/abs/2305.16527), Theorems 1 to 3, Figure 2, Appendix B.2.

A two scale method

$$f(x) = \underbrace{\hat{f}(x)}_{\text{smooth, coarse scale}} + \underbrace{r(x)}_{\text{fine scale residual}}.$$

Smooth component



Function approximation

Small scale, high frequency residual



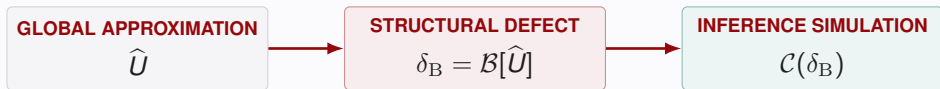
Monte Carlo

$$\hat{I} = \int_{\Omega} \hat{f} d\mu + \frac{1}{N} \sum_{i=1}^N r(X_i), \quad X_i \stackrel{\text{iid}}{\sim} \mu.$$

$$\text{RMSE}(\hat{I} | \hat{f}) = \frac{\sqrt{\text{Var}_{\mu}(r)}}{\sqrt{N}} \leq \frac{\|r\|_{L^2(\mu)}}{\sqrt{N}}.$$

Small residual energy controls the variance, not frequency alone.

Bellman residuals lift the same identity from integrals to paths



$$U_{\text{corr}} = \hat{U} + \mathcal{C}(\delta_B).$$

INTEGRATION

$\hat{f}$
 $f - \hat{f}$
 residual Monte Carlo

SCIENTIFIC ML

$\hat{u}$
 PDE Bellman residual
 Feynman Kac or MLP

GENERATIVE AI

$\hat{V}$ or H
 Bellman or Doob defect
 SMC, Girsanov, or two queries

The defect is computable because the governing structure is known.

Scientific machine learning

SCaSML: simulate the model's defect.

Linear case: subtraction gives the defect law

EQUATION 1 GOVERNING EQUATION: THE BELLMAN EQUATION

$$\partial_t u + \mathcal{L}u = f, \quad u(T, \cdot) = g,$$

EQUATION 2 ANOTHER BELLMAN EQUATION THAT THE SURROGATE SATISFIED

$$\partial_t \hat{u} + \mathcal{L}\hat{u} = \hat{f}, \quad \hat{u}(T, \cdot) = \hat{g}.$$

EQUATION 1 MINUS EQUATION 2

$$\underbrace{(\partial_t + \mathcal{L})(u - \hat{u})}_{(\partial_t + \mathcal{L})\check{u}} = \underbrace{f - \hat{f}}_{\text{Bellman error } \delta_B}$$

Law of Defect: Value evaluation with Bellman Error

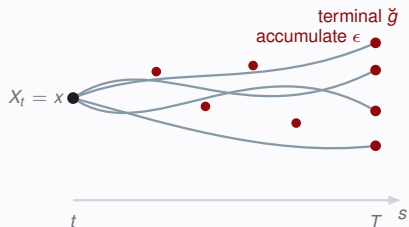
$$\check{u} := u - \hat{u}, \quad \check{g} := g - \hat{g}, \quad \delta_B := f - \hat{f} = -\epsilon,$$

$$(\partial_t + \mathcal{L})\check{u} = \delta_B, \quad \check{u}(T, \cdot) = \check{g}.$$

Feynman Kac simulates only the defect

SAMPLE THE GENERATOR

$$X_t = x, \quad dX_s = \mu(s, X_s) ds + \sigma(s, X_s) dW_s.$$



FEYNMAN KAC REPRESENTATION

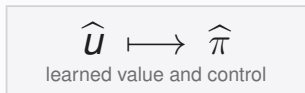
$$\begin{aligned} \Delta &:= \check{g}(X_T) - \int_t^T \delta_B(s, X_s) ds \\ &= \check{g}(X_T) + \int_t^T \epsilon(s, X_s) ds, \\ \check{u}(t, x) &= \mathbb{E}_{t,x}[\Delta]. \end{aligned}$$

MONTE CARLO REALIZATION

$$\hat{u}_{\text{CORR}}^{(N)}(t, x) = \hat{u}(t, x) + \frac{1}{N} \sum_{n=1}^N \Delta^{(n)}.$$

Simulator feedback for a learned controller

Learn a controller



current solution

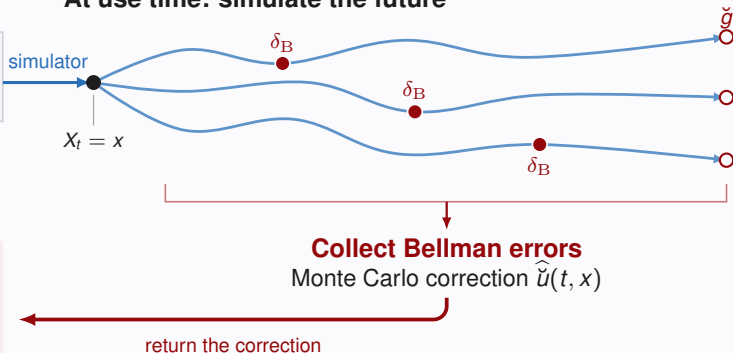
$$\hat{u}(t, x)$$

Correct the current value

$$\hat{U}_{\text{corr}} = \hat{U} + \hat{\tilde{U}}$$

update the control from $\hat{U}_{\text{corr}}$

At use time: simulate the future



Source: SCaSML, Law of Defect and stochastic correction.

Defect size controls the Monte Carlo variance

Monte Carlo Factor

$$\text{Var}(\hat{u}_{\text{corr}}^{(N)} | \hat{u}) = \frac{\text{Var}(\Delta | \hat{u})}{N}.$$

Defect Factor Condition on $\hat{u}$ and assume $\|\check{g}\|_{\infty}, \|\epsilon\|_{\infty} < \infty$.

$$|\Delta| \leq \|\check{g}\|_{\infty} + (T - t)\|\epsilon\|_{\infty},$$

The errors multiply

$$\|\check{g}\|_{\infty} \leq C_T e(\hat{u}), \|\epsilon\|_{\infty} \leq C_{\epsilon} e(\hat{u}) \implies \boxed{\text{RMSE}_{\hat{u}} \leq (C_T + (T - t)C_{\epsilon}) \frac{e(\hat{u})}{\sqrt{N}}}.$$

Use the surrogate as a control variate.

Better surrogate model $\rightarrow$ Lower correction variance

Nonlinear case: subtraction gives the defect law

Semi-Linear Equation

$$\mathcal{N}[v] := \partial_t v + \mathcal{L}v + F(v, \sigma^\top \nabla v).$$

EQUATION 1

$$\mathcal{N}[u] = f, \quad u(T) = g.$$

EQUATION 1 MINUS EQUATION 2

$$\begin{aligned} \partial_t \check{u} + \mathcal{L}\check{u} + F(\hat{u} + \check{u}, \hat{z} + \check{z}) - F(\hat{u}, \hat{z}) &= \underbrace{f - \hat{f}}_{\text{Bellman error } \delta_B} = -\epsilon, \\ \check{u}(T, \cdot) &= g - \hat{g}. \end{aligned}$$

EQUATION 2

$$\mathcal{N}[\hat{u}] = \hat{f}, \quad \hat{u}(T) = \hat{g}.$$

$$\text{Law of Defect: } \partial_t \check{u} + \mathcal{L}\check{u} + \check{F}(\check{u}, \check{z}) = 0.$$

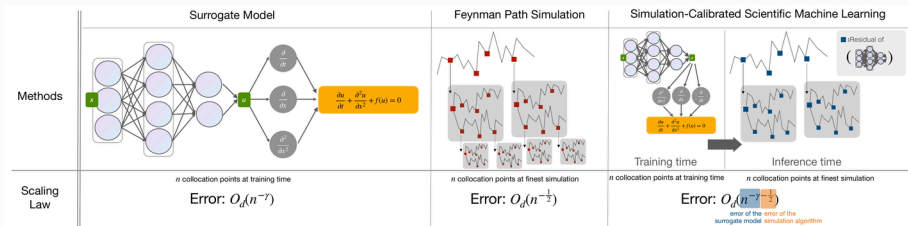
Key Observation: Closeness

$$\check{F}(v, z) := F(\hat{u} + v, \hat{z} + z) - F(\hat{u}, \hat{z}) + \epsilon$$

$$\check{u} := u - \hat{u}, \hat{z} := \sigma^\top \nabla \hat{u}, \check{z} := \sigma^\top \nabla \check{u}.$$

Training and inference exponents add

$$e(\hat{u}) = O(m^{-\gamma}), \quad N \asymp m \implies \text{Err}_{\text{SCaSML}} = O\left(m^{-\gamma} N^{-1/2+o(1)}\right) = O\left(m^{-\gamma-1/2+o(1)}\right).$$



TRAINING

 γ

+

INFERENCE

 $\frac{1}{2}$

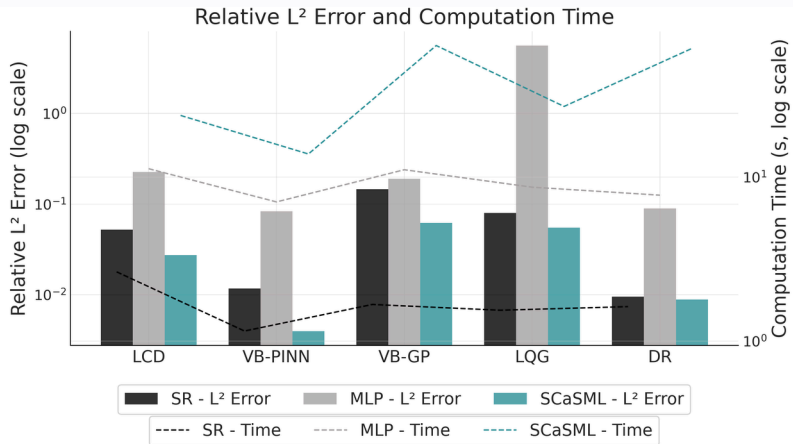
=

CORRECTED EXPONENT

 $\gamma + \frac{1}{2}$

Monte Carlo multiplies the surrogate error; it does not replace it.

SCaSML spends inference compute to buy accuracy



More inference compute lowers relative L^2 error across all five benchmarks.

Diffusion model

URGE: correct approximate guidance with path weights.

Guidance turns “at work” into a sampling direction

Suppose a diffusion model already knows how José looks.

BASE DISTRIBUTION $X_t \sim p_t$, public images of José



REWARD MODEL

Is José at work?

$G = \mathbb{P}(\text{José at work})$

extra drift follows predicted reward

$V(t)^2 \nabla G_t(X_t)$



GUIDED SAMPLE synthetic target



synthetic illustration • José at work

The base model supplies identity; the reward specifies the task; guidance changes the proposal.

Guidance follows the value, not the reward

BACKWARD VALUE

The terminal reward is propagated to time t through

$$h(x, t) = \mathbb{E}_{\mathbb{P}}[e^{r(X_T)} \mid X_t = x],$$

— Intractable in general

Ideal Doob guidance

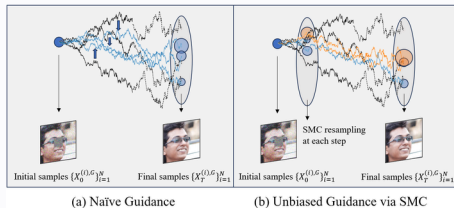
$$V(t)^2 \nabla \log h(x, t)$$

value gradient, **not** reward gradient $V(t)^2 \nabla_x r(x, t)$

PRACTICAL PROPOSAL

$$G \approx \log h \implies V(t)^2 \nabla \log h \longrightarrow V(t)^2 \nabla G.$$

$G \neq \log h \implies \mathbb{P}^G \neq \mathbb{Q}$: **correction is required.**



Reward scores the endpoint.

Value guides the path.

Girsanov gives the exact path correction

GUIDED PROPOSAL TO BASE DYNAMICS

$$\frac{d\mathbb{P}}{d\mathbb{P}^G}(X_{0:t}^G) = \exp\left(-\int_0^t V \nabla G^\top dW - \frac{1}{2} \int_0^t V^2 \|\nabla G\|_2^2 ds\right).$$

REWARD TILT

$$\frac{d\mathbb{Q}}{d\mathbb{P}}(X_{0:t}) \propto e^{r(X_t, t) - r(X_0, 0)}.$$

COMBINED PATH CORRECTION

$$\frac{d\mathbb{Q}}{d\mathbb{P}^G}(X_{0:t}^G) \propto \exp\left[r(X_t^G, t) - r(X_0^G, 0) - \int_0^t V \nabla G^\top dW - \frac{1}{2} \int_0^t V^2 \|\nabla G\|_2^2 ds\right].$$

Key advantage

The correction weight needs no reward gradient, reward Laplacian, or score.

URGE corrects diffusion paths sequentially

$$X_{k+1}^{i,G} = X_k^{i,G} + (v_k^i + V_k^2 \nabla G_k^i) \Delta t_k + V_k \sqrt{\Delta t_k} \xi_k^i, \quad \xi_k^i \sim \mathcal{N}(0, I).$$

$$\log \beta_k^i = r(X_{k+1}^{i,G}, t_{k+1}) - r(X_k^{i,G}, t_k) - V_k (\nabla G_k^i)^\top \sqrt{\Delta t_k} \xi_k^i - \frac{1}{2} V_k^2 \|\nabla G_k^i\|_2^2 \Delta t_k.$$

Sample particles

Update path weights

Normalize and resample

Particles at t_k

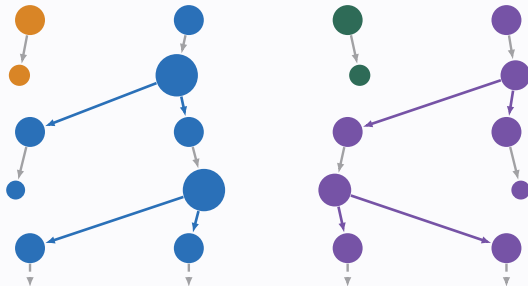
Sample and weight

Resample

Sample and weight

Resample

Color = ancestry



$N = 4$

Area = weight

Equal weights

Fresh noise per copy

Equal weights

Illustrative resampling outcomes

URGE is exact in an iterated limit

Iterated consistency

Under the paper's regularity assumptions and correct initial law,

$$\lim_{N \rightarrow \infty} \lim_{\Delta t \rightarrow 0} \mathbb{E}_{\mathbb{P}}[M_T^{N, \Delta t}(\varphi)] = \frac{\mathbb{E}_{\mathbb{P}}[e^{r(X_T) - r(X_0, 0)} \varphi(X_T, T)]}{\mathbb{E}_{\mathbb{P}}[e^{r(X_T) - r(X_0, 0)}]}.$$

Scope of the theorem

Asymptotic consistency only: the result gives no finite N total variation or sample complexity bound.

Path weights recover the particle correction rate

Define the terminal conditioned instantaneous path weight rate

$$\lambda(x, t) = \lim_{s \uparrow t} \frac{\mathbb{E}_{\mathbb{P}^G}[w_{s,t}^{\text{URGE}} \mid X_t^G = x] - 1}{t - s}.$$

If the guided marginal density is sufficiently smooth, then $\lambda(x, t) = w_{\text{particle}}(x, t)$.

Previous Works:

$$w_{\text{particle}}(x, t) = \partial_t r - \frac{V^2}{2}(\Delta r + \|\nabla r\|^2) + \nabla r^\top (v + V^2 \nabla G) + V^2 \nabla_x \log p^\top \nabla_x (G - r) + V^2 \Delta G$$

PATH SPACE

A weight for the simulated path segment

$w_{s,t}^{\text{URGE}}$

Path weights; URGE still resamples particles.

$\lambda = w_{\text{particle}}$
only as $s \uparrow t$

$\neq$
finite $\Delta t, N$

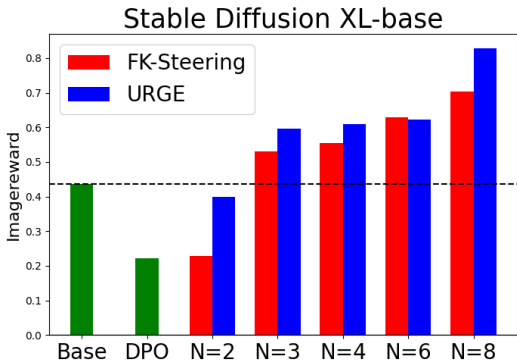
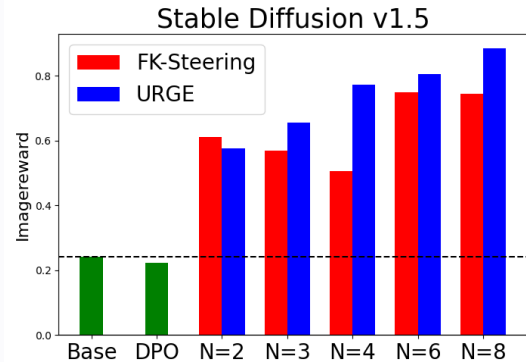
PARTICLE RESAMPLING

Local selection at each solver step

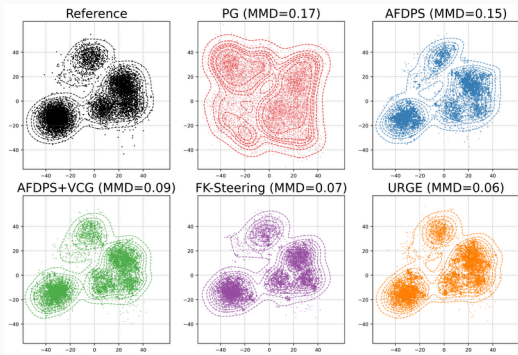
$w_{\text{particle}}(x, t)$

Kills or duplicates particles.

Inference scaling with more particles



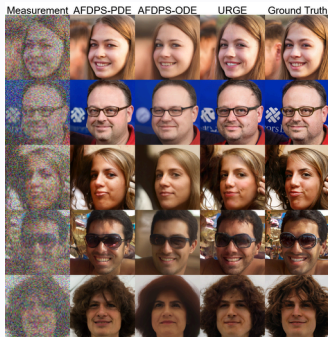
Pathwise correction recovers the target modes



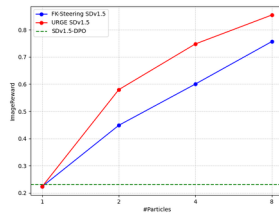
On the reported 30 dimensional Gaussian mixture, URGE gives the closest sample cloud to the analytical target, with **MMD 0.06**.

From inverse problems to neural rewards

FFHQ 256 SUPER RESOLUTION

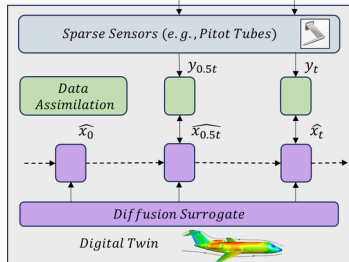
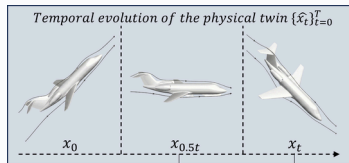


SD V1.5 PARTICLE SCALING



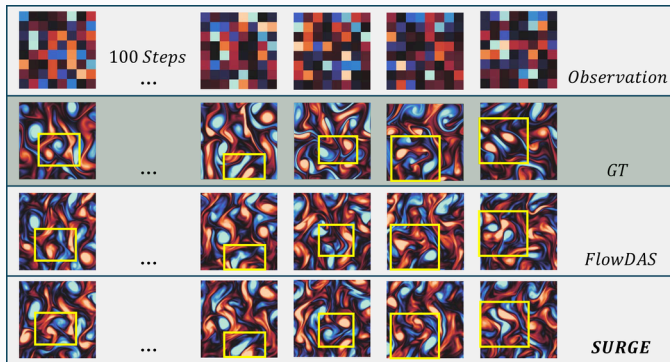
URGE matches derivative based restoration;
ImageReward exceeds FK Steering for
 $N = 2, 4, 8$.

Data Assimilation



REWARD

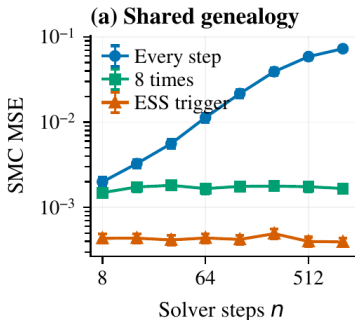
Reward measures **alignment with reality**.



From SMC to Stein guidance

SMC numerics reveal shared ancestry; sequential rejection removes copying; value-only correction removes the remaining bias.

SMC numerics under mesh refinement



ONLY SMC VARIANTS ARE COMPARED Same target and value approximation; only the numerical mesh and resampling schedule change.

What the plot says

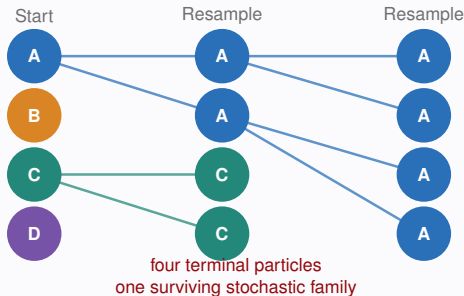
- Resampling at every solver step makes the MSE grow rapidly with depth n .
- A fixed number of resampling times or an ESS trigger remains nearly flat.

Instability tracks population interaction, not merely the local diffusion step.

Source: Supplied numerical experiment, `fig_stein_experiments.pdf`, panel (a).

Shared ancestry explained the SMC error

REPEATED RESAMPLING COPIES ENTIRE PREFIXES



WHY LOCAL DIAGNOSTICS MISS IT

Everything can look healthy

Flat weights can give $ESS = N$; even every one-particle marginal can be correct.

But the empirical mean sees pair covariance

$$\text{For } \bar{f}_N = N^{-1} \sum_i f(X_n^i),$$

$$\text{Var}(\bar{f}_N) = \frac{1}{N} \text{Var}(f(X_n^1)) + \frac{N-1}{N} \text{Cov}(f(X_n^1), f(X_n^2)).$$

N endpoints do not mean N independent histories.

Source: Supplied manuscript, neutral Brownian SMC construction and variance decomposition.

Perfect ESS can hide nonconvergence

Full multinomial resampling at every microstep; $X_0^{i,N} = 0$, $n, N \geq 2$.

NEUTRAL BROWNIAN CHAIN

$$X_p^{i,N} = X_{p-1}^{A_p^i,N} + n^{-1/2} \xi_p^i,$$

$$A_p^i \sim \text{Unif}\{1, \dots, N\}, \quad \xi_p^i \sim \mathcal{N}(0, 1).$$

Every local check passes

$\phi \equiv 1$, $\text{ESS}_p = N$, and

$X_n^{i,N} \sim \mathcal{N}(0, 1)$ exactly.

Every one particle marginal is exact.

Yet the empirical mean does not stabilize

For $g(x) = \sin(x)/2$,

$$\mathbb{E} \left[\left(\frac{1}{N} \sum_{i=1}^N g(X_n^{i,N}) \right)^2 \right] \geq \frac{1}{32e^2} \min \left\{ \frac{n}{N}, 1 \right\}.$$

For fixed N , the error stays nonzero as $n \rightarrow \infty$.

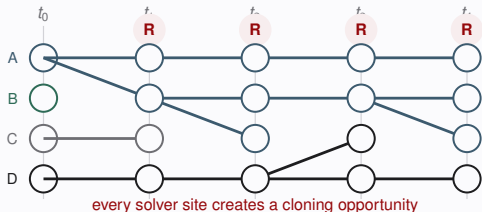
RMSE $\leq \varepsilon$ forces $N = \Omega(n\varepsilon^{-2})$ and work $= \Omega(n^2\varepsilon^{-2})$.

Source: User supplied manuscript, *Two Queries, No Gradients*, Theorem 1 and Corollary 2; counterexample restored from the original talk.

Sequential rejection retries children, not paths

FULL MICRO RESAMPLING

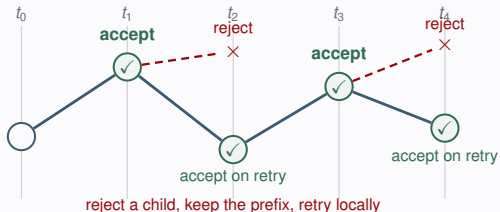
one parent pool



**categorical ancestor draw
from one shared particle pool**

EXACT CHILD REJECTION

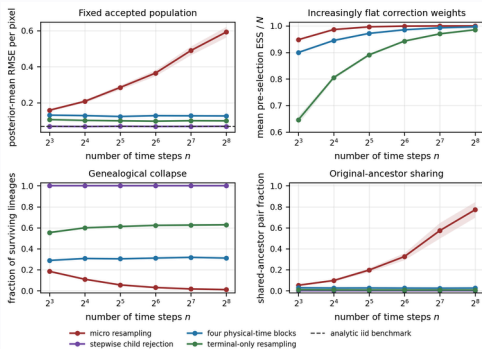
one proposal stream



**one accepted trajectory
with local candidate rejections**

Resampling moves mass **across lineages**. Rejection spends serial trials **within a lineage**.

Rejection stays stable as SMC genealogy collapses



CONTROLLED LINEAR IMAGING

$N = 160$ accepted paths $\cdot$ 36 runs

$n = 8, 16, 32, 64, 128, 256$

Sequential rejection

RMSE : 0.0683 to 0.0696

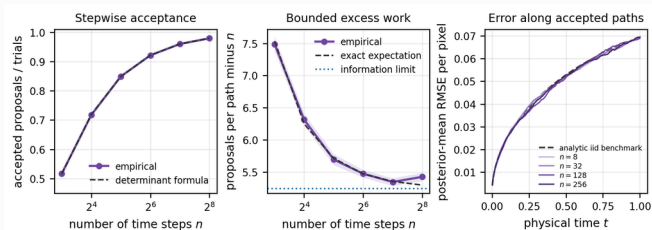
Micro-resampling SMC

RMSE : 0.159 to 0.593

At the finest mesh, ESS is $0.99994N$, yet only 1.02% of lineages survive and 77.2% of terminal pairs share an ancestor.

Rejection stays on the iid benchmark while repeated SMC cloning collapses genealogy.

Rejection trades shared ancestry for local retries



$$W_n = \sum_{k=1}^n \alpha_{k,n}^{-1} = n + I(X_1; Y) + O(n^{-1})$$

Information controls retry overhead; population controls independent sampling variance.

Acceptance: 0.517 $\rightarrow$ 0.979

Excess proposals: 7.54 $\rightarrow$ 5.30

Information limit: $I(X_1; Y) = 5.239$ nats

Source: Supplied diffusion imaging manuscript, rejection diagnostics and exact Gaussian work formula.

Rejection loses SMC's running correction weight

SMC CARRIES PROBABILITY COMPENSATION

Particle state

$$(X_{0:k}^i, W_k^i) \quad W_{k+1}^i = W_k^i r_k^i$$

The incremental factor r_k^i records how far the approximate value function is from Bellman consistency along that path. **Path and correction weight travel together.**

PURE REJECTION RETURNS AN UNWEIGHTED PATH

The local normalizer is hidden

$$G_k(x) = (P_k H_{k+1})(x),$$

$$Q_k^H(x, dy) = \frac{P_k(x, dy) H_{k+1}(y)}{G_k(x)}.$$

The accept/retry decision samples Q_k^H , but it does not retain G_k as a path weight. Approximate local tilts therefore need not compose to the target.

The missing whole-path compensation

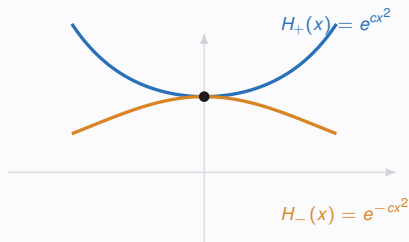
$$R_H = H_0(x_0) \prod_{k=0}^{n-1} \frac{G_k(X_k)}{H_k(X_k)}$$

Generate first; correct once at the end.

Source: Supplied manuscript and Lu (2026), path-law factorization for approximate Doob values.

A direct final correction asks for derivatives

**SAME VALUE AND GRADIENT, DIFFERENT
AVERAGE**



$$H_{\pm}(0) = 1, \quad H'_{\pm}(0) = 0, \quad c > 0$$

$$\mathbb{E}H_{\pm}(\sqrt{h}Z) = 1 \pm ch + O(h^2)$$

GUIDANCE USES A FIRST DERIVATIVE

$$dX_t = (b_t + a_t \nabla \log H_t) dt + \sigma_t dB_t.$$

The direct residual route

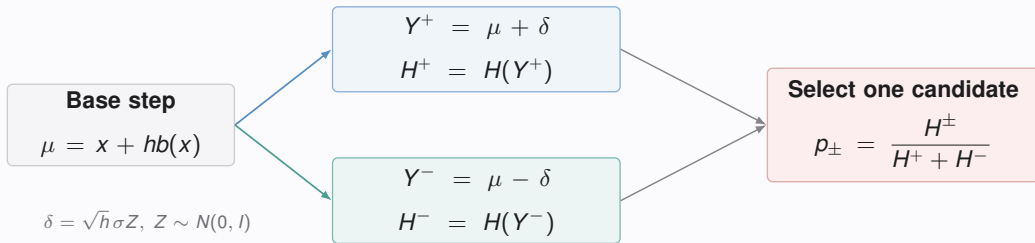
$$\mathcal{E}_t = \frac{\partial_t H_t + b_t \cdot \nabla H_t + \frac{1}{2} a_t : \nabla^2 H_t}{H_t}.$$

$$R_H = H_0(x_0) \exp\left(\int_0^T \mathcal{E}_t(X_t) dt\right).$$

Goal: use function values only—no ∇H and no $\nabla^2 H$.

Source: Lu (2026), blog, Sec. 5; Zhu et al. (2026), Sec. 3 and 4. Continuous identity assumes a valid change of measure.

Two symmetric queries choose a direction



Value contrast controls selection

$$\mathbb{E}[Y^J - \mu \mid Z] = \delta \frac{H^{+} - H^{-}}{H^{+} + H^{-}}.$$

The even part gives an average

$$S = \frac{H^{+} + H^{-}}{2}.$$

Keep S for the later path correction.

Source: Lu (2026), blog, Sec. 6; Zhu et al. (2026), Sec. 3.1.1, Eq. (5).

Why the direction is a log gradient

SUBTRACT: RECOVER A DIRECTIONAL SLOPE

Odd part

$$H^+ - H^- \approx 2\delta^\top \nabla H(\mu).$$

ADD: NORMALIZE BY THE VALUE

Even part

$$H^+ + H^- \approx 2H(\mu).$$

SELECTION COMBINES BOTH OPERATIONS

$$\frac{H^+ - H^-}{H^+ + H^-} \approx \delta^\top \nabla \log H(\mu).$$

Then average over the random direction:

$$\mathbb{E}[\delta\delta^\top] = h\sigma\sigma^\top.$$

Two value queries recover the leading Doob drift

$$\mathbb{E}[Y^J \mid x] = \mu + h\sigma\sigma^\top \nabla \log H(\mu) + o(h).$$

This is a small-step mean-drift statement for positive smooth H , not exact finite-step sampling.

Source: [Zhu et al. \(2026\)](#), Sec. 3.1.1, Eq. (5); [Lu \(2026\)](#), [blog](#), Sec. 6.

Stein links value selection to gradient guidance

GAUSSIAN INTEGRATION BY PARTS

Stein identity

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[\nabla f(Z)], \quad Z \sim N(0, I).$$

Set $Y = \mu + \sqrt{h}\sigma Z$, $f(Z) = H(Y)$:

$$\mathbb{E}[ZH(Y)] = \sqrt{h}\sigma^\top \mathbb{E}[\nabla H(Y)].$$

Random direction $\times$ value
reveals an average slope.

Finite pools replace these expectations with parallel value queries.

EXACT LOCAL REJECTION, WHEN AN ENVELOPE EXISTS

Propose Y from the base Gaussian. Accept with probability $H(Y)/M$, where $0 < H \leq M$.

$$G_h(\mu) = \mathbb{E}H(\mu + \sqrt{h}\sigma Z).$$

The accepted mean

$$\mathbb{E}[Y \mid \text{accept}] = \mu + h\sigma\sigma^\top \nabla_\mu \log G_h(\mu).$$

The exact identity uses the smoothed value G_h . It matches a mean, not a shifted Gaussian law.

Source: Lu (2026), blog, Sec. 7; Zhu et al. (2026), Eq. (6), Eq. (23), and Proposition 6.

A centered pool extends two queries to any K

A BALANCED SET OF RANDOM DIRECTIONS

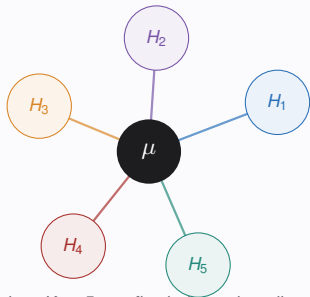


Illustration: $K = 5$, not five independent directions
Each $Z_i \sim N(0, I)$; pool directions are dependent.

$K = 2$: $Z_1 = -Z_2$. **Odd K also works.**

DRAW, CENTER, AND RESTORE MARGINAL VARIANCE

$$G_i \stackrel{\text{iid}}{\sim} N(0, I), \quad \bar{G} = \frac{1}{K} \sum_i G_i,$$

$$Z_i = \sqrt{\frac{K}{K-1}} (G_i - \bar{G}), \quad \boxed{\sum_i Z_i = 0}.$$

Query in parallel, then select

$$Y_i = \mu + \sqrt{h} \sigma Z_i, \quad H_i = H(Y_i),$$

$$\Pr(J = i \mid U) = \frac{H_i}{\sum_j H_j}, \quad S = \frac{1}{K} \sum_i H_i.$$

Source: [Zhu et al. \(2026\)](#), Sec. 3.1.2, Eqs. (7) and (8); [Lu \(2026\)](#), [blog](#), Sec. 7.

Centering cancels the leading pool noise

THE SAME POOL AVERAGE WILL ENTER THE PATH WEIGHT

$$S = \frac{1}{K} \sum_i H(\mu + \sqrt{h} \sigma Z_i) = H(\mu) + \underbrace{\sqrt{h} g^\top \bar{Z}}_{\text{linear random term}} + O_{L^2}(h), \quad g = \sigma^\top \nabla H(\mu).$$

Independent candidates

$\mathbb{E} \bar{Z} = 0$, but $\bar{Z}$ varies from pool to pool.

$\sqrt{h}$ scale fluctuations

The linear term vanishes only after averaging.

Centered candidates

$\bar{Z} = 0$ for every pool.

h scale fluctuations

The linear term cancels before averaging.

A better arrangement of queries, not an extra derivative evaluation.

Small-step comparison with fixed K , positive smooth values, and controlled second-order remainder. The smoothing bias is a separate issue.

Source: [Zhu et al. \(2026\)](#), Remark 2; [Lu \(2026\)](#), [blog](#), Sec. 7.

The same queries provide the path weight

AT EACH SOLVER STEP

One pool, two outputs

$$S_k = \frac{1}{K} \sum_{i=1}^K H_{k+1}(Y_{k,i})$$

Choose X_{k+1} with probability proportional to its value.

Keep both S_k and $H_{k+1}(X_{k+1})$.

No extra value query. No value derivative.

ACCUMULATE THE PROBABILITY COMPENSATION

Weight for a complete proposal

$$R = \phi(X_n) \prod_{k=0}^{n-1} \frac{S_k}{H_{k+1}(X_{k+1})}.$$

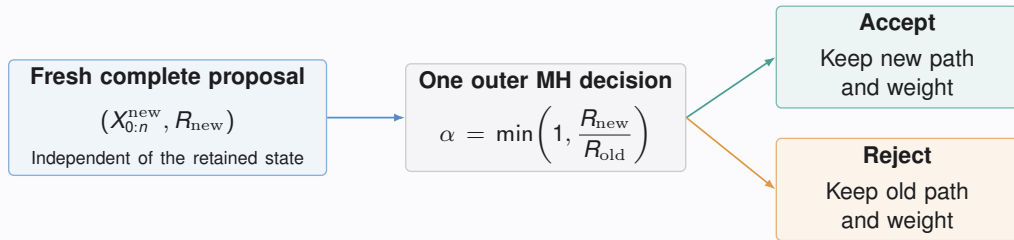
$$\begin{aligned} \log R &= \log \phi(X_n) \\ &+ \sum_{k=0}^{n-1} \log \frac{S_k}{H_{k+1}(X_{k+1})}. \end{aligned}$$

Valid on the augmented space of pools and choices, not by an arbitrary density plug-in.

Relative values guide. The same pool average corrects.

Source: Lu (2026), blog, Sec. 8; Zhu et al. (2026), Algorithm 1 and Theorem 12.

Whole-path MH corrects after generation



What is independent?

Complete proposals can be generated independently and batched in parallel.

What is not independent?

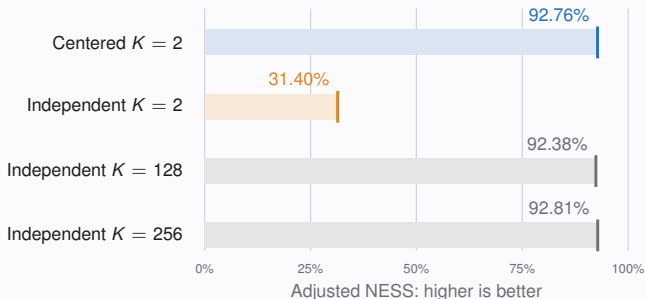
Consecutive outputs of one MH chain may repeat and remain correlated.

The target is invariant; a finite MH run can still have initialization error.

Source: Lu (2026), [blog](#), Sec. 8; Zhu et al. (2026), Algorithm 2.

Two centered candidates stabilize weights

BROWNIAN BENCHMARK: PERTURBED VALUE, $n = 192$



Two queries suffice in this test

92.76%

Independent pools require 128 or 256 candidates to reach a similar level in this test.

Adjusted NESS evaluates normalizer estimation. It is not image quality or the number of independent MH outputs.

Source: [Zhu et al. \(2026\)](#), Sec. 5.1 and Figure 2; [Lu \(2026\)](#), [blog](#), Sec. 10. Values redrawn from the reported results.

Whole-path correction improves the target match

Hopper: fixed diffusion policy and critic, averaged over eight fixed states.

CONDITIONAL CRITIC-SCORE KS DISTANCE: LOWER IS BETTER

Target mass	Guided only	With path MH
30%	0.2174	0.0306
60%	0.2804	0.0697
80%	0.4041	0.1128

Target mass is placed on the original policy's top 10% of actions by critic score.

Same two-candidate guidance

$K = 2$ at every diffusion step.

Correction uses 32 complete proposals:

1 initialization + 31 outer updates.

What the experiment measures

Critic-score distribution matching, not realized return.

A useful direction does not remove the need for probability correction.

Source: [Zhu et al. \(2026\)](#), Sec. 5.2, Table 2, Appendix H.3; [Lu \(2026\)](#), [blog](#), Sec. 10.

One correction principle, three numerical languages

Setting	Learned object	Structural defect	Inference correction
Integration	$\hat{f}$	$f - \hat{f}$	residual Monte Carlo
Scientific ML	$\hat{u}$	PDE or Bellman residual	Feynman Kac or multilevel Picard
Diffusion	G or H	Doob or path law defect	Girsanov weights or Stein pools with path MH

Inference time scaling law

$$U_{\text{corr}} = \underbrace{\hat{U}}_{\text{learned bulk}} + \underbrace{\mathcal{C}(\delta_B)}_{\text{simulate the structural miss}}.$$

Learn the bulk once; spend inference compute only on the miss.

Thank you