# RISK-CALIBRATED PROPOSAL TRANSPORT FOR FINITE-PARTICLE DIFFUSION STEERING

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Inference-time steering combines pretrained diffusion experts or rewards without retraining by changing the dynamics that transport noise to data. Feynman–Kac correction compensates for proposal mismatch through importance-weighted sequential Monte Carlo (SMC), whose finite-particle behavior depends on the proposal. Variance-controlling guidance (VCG) improves that proposal by fitting a linear drift correction to minimize empirical log-weight-rate variance. Although its population optimum cannot worsen residual variance, finite-particle VCG can nearly eliminate its fitting residual while increasing residual risk on new states by orders of magnitude. The resulting update can degrade unweighted generation or accelerate particle collapse. We show that the centered Feynman–Kac rate is the normalized transport residual and that expected out-of-fit benefit is exactly population headroom minus coefficient-estimation penalty. Under regularity assumptions, a Wasserstein analysis bounds the unweighted proposal's terminal error using this residual. These results motivate Risk-Calibrated Proposal Transport (RCPT), which uses deletion leave-one-out residuals to calibrate the retained fraction of the VCG update, adding no model calls and only small linear-algebra overhead. Experiments on 2D checker distributions, scaffold decoration, molecular property optimization, and class-conditional CIFAR-10 generation demonstrate recovery from harmful fitted updates. Across molecular and image domains, RCPT consistently lowers mean residual risk and improves a broad range of terminal metrics relative to uncalibrated VCG.

## 1 INTRODUCTION

Inference-time steering adapts pretrained diffusion or flow models to targets assembled from experts or rewards. These models transport noise toward data through a stochastic differential equation (SDE) (Song et al., 2021),

$$\mathrm{d}X_t = v_t(X_t)\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t, \qquad X_0 \sim p_0,$$

where $v_t$ is the drift, $\sigma_t$ is the noise amplitude, and $W_t$ is Brownian motion. An ordinary differential equation (ODE) sampler uses zero noise and an appropriate drift; steering changes the drift while keeping the pretrained models fixed.

A cheap guided proposal need not follow the requested target path. Combining expert fields, as in classifier-free guidance (Ho & Salimans, 2021), gives an unweighted Anchor, but matching an instantaneous score does not ensure that the prescribed marginals satisfy the sampler's forward equation. Feynman–Kac correction (FKC) and Adaptive Path Correction with Exponents (ACE) account for this mismatch through importance weights and SMC (Skreta et al., 2025; Lee et al., 2026). Their target recovery is asymptotic: finite-particle resampling can deplete ancestral diversity, while unweighted deployment retains proposal bias.

Proposal fitting aims to improve both deployment modes by estimating a drift correction from an ensemble of samples, or particles. DriftLite's variance-controlling guidance (VCG) minimizes FK log-weight-rate variance in a linear field class, while energy-controlling guidance (ECG) fits gradient fields through an energy objective (Ren et al., 2026). Population VCG cannot increase residual variance relative to Anchor because its class includes the zero correction. With finite particles, that benefit depends on whether the estimated coefficients generalize.



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.

Limited fitting information can reverse this benefit. We focus on budgets below 100 fitting particles, with larger populations testing how calibration responds as information grows. When effective information approaches controller rank, VCG can nearly eliminate its fitting residual yet increase independent residual risk by orders of magnitude (Figure 1). Increasing the fitting population costs additional expert evaluations, motivating calibration of the available update.

We treat proposal fitting as a prediction problem whose benefit must survive deployment. The centered FK rate is the normalized transport residual, and expected out-of-fit improvement equals population headroom minus coefficient-estimation penalty. This separates the improvement available to an ideal controller from its estimation cost. A Wasserstein analysis bounds unweighted terminal proposal error through the residual, which also controls leading local correction-weight inflation.

Risk-Calibrated Proposal Transport (RCPT) estimates this tradeoff and retains a scalar fraction of the VCG update. Deletion leave-one-out residuals estimate the residual scale, while a structured penalty accounts for controller complexity relative to fitting information. RCPT adds no model calls and only small linear-algebra overhead to VCG proposal fitting. It supports frozen or online deployment and retains more of VCG as fitting information grows.

Across molecular and image domains, RCPT lowers mean residual risk and improves terminal metrics relative to baselines including uncalibrated VCG (Tables 3–5).

Our contributions are:

- **A finite-particle failure mode hidden by population guarantees.** We formulate proposal fitting as an out-of-fit prediction problem and expose fit-to-deployment inversion: empirical VCG can nearly eliminate fitting residual while increasing independent risk and degrading unweighted or corrected generation. We connect the centered FK residual to a Wasserstein terminal-error bound and leading local importance-weight inflation.
- **Deployment-risk theory and nearly cost-free calibration.** We derive an exact headroom–penalty decomposition, an asymptotic estimation penalty, and a bias-aware scalar oracle. RCPT instantiates this principle with structured deletion-LOO calibration, no additional expert calls, and a conditional bound on its decision error.
- **Deployment-scale evidence across inference modes and domains.** Controlled 10,000-particle deployments and common-fit regularization audits test generalization directly. A complete 76-pocket scaffold benchmark, eight-seed molecular SMC protocols, and CIFAR-10 generation show that calibration repairs harmful VCG fits and improves downstream quality, with restored genealogy under corrected molecular sampling.

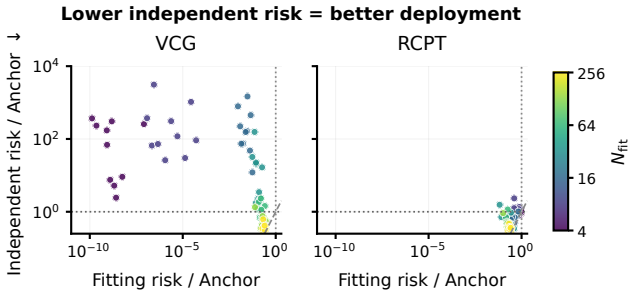


Figure 1: **Low fitting risk can hide high deployment risk.** Lower on the vertical axis means better independent risk; a smaller fitting residual alone is insufficient. Both axes show log-scaled ratios to Anchor; dotted lines mark Anchor. Color denotes fitting population.

## 2 Related Work

**Diffusion steering and Anchor proposals.** Score composition and reward guidance construct cheap proposals from pretrained models (Ho & Salimans, 2021; Chung et al., 2025; Du et al., 2023; Bradley et al., 2025; Song et al., 2023; Yu et al., 2023; Bansal et al., 2024; Ye et al., 2024). FKC

and ACE correct prescribed paths through importance weighting and SMC (Skreta et al., 2025; Lee et al., 2026; Doucet et al., 2000; Naesseth et al., 2019). Related particle methods support diffusion sampling, composition, and discrete correction (Wu et al., 2023; Singhal et al., 2025; Thornton et al., 2025; Wu et al., 2025; Hasan et al., 2025; Kit et al., 2025). Finite-particle costs and ancestral collapse motivate improving the proposal before correction (Chatterjee & Diaconis, 2018; Beskos et al., 2014; Agapiou et al., 2017; Koskela et al., 2020). RCPT keeps the target fixed and calibrates a fitted update to Anchor; Table 2 compares both unweighted and SMC deployment.

**Proposal fitting: VCG and ECG.** Our closest baselines are DriftLite's **VCG**, which minimizes FK rate variance, and **ECG**, which fits gradient corrections through an energy objective (Ren et al., 2026). Proposal fitting improves sampling through adaptive importance sampling, controlled or twisted SMC, and diffusion or incremental-divergence objectives (Cornebise et al., 2008; Snyder et al., 2015; Heng et al., 2020; Whiteley & Lee, 2014; Vargas et al., 2024; Phillips et al., 2024; Kim et al., 2025). Optimal discrete proposals, tree sampling, trust-region twists, and guidance-variance control offer complementary mechanisms (Ou et al., 2026; Jain et al., 2025; Wang et al., 2026; Kim, 2026); biased-proposal theory separates approximation and particle errors (Strasman et al., 2026). RCPT addresses the estimation error of VCG's fitted coefficient; Table 2 compares VCG, ECG, and RCPT on the same Checker deployments.

**Regression regularization and calibration.** Regularized and selected Stein control variates address unstable regressions (South et al., 2023; Leluc et al., 2021), while classical shrinkage and covariance penalties quantify prediction error (Efron & Morris, 1973; Efron, 2004). RCPT's contribution is to identify and control the deployment risk of a regression coefficient that changes generative dynamics. Ridge penalizes coefficients and principal-component regression (PCR) truncates fitted directions; RCPT instead calibrates one amplitude using structured headroom and penalty estimates. Direct LOO substitutes a cross-validated amplitude for this structured rule and serves as a calibration ablation. Table 1 and the paired comparisons in Section 6 show lower mean reference risk for RCPT than Direct LOO and LOO-selected ridge/PCR in the common-fit Checker audit.

## 3 Finite-Particle Proposal Risk

**Target path and Markov proposal.** Let $\widetilde{p}_t > 0$ be a differentiable unnormalized density, $Z_t = \int \widetilde{p}_t(x)\ \mathrm{d}x$, and $p_t = \widetilde{p}_t/Z_t$ for $t \in [0,1]$. In the applications, where $q_t^{(i)}$ are expert densities, $R_{j,t}$ are reward functions, and $\gamma_i(t), \beta_{j,t}$ are time-varying exponents, the target path is

$$\widetilde{p}_t(x) = \prod_{i=1}^{I} q_t^{(i)}(x)^{\gamma_i(t)} \exp\left\{ \sum_{j=1}^{J} \beta_{j,t} R_{j,t}(x) \right\}, \tag{1}$$

which includes products, ratios, time-varying exponents, and reward tilting. Let $\mathcal{L}_{t,v}^*$ be the conservative forward Kolmogorov operator of a proposal with selectable drift $v_t$, and let $U_{s,t}^v$ be its Markov evolution on observables. The unweighted proposal law $\rho_t^v$ satisfies $\partial_t \rho_t^v = \mathcal{L}_{t,v}^* \rho_t^v$ and $\rho_0^v = p_0$. The corresponding raw Feynman–Kac rate is

$$g_t^v(x) = \frac{\partial_t \widetilde{p}_t(x) - \mathcal{L}_{t,v}^* \widetilde{p}_t(x)}{\widetilde{p}_t(x)}. \tag{2}$$

Proposal-Only deployment discards this correction; FKC/ACE retains its incremental weights and a stated resampling rule. RCPT changes neither the target path nor the exponent schedule.

**Population VCG is a projection.** For Anchor drift $v_t^0$ and $m$ fixed fields $\phi_{j,t}$, fit an additive drift $v_t^\lambda = v_t^0 + \sum_{j=1}^m \lambda_{j,t}\phi_{j,t}$. With unchanged diffusion, the FK rate becomes $g_t^\lambda = g_t^0 + a_t^\top \lambda_t$, where $a_{j,t} = \mathcal{T}_{p_t}\phi_{j,t}$ and $\mathcal{T}_{p_t}\phi = p_t^{-1}\nabla\cdot(p_t\phi)$ is the Stein transport operator. Define the population residual risk $V_t(\lambda) = \mathrm{Var}_{p_t}(g_t^\lambda)$, $C_t = \mathrm{Cov}_{p_t}(a_t, a_t)$, and $d_t = \mathrm{Cov}_{p_t}(a_t, g_t^0)$. The covariance-weighted seminorm $\|z\|_{C_t}^2 = z^\top C_t z = \mathrm{Var}_{p_t}(a_t^\top z)$ measures how much a coefficient change affects the residual. This is the $C_t$ *geometry*: errors are weighted by their effect on residual variance, rather than by their Euclidean size. The minimum-norm population VCG coefficient is $\lambda_t^\star = -C_t^\dagger d_t$, where $\dagger$ denotes the Moore–Penrose pseudoinverse (Ren et al., 2026). Writing $V_{t,\star} = V_t(\lambda_t^\star)$ gives

$$V_t(\lambda) = V_{t,\star} + \|\lambda - \lambda_t^\star\|_{C_t}^2. \tag{3}$$

Appendix D.1 gives the derivation, including singular $C_t$.

### 3.1 Headroom and coefficient-estimation error

We define out-of-fit risk by evaluating on the target law and averaging over fitting samples. At time $t$, sample $D$ produces $\widehat{\lambda}_t(D)$ with expected prediction risk $\mathcal{R}_t(\widehat{\lambda}_t) = \mathbb{E}_D V_t(\widehat{\lambda}_t(D))$. Thus $V_t$ evaluates a fixed coefficient, while $\mathcal{R}_t$ also averages its estimation variability. We now fix $t$ and omit time subscripts in the coefficient-risk analysis below.

**Theorem 1** (Exact out-of-fit risk decomposition)**.** *For a fixed target law and field dictionary, assume $\mathbb{E}_D\|\widehat{\lambda}\|_C^2 < \infty$. Define population headroom $H = \|\lambda^\star\|_C^2$ and coefficient-estimation penalty $P = \mathbb{E}_D\|\widehat{\lambda} - \lambda^\star\|_C^2$. Then*

$$\mathcal{R}(\widehat{\lambda}) - V(0) = P - H. \tag{4}$$

Taking expectation in Eq. (3) proves the identity. The penalty includes both coefficient variance and squared bias, so it does not require unbiased fitting. The full estimated update harms expected residual risk precisely when $P > H$. Here, estimation error exceeds the available improvement.

### 3.2 Scaling the fitted correction to minimize risk

Scaling $\widehat{\lambda}$ by $\alpha \in [0, 1]$ interpolates between Anchor and the full VCG fit. The optimal fixed amplitude balances retained signal against estimation error.

**Theorem 2** (Expected-risk oracle and the structured case)**.** *Let $B = (\lambda^\star)^\top C\mathbb{E}_D(\widehat{\lambda} - \lambda^\star)$ measure alignment between mean estimation error and the population coefficient. For deterministic $\alpha \in [0, 1]$,*

$$\mathcal{R}(\alpha\widehat{\lambda}) = V_\star + (1-\alpha)^2 H + \alpha^2 P - 2\alpha(1-\alpha)B. \tag{5}$$

*If $H + P + 2B > 0$, the minimizing amplitude is*

$$\alpha^\star = \Pi_{[0,1]}\left(\frac{H+B}{H+P+2B}\right), \tag{6}$$

*where $\Pi_{[0,1]}$ clips to $[0, 1]$; if $H + P + 2B = 0$, all amplitudes have the same risk. When $B = 0$ and $H + P > 0$, this becomes $\alpha_0 = H/(H + P)$, with*

$$\mathcal{R}(\alpha_0\widehat{\lambda}) - V(0) = -\frac{H^2}{H+P}. \tag{7}$$

The oracle retains more of the fit when headroom is large relative to estimation error and $B = 0$. Unbiased fitting is sufficient for $B = 0$; more generally, only its mean error's alignment with $\lambda^\star$ under $C$ must vanish. The fraction $H/(H+P)$ motivates RCPT's structured rule. Section 4 analyzes data-selected amplitudes, and Appendix D.5 quantifies sensitivity to nonzero $B$.

## 4 Risk-Calibrated Proposal Transport

RCPT uses a structured estimate of prediction risk to select the strength of the fitted VCG correction.

**Candidate fit.** Given $\{a_n, g_n^0, \omega_n\}_{n=1}^N$ with $0 \le \omega_n < 1$ and $\sum_n \omega_n = 1$, let $\bar{a} = \sum_n \omega_n a_n$, $\bar{g}^0 = \sum_n \omega_n g_n^0$, and form the weighted centered moments $\widehat{C}$ and $\widehat{d}$. The coefficient

$$\widehat{\lambda} = -(\widehat{C} + \eta I)^\dagger \widehat{d} \tag{8}$$

minimizes $\sum_n \omega_n(g_n^0 + a_n^\top\lambda - c)^2 + \eta\|\lambda\|_2^2$ over $(\lambda, c)$. This is the VCG fit reused by RCPT.

**Structured penalty and headroom.** Delete observation $n$, renormalize $\omega_i^{(-n)} = \omega_i/(1-\omega_n)$, and refit with the same ridge convention. With the deleted weighted means,

$$e_n = g_n^0 + a_n^\top\widehat{\lambda}^{(-n)} - \widehat{c}^{(-n)}, \qquad \widehat{c}^{(-n)} = \bar{g}_{-n}^0 + \bar{a}_{-n}^\top\widehat{\lambda}^{(-n)}. \tag{9}$$

**Algorithm 1** Risk-Calibrated Proposal Transport (structured $H, P$)

**Input:** target path, Anchor, basis, weighted ensemble, ridge $\eta$.

1. Fit $\widehat{\lambda}$ by Eq. (8).
2. Compute deletion residuals by Eq. (9).
3. Form $\widehat{P}$, $\widehat{H}$, and $\widehat{\alpha}$ by Eq. (10).
4. Deploy $v^0 + \sum_j \widehat{\alpha}\widehat{\lambda}_j \phi_j$ online, or freeze the schedule for independent deployment.

Set $\widehat{\sigma}^2_{\mathrm{LOO}} = \sum_n \omega_n e_n^2$, $\widehat{N}_{\mathrm{eff}} = (\sum_n \omega_n^2)^{-1}$, $\widehat{r}_{\mathrm{eff}} = \mathrm{tr}[\widehat{C}(\widehat{C} + \eta I)^{\dagger}]$, and $S = \widehat{\lambda}^{\top}\widehat{C}\widehat{\lambda}$. RCPT uses

$$\widehat{P} = \widehat{\sigma}^2_{\mathrm{LOO}} \frac{\widehat{r}_{\mathrm{eff}}}{\widehat{N}_{\mathrm{eff}}}, \qquad \widehat{H} = [S - \widehat{P}]_+, \qquad \widehat{\alpha} = \frac{\widehat{H}}{\widehat{H} + \widehat{P} + \varepsilon_{\mathrm{num}}}, \qquad \widetilde{\lambda} = \widehat{\alpha}\widehat{\lambda}, \tag{10}$$

where $\varepsilon_{\mathrm{num}} > 0$ stabilizes a zero denominator. The rank-over-information penalty is motivated by the classical regression expansion $P = N^{-1}\,\mathrm{tr}(\Omega C^{\dagger}) + o(N^{-1})$, which reduces to $\sigma_\epsilon^2\,\mathrm{rank}(C)/N$ when $\Omega = \sigma_\epsilon^2 C$ (Lemma D.1, with proof in Appendix D.6). The effective-rank and normalized-weight factors give a scalar approximation to this covariance penalty. They retain a low-dimensional decision when the fitting ensemble is small.

**The criterion implemented by RCPT.** Equation (10) is exactly

$$\widehat{\alpha} = \operatorname*{arg\,min}_{0 \le \alpha \le 1} \widehat{\Delta}_{\mathrm{S}}(\alpha), \qquad \widehat{\Delta}_{\mathrm{S}}(\alpha) = (S + \varepsilon_{\mathrm{num}})\alpha^2 - 2(S - \widehat{P})\alpha. \tag{11}$$

This quadratic gives a direct operational meaning to the rule. $S$ measures the fitted correction's empirical variation, and $\widehat{P}$ discounts the supported improvement using deleted prediction residuals. The positive part returns Anchor when the estimated penalty exhausts the empirical gain. As supported gain grows relative to penalty, the rule retains more of the update. **Deployment and computation.** A frozen schedule can be reused on independent trajectories, while an online controller recalibrates at each fitted step. Proposal-Only sampling uses the selected drift directly; corrected SMC additionally retains FK weights and resampling. The implementation uses the same feature values and divergences as VCG. Additional work consists of small dense deletion-refit systems. Coefficient clipping and numerical fallbacks are specified separately in Appendix H.

The fitted correction defines a quadratic risk along its direction. RCPT approximates this quadratic using $S$ and $\widehat{P}$, and their errors control the selected risk.

**Theorem 3** (Calibration error of the implemented rule)**.** *Freeze a realized fit with finite residual second moments and write $f = a^{\top}\widehat{\lambda}$, $q = \mathrm{Var}_p(f)$, and $\kappa = \mathrm{Cov}_p(g^0 + f, f)$. If $|S - q| \le e_q$ and $|\widehat{P} - \kappa| \le e_\kappa$, then RCPT satisfies*

$$V(\widehat{\alpha}\widehat{\lambda}) - V(0) \le e_q + 2e_\kappa, \tag{12}$$

$$V(\widehat{\alpha}\widehat{\lambda}) - \min_{\alpha \in [0,1]} V(\alpha\widehat{\lambda}) \le 2e_q + 4e_\kappa + \varepsilon_{\mathrm{num}}. \tag{13}$$

For $q > 0$, the ray oracle is $\alpha_{\mathrm{ray}} = \Pi_{[0,1]}(1 - \kappa/q)$; when it is interior, excess risk is $q(\widehat{\alpha} - \alpha_{\mathrm{ray}})^2$ (Appendix D.4). This curvature-weighted regret measures amplitude-selection error. The Checker audit evaluates it against an independent corrected-SMC reference law (Appendix I.2).

### 4.1 Residual risk upper-bounds Wasserstein error

The centered FK rate is the source that separates the proposal from the target path. Its minimum correction energy preserves the spatial geometry of this source and gives a direct $W_2$ bound.

**Theorem 4** (FK residual and Wasserstein transport)**.** *Under Appendix C.1's conservative evolution and regularity assumptions, define $r_t^v = g_t^v - \mathbb{E}_{p_t} g_t^v$. Then $\partial_t \log Z_t = \mathbb{E}_{p_t} g_t^v$ and $\partial_t p_t = \mathcal{L}_{t,v}^* p_t + p_t r_t^v$. Suppose additionally that the diffusion on $\mathbb{R}^d$ is state-independent, the laws have finite second*

*moments, and* $\langle x-y, v_t(x)-v_t(y)\rangle \le L_t\|x-y\|^2$ *with* $L \in L^1(0,1)$. *Let* $-\nabla\cdot(p_t\nabla\psi_t) = p_t r_t^v$ *have finite energy* $\mathcal{E}_t(v) = \mathbb{E}_{p_t}\|\nabla\psi_t\|^2$, *with its corrected dynamics well posed. For* $\rho_0^v = p_0$,

$$W_2(p_1,\rho_1^v) \le \int_0^1 K_v(t,1)\sqrt{\mathcal{E}_t(v)}\,\mathrm{d}t, \qquad K_v(t,1) = \exp\left\{\int_t^1 L_s\,\mathrm{d}s\right\}. \tag{14}$$

*If* $p_t$ *has Poincaré constant* $C_P(t)$, *then* $\mathcal{E}_t(v) \le C_P(t)V_t(v)$, *where* $V_t(v) = \mathbb{E}_{p_t}(r_t^v)^2$.

The proof in Appendix C.3 adds the minimum-energy velocity that removes the residual and couples the corrected and proposal processes with common noise. The integral retains the residual's time profile before the optional Poincaré conversion to variance. This specializes the transport geometry of Benamou & Brenier (2000); Peyre (2018) to fitted FK proposals.

The same residual also controls the leading local cost of FK weighting.

**Proposition 1** (Local correction-weight second moment)**.** *Initialize a proposal segment at* $X_t \sim p_t$. *Let* $\overline{W}^v_{t,\Delta} = \mathrm{d}\mathbb{P}^v_{t,\Delta}/\,\mathrm{d}\mathbb{Q}^v_{t,\Delta}$ *be its exact Feynman–Kac segment weight, normalized to mean one. If* $\overline{W}^v_{t,\Delta} = 1 + \Delta r_t^v(X_t) + o_{L^2}(\Delta)$, *then*

$$\mathbb{E}_{\mathbb{Q}}[(\overline{W}^v_{t,\Delta})^2] = 1 + \chi^2(\mathbb{P}^v_{t,\Delta}\|\mathbb{Q}^v_{t,\Delta}) = 1 + \Delta^2 V_t(v) + o(\Delta^2). \tag{15}$$

The expansion follows by squaring the normalized local weight and taking expectation (Appendix C.4). **Residual risk therefore has two concrete roles:** it controls a transport upper bound for unweighted deployment and the leading local correction-weight variance for corrected sampling. We evaluate terminal discrepancies and genealogy separately to measure this behavior.

## 5 Experiments

**Evaluation Design.** We separate fitting from evaluation and never use terminal metrics to select a controller. Frozen schedules deploy on fresh trajectories; online controllers refit the live population. Common-state risk compares methods on shared evaluation states, whereas active-path risk follows each method's trajectory. The four experiments test independent risk, terminal generation, and particle genealogy; Appendix I gives implementation details and cohort definitions.

### 5.1 Synthetic Checker Distribution

Table 1: **Independent risk on common Checker fits.** Mean integrated Risk/Anchor over ten paired fits; lower is better. Direct LOO ablates calibration; ridge and PCR select regularization.

| $N_{\mathrm{fit}}$ | Anchor | VCG | Direct LOO | LOO ridge | LOO PCR | **RCPT** |
|---|---|---|---|---|---|---|
| 8 | 1.0000 | 244.5 | 0.9239 | 0.9211 | 1.0300 | **0.8013** |
| 16 | 1.0000 | 174.0 | 1.1260 | 0.8306 | 1.1740 | **0.4931** |
| 32 | 1.0000 | 2.797 | 0.3821 | 0.3705 | 0.6127 | **0.3589** |

Table 2: **Checker transport.** Ten paired seeds, fit 16/deploy 2,000; all endpoints are finite. VCG/RCPT use frozen schedules and ECG the raw configuration. Lineages counts terminal ancestors among 2,000 SMC particles. Bold marks the best mean; Appendix Table I.2 ablates calibration.

| | Proposal-Only | | | Corrected SMC | | | |
|---|---|---|---|---|---|---|---|
| Controller | $W_1\downarrow$ | $W_2\downarrow$ | MMD$\downarrow$ | $W_1\downarrow$ | $W_2\downarrow$ | MMD$\downarrow$ | Lineages$\uparrow$ |
| Anchor | 0.474 | 0.646 | **0.048** | 0.367 | 0.510 | 0.048 | 338.7 |
| VCG | 0.814 | 1.019 | 0.226 | 2.637 | 3.033 | 5.123 | 381.6 |
| ECG | 1.394 | 1.636 | 0.325 | 0.833 | 1.113 | 0.214 | 233.7 |
| RCPT | **0.442** | **0.575** | 0.073 | **0.316** | **0.464** | **0.042** | **568.5** |

Checker isolates finite-particle fitting from independent deployment. A common-fit audit uses 12 fields, $N = 8, 16, 32$, and ten paired fits per budget; independent corrected-SMC reference systems evaluate residual risk, and 10,000 fresh trajectories per fit test deployment. A separate matched cohort fits 16 particles and deploys 2,000 per method in both Proposal-Only and SMC modes. Appendix I.2 details the reference construction and common-fit replay, and Appendix I specifies the matched endpoint cohort.

RCPT has the lowest mean reference risk at every fitting budget (Table 1) and lower mean $W_1$ and $W_2$ than Anchor, VCG, and ECG in both matched deployment modes (Table 2). Under SMC, it also has the lowest MMD and retains 568.5 terminal lineages, versus 338.7 for Anchor. Proposal-Only MMD favors Anchor, illustrating that endpoint rankings depend on the metric. Figure 2 visualizes generated distributions; Appendix Table I.4 reports common-fit endpoints and completion counts.

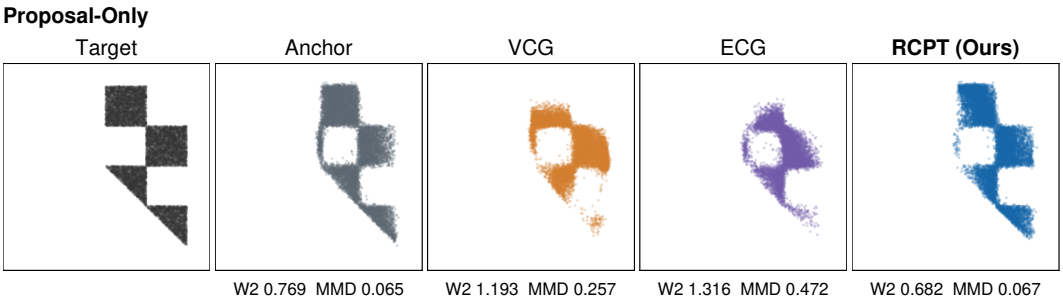


Figure 2: **Separate Checker Proposal-Only replay with ECG.** Frozen fit-16 schedules, 10,000 fresh trajectories, and no resampling. Labels use a fixed 2,000-sample endpoint audit. This illustrative replay is separate from Table 2. RCPT gives the lowest $W_2$; the corresponding SMC results are in Appendix Table I.1.

### 5.2 Scaffold Decoration on CrossDocked2020

We evaluate all 76 CrossDocked2020 scaffold–pocket tasks of Lee et al. (2026). Four expert paths supply fragment, ligand, conformation, and pocket information. Sixteen particles fit each 12-field schedule, which is frozen for 16 fresh paired Proposal-Only outputs. Chemistry and docking are evaluated afterward; unique useful yield counts distinct valid, scaffold-retaining molecules that beat the reference docking score per generation slot. Appendix I.3 specifies the models, fields, and evaluation protocol.

RCPT restores useful molecular generation after the expanded VCG fit (Table 3). Validity rises from 0.229 to 0.434 and unique useful yield from 0.118 to 0.251, with paired 95% pocket-bootstrap improvement intervals $[0.151, 0.260]$ and $[0.086, 0.181]$. A separate 64-output deployment of the same schedules reproduces the VCG repair across 4,864 candidates per arm (Appendix Table I.5). Figure 3 shows matched molecular examples.

Table 3: **Scaffold decoration.** Means over 76 pockets, fit 16/deploy 16. Yields use fixed generation denominators; risk is active-path variance normalized by Anchor. Unique useful counts distinct valid, scaffold-retaining, reference-beating outputs. †ECG is a separate 76-pocket run.

| Controller | Validity↑ | Scaffold↑ | Risk/Anchor↓ | Dock success↑ | Affinity↑ | Unique useful↑ |
|---|---|---|---|---|---|---|
| Anchor | 0.425 | 0.416 | 1.000 | 0.890 | 0.282 | 0.236 |
| VCG | 0.229 | 0.225 | 26.275 | 0.569 | 0.157 | 0.118 |
| ECG† | 0.221 | 0.215 | 3.593 | 0.465 | 0.140 | 0.126 |
| RCPT | **0.434** | **0.419** | **0.877** | **0.903** | **0.297** | **0.251** |

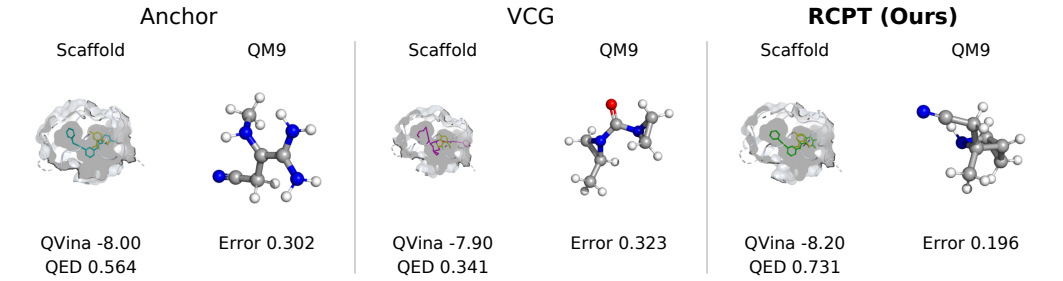


Figure 3: **Molecular generation.** Matched scaffold (left: lower QVina, higher QED) and QM9 (right: lower property error) examples for each proposal. Tables 3 and 4 report performance.

### 5.3 Multi-Property Molecular Reward Steering on QM9

We steer pretrained EEGSDE (Bao et al., 2023) toward polarizability and HOMO–LUMO-gap targets on QM9 (Ramakrishnan et al., 2014) using 12 gated reward-gradient fields. Each of eight split seeds fits on 16 particles and deploys eight fresh SMC batches of 16; eight online seeds instead refit 144 live particles. Independent property evaluators and canonical molecular identities measure attainment and diversity. Appendix I.4 gives the reward, controller, and resampling details.

RCPT improves property attainment and genealogy in split deployment (Table 4): unique successful molecules increase from $13.88$ for VCG to $23.00$ per 128 outputs, while the surviving-lineage fraction rises from $0.1611$ to $0.3525$. Online, both VCG and RCPT preserve all 144 lineages; RCPT retains a mean amplitude of $0.945$ and achieves slightly better mean evaluator error and unique success. These protocols demonstrate both repair of an unstable fit and retention of a well-supported update; Appendix Table I.6 reports seed variability.

Table 4: **Multi-property molecular steering.** Eight-seed means. Unique success counts distinct valid target-meeting identities per 128 split or 144 online outputs. Lineages are normalized by batch size: $N = 16$ split, $N = 144$ online.

| Protocol | Controller | Eval. MAE↓ | Unique success↑ | Lineages/$N$ ↑ |
|---|---|---|---|---|
| Split $16 + 8 \times 16$ | Anchor | 0.9171 | 20.75/128 | 0.3311 |
| | VCG | 0.9281 | 13.88/128 | 0.1611 |
| | RCPT | **0.8997** | **23.00/128** | **0.3525** |
| Online $N = 144$ | Anchor | 1.0570 | 0.75/144 | 0.0069 |
| | VCG | 0.8027 | 45.25/144 | **1.0000** |
| | RCPT | **0.7994** | **45.50/144** | **1.0000** |

### 5.4 Improving Guided Image Generation on CIFAR-10

The image experiment uses pretrained DLSM (Chao et al., 2022) and ten classifier-gradient fields. A frozen fit-16 schedule generates 32 images per class, giving 320 outputs per method. Independent ResNet and DLSM classifiers measure class alignment; Inception, DINO, and CLIP features measure terminal quality. Appendix I.5 details sampling, evaluators, and the separate independent-state risk audit.

RCPT improves all seven reported endpoint metrics over Anchor and VCG in this evaluation (Table 5). Relative to VCG, independent accuracy increases from $0.6063$ to $0.6438$, KID falls from $0.2370$ to $0.1595$, and Clean-FID falls from $79.692$ to $78.771$. A separate all-class mechanism study finds lower residual risk than VCG in 48/50, 47/50, and 42/50 class–seed cells at fitting populations 8, 16, and 32. Appendix Figure I.1 gives qualitative examples.

Table 5: **Class-conditional Image Generation on CIFAR-10.** All 320 outputs enter evaluation. KID uses independent ResNet features; FID and DINO FD use their corresponding feature spaces. Sampling and evaluator details are in Appendix I.5.

| Controller | Ind. acc.↑ | DLSM acc.↑ | KID↓ | FID↓ | DINO FD↓ | DINO MMD↓ | CLIP margin↑ |
|---|---|---|---|---|---|---|---|
| Anchor | 0.6406 | 0.9188 | 0.1707 | 79.048 | 0.19474 | 0.01205 | 0.00735 |
| VCG | 0.6063 | 0.8938 | 0.2370 | 79.692 | 0.20130 | 0.01493 | 0.00460 |
| RCPT | **0.6438** | **0.9281** | **0.1595** | **78.771** | **0.19380** | **0.01034** | **0.00774** |

## 6 Discussion

**Do existing regularization methods already solve the problem?** The ablations separate changes to the fitted direction through ridge or PCR from changes to the deployed amplitude through direct or structured calibration. On common Checker fits at $N = 16$, RCPT's mean Risk/Anchor is $0.4931$, versus $0.8306$ for LOO-selected ridge and $1.1740$ for PCR (Table 1); both paired 95% bootstrap intervals favor RCPT. RCPT captures $80.7\%$ of reference-ray improvement (Appendix Table B.1); Appendix B gives the paired intervals.

**Does reducing controller capacity solve the problem?** Reducing capacity can stabilize residual fitting but discard useful corrections: VCG-4 has lower scaffold risk, while full-class RCPT has higher validity and unique useful yield and improves all seven image metrics in the reported run

(Appendix Tables B.2 and B.5). Although oracle directionwise shrinkage can improve on a common amplitude, estimating additional amplitudes from finite particles can overfit deletion noise and increase deployment risk, motivating RCPT's scalar calibration (Appendix F).

**The structured decision adds value beyond direct calibration.** Relative to the Direct LOO amplitude ablation, RCPT improves paired Checker risk at $N = 16$, mean online QM9 unique success, and all seven CIFAR terminal metrics. On Scaffold, Direct LOO has lower common-state risk, while RCPT has stronger downstream means with paired intervals including zero (Appendix Tables B.3, B.4, and B.5).

**Do simple baselines already solve the problem?** Fixed shrinkage, ridge, averaging, and norm caps can repair particular deployments, but their settings transfer unevenly. Fixed $\alpha = 0.50$ exceeds RCPT's molecular yields on the ten-pocket scaffold subset, yet worsens mean QM9 evaluator error relative to VCG in the four-seed audit. Fixed ridge is similarly setting-sensitive. RCPT supplies one data-dependent rule across these tasks; the evidence supports adaptive protection, not uniform superiority over every tuned heuristic (Appendix Tables I.12 and I.13).

**Supported updates remain inexpensive.** Figure 4 shows conservative scaffold and image amplitudes and stronger retention on Checker and online QM9. In the scaffold capacity audit, increasing the fitting budget from 16 to 64 raises mean RCPT amplitude from $0.194$ to $0.906$. Calibration adds no expert calls; measured QM9 overhead is $3.13$ seconds, or $0.081\%$, with equal probed peak memory. Appendix Table B.6 and Appendix K separate application cost from selector timings.

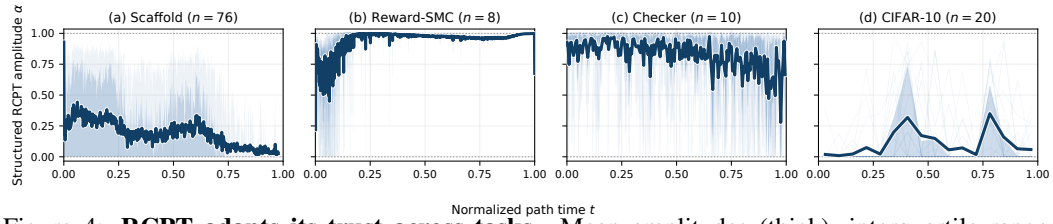


Figure 4: **RCPT adapts its trust across tasks.** Mean amplitudes (thick), interquartile ranges (bands), and individual units (thin) against normalized native time. Zero recovers Anchor; one retains VCG. See Appendix J.2 for more details.

**Why not refit VCG online?** Online refitting tracks changes in the particle distribution, but each update still estimates a controller from a finite ensemble. With $m$ particles, the centered empirical feature covariance has rank at most $m - 1$. Thus, repeated fitting does not by itself increase the independent information available at a given time. RCPT applies the same calibration principle to frozen and online fits. In the online QM9 experiment with 144 particles, it retains most of the VCG update, achieves slightly better terminal performance, and preserves all lineages (Table 4). This illustrates how calibration accommodates informative online fits while controlling the amplitude supported by the available particles.

**Limitations and future directions.** The structured guarantee is conditional on calibration error, and residual upper bounds do not impose a universal ordering of endpoints or genealogy. The common-fit $N = 32$ SMC audit favors Anchor over RCPT in $W_2$ by $0.0817$ (paired 95% interval $[0.0051, 0.1629]$), while the larger scaffold deployment is essentially tied with Anchor. Extending frozen-schedule guarantees to online interacting particles requires dependence and target-reference controls; lineage-aware calibration and stable grouped shrinkage are natural next steps. Multiple fitting seeds per scaffold task and broader matched regularization cohorts would further test robustness. Appendix H.4 states the relevant transfer conditions.

## 7 Conclusion

Finite-particle proposal fitting creates a deployment decision that population control alone does not resolve. RCPT selects how much of an estimated FK correction to apply using the existing regression information. The headroom–penalty identity explains the statistical failure, and realized-ray analysis makes the quality of the selected amplitude directly testable. Common-fit comparisons show a substantial advantage over adaptive regularization at a scarce-particle budget, while controlled and molecular deployments establish practical repair of harmful VCG fits. The resulting method provides an inexpensive way to retain supported adaptation when fitting information is limited.

### AI USE STATEMENT

Large language models (LLMs) were used to revise author-written text for clarity, concision, and organization, and to assist with literature retrieval, theoretical formulation, mathematical derivations and proof writing, feedback on experimental design, code development, and interpretation of results. LLMs also assisted with drafting and revising manuscript passages and preparing figures, tables, references, and LaTeX. The authors determine the final research claims and take responsibility for the correctness, originality, and verification of all AI-assisted text, proofs, code, and interpretations included in this work.

### ETHICS STATEMENT

Generated molecules are computational candidates, not validated compounds. Docking and learned-property scores do not establish biological activity, synthesizability, or safety. The method inherits the limitations of the pretrained generators and evaluators used for steering.

### REPRODUCIBILITY STATEMENT

We provide supplementary source code for the structured RCPT estimator, the Direct LOO calibration ablation, and integrations for Checker, scaffold decoration, QM9 reward steering, and CIFAR-10. The package includes application drivers, configuration and dependency files, and CPU unit tests for the core estimator; its README describes setup commands, entrypoints, and external asset requirements. Appendices C–F state the theoretical assumptions and proofs. Appendix H specifies weighted fitting, deletion refits, and numerical guards, while Appendix I documents datasets, fitting and deployment protocols, evaluation metrics, and statistical units. Pretrained weights, datasets, generated samples, and experiment outputs are not bundled. These materials support inspection and testing of the calibration rule and document the setup required for application-scale reproduction.

## REFERENCES

S. Agapiou, O. Papaspiliopoulos, D. Sanz-Alonso, and A. M. Stuart. Importance sampling: Intrinsic dimension and computational cost. *Statistical Science*, 32(3):405–431, 2017. ISSN 08834237, 21688745. URL http://www.jstor.org/stable/26408299.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Roni Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. In B. Kim, Y. Yue, S. Chaudhuri, K. Fragkiadaki, M. Khan, and Y. Sun (eds.), *International Conference on Learning Representations*, volume 2024, pp. 51304–51323, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/dfbb3d1807b21dadee735eb75069ada4-Paper-Conference.pdf.

Fan Bao, Min Zhao, Zhongkai Hao, Peiyao Li, Chongxuan Li, and Jun Zhu. Equivariant energy-guided SDE for inverse molecular design. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=r0otLtOwYW.

Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numerische Mathematik*, 84:375–393, 2000. doi: 10.1007/s002110050002.

Alexandros Beskos, Dan Crisan, and Ajay Jasra. On the stability of sequential Monte Carlo methods in high dimensions. *The Annals of Applied Probability*, 24(4):1396 – 1445, 2014. doi: 10.1214/13-AAP951. URL https://doi.org/10.1214/13-AAP951.

Arwen Bradley, Preetum Nakkiran, David Berthelot, James Thornton, and Joshua M. Susskind. Mechanisms of projective composition of diffusion models. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 5319–5354. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/bradley25a.html.

Chen-Hao Chao, Wei-Fang Sun, Bo-Wun Cheng, Yi-Chen Lo, Chia-Che Chang, Yu-Lun Liu, Yu-Lin Chang, Chia-Ping Chen, and Chun-Yi Lee. Denoising likelihood score matching for conditional score-based data generation. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=LcF-EEt8cCC.

Sourav Chatterjee and Persi Diaconis. The sample size required in importance sampling. *The Annals of Applied Probability*, 28(2):1099–1135, 2018. ISSN 10505164, 21688737. URL https://www.jstor.org/stable/26542331.

Hyungjin Chung, Jeongsol Kim, Geon Yeong Park, Hyelin Nam, and Jong Chul YE. Cfg++: Manifold-constrained classifier free guidance for diffusion models. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 30824–30850, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/4c9477b9e2c7ec0ad3f4f15077aaf85a-Paper-Conference.pdf.

Julien Cornebise, Éric Moulines, and Jimmy Olsson. Adaptive methods for sequential importance sampling with application to state space models. *Statistics and Computing*, 18(4):461–480, 2008.

Pierre Del Moral and Emmanuel Rio. Concentration inequalities for mean field particle models. *The Annals of Applied Probability*, 21(3):1017–1052, 2011. doi: 10.1214/10-AAP716. URL https://arxiv.org/abs/1211.1837.

Arnaud Doucet, Simon Godsill, and Christophe Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

Yilun Du, Conor Durkan, Robin Strudel, Joshua B. Tenenbaum, Sander Dieleman, Rob Fergus, Jascha Sohl-Dickstein, Arnaud Doucet, and Will Sussman Grathwohl. Reduce, reuse, recycle: Compositional generation with energy-based diffusion models and MCMC. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 8489–8510. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/du23a.html.

Bradley Efron. The estimation of prediction error: Covariance penalties and cross-validation. *Journal of the American Statistical Association*, 99(467):619–632, 2004. doi: 10.1198/016214504000000692.

Bradley Efron and Carl Morris. Stein's estimation rule and its competitors: An empirical Bayes approach. *Journal of the American Statistical Association*, 68(341):117–130, 1973. doi: 10.1080/01621459.1973.10481350.

Paul G. Francoeur, Tomohide Masuda, Jocelyn Sunseri, Andrew Jia, Ian Iovanisci, Ian Snyder, and David R. Koes. Three-dimensional convolutional neural networks and a cross-docked data set for structure-based drug design. *Journal of Chemical Information and Modeling*, 60(9):4200–4215, 2020. doi: 10.1021/acs.jcim.0c00411.

Mohsin Hasan, Marta Skreta, Alan Aspuru-Guzik, Yoshua Bengio, and Kirill Neklyudov. Discrete feynman-kac correctors. In *2nd AI for Math Workshop @ ICML 2025*, 2025. URL https://openreview.net/forum?id=CafvO7fIdq.

Jeremy Heng, Adrian N. Bishop, George Deligiannidis, and Arnaud Doucet. Controlled sequential monte carlo. *The Annals of Statistics*, 48(5):pp. 2904–2929, 2020. ISSN 00905364, 21688966. URL https://www.jstor.org/stable/27028726.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. In *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications*, 2021. URL https://openreview.net/forum?id=qw8AKxfYbI.

Emiel Hoogeboom, Víctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8867–8887. PMLR, 2022. URL https://proceedings.mlr.press/v162/hoogeboom22a.html.

Vineet Jain, Kusha Sareen, Mohammad Pedramfar, and Siamak Ravanbakhsh. Diffusion tree sampling: Scalable inference-time alignment of diffusion models. In *Advances in Neural Information Processing Systems*, volume 38, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/d6484394c4cb5e1f4ecad8d90b912025-Abstract-Conference.html.

Kwanyoung Kim. VGAS: Variance-reduced guidance and adaptive selection for training-free reward alignment in discrete diffusion. *arXiv preprint arXiv:2608.26585*, 2026. URL https://arxiv.org/abs/2608.26585v2.

Kyurae Kim, Zuheng Xu, Jacob R. Gardner, and Trevor Campbell. Tuning sequential Monte Carlo samplers via greedy incremental divergence minimization. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 30826–30878, 2025. URL https://proceedings.mlr.press/v267/kim25ak.html.

Lee Cheuk Kit, Paul Jeha, Jes Frellsen, Pietro Lio, Michael Samuel Albergo, and Francisco Vargas. Debiasing guidance for discrete diffusion with sequential monte carlo. In *Frontiers in Probabilistic Inference: Learning meets Sampling*, 2025. URL https://openreview.net/forum?id=vg0dOu3w9g.

Jere Koskela, Paul A. Jenkins, Adam M. Johansen, and Dario Spanò. Asymptotic genealogies of interacting particle systems with an application to sequential monte carlo. *The Annals of Statistics*, 48(1):pp. 560–583, 2020. ISSN 00905364, 21688966. URL https://www.jstor.org/stable/26923113.

Ziseok Lee, Minyeong Hwang, Wooyeol Lee, Sanghyun Jo, Jihyung Ko, Young Bin Park, Jae-Mun Choi, Eunho Yang, and Kyungsu Kim. On the collapse of generative paths: A criterion and correction for diffusion steering. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=emv2qsi3TG.

Rémi Leluc, François Portier, and Johan Segers. Control variate selection for Monte Carlo integration. *Statistics and Computing*, 31(4):50, 2021. doi: 10.1007/s11222-021-10011-z. URL https://arxiv.org/abs/1906.10920.

Christian A. Naesseth, Fredrik Lindsten, and Thomas B. Schön. Elements of sequential monte carlo. *Foundations and Trends in Machine Learning*, 12(3):187–306, 11 2019. ISSN 1935-8237. doi: 10.1561/2200000074. URL https://doi.org/10.1561/2200000074.

Zijing Ou, Chinmay Pani, and Yingzhen Li. Inference-time scaling of discrete diffusion models via importance weighting and optimal proposal design. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/3df1eca840e82b11bbc33f68c773c38e-Abstract-Conference.html.

Rémi Peyre. Comparison between W2 distance and $\dot{h}^{-1}$ norm, and localization of Wasserstein distance. *ESAIM: Control, Optimisation and Calculus of Variations*, 24(4):1489–1501, 2018. doi: 10.1051/cocv/2017050. URL https://www.numdam.org/articles/10.1051/cocv/2017050/.

Angus Phillips, Hai-Dang Dau, Michael John Hutchinson, Valentin De Bortoli, George Deligiannidis, and Arnaud Doucet. Particle denoising diffusion sampler. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 40688–40724, 2024. URL https://proceedings.mlr.press/v235/phillips24a.html.

Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1:140022, 2014. doi: 10.1038/sdata.2014.22.

Yinuo Ren, Wenhao Gao, Lexing Ying, Grant M. Rotskoff, and Jiequn Han. Driftlite: Lightweight drift control for inference-time scaling of diffusion models. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=l01eG3Qikl.

Arne Schneuing, Charles Harris, Yuanqi Du, Kieran Didi, Arian Jamasb, Ilia Igashov, Weitao Du, Carla Gomes, Tom L. Blundell, Pietro Liò, Max Welling, Michael Bronstein, and Bruno Correia. Structure-based drug design with equivariant diffusion models. *Nature Computational Science*, 4 (12):899–909, 2024. doi: 10.1038/s43588-024-00737-x.

Raghav Singhal, Zachary Horvitz, Ryan Teehan, Mengye Ren, Zhou Yu, Kathleen Mckeown, and Rajesh Ranganath. A general framework for inference-time scaling and steering of diffusion models. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 55810–55827. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/singhal25b.html.

Marta Skreta, Tara Akhound-Sadegh, Viktor Ohanesian, Roberto Bondesan, Alan Aspuru-Guzik, Arnaud Doucet, Rob Brekelmans, Alexander Tong, and Kirill Neklyudov. Feynman-kac correctors in diffusion: Annealing, guidance, and product of experts. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 55906–55949. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/skreta25a.html.

Chris Snyder, Thomas Bengtsson, and Mathias Morzfeld. Performance bounds for particle filters using the optimal proposal. *Monthly Weather Review*, 143(11):4750 – 4761, 2015. doi: 10.1175/MWR-D-15-0144.1. URL https://journals.ametsoc.org/view/journals/mwre/143/11/mwr-d-15-0144.1.xml.

Jiaming Song, Qinsheng Zhang, Hongxu Yin, Morteza Mardani, Ming-Yu Liu, Jan Kautz, Yongxin Chen, and Arash Vahdat. Loss-guided diffusion models for plug-and-play controllable generation. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 32483–32498. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/song23k.html.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

Leah F. South, Chris J. Oates, Antonietta Mira, and Christopher Drovandi. Regularized zero-variance control variates. *Bayesian Analysis*, 18(3), 2023. doi: 10.1214/22-BA1328. URL https://arxiv.org/abs/1811.05073.

Stanislas Strasman, Gabriel Victorino Cardoso, Sylvain Le Corff, Vincent Lemaire, and Antonio Ocello. Non-asymptotic error bounds for SMC with biased proposals: Application to conditional diffusion sampling. *arXiv preprint arXiv:2607.04780*, 2026. URL https://arxiv.org/abs/2607.04780.

James Thornton, Louis Béthune, Ruixiang ZHANG, Arwen Bradley, Preetum Nakkiran, and Shuangfei Zhai. Composition and control with distilled energy diffusion models and sequential monte carlo. In Yingzhen Li, Stephan Mandt, Shipra Agrawal, and Emtiyaz Khan (eds.), *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 3259–3267. PMLR, 03–05 May 2025. URL https://proceedings.mlr.press/v258/thornton25a.html.

Francisco Vargas, Shreyas Padhy, Denis Blessing, and Nikolas Nüsken. Transport meets variational inference: Controlled Monte Carlo diffusions. In *International Conference on Learning Representations*, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/hash/f3716db40060004d0629d4051b2c57ab-Abstract-Conference.html.

Weixin Wang, Yu Yang, Wei Deng, and Pan Xu. Inference-time alignment of diffusion models via trust-region iterative twisted sequential Monte Carlo. *arXiv preprint arXiv:2605.25123*, 2026. URL https://arxiv.org/abs/2605.25123.

Nick Whiteley and Anthony Lee. Twisted particle filters. *The Annals of Statistics*, 42(1):115 – 141, 2014. doi: 10.1214/13-AOS1167. URL https://doi.org/10.1214/13-AOS1167.

Luhuan Wu, Brian Trippe, Christian Naesseth, David Blei, and John Cunningham. Practical and asymptotically exact conditional sampling in diffusion models. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 31372–31403. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/63e8bc7bbf1cfea36d1d1b6538aecce5-Paper-Conference.pdf.

Luhuan Wu, Yi Han, Christian Andersson Naesseth, and John Cunningham. Reverse diffusion sequential monte carlo samplers. In D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (eds.), *Advances in Neural Information Processing Systems*, volume 38, pp. 69445–69459. Curran Associates, Inc., 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/6490ab6ef8eb7116c6c1889795d21748-Paper-Conference.pdf.

Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=PzcvxEMzvQC.

Haotian Ye, Haowei Lin, Jiaqi Han, Minkai Xu, Sheng Liu, Yitao Liang, Jianzhu Ma, James Zou, and Stefano Ermon. Tfg: Unified training-free guidance for diffusion models. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 22370–22417. Curran Associates, Inc., 2024. doi: 10.52202/079017-0704. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/2818054fc6de6dacdda0f142a3475933-Paper-Conference.pdf.

Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem, and Jian Zhang. Freedom: Training-free energy-guided conditional diffusion model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 23174–23184, October 2023.

**RCPT: Risk-Calibrated Proposal Transport for Finite-Particle Diffusion Steering**
**Technical Supplement**

# A Supplementary Overview

The supplement first presents additional ablations and cost measurements, then the proofs needed by the main paper, implementation details, and additional evaluations. RCPT denotes the structured $H, P$ rule in all reported tables and figures. Direct LOO is defined as a calibration ablation. Time runs from noise at $t = 0$ to the terminal law at $t = 1$, unless an experiment states its native reverse-time coordinate.

**Scope of the guarantees.** The transport bounds concern a fixed proposal schedule and the specified target path. The $H, P, B$ identities average over fitting samples, while the structured calibration theorem controls a realized decision through its quadratic approximation errors. The directional oracle and counterexample explain the choice of one amplitude. None of these results orders all terminal metrics, certifies the online interacting-particle loop, or identifies weight ESS with independent regression information.

**Broader impact.** RCPT may reduce the particle and memory cost of inference-time steering. It inherits the limitations and potential misuse of the underlying generators: generated molecules remain computational hypotheses, and image steering can also produce misleading content. Proposal calibration supplies no scientific validity or safety guarantee.

# B Additional Ablation and Cost Tables

## B.1 Common-fit calibration and paired comparisons

The paired $N = 16$ comparison supports the structured decision. RCPT-minus-comparator Risk/Anchor differences are $-0.6331$ for Direct LOO (95% CI $[-1.363, -0.0669]$), $-0.3375$ for LOO-selected ridge (95% CI $[-0.7484, -0.0510]$), and $-0.6804$ for LOO-selected PCR (95% CI $[-1.182, -0.2518]$). All three paired fitting-seed bootstrap intervals exclude zero, indicating lower mean reference risk at the pairwise 95% level, conditional on the reference construction.

The reference-ray oracle quantifies the useful correction available along the fitted direction. For mean normalized risks $\bar{R}$, define the descriptive fraction

$$\frac{1 - \bar{R}(M)}{1 - \bar{R}(\text{ray oracle})}.$$

Anchor gives zero and the reference-ray oracle gives one. The diagnostic oracle risks are $0.5758$, $0.3717$, and $0.2570$ at $N = 8, 16, 32$, respectively. Applying this expression to the reported aggregate risks gives RCPT approximately $46.8\%$, $80.7\%$, and $86.3\%$ at $N = 8, 16, 32$. This is a ratio of aggregate improvements, rather than a mean of per-fit fractions. Table B.1 pairs it with the reported curvature-weighted ray regret. The oracle uses evaluation states and serves only as a diagnostic.

Table B.1: **Same-ray calibration ablation.** Captured improvement uses the diagnostic oracle risks specified above and the mean normalized risks in Table 1. Regret is the mean integrated excess risk over the reference-ray oracle. All deployable rows use the same full-data VCG direction.

| | $N = 8$ | | $N = 16$ | | $N = 32$ | |
|---|---|---|---|---|---|---|
| Decision | Captured | Regret | Captured | Regret | Captured | Regret |
| Anchor | 0% | 15.720 | 0% | 23.290 | 0% | 27.540 |
| Direct LOO | 17.9% | 12.900 | −20.1% | 27.960 | 83.2% | 4.635 |
| **RCPT** | **46.8%** | **8.359** | **80.7%** | **4.500** | **86.3%** | **3.776** |

### B.2 SCAFFOLD CAPACITY AND CALIBRATION

The capacity study and matched calibration cohort answer different questions and retain separate means. Appendix I.3 gives the fitting and deployment protocols.

Table B.2: **Scaffold controller-capacity ablation.** The same 76 pockets and 16 deployment slots per arm. VCG-4 retains the ungated fields; VCG-12 and RCPT use all 12 fields. Mean RCPT amplitude is 0.194 at fit 16 and 0.906 at fit 64. All yields use the fixed generation denominator.

| Fit | Controller | Validity↑ | Scaffold↑ | Risk/Anchor↓ | Dock success↑ | Affinity↑ | Unique useful↑ |
|---|---|---|---|---|---|---|---|
| 16 | VCG-4 | 0.383 | 0.374 | 0.650 | 0.867 | 0.257 | 0.225 |
| 16 | VCG-12 | 0.229 | 0.225 | 26.275 | 0.569 | 0.157 | 0.118 |
| 16 | RCPT | 0.434 | 0.419 | 0.877 | 0.903 | 0.297 | 0.251 |
| 64 | VCG-4 | 0.354 | 0.335 | 0.297 | 0.936 | 0.243 | 0.206 |
| 64 | VCG-12 | 0.335 | 0.322 | 1.069 | 0.915 | 0.236 | 0.200 |
| 64 | RCPT | 0.441 | 0.422 | 0.548 | 0.911 | 0.312 | 0.271 |

Table B.3: **Matched Scaffold calibration ablation.** A separate fit-16/deploy-16 cohort on all 76 pockets. Anchor, Direct LOO, and RCPT isolate the addition of calibration and its structured form; the full VCG fit is shown for reference. These means are kept separate from the earlier cohort. Risk is the active-path diagnostic; paired uncertainty is reported in Appendix I.3.

| Controller | Validity↑ | Scaffold↑ | Risk/Anchor↓ | Dock success↑ | Affinity↑ | Unique useful↑ |
|---|---|---|---|---|---|---|
| Anchor | 0.438 | 0.415 | 1.000 | 0.898 | 0.294 | 0.251 |
| VCG-12 | 0.220 | 0.213 | 21.184 | 0.562 | 0.143 | 0.111 |
| Direct LOO | 0.418 | 0.400 | 0.756 | 0.914 | 0.289 | 0.247 |
| RCPT | 0.438 | 0.423 | 0.876 | 0.906 | 0.303 | 0.262 |

### B.3 QM9 AND IMAGE CALIBRATION

These ablations use the online QM9 budget and the CIFAR terminal-generation cohort described in Appendices I.4 and I.5.

Table B.4: **QM9 calibration ablation.** Eight online seeds, $N = 144$. Anchor has no fitted correction; Direct LOO uses direct deleted-loss calibration; RCPT uses the structured headroom–penalty decision. Unique success counts distinct valid target-meeting identities per 144 outputs.

| Calibration | MAE↓ | Unique success /144↑ | Lineages/$N$ ↑ |
|---|---|---|---|
| Anchor | 1.0570 | 0.75 | 0.0069 |
| Direct LOO | 0.8019 | 43.13 | 0.9479 |
| RCPT | **0.7994** | **45.50** | **1.0000** |

### B.4 APPLICATION COST

In the common-fit Checker implementation, mean selector times per 200-step schedule are 0.132 seconds for RCPT, 1.31 for Direct LOO, 9.34 for selected ridge, and 13.03 for selected PCR. These implementation-specific selector timings are separate from the hardware-identified generation measurements below. Appendix K gives the instrumentation.

Table B.5: **CIFAR-10 calibration and controller-capacity ablations.** The same 320-output seed as Table 5. The upper block adds direct calibration to Anchor, then restores RCPT's structured decision. The lower block tests a rank-four restriction against the full ten-field fit and calibrated full fit. The shared RCPT row is repeated for readability.

| Controller | Ind. acc.↑ | DLSM acc.↑ | KID↓ | FID↓ | DINO FD↓ | DINO MMD↓ | CLIP margin↑ |
|---|---|---|---|---|---|---|---|
| *Calibration ablation* | | | | | | | |
| Anchor | 0.6406 | 0.9188 | 0.1707 | 79.048 | 0.19474 | 0.01205 | 0.00735 |
| Direct LOO | 0.6313 | 0.9219 | 0.2052 | 78.898 | 0.19398 | 0.01202 | 0.00646 |
| RCPT | **0.6438** | **0.9281** | **0.1595** | **78.771** | **0.19380** | **0.01034** | **0.00774** |
| *Controller-capacity ablation* | | | | | | | |
| VCG-4 | 0.6156 | 0.9250 | 0.2279 | 79.468 | 0.19465 | 0.01228 | 0.00523 |
| VCG-10 | 0.6063 | 0.8938 | 0.2370 | 79.692 | 0.20130 | 0.01493 | 0.00460 |
| RCPT | **0.6438** | **0.9281** | **0.1595** | **78.771** | **0.19380** | **0.01034** | **0.00774** |

Table B.6: **Measured application cost.** Scaffold means use all 76 pockets on an RTX A6000 (48 GB). QM9 uses four recorded A100-SXM4 (80 GB) seeds; peak memory is from a matched 10-step probe. Times exclude offline evaluation. Full instrumentation is in Appendix K.

| Setting | Controller | Generation time (s) | Peak memory (GB) |
|---|---|---|---|
| Scaffold, 16 particles | VCG | 201.715 | 5.070 |
| | RCPT | 197.457 | 4.742 |
| QM9, 144 particles | VCG | 3879.518 | 40.204 |
| | RCPT | 3882.645 | 40.204 |

# C Transport Residual and Correction Weights

## C.1 Analytic setup

Let $\widetilde{p}_t > 0$ have finite normalizer $Z_t$ and sufficient space-time regularity to differentiate under the integral. For each fixed admissible drift schedule, assume the conservative operator $\mathcal{L}^*_{t,v}$ generates a Markov evolution $U^{v,*}_{s,t}$, the forward equation has a unique mild solution, and $\int_0^1 \|p_t r_t^v\|_{L^1}\,\mathrm{d}t < \infty$. Boundary fluxes vanish. Assume $r_t^v \in L^2(p_t)$ for the stated risk bounds. Random schedules fitted independently of deployment can be conditioned on throughout. These assumptions do not identify a particle-dependent online drift with a fixed single-particle Markov process.

Nonsmooth or vanishing terminal densities require a limiting argument. One may apply the dynamic result up to $1-\tau$ on a regular positive path, then let $\tau \downarrow 0$ if the endpoint laws converge in $W_2$ and the energy bound remains uniformly integrable. The triangle inequality and this convergence pass the upper bound to the limit.

## C.2 Proof of Theorem 4

*Proof.* The definition of $g_t^v$ gives $\partial_t \widetilde{p}_t = \mathcal{L}^*_{t,v}\widetilde{p}_t + g_t^v \widetilde{p}_t$. Integrating and using conservativity yields $\dot{Z}_t/Z_t = \mathbb{E}_{p_t} g_t^v$. Substituting $\widetilde{p}_t = Z_t p_t$ proves

$$\partial_t p_t = \mathcal{L}^*_{t,v} p_t + p_t r_t^v, \qquad r_t^v = g_t^v - \mathbb{E}_{p_t} g_t^v. \tag{C.1}$$

Since $p_0 = \rho_0^v$, variation of constants gives

$$p_1 - \rho_1^v = \int_0^1 U^{v,*}_{t,1}(p_t r_t^v)\,\mathrm{d}t. \tag{C.2}$$

Markov kernels contract the total-variation norm of signed measures. Therefore, with $d_{\mathrm{TV}}(p,q) = \frac{1}{2}\|p-q\|_1$,

$$d_{\mathrm{TV}}(p_1, \rho_1^v) \le \frac{1}{2}\int_0^1 \mathbb{E}_{p_t}|r_t^v|\,\mathrm{d}t \le \frac{1}{2}\int_0^1 \sqrt{V_t(v)}\,\mathrm{d}t.$$

This proves the normalized source identity. The Wasserstein part follows in Appendix C.3. For completeness, the optional final time relaxation is

$$d_{\mathrm{TV}}(p_1, \rho_1^v) \le \tfrac{1}{2} \int_0^1 \sqrt{V_t(v)}\, \mathrm{d}t \le \tfrac{1}{2} \sqrt{\int_0^1 V_t(v)\, \mathrm{d}t}. \tag{C.3}$$

□

For a statistic $m \in [a, b]$, the same argument implies $|\mathbb{E}_{p_1} m - \mathbb{E}_{\rho_1^v} m| \le (b-a) d_{\mathrm{TV}}(p_1, \rho_1^v)$. Zero residual along the entire path implies exact unweighted marginals by uniqueness. A reduced residual bound alone does not imply exactness or a strict reduction in terminal discrepancy: the propagated signed defects in Eq. (C.2) can cancel.

### C.3 Proof of the Wasserstein upper bound

Transport energy retains spatial information that residual variance discards. For a zero-mass signed density $h$, define

$$\|h\|_{-1,p} = \sup_{\int \|\nabla f\|^2 p \le 1} \left| \int f h \right|, \qquad \mathcal{A}_p = -\nabla \cdot (p \nabla). \tag{C.4}$$

We use weak solutions modulo constants, with zero boundary flux or sufficient decay at infinity. If $\mathcal{A}_p \psi = h$ has finite energy, testing its weak equation and using Cauchy–Schwarz gives $\|h\|_{-1,p}^2 = \int \|\nabla \psi\|^2 p = \langle h, \mathcal{A}_p^{-1} h \rangle$. In particular, $\mathcal{E}_t(v) = \|p_t r_t^v\|_{-1,p_t}^2$.

**Lemma C.1** (Minimum correction energy)**.** *Among velocities satisfying $-\nabla \cdot (p_t w) = p_t r_t^v$, the gradient $w = \nabla \psi_t$ minimizes $\mathbb{E}_{p_t} \|w\|^2$. If $p_t$ has Poincaré constant $C_P(t)$, then $\mathcal{E}_t(v) \le C_P(t) V_t(v)$.*

*Proof.* Every admissible $w$ differs from $\nabla \psi_t$ by a field $z$ with $\nabla \cdot (p_t z) = 0$. Integration by parts gives $\mathbb{E}_{p_t} \nabla \psi_t \cdot z = 0$, proving minimality by the Pythagorean identity. Choose the mean-zero representative of $\psi_t$. Then

$$\mathcal{E}_t(v) = \mathbb{E}_{p_t}(r_t^v \psi_t) \le \sqrt{V_t(v) C_P(t) \mathcal{E}_t(v)}.$$

Division proves the claim when the energy is positive; the zero case is immediate. □

**Proof of the Wasserstein bound in Theorem 4.** The corrected drift $v_t + \nabla \psi_t$ has marginals $p_t$ by Eq. (C.1) and uniqueness. Couple its process $X_t$ with the proposal $Y_t$ using the same additive Brownian noise and an optimal initial coupling. The noise cancels. With $D_t = (\mathbb{E}\|X_t - Y_t\|^2)^{1/2}$, the one-sided Lipschitz condition implies

$$\tfrac{1}{2} \partial_t D_t^2 \le L_t D_t^2 + D_t \sqrt{\mathcal{E}_t(v)}.$$

Regularization at $D_t = 0$ and Gronwall give the slightly more general bound

$$W_2(p_1, \rho_1^v) \le K_v(0,1) W_2(p_0, \rho_0^v) + \int_0^1 K_v(t,1) \sqrt{\mathcal{E}_t(v)}\, \mathrm{d}t. \tag{C.5}$$

This proves Eq. (14). Poincaré is needed only to replace energy by variance. The energy bound itself does not require it.

**What correction changes.** Exact FK weighting targets $p_t$ despite proposal mismatch. The preceding terminal bounds concern the unweighted proposal law $\rho_t^v$. They do not bound the error of a finite corrected particle system. The local weight result below explains a separate benefit of reducing residual variance. Particle error additionally depends on weights, dependence, resampling, and mixing (Del Moral & Rio, 2011).

### C.4 Proof of Proposition 1

*Proof.* Write $\overline{W} - 1 = \Delta r_t^v(X_t) + R_\Delta$ with $\|R_\Delta\|_2 = o(\Delta)$. Then

$$\mathbb{E}(\overline{W} - 1)^2 = \Delta^2 \mathbb{E}_{p_t}(r_t^v)^2 + 2\Delta \mathbb{E}[r_t^v(X_t) R_\Delta] + \mathbb{E}R_\Delta^2 = \Delta^2 V_t(v) + o(\Delta^2).$$

Because $\mathbb{E}\overline{W} = 1$, this equals both $\mathbb{E}\overline{W}^2 - 1$ and $\chi^2(\mathbb{P}\|\mathbb{Q})$. □

One sufficient route to the assumed expansion is $A_\Delta = \int_t^{t+\Delta} g_s^v(X_s)\,\mathrm{d}s = \Delta g_t^v(X_t) + o_{L^2}(\Delta)$ together with $\|e^{A_\Delta} - 1 - A_\Delta\|_2 = o(\Delta)$. Normalizing $e^{A_\Delta}$ then subtracts $\mathbb{E}_{p_t} g_t^v$ at first order. For example, bounded rates with local $L^2$ continuity satisfy these conditions. Uniform conclusions require uniform remainder control.

### C.5 Why local variance does not order genealogy

Let $Z$ be uniform on $\{-1, 1\}$ and $0 < \epsilon < 1$. Two weighting schemes use $G_1 = 1 + \epsilon Z$ and respectively $G_2 = 1 - \epsilon Z$ or $G_2 = 1 + \epsilon Z$. Every increment has mean one and squared coefficient of variation $\epsilon^2$ under this common reference law. However,

$$\mathrm{CV}^2(G_1 G_2) = 0 \quad \text{or} \quad \frac{4\epsilon^2}{(1+\epsilon^2)^2},$$

respectively. Thus marginal incremental second moments do not determine cumulative weight variation. Resampling adds further dependence on timing. For exactly Gaussian log weight $L$ with variance $s^2$,

$$\frac{\mathbb{E}e^{2L}}{(\mathbb{E}e^L)^2} = e^{s^2}, \qquad \mathrm{ESS}_{\mathrm{pop}}/N = e^{-s^2}, \tag{C.6}$$

but the variance of a sum of log increments includes their cross-covariances.

For multinomial resampling with weights $w_i$, the probability that all children share one parent is $\sum_i w_i^N$. Resampling resets numerical weights without restoring ancestors. We therefore report both the number of surviving initial ancestors and lineage ESS:

$$N_{\mathrm{anc}} = |\{a : \pi_a > 0\}|, \qquad \mathrm{ESS}_{\mathrm{anc}} = (\sum_a \pi_a^2)^{-1}, \tag{C.7}$$

where $\pi_a$ is the fraction of descendants of ancestor $a$.

### C.6 Time discretization and approximate residuals

On a grid $0 = t_0 < \cdots < t_K = 1$, hold $\lambda_k$ fixed on each interval. If $|\partial_t V_t(\lambda_k)| \le L_k$, then

$$\left| \int_{t_k}^{t_{k+1}} V_t(\lambda_k)\,\mathrm{d}t - \Delta_k V_{t_k}(\lambda_k) \right| \le \tfrac{1}{2} L_k \Delta_k^2. \tag{C.8}$$

Consequently the continuous-time proposal with that schedule satisfies

$$d_{\mathrm{TV}}(p_1, \rho_1^{\lambda_{0:K-1}}) \le \tfrac{1}{2} \sqrt{\sum_k \Delta_k V_{t_k}(\lambda_k) + \tfrac{1}{2} \sum_k L_k \Delta_k^2}. \tag{C.9}$$

The discretized sampler has an additional error $d_{\mathrm{TV}}(\rho_1^\lambda, \rho_{1,h}^\lambda)$; the bound does not assert a universal rate for it. If $r_t = \widehat{r}_t + \delta_t$, then

$$\|r\|_{L^2(\,\mathrm{d}t\,p_t)} \le \|\widehat{r}\|_{L^2(\,\mathrm{d}t\,p_t)} + \|\delta\|_{L^2(\,\mathrm{d}t\,p_t)}. \tag{C.10}$$

Score, divergence, and tracked-density errors therefore remain separate from coefficient-estimation error.

**Wasserstein model and numerical errors.** The direct transport argument also separates approximation errors without converting them to TV. If the true centered source is $r_t = \widehat{r}_t + \delta r_t$, then

$$\|p_t r_t\|_{-1,p_t} \le \|p_t \widehat{r}_t\|_{-1,p_t} + \|p_t \delta r_t\|_{-1,p_t} \le \sqrt{\widehat{\mathcal{E}}_t} + \sqrt{C_P(t)}\|\delta r_t\|_{L^2(p_t)}.$$

The last step alone uses Poincaré. A drift error $\delta v_t$ changes the source by a weighted divergence, whose negative norm is at most $\|\delta v_t\|_{L^2(p_t)}$ by the dual definition. Apply Eq. (C.5) to the actual drift with its corresponding stability factor. These are deterministic model-error terms, not extra coefficient-estimation penalties.

For a piecewise schedule, let $F(t) = K_v(t,1)\sqrt{\mathcal{E}_t(v)}$ and suppose $|F(t) - F(t_k)| \le \omega_k(t - t_k)$ on its $k$th interval. Then

$$W_2(p_1, \rho_1^v) \le \sum_k \Delta_k F(t_k) + \sum_k \int_0^{\Delta_k} \omega_k(s)\,\mathrm{d}s.$$

This follows by integrating the modulus bound and preserves the energy profile until the final quadrature. A discretized law $\rho_{1,h}^v$ and an empirical output $\widehat{\rho}_{1,h}^v$ add, by the triangle inequality, $W_2(\rho_1^v, \rho_{1,h}^v)$ and $W_2(\rho_{1,h}^v, \widehat{\rho}_{1,h}^v)$. Finite reference samples add their own error. For SMC, this last term is an interacting-particle error and cannot be replaced by an iid sampling rate without justification.

# D Finite-Particle Risk and Structured Calibration

## D.1 Population projection

The linear residual formula follows from $\mathcal{L}_{v+u}^* p = \mathcal{L}_v^* p - \nabla \cdot (pu)$. Expanding its variance gives $V(\lambda) = V(0) + 2d^\top \lambda + \lambda^\top C \lambda$. For any $z \in \ker C$, $z^\top(a - \mathbb{E}a) = 0$ almost surely, so $z^\top d = 0$. Thus $d \in \mathrm{range}(C)$ and completing the square proves Eq. (3), including singular $C$. This is the population VCG projection of Ren et al. (2026).

## D.2 Proof of Theorem 1

*Proof.* Apply Eq. (3) to each realized fit and take expectation: $\mathcal{R}(\widehat{\lambda}) = V_\star + P$. At Anchor, $V(0) = V_\star + H$. Subtraction proves the identity. □

## D.3 Proof of Theorem 2

*Proof.* Set $e = \widehat{\lambda} - \lambda^\star$. Then

$$\mathcal{R}(\alpha\widehat{\lambda}) - V_\star = \mathbb{E}\| - (1-\alpha)\lambda^\star + \alpha e\|_C^2 = (1-\alpha)^2 H + \alpha^2 P - 2\alpha(1-\alpha)B.$$

The coefficient of $\alpha^2$ is $D_0 = H + P + 2B = \mathbb{E}\|\widehat{\lambda}\|_C^2 \ge 0$. For $D_0 > 0$, minimizing the convex quadratic gives Eq. (6); clipping enforces the interval constraint. When $B = 0$, substitution of $H/(H+P)$ gives Eq. (7). If $D_0 = 0$, the fitted correction is constant $p$-almost surely for almost every fit, so its amplitude has no effect on variance. □

## D.4 Proof of Theorem 3

*Proof.* Since $S, \widehat{P} \ge 0$ and $\varepsilon_{\mathrm{num}} > 0$, the minimizer of Eq. (11) is $\Pi_{[0,1]}((S - \widehat{P})/(S + \varepsilon_{\mathrm{num}}))$. It equals Eq. (10), including $S \le \widehat{P}$. For the realized field, write

$$\Delta(\alpha) = V(\alpha\widehat{\lambda}) - V(0) = q\alpha^2 - 2(q - \kappa)\alpha. \tag{D.1}$$

For $q > 0$, the true minimizer is $\alpha_{\mathrm{ray}} = \Pi_{[0,1]}(1 - \kappa/q)$; for $q = 0$, the risk is constant. Let $e = e_q + 2e_\kappa$. Since $|\alpha^2 - 2\alpha| \le 1$ on $[0,1]$,

$$-e \le \widehat{\Delta}_{\mathrm{S}}(\alpha) - \Delta(\alpha) - \varepsilon_{\mathrm{num}}\alpha^2 \le e.$$

Optimality and $\widehat{\Delta}_{\mathrm{S}}(0) = 0$ imply $\Delta(\widehat{\alpha}) \le \widehat{\Delta}_{\mathrm{S}}(\widehat{\alpha}) + e \le e$. Comparing instead to $\alpha_{\mathrm{ray}}$ gives

$$\Delta(\widehat{\alpha}) \le \widehat{\Delta}_{\mathrm{S}}(\alpha_{\mathrm{ray}}) + e \le \Delta(\alpha_{\mathrm{ray}}) + 2e + \varepsilon_{\mathrm{num}}.$$

Finally, for a fixed dictionary, $d = -C\lambda^\star$ gives $\kappa = \widehat{\lambda}^\top C(\widehat{\lambda} - \lambda^\star)$. Taking expectation yields $\mathbb{E}_D \kappa = P + B$. □

If the estimation-error event has probability at least $1 - \delta$, the same conclusions hold with that probability. Establishing such an event requires control of the sampling law, fitted field, and residual estimator. The algebraic bound does not supply this control automatically.

### D.5 What the structured model assumes

**Bias sensitivity of the deterministic oracle.** Write $\alpha_0 = H/(H+P)$ for $H+P > 0$. For nonzero $B$, let $\alpha_B$ be the constrained oracle in Eq. (6). Then

$$0 \le \mathcal{R}(\alpha_0 \widehat{\lambda}) - \mathcal{R}(\alpha_B \widehat{\lambda}) \le |B|/2. \tag{D.2}$$

Indeed, the omitted term $-2B\alpha(1-\alpha)$ has range width $|B|/2$, and $\alpha_0$ minimizes the remaining quadratic. When the bias-aware oracle is interior and $D_0 = H + P + 2B > 0$, the exact regret is

$$\frac{B^2(P-H)^2}{(H+P)^2 D_0}, \qquad \frac{H+B}{D_0} - \frac{H}{H+P} = \frac{B(P-H)}{(H+P)D_0}.$$

These statements concern population constants and deterministic amplitudes. They do not transfer unchanged to an estimated $\widehat{\alpha}$. In particular, $\mathbb{E}\|\widehat{\lambda}\|_C^2 = H+P+2B$, so subtracting a penalty estimate does not universally give an unbiased estimate of $H$.

**An exact normal-means interpretation.** In whitened coordinates, suppose

$$Z \mid \theta \sim \mathcal{N}(\theta, \sigma^2 I_r), \qquad \theta \sim \mathcal{N}(0, \tau^2 I_r).$$

Completing the square gives $\mathbb{E}[\theta \mid Z] = \tau^2(\tau^2+\sigma^2)^{-1} Z$. This posterior mean minimizes posterior squared loss. With $H_{\mathrm{prior}} = r\tau^2$ and $P = r\sigma^2$, its amplitude is exactly $H_{\mathrm{prior}}/(H_{\mathrm{prior}} + P)$. For known $\sigma^2$, maximizing the marginal normal likelihood gives $\widehat{\tau}^2 = [\|Z\|^2/r - \sigma^2]_+$, whose plug-in amplitude is $[1 - r\sigma^2/\|Z\|^2]_+$ (zero when $Z = 0$). This is the RCPT algebra with $S = \|Z\|^2$, known penalty $P$, and $\varepsilon_{\mathrm{num}} = 0$. It is a classical empirical-Bayes shrinkage interpretation (Efron & Morris, 1973), not a claim that the fitted diffusion controllers follow this model. Estimating the hyperparameters, geometry, and penalty does not inherit exact posterior optimality.

**Why Direct LOO need not improve the decision.** For an interior realized oracle, $\Delta(\alpha) - \Delta(\alpha_{\mathrm{ray}}) = q(\alpha - \alpha_{\mathrm{ray}})^2$. A lower-variance estimate of the amplitude can therefore be preferable even if its criterion is biased. At a boundary, the exact regret also includes $2(q\alpha_{\mathrm{ray}} - (q - \kappa))(\alpha - \alpha_{\mathrm{ray}}) \ge 0$. The criterion-estimation tradeoff is consistent with established comparisons of covariance penalties and cross-validation (Efron, 2004). It is not a universal dominance result for structured RCPT. The common-fit audit in Appendix I.2 illustrates the distinction: smaller unweighted amplitude error need not give smaller curvature-weighted ray regret. Accurate quadratic coefficients are sufficient, but not necessary, for a good decision; multiplying the entire criterion by a positive constant leaves its minimizer unchanged. Likewise, a zero-mean criterion correction at each fixed $\alpha$ need not have zero mean after minimizing over $\alpha$.

### D.6 Finite-particle penalty and its proof

**Lemma D.1** (Finite-particle estimation penalty)**.** *Let $X_1, \dots, X_N \overset{\text{iid}}{\sim} p$ with a fixed dictionary, $A = a - \mathbb{E}_p a$, $\epsilon = g^0 - \mathbb{E}_p g^0 + A^\top \lambda^\star$, and $\Omega = \mathbb{E}_p[\epsilon^2 A A^\top]$. Suppose $C = \mathbb{E}_p A A^\top$ has fixed rank, its nonzero eigenvalues are positive, $\mathbb{E}\|A\|^4 < \infty$, $\mathbb{E}\epsilon^2 < \infty$, $\mathbb{E}[\epsilon^2 \|A\|^2] < \infty$, and empirical rank is eventually equal to population rank in probability. Use minimum-norm least squares, or ridge $\eta_N = o(N^{-1/2})$ in the normalized objective. On* $\mathrm{range}(C)$*,*

$$\sqrt{N}(\widehat{\lambda} - \lambda^\star) \xrightarrow{d} \mathcal{N}(0, C^\dagger \Omega C^\dagger). \tag{D.3}$$

*If $\{N\|\widehat{\lambda} - \lambda^\star\|_C^2\}$ is also uniformly integrable, then $P = N^{-1} \operatorname{tr}(\Omega C^\dagger) + o(N^{-1})$.*

*Proof.* The population normal equation is $\mathbb{E}A\epsilon = 0$. The centered empirical normal equations and consistency of the covariance inverse on its range give

$$\widehat{\lambda} - \lambda^\star = -C^\dagger \frac{1}{N} \sum_n A_n \epsilon_n + o_p(N^{-1/2}). \tag{D.4}$$

Sample centering contributes $O_p(N^{-1})$; the stipulated ridge contributes $o_p(N^{-1/2})$. Apply the multivariate CLT. The quadratic continuous-mapping theorem and uniform integrability then yield $NP \to \operatorname{tr}(CC^\dagger \Omega C^\dagger) = \operatorname{tr}(\Omega C^\dagger)$. □

The reduction to $\sigma_\epsilon^2 \operatorname{rank}(C)/N$ requires $\Omega = \sigma_\epsilon^2 C$. RCPT further substitutes a deleted prediction residual for $\sigma_\epsilon^2$, effective rank for rank, and weight ESS for sample count. These choices define its low-particle model; the asymptotic lemma does not assert that they are accurate near interpolation.

**Ridge and dependent particles.** Fixed population ridge targets $\lambda_\eta = -(C + \eta I)^{-1} d$ and incurs

$$V(\lambda_\eta) - V(\lambda^\star) = \eta^2 (\lambda^\star)^\top (C + \eta I)^{-1} C (C + \eta I)^{-1} \lambda^\star. \tag{D.5}$$

Thus the preceding centered CLT is not a fixed-ridge theorem. For independent observations with deterministic normalized weights, analogous weighted normal equations give leading variance $(\sum_n \omega_n^2) \operatorname{tr}(\Omega C^\dagger)$ under weighted consistency, a weighted CLT, and uniform integrability. Random importance weights and interacting particles require their own covariance analysis. If a particle-system CLT for the regression score has covariance $\Gamma_t$, the same linearization, covariance consistency, and uniform integrability give $N^{-1} \operatorname{tr}(\Gamma_t C^\dagger)$ instead. Conventional weight ESS does not estimate this covariance in general; equal weights after resampling do not imply independent observations.

**Frozen-schedule consequence.** For a schedule fitted independently of deployment, combine the projection identity with the dynamic Wasserstein bound. If $\widetilde{\lambda}_k$ is held on $[t_k, t_{k+1})$ and the assumptions of Theorem 4 hold, then

$$W_2(p_1, \rho_1^{\widetilde{\lambda}}) \le \sum_k \int_{t_k}^{t_{k+1}} K_v(t,1) \sqrt{C_P(t) \{ V_{t,\star} + \|\widetilde{\lambda}_k - \lambda_t^\star\|_{C_t}^2 \}} \, \mathrm{d}t. \tag{D.6}$$

This retains the time profile of the estimation error. Theorem 3 controls the selected coefficient at each grid point under its calibration assumptions. Interpolation, model error, and numerical integration require the separate controls above. A smaller upper bound alone does not order endpoint errors. A guarantee for the same-particle online loop additionally requires analysis of that interacting process.

# E Calibration Ablation and ECG Baseline

## E.1 Direct LOO

Using the deleted means and fits in the main text, set $u_n = g_n^0 - \bar{g}_{-n}^0$ and $z_n = (a_n - \bar{a}_{-n})^\top \widehat{\lambda}^{(-n)}$. Define $A_D = \sum_n \omega_n u_n^2$, $b_D = \sum_n \omega_n u_n z_n$, and $c_D = \sum_n \omega_n z_n^2$. Then

$$\widehat{L}_D(\alpha) = A_D + 2 b_D \alpha + c_D \alpha^2, \qquad \widehat{\alpha}_D = \begin{cases} \Pi_{[0,1]}(-b_D/c_D), & c_D > 0, \\ 0, & c_D = 0. \end{cases} \tag{E.1}$$

The formula follows by minimizing a scalar quadratic. The intercept must be $\bar{g}_{-n}^0 + \alpha \bar{a}_{-n}^\top \widehat{\lambda}^{(-n)}$ at each amplitude; holding the unscaled intercept fixed changes the objective. Both selectors deploy the full-data coefficient, not an average of deleted coefficients. Deletion weights are renormalized and the same ridge normalization is used in every fold.

Direct LOO estimates deleted prediction loss without a scalar covariance model. It changes only the calibration criterion, while retaining the full-data VCG direction. Appendix H.4 distinguishes target-law risk from the reference-law diagnostics and numerical guards used in deployment.

### E.2 Residual fitting versus energy fitting

VCG minimizes $\mathrm{Var}_p(g^0 + \mathcal{T}_p u)$ in a vector-field class. For gradient controls $u = \nabla\psi$, energy-controlling guidance (ECG) instead uses the weak Poisson objective (Ren et al., 2026)

$$\min_{\psi} \ \tfrac{1}{2}\mathbb{E}_p\|\nabla\psi\|^2 - \mathbb{E}_p[r^0\psi]. \tag{E.2}$$

Its first variation is $\mathbb{E}_p\nabla\psi\cdot\nabla h = \mathbb{E}_p r^0 h$, equivalent under integration by parts to $-\nabla\cdot(p\nabla\psi) = pr^0$. For $\psi = \sum_j \theta_j\psi_j$, ECG solves $M\theta = b$, with $M_{ij} = \mathbb{E}_p\nabla\psi_i\cdot\nabla\psi_j$ and $b_i = \mathrm{Cov}_p(\psi_i, g^0)$. The two finite-dimensional objectives need not select the same field.

ECG provides a distinct proposal-fitting baseline. Its scalar-potential class does not automatically contain arbitrary gated vector fields used by VCG and RCPT. Conversely, ECG's energy objective avoids squared Stein features in its normal matrix. These are representation and objective differences, not a general superiority result. ECG also estimates its moments from finite particles. Comparisons should state the available field class, fitting budget, and whether correction weights and resampling are retained; no resampling alone does not mean unweighted Proposal-Only sampling.

## F Scalar and Directional Calibration

### F.1 The oracle benefit of directional shrinkage

Scalar shrinkage is a restricted decision class, not an isotropy assumption. On the range of $C$, write $\theta = C^{1/2}\lambda^\star$ and $Z = C^{1/2}\widehat{\lambda} = \theta + \xi$. Suppose $\mathbb{E}\xi = 0$ and define $H_j = \theta_j^2$, $P_j = \mathbb{E}\xi_j^2$. For fixed diagonal amplitudes,

$$\mathbb{E}\|\operatorname{diag}(\alpha_j)Z - \theta\|^2 = \sum_j\{(1-\alpha_j)^2 H_j + \alpha_j^2 P_j\}. \tag{F.1}$$

No independence between the coordinates is needed. Cross-covariances do not enter a sum of coordinatewise squares.

**Proposition F.1** (Exact directional oracle gain)**.** *Write $T_j = H_j + P_j$. If all $T_j = 0$, every amplitude has the same risk. Otherwise omit coordinates with $T_j = 0$. The directional oracle is $\alpha_j^* = H_j/T_j$, the scalar oracle is $\alpha_s^* = \sum_j H_j / \sum_j T_j$, and their risk difference is*

$$\mathcal{R}^*_{\mathrm{scalar}} - \mathcal{R}^*_{\mathrm{directional}} = \sum_j T_j(\alpha_j^* - \alpha_s^*)^2 \geq 0. \tag{F.2}$$

*Proof.* Complete each square in Eq. (F.1). The unconstrained minimizers lie in $[0,1]$. Under a common amplitude, minimizing the sum gives $\alpha_s^*$. Subtract the coordinatewise minima to obtain the identity. □

For example, let $\theta = (1,0)$ and $\xi = (0,\varepsilon)$ with $\mathbb{E}\varepsilon = 0$, $\mathbb{E}\varepsilon^2 = L > 0$. The directional oracle keeps the first coordinate and removes the second, giving zero coefficient risk. The scalar oracle has risk $L/(1+L)$. Grouping coordinates replaces $H_j, P_j$ by their group sums. A zero-amplitude group is a controller truncation. Thus reduced rank and shrinkage are complementary decisions. These are oracle comparisons; estimating additional amplitudes has a cost.

### F.2 An exact sample on which directional LOO is worse

Extra amplitudes can fit deletion noise even when they use the same fits. Let $a$ be uniform on $\{e_1, -e_1, e_2, -e_2\}$, let $g$ be an independent uniform sign, and use $\eta = 1$, $N = 4$. The ordered sample

$$D = ((e_1,-1),(e_1,-1),(e_2,1),(-e_2,1))$$

has probability $8^{-4}$. Direct calculation gives $\widehat{\lambda}_D = (2/5, 0)$ and the deleted residual representation $u_i + v_i^\top\alpha$ with

$$u = (-4/3, -4/3, 4/3, 4/3)^\top, \qquad (v_i^\top)_{i=1}^4 = \begin{pmatrix} 8/33 & 0 \\ 8/33 & 0 \\ -8/39 & 16/39 \\ -8/39 & 16/39 \end{pmatrix}. \tag{F.3}$$

The first two deletion coefficients are $(4/11, 0)$; the last two are $(4/13, 4/13)$ and $(4/13, -4/13)$, which verifies the matrix using the deleted means. The scalar criterion has minimizer $143/290$. For the two-coordinate criterion, at $(1, 0)$ both residual types are nonzero, the derivative in coordinate one is negative, and the derivative in coordinate two is positive. These are the constrained optimality conditions at that corner. The two distinct row vectors span the plane, so strict convexity makes $(1, 0)$ the unique minimizer.

Population risk is $V(\lambda) = 1 + \|\lambda\|^2/2$. Deploying the full fit with these amplitudes therefore gives

$$V_{\text{scalar}} = 1 + \frac{2}{25}\left(\frac{143}{290}\right)^2 \simeq 1.01945, \qquad V_{\text{directional}} = 1 + \frac{2}{25} = 1.08.$$

This is a positive-probability counterexample to samplewise dominance of directional LOO. It is not a comparison of their risks averaged over all possible fitting samples. Together with Proposition F.1, it separates the benefit of a richer oracle from the reliability of estimating it.

### F.3 When the bias term can be omitted

Projection orthogonality alone does not imply $B = 0$. One sufficient condition is a correctly specified conditional linear model $g = c - a^\top \lambda^\star + \varepsilon$, with $\mathbb{E}[\varepsilon \mid a] = 0$, full empirical rank, unregularized least squares, and integrable coefficients. Conditioning on the design gives $\mathbb{E}[\widehat{\lambda} \mid a_{1:N}] = \lambda^\star$, so $B = 0$. This is stronger than the population normal equation.

A centered first-order CLT is also insufficient by itself. A coefficient bias of order $N^{-1}$ can make $B$ the same order as the estimation penalty $P$. The general bound $|B| \le \sqrt{H}\|\mathbb{E}\widehat{\lambda} - \lambda^\star\|_C \le \sqrt{HP}$ follows from Cauchy–Schwarz, but need not be small relative to $P$. In a noiseless scalar example with ridge contraction $\widehat{\lambda} = c\lambda^\star$, $0 < c < 1$, one has $P = (1-c)^2 H$ and $B = -(1-c)H$. The constrained risk oracle retains the entire fitted update, whereas $H/(H+P) < 1$ shrinks it again. This proves that ridge can invalidate the zero-bias interpretation even without sampling noise. The paper's bias audit is empirical evidence for its audited regime, not a universal justification for dropping $B$.

## G Specialization to FKC and ACE

Let expert $i$ provide density $q_t^{(i)}$, score $s_t^{(i)} = \nabla \log q_t^{(i)}$, and probability-flow velocity $v_t^{(i)}$. For a valid exponent schedule $\gamma_i(t)$,

$$p_t(x) = \frac{1}{Z_t}\prod_{i=1}^{I} q_t^{(i)}(x)^{\gamma_i(t)}, \qquad s_t(x) = \sum_i \gamma_i(t) s_t^{(i)}(x). \tag{G.1}$$

In the implementation's SDE parameterization, the selected probability-flow component enters

$$g_t^v = \nabla\cdot v_t + v_t \cdot s_t + b_t, \tag{G.2}$$

up to a spatial constant, where $\eta_t$ is the scalar linear-drift coefficient fixed by the underlying diffusion schedule and

$$\begin{aligned} b_t(x) = \eta_t x \cdot s_t(x) &+ \sum_i \dot{\gamma}_i(t) \log q_t^{(i)}(x) \\ &- \sum_i \gamma_i(t)\left\{\nabla\cdot v_t^{(i)}(x) + v_t^{(i)}(x)\cdot s_t^{(i)}(x)\right\}. \end{aligned} \tag{G.3}$$

The omitted divergence $\nabla\cdot(\eta_t x) = d\eta_t$ is spatially constant and irrelevant to normalized weights.

The generic theory requires a valid path. VCG applies directly when the fixed-exponent path is valid. If fixed exponents create an invalid or singular intermediate path, ACE first repairs the schedule while preserving the desired endpoints (Lee et al., 2026). Proposal fitting then acts on the repaired path. RCPT does not substitute for path repair and does not change Eq. (G.1).

# H Implementation Details

## H.1 Stable weighted least squares

For normalized current particle weights $\omega_n$, the code minimizes

$$\min_{\lambda,c} \sum_{n=1}^{N} \omega_n \left(g_n^0 + A_n\lambda - c\right)^2 + \eta\|\lambda\|_2^2. \tag{H.1}$$

The intercept is eliminated by weighted centering. Covariances and the $m \times m$ linear solve are computed in float64; singular systems use a pseudoinverse. A coefficient clip is available in the high-dimensional experiments. After clipping, the code recomputes the selected empirical variance and uses the anchor coefficient $\lambda = 0$ if it increased beyond a floating-point tolerance. The chosen proposal is used consistently for state propagation and all tracked $\log q_i$ updates.

Theorem 3 analyzes the scalar-selected field before an additional coefficient clip. If clipping changes it to $\lambda_{\text{final}}$, let $e_{\text{clip}} = \|\lambda_{\text{final}} - \widehat{\alpha}\widehat{\lambda}\|_C$. The triangle inequality for centered residuals gives

$$\sqrt{V(\lambda_{\text{final}})} \le \sqrt{V(\widehat{\alpha}\widehat{\lambda})} + e_{\text{clip}}. \tag{H.2}$$

An Anchor fallback has exactly the Anchor risk, but an empirical-variance guard alone is not an out-of-fit guarantee. The realized deployed field and any clipping therefore need to be distinguished in applying the theory. The logged amplitude is the scalar selected before componentwise clipping; if an empirical guard falls back to Anchor, it remains a diagnostic of the pre-guard decision rather than the final coefficient.

The logged diagnostics are the anchor variance, selected variance, regularized within-class floor, removable fraction

$$R_t^2 = \frac{\widehat{\mathrm{Var}}(g_t^0) - \widehat{V}_{t,*}}{\widehat{\mathrm{Var}}(g_t^0)}, \tag{H.3}$$

coefficient norm, conditional ESS before and after selection, conventional ESS, and resampling events. Same-particle floors are optimization diagnostics, not held-out estimates of the population floor.

## H.2 Basis and divergence by domain

**Two dimensions.** The controlled ACE experiment anchors at its standard velocity and uses every coordinate vector field whose scalar coefficient is a monomial of total degree at most two. Exact automatic differentiation computes divergence in two dimensions.

**Images.** We use the official pretrained DLSM CIFAR-10 score model and noise-conditioned classifier at upstream commit `259c767`. For class $c$, the target path is $q_t(x)p_\phi(c \mid x, t)$ and the Anchor is its standard classifier-guided VE probability-flow velocity. Writing $h_c = \log p_\phi(c \mid x, t)$ and $\mathsf{a}(t)$ for the diffusion variance, the Anchor response in generation time $s = 1 - t$ is

$$g_A = -\partial_t h_c + \tfrac{1}{2}\mathsf{a}(t)\left(\Delta h_c + \|\nabla h_c\|_2^2\right). \tag{H.4}$$

The ten controller fields are $\phi_j = \frac{1}{2}\mathsf{a}(t)\nabla \log p_\phi(j \mid x, t)$, one for each CIFAR-10 class, with affine residual feature $\mathcal{T}_{p_t}\phi_j = \nabla\cdot\phi_j + \phi_j^\top s_t$. One realized Rademacher Hutchinson probe estimates all trace terms and is shared by Anchor, VCG, and RCPT in each class–seed–time cell. Fit states are forward-noised training images; held-out states are independently selected and noised test images. Ridge is $10^{-4}$. Schedules are fitted at 16 midpoint times and linearly interpolated during generation. In the matched predictor–corrector audit, every method uses 1,000 reverse-VE predictor steps, one Langevin corrector step per time level, and SNR 0.16. The corrector stays at the cohort's common Anchor scale for all methods; the predictor applies half of the probability-flow schedule coefficient so that its reverse-SDE drift perturbation matches the fitted physical field. Neither terminal images nor evaluation metrics enter fitting or RCPT calibration. Forward-noised labeled data are an empirical proxy for the model-defined conditional path; their risk is not a direct numerical evaluation of the exact $p_t$ expectation in the theory.

**Molecules.** States are flattened padded molecular coordinates and attributes, with all existing projections and masks retained. In scaffold decoration, each of the four expert probability-flow velocities is crossed with degree zero, one, and two bounded relative-density Legendre gates, giving the fixed 12-field basis in Eq. (I.2). Sixteen particles fit the VCG or RCPT schedule, and that schedule is frozen for independent deployment. Finite controller coefficients, feature matrices, state increments, and residuals are checked before every accepted update; these VCG-based arms share the stated numerical guards.

**Reward-tilted molecules.** The EEGSDE basis (Bao et al., 2023) begins with $\nu_t \nabla R_{\mathrm{polarizability,t}}$ and $\nu_t \nabla R_{\mathrm{gap,t}}$, where $\nu_t$ is the one-step reverse-transition variance, and crosses these families with the fixed bounded Legendre gates. Each resulting field is normalized by its fit-batch RMS over active state coordinates before regression; the frozen schedule stores these scales. The base conditional diffusion drift and scaled FKC Anchor are fixed parts of the proposal and are not counted as controller fields. Autodifferentiation supplies reward gradients and Stein divergences. A same-state time finite difference evaluates both $\dot{\beta}_{j,t} R_{j,t}$ and $\beta_{j,t} \partial_t R_{j,t}$ in the reward-path potential. Generation uses the training reward networks; all reported properties use independently trained evaluators, and molecular identities use RDKit canonical isomeric SMILES.

### H.3 Deletion scope and computational complexity

Deletion refits in this description act on the evaluated regression features. When gate statistics or field normalizations are estimated from the same cohort and then held fixed across folds, this is feature-level deletion rather than deletion of the entire field-construction pipeline. The fixed-dictionary statistical interpretation requires that construction to be fixed or conditioned on with a justified sampling law. The realized-field bound in Theorem 3 still applies, but its calibration errors must include any resulting adaptivity. No iid claim is made solely because the feature matrix has been frozen.

After field values and divergences are available, weighted covariance formation and the VCG solve cost $O(Nm^2 + m^3)$. RCPT reuses those features and adds only dense leave-one-out algebra, with no expert or denoiser call. Hardware-identified timing, peak VRAM, and measured particle scaling are reported in Appendix K. CIFAR timing is reported separately from those controlled audits.

**Algorithm H.1** Online VCG/RCPT Sampling Loop

**Input:** valid path, Anchor, basis, ridge, correction flag, particles $X_0^{[n]} \sim p_0$. Initialize log weights and tracked expert log densities. At each step:

1. Evaluate target score, Anchor, fields, and divergences; form the rate and feature matrix.
2. Normalize weights and solve Eq. (H.1). For RCPT, compute deletion residuals and the structured amplitude.
3. Apply the configured coefficient clip and empirical-variance fallback.
4. Propagate under the selected drift and update tracked densities using that same drift.
5. If correction is retained, update FK weights. If the resampling rule triggers, resample states and tracked quantities with common ancestry and reset weights.

Return particles, weights (if retained), and diagnostics.

### H.4 Evaluation laws and dependent particles

The theoretical risk is defined under the target marginal $p_t$. Common-state evaluation compares methods on identical states but estimates risk under the reference law; active-path evaluation also changes the state distribution with the proposal. Neither is automatically a target-law expectation. The Checker corrected-SMC reference is an approximation, and the molecular common-state audit does not quantify its gap to the target law.

Equal weights after resampling do not imply independent fitting information. Particles can share ancestors, and deleting one descendant can leave closely related states in the fitting ensemble. Weight ESS therefore does not by itself quantify regression information or justify an online guarantee.

Lineage-aware deletion and particle-system concentration require a separate analysis (Del Moral & Rio, 2011).

Data-derived features, coefficient clipping, and empirical fallbacks are part of the deployed estimator. The conditional structured bound evaluates a fixed fitted direction; additional transformations need their own error control, and an empirical Anchor fallback is not an independent out-of-fit certificate. These distinctions delimit the interpretation of the reported risk audits.

# I Additional Evaluations

## I.1 Controlled targets

**Checker.** Draw $U \sim \mathrm{Unif}[-2,2]$, $V \sim \mathrm{Unif}[0,1]$, and $K \sim \mathrm{Bernoulli}(1/2)$, set

$$X_{\mathrm{raw}} = 2U, \quad Y = 2\{V - 2K + (\lfloor U \rfloor \bmod 2)\}, \quad X = (X_{\mathrm{raw}} + 2 \bmod 8) - 4, \tag{I.1}$$

and condition on $A = \mathbf{1}\{X \geq 0\} = 1$ and $B = \mathbf{1}\{X + Y \geq 0\} = 1$. This gives the exact alternating checker law on $[-4,4]^2$. Three separately trained expert paths define the ACE target. Anchor is only the learned joint expert velocity $v_{XY|B}$; it is not the sum of all three expert velocities. The other paths enter the time-dependent multiplicative target, while VCG and RCPT add a 12-field coordinate-polynomial correction (six monomials of total degree at most two for each of two velocity coordinates) to the shared Anchor. At every fitted time, Anchor, VCG, and RCPT share the same target, expert evaluations, integration grid, ridge convention, and random streams. A schedule fitted on 16 particles is frozen before 2,000 fresh Proposal-Only or SMC trajectories are generated in the primary comparison (Table 2). All runs use 200 integration steps, ridge $10^{-4}$, SMC ESS threshold 0.7, and the expert schedule `default_linear / 1-t**2 / cos_t`. The calibration ablation uses these same seeds and deployments. Its residual diagnostic is an active-path integral, not an independent target-law risk estimate.

**Illustrative replay and calibration ablation.** The main Checker figure uses one seed, 16 fitting particles, 200 integration steps, and the exact common-endpoint schedule `default_linear / 1-t**2 / cos_t`. Each frozen schedule generates 10,000 particles. Endpoint estimates use fixed 2,000-particle audit indices; Proposal-Only has no FK weights, resampling, or terminal reweighting. The accompanying SMC replay retains correction.

Table I.1: **Single-seed Checker figure audit.** This replay is illustrative, not an additional ten-seed estimate. Residual is an active-path integral and is not pooled across deployment modes.

| Mode | Controller | $W_2$ | MMD | Residual |
|---|---|---|---|---|
| Proposal | Anchor | 0.7689 | 0.0645 | 668.21 |
| Proposal | VCG | 1.1933 | 0.2572 | 3087.30 |
| Proposal | ECG | 1.3156 | 0.4725 | 291.44 |
| Proposal | RCPT | 0.6819 | 0.0670 | 30.03 |
| SMC | Anchor | 1.3888 | 0.1387 | 48.64 |
| SMC | VCG | 5.7207 | 6.9463 | 5673.08 |
| SMC | ECG | 1.7622 | 0.4858 | 152.01 |
| SMC | RCPT | 0.9445 | 0.1235 | 40.87 |

In this calibration ablation, RCPT has lower mean $W_1$ and $W_2$ than Direct LOO in both deployment modes and lower mean MMD under SMC. Direct LOO has lower Proposal-Only MMD and retains more SMC ancestors on average (630.7 versus 568.5). These distinct rankings separate endpoint accuracy from genealogy.

## I.2 Common-fit realized-ray and adaptive-regularization audit

**Protocol and evaluation law.** This audit isolates the calibration decision on identical fitting states. Thirty raw-VCG pilot fits use the 12-field Checker dictionary, fitting populations $N \in \{8, 16, 32\}$, and ten paired seeds per budget. At each of 200 native steps, RCPT and Direct LOO use the same

Table I.2: **Checker calibration ablation.** The same ten paired seeds, 16 fitting particles, fixed schedule, and 2,000-particle deployments as Table 2. Values are means. Direct LOO replaces the structured amplitude criterion. Residual is the integrated active-path diagnostic, not independent target-law risk; its Direct LOO Proposal-Only mean is driven by one severe run. Ancestors counts terminal SMC lineages out of 2,000; without resampling every trajectory retains its own ancestor.

| Mode | Controller | $W_1 \downarrow$ | $W_2 \downarrow$ | MMD$\downarrow$ | Residual | Ancestors |
|---|---|---|---|---|---|---|
| Proposal | Anchor | 0.474 | 0.646 | 0.048 | 614.37 | 2000.0 |
| | Direct LOO | 0.465 | 0.636 | 0.068 | 1777.69 | 2000.0 |
| | RCPT | 0.442 | 0.575 | 0.073 | 103.24 | 2000.0 |
| SMC | Anchor | 0.367 | 0.510 | 0.048 | 31.56 | 338.7 |
| | Direct LOO | 0.435 | 0.631 | 0.059 | 25.62 | 630.7 |
| | RCPT | 0.316 | 0.464 | 0.042 | 25.09 | 568.5 |

fitted direction. LOO-selected ridge and PCR use the same fitting states; PCR recomputes its basis within deletion folds. Selection uses fitting residuals only. No endpoint metric selects a ridge level, rank, or amplitude. Coefficient clipping and calibration fallback are disabled in this audit, so scalar rules remain on the fitted ray. This is a separate cohort from the primary fixed-schedule comparison and its matched calibration ablation.

Independent reference states separate calibration from evaluation. Intermediate normalized target densities are unavailable for the learned expert paths. We therefore use two independent corrected Anchor-SMC systems of 16,384 particles each, with separately normalized weights. Both systems resample five times; minimum pre-resampling ESS is 9,102–9,229. Their combined 32,768-particle weighted law $\mu_t$ defines the reference risk. This is neither exact target integration nor iid target sampling. Disagreement between systems measures simulation sensitivity; it does not bound shared model bias or the target-reference covariance gap in Appendix H.4.

The audit uses 21 fixed rounded indices from $\mathrm{linspace}(0, 199, 21)$. Each native step is assigned to its nearest audit node; the resulting cell counts divided by 200 give time weights $w_k$. Reported Risk/Anchor is the ratio of $\sum_k w_k \mathrm{Var}_{\mu_{t_k}}(g^0_{t_k} + a^\top_{t_k}\lambda_{t_k})$ to the corresponding Anchor risk, summarized over the ten fitting seeds. Bootstrap intervals resample paired fitting seeds, not time nodes. They describe variation over fits conditional on this reference construction.

Table 1 in the main text reports this comparison.

**Adaptive regularization.** RCPT has the lowest mean reference risk among the implemented rules at all three fitting populations (Table 1). At $N = 16$, its paired Risk/Anchor differences from Direct LOO, ridge, and PCR are $-0.6331$, $-0.3375$, and $-0.6804$, with respective 95% bootstrap intervals $[-1.363, -0.0669]$, $[-0.7484, -0.0510]$, and $[-1.182, -0.2518]$. At $N = 8$ and 32, the intervals against Direct LOO and ridge include zero; the intervals against PCR remain below zero. These are pairwise intervals over ten fits, not simultaneous guarantees. The results support scalar calibration at low particle budgets without establishing dominance over all directional estimators. PCR changes a learned basis and rank; it does not exhaust grouped or coordinatewise shrinkage.

**Realized calibration.** For each fit and node, replace $p$ by $\mu_t$ in Eq. (D.1) to obtain $q_\mu$, $\kappa_\mu$, and the reference-ray oracle. Direct variance evaluation agrees with the quadratic at $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$ to maximum absolute error $2.328 \times 10^{-10}$. None of the 630 fit–time cells requires the near-zero-curvature convention, whose tolerance is $\max\{10^{-14}, 10^{-10}\max(V_\mu(0), 1)\}$.

The structured ingredients are not quantitatively calibrated pointwise. Median $S/q_\mu$ rises from $0.105$ to $0.276$ and $0.606$ as $N$ grows; $\kappa_\mu < 0$ in 56/630 cells (Table I.3). A nonnegative penalty cannot match those negative covariances exactly. Nevertheless, mean integrated ray regret is $8.359$, $4.500$, and $3.776$ for RCPT, versus $12.900$, $27.960$, and $4.635$ for Direct LOO. Direct LOO has smaller mean absolute amplitude error at $N = 8$ and 32, but not smaller regret. Curvature weights amplitude errors differently across times, and boundary optima add a one-sided term. This distinction explains why unweighted amplitude accuracy and decision quality need not have the same ordering.

Table I.3: **Realized quadratic diagnostics.** All quantities use the independent corrected-SMC reference law. Spearman coefficients summarize the seed-level $S$–$q_\mu$ and $\widehat{P}$–$\kappa_\mu$ diagnostics. Their limited agreement precludes interpreting the structured ingredients as accurate pointwise estimates.

| $N$ | Median $S/q_\mu$ | $\kappa_\mu < 0$ | $\rho(S, q_\mu)$ | $\rho(\widehat{P}, \kappa_\mu)$ |
|---|---|---|---|---|
| 8 | 0.105 | 4.8% | 0.418 | 0.406 |
| 16 | 0.276 | 6.2% | -0.188 | 0.103 |
| 32 | 0.606 | 15.7% | 0.115 | 0.673 |

**What the oracle calculation establishes.** The expected-risk oracle and the realized calibration bound answer different questions. Applying the $H, P, B$ identity to the empirical distribution of ten fits gives an oracle gain over the better endpoint larger than the two-system sensitivity in 22/63 population-size–time cells: 6/21, 6/21, and 10/21. This diagnoses Theorem 2's expected-risk calculation on the reference law. It is not the frequency with which Theorem 3's plug-in error bound is informative. The latter additionally requires control of realized approximation errors and target-reference transfer. No finite-sample certificate for the implemented selector follows from this audit.

**Endpoint deployment and completeness.** Each frozen schedule generates 10,000 paired trajectories in Proposal-Only and corrected-SMC modes. Exact discrete optimal transport estimates $W_2$ on a predeclared paired 2,000-sample metric subset. Table I.4 retains all attempted denominators. RCPT completes 60/60 deployments, compared with 45/60 for VCG and 59/60 for PCR. Means for incomplete arms condition on completion and are not unconditional performance comparisons.

Table I.4: **Common-fit Checker endpoints.** Mean $W_2 \pm$ standard error across completed seeds. Finite gives the completion count over every planned seed; lineages are SMC terminal ancestors out of 10,000, averaged over completed runs. The reference-risk analysis includes all 30 fits per method, including those whose endpoint deployment failed.

| | | Proposal-Only | | Corrected SMC | | |
|---|---|---|---|---|---|---|
| $N$ | Controller | $W_2 \downarrow$ | Finite | $W_2 \downarrow$ | Finite | Lineages |
| 8 | Anchor | $0.7725 \pm 0.0136$ | 10/10 | $0.5773 \pm 0.0290$ | 10/10 | 1723 |
| | VCG | $1.594 \pm 0.1779$ | 5/10 | $2.316 \pm 0.3144$ | 8/10 | 613.9 |
| | Direct LOO | $0.9814 \pm 0.0422$ | 10/10 | $0.7267 \pm 0.0735$ | 10/10 | 1642 |
| | LOO ridge | $1.082 \pm 0.0467$ | 10/10 | $0.7458 \pm 0.0871$ | 10/10 | 1631 |
| | LOO PCR | $1.073 \pm 0.0755$ | 9/10 | $0.8080 \pm 0.1280$ | 10/10 | 1559 |
| | RCPT | $0.8360 \pm 0.0447$ | 10/10 | $0.6417 \pm 0.0372$ | 10/10 | 1873 |
| 16 | Anchor | $0.7973 \pm 0.0159$ | 10/10 | $0.6602 \pm 0.0337$ | 10/10 | 1893 |
| | VCG | $1.423 \pm 0.5909$ | 3/10 | $3.744 \pm 0.2502$ | 9/10 | 657.2 |
| | Direct LOO | $0.9174 \pm 0.1168$ | 10/10 | $0.7979 \pm 0.1391$ | 10/10 | 1542 |
| | LOO ridge | $0.9675 \pm 0.1051$ | 10/10 | $0.6963 \pm 0.0639$ | 10/10 | 1832 |
| | LOO PCR | $0.9682 \pm 0.1157$ | 10/10 | $0.7585 \pm 0.0634$ | 10/10 | 1804 |
| | RCPT | $0.9062 \pm 0.1097$ | 10/10 | $0.7341 \pm 0.1011$ | 10/10 | 1592 |
| 32 | Anchor | $0.8044 \pm 0.0188$ | 10/10 | $0.5990 \pm 0.0310$ | 10/10 | 1895 |
| | VCG | $0.7991 \pm 0.0545$ | 10/10 | $1.731 \pm 0.5101$ | 10/10 | 2400 |
| | Direct LOO | $0.7918 \pm 0.0486$ | 10/10 | $0.6553 \pm 0.0218$ | 10/10 | 2861 |
| | LOO ridge | $0.8054 \pm 0.0455$ | 10/10 | $0.6068 \pm 0.0463$ | 10/10 | 3203 |
| | LOO PCR | $0.8023 \pm 0.0390$ | 10/10 | $0.6067 \pm 0.0463$ | 10/10 | 3138 |
| | RCPT | $0.7851 \pm 0.0482$ | 10/10 | $0.6806 \pm 0.0299$ | 10/10 | 2671 |

Residual improvement does not uniformly improve terminal transport. RCPT has lower mean Proposal-Only $W_2$ than Direct LOO, ridge, and PCR at each budget, but SMC favors ridge at $N = 16$ and both ridge and PCR at $N = 32$. Paired RCPT-minus-Anchor $W_2$ intervals include zero in Proposal-Only at every budget and in SMC at $N = 8, 16$. At $N = 32$, the SMC difference is $+0.0817$, with 95% bootstrap interval $[+0.0051, +0.1629]$. Reference risk, propagated transport energy, and corrected finite-particle endpoints are different quantities; the one-sided transport bounds do not order these observed metrics.

**Selection cost.** Mean selector time per 200-step schedule is $0.132$ seconds for RCPT, $1.31$ for Direct LOO, $9.34$ for ridge selection, and $13.03$ for PCR. These are selector timings within this implementation, not total sampling times or hardware-independent complexity comparisons. Together with the risk results, they support scalar shrinkage as an inexpensive decision under limited fitting information. They do not establish its global optimality among regularizers.

### I.3 SCAFFOLD PROTOCOL

The fixed benchmark contains all 76 CrossDocked2020 scaffold–pocket tasks (Francoeur et al., 2020). Its four pretrained expert paths are EDM for the fragment and full ligand, GeoDiff for fragment conformations, and pocket-conditioned DiffSBDD for the full ligand (Hoogeboom et al., 2022; Xu et al., 2022; Schneuing et al., 2024). For each expert velocity $v_i$, the controller contains

$$\begin{aligned} \phi_{i,0} = v_i, \qquad \phi_{i,1} = z_i v_i, \qquad \phi_{i,2} = \tfrac{1}{2}(3z_i^2 - 1)v_i, \\ r_i = \log q_i - I^{-1} \sum_{\ell=1}^{I} \log q_\ell, \\ z_i = \tanh\left( \frac{r_i - \mu_i}{\sigma_i} \right), \end{aligned} \tag{I.2}$$

giving 12 relative-density-gated fields with bounded gates. The degree-zero fields are exactly the original mechanistic expert velocities. The gate statistics are fixed from the fitting cohort and never selected using decoded chemistry.

Each pocket uses 16 fitting particles with one fitting seed. Anchor is the natural ACE proposal; VCG is the ordinary 12-field regression; RCPT applies exact weighted leave-one-out refitting at the same fitted times. The resulting schedules are frozen and deployed on 16 fresh paired trajectories per arm. Proposal-Only deployment retains neither correction weights nor resampling. RDKit decoding, scaffold retention, uniqueness, diversity, QED, SA, Lipinski summaries, and QuickVina docking are computed only after deployment and never enter controller selection. Affinity-improved yield uses the fixed generation denominator; "unique useful" additionally deduplicates valid, scaffold-retaining, reference-beating molecules. Docking success alone counts any candidate with a successful docking result, even if it fails the separate chemical-validity check, divided by all generated slots. Paired pocket bootstrap intervals treat the 76 pockets as statistical units. The main table uses the original complete fit-16 deployment; an independent 64-trajectory replication reuses the same frozen schedules and increases evaluation to 4,864 candidates per arm.

The common-state audit evaluates all 76 fitted schedules on identical shared reference states, rather than their active deployment paths. VCG exceeds Anchor risk in 75/76 pockets. RCPT improves on VCG in all 76 and repairs 72 of those 75 harmful updates; Direct LOO repairs 73. Mean Risk/Anchor is 25.9641 for VCG, 0.8935 for RCPT, and 0.7958 for Direct LOO. Direct LOO beats RCPT in 73/76 pockets; the paired RCPT-minus-Direct difference is 0.0977 (95% bootstrap CI [0.0622,0.1197]). The worst ratios are 1491.65, 2.00, and 3.02, respectively: the mean improvement and tail comparison differ.

Table I.5: **High-precision 76-pocket scaffold deployment.** The fixed fit-16 schedules are independently deployed with 64 candidate slots per arm and pocket (4,864 per arm). This table isolates the VCG-to-RCPT comparison; downstream means remain close to Anchor. Risk is normalized by Anchor. Dock counts successful docking over all generated candidates, irrespective of chemical validity. Bold marks the better of the two reported controllers.

| Controller | Validity↑ | Scaffold↑ | Risk/Anchor↓ | Dock success↑ | Affinity↑ | Unique useful↑ |
|---|---|---|---|---|---|---|
| VCG | 0.216 | 0.212 | 23.822 | 0.569 | 0.148 | 0.101 |
| RCPT | **0.433** | **0.415** | **0.878** | **0.904** | **0.292** | **0.224** |

Table B.3 reports this comparison.

Direct LOO lowers active-path Risk/Anchor by 0.1199 relative to RCPT (95% CI $[-0.1293, -0.1099]$), but its validity difference is $-0.0206$ ($[-0.0485, 0.0074]$) and unique useful yield difference is $-0.0148$ ($[-0.0419, 0.0115]$). All reported downstream paired intervals in-

clude zero. The molecular implementation includes coefficient clipping and an empirical fallback, so these deployed arms include the guards in Appendix H. The bundled three-later-seed Direct summary reports validity 0.4252, scaffold yield 0.4068, Risk/Anchor 0.7576, docking 0.9191, affinity yield 0.2944, and unique useful yield 0.2536. It is a different aggregation from Table B.3; matched three-seed control summaries are unavailable, so it is not used for paired comparisons.

The separate ECG run covers the same 76-pocket manifest with 16 fitting particles and 16 fresh outputs per pocket. The implementation requires the `scores` basis: $\phi_i = \log q_t^{(i)}$ and $\nabla\phi_i = \nabla \log q_t^{(i)}$, one potential per expert. It solves the weighted Ritz system with centered potential–response covariance and the uncentered gradient Gram matrix, then applies the configured coefficient clip. It does not reject a finite ECG coefficient merely because its empirical VCG residual increases. This four-potential class differs from the 12 gated velocity fields; the endpoint comparison therefore assesses the instantiated proposal methods, not two optimizers on an identical field class. Table B.2 reports the separate controller-capacity ablation.

### I.4 QM9 STEERING PROTOCOL

The pretrained EEGSDE model is conditioned on dipole moment and steered toward polarizability and HOMO–LUMO-gap targets. The 12-field class crosses the two reward-gradient families with bounded time/state Legendre gates; all arms use the same model, target, evaluator, resampling rule, and total trajectory budget. Reported properties come from independently trained evaluators, not the guidance networks. Molecular identities are RDKit canonical isomeric SMILES.

The *split* protocol uses 16 particles to fit one schedule and eight independent SMC deployment batches of 16, for 144 trajectories in total. The *online* protocol uses one live population of 144: coefficients are refitted from that current population, and the same particles continue and become outputs. Fitting uses all 144 current particles before the next stochastic transition. The serialized `fit_particles=16` option is inactive in online mode. Hence online is a strong adaptive batch-generation baseline, whereas split measures out-of-fit reuse of a frozen controller. Minimum ESS and surviving lineage counts are divided by batch size in the combined table, but one 144-particle genealogy is not statistically identical to eight independent 16-particle genealogies. Both protocols contain eight paired fit seeds and use leave-one-out refitting (mean trust 0.203 split and 0.945 online).

Table I.6: QM9 steering split means ± sample SD over eight paired fits. Unique success counts distinct valid identities meeting both targets at the unchanged one-MAD threshold per 128 generation slots. Minimum ESS and lineages are reported per 16-particle deployment batch.

| Split method | Eval. MAE | Unique success /128 | Min ESS | Lineages |
|---|---|---|---|---|
| Anchor | $0.9171 \pm 0.0413$ | $20.75 \pm 6.27$ | $4.023 \pm 0.656$ | $5.297 \pm 1.44$ |
| VCG | $0.9281 \pm 0.114$ | $13.88 \pm 5.79$ | $2.624 \pm 1.14$ | $2.578 \pm 1.49$ |
| RCPT | $0.8997 \pm 0.0746$ | $23.00 \pm 8.05$ | $4.207 \pm 1.54$ | $5.641 \pm 2.65$ |

**ECG proposal comparison.** The frozen ECG baseline uses the scalar potentials

$$\phi_{r,\ell}(x,t) = \nu_t \int_0^{e_r(x,t)} -sP_\ell(\tanh s)\,\mathrm{d}s, \qquad r \in \{\text{polarizability}, \text{gap}\}, \quad \ell = 0,\dots,5, \tag{I.3}$$

where $e_r$ is target minus prediction in normalized property units and $\nu_t$ is the one-step reverse-transition variance. Thirty-two-node Gauss–Legendre quadrature evaluates the potentials; analytic gradients supply the 12 gated fields. Potentials and fields share fit-batch RMS normalization. At each fit, $(\widehat{G} + 10^{-8}I)\theta = \widehat{\mathrm{Cov}}_w(g^0, \phi)$, with $\widehat{G}_{ij} = \mathbb{E}_w[\nabla\phi_i^\top \nabla\phi_j]$. Coefficients are clipped to $\pm 2 \times 10^{-4}$ without a residual-risk acceptance rule. Sixteen particles fit a frozen schedule before eight independent 16-particle deployment batches, using eight seeds and 1,000 steps.

ECG's property scores conceal severe collapse in frozen deployment. Every batch retains exactly one ancestor. Mean uniqueness is 0.0761, and mean resampling count is 1,105.5 per seed across eight batches, or 138.19 per batch. The reported deployed residual ratio is approximately $3.58 \times 10^5$. The 0.4049 evaluator success statistic counts repeated outputs; it cannot be substituted for the number of distinct successful molecules. Table 4 therefore reports distinct successful identities and

Table I.7: **ECG completion and diversity diagnostics.** Frozen/deploy values use eight seeds; $\pm$ denotes sample SD. The reduced online DriftLite-style configuration has one completed seed. ESS and lineages are absolute counts per batch. Success is distinct from identity-level unique success.

| ECG configuration | Seeds | MAE↓ | Success↑ | Valid | Unique | Min ESS | Lineages |
|---|---|---|---|---|---|---|---|
| Frozen/deploy, 12 fields | 8 | $0.8500 \pm 0.1858$ | $0.4049 \pm 0.2276$ | 0.9150 | 0.0761 | 1.0 | 1.0 |
| Reduced online | 1 | 1.2414 | 0.0000 | 1.0000 | 0.0347 | 1.0 | 1 |

lineage fractions, while Table I.6 retains the available identity-level counts for Anchor, VCG, and RCPT.

The original 12-field online ECG did not complete because its controller geometry became nonfinite. A reduced implementation closest to the DriftLite online construction completed one seed, shown in Table I.8. Its error exceeds Anchor and it has zero success. This is a configuration-specific feasibility result, not an eight-seed estimate or a general ordering of ECG and VCG objectives.

Table I.8: **Completed online ECG audit.** One paired seed, $N = 144$. The Anchor, VCG, and RCPT values come from the same evaluation run; the reduced ECG result is from the accompanying online audit. The failed 12-field ECG run is not assigned a finite score.

| Controller | Eval. MAE↓ | Success↑ | Uniqueness↑ |
|---|---|---|---|
| Anchor | 1.1307 | 0.0000 | 0.0069 |
| VCG | 0.8422 | 0.3736 | **1.0000** |
| ECG | 1.2414 | 0.0000 | 0.0347 |
| RCPT | **0.8379** | **0.3788** | **1.0000** |

**Online calibration ablation.** Direct LOO replaces the structured amplitude in the same 12-field online pipeline. Its eight-seed cohort has no matched frozen/deploy run. Table B.4 isolates this change from the main proposal comparisons. One Direct LOO seed resamples once and retains 84 of 144 ancestors; the other seven retain all 144. Direct-minus-RCPT MAE is 0.0025 (95% paired bootstrap CI $[-0.0017, 0.0074]$). The endpoint difference is unresolved at this precision. Runtime is not compared across separately executed cohorts.

Table B.4 reports the online calibration ablation.

**Regularized Direct ablation.** Write the exact deleted-loss quadratic as $L_D(\alpha) = A + 2b\alpha + c\alpha^2$ and the structured amplitude as $\alpha_S$. The exploratory selector minimizes

$$\alpha_R = \arg\min_{\alpha \in [0,1]} \{L_D(\alpha) + \gamma c(\alpha - \alpha_S)^2\}. \tag{I.4}$$

For $c > 0$, its unconstrained solution is $(-b/c + \gamma\alpha_S)/(1 + \gamma)$; clipping to $[0, 1]$ is performed once. The run uses $\gamma = 1$ and disables componentwise coefficient clipping. Numerically vanishing curvature selects Anchor.

Table I.9: **Exploratory regularized Direct LOO.** One fresh online seed, $N = 144$. This tests an additional regularized selector, not the main RCPT method; its small error advantage is not replicated evidence.

| Controller | MAE | Success | Unique /144 | Resamples | Ancestors |
|---|---|---|---|---|---|
| Anchor | 1.3244 | 0.0000 | 0 | 70 | 1 |
| VCG | 0.8132 | 0.4541 | 45 | 0 | 144 |
| Direct LOO | 0.8080 | 0.4441 | 46 | 0 | 144 |
| Reg. Direct ($\gamma = 1$) | 0.8044 | 0.4538 | 46 | 0 | 144 |
| RCPT | 0.8099 | 0.4594 | 46 | 0 | 144 |

### I.5 CIFAR-10 CLASS-CONDITIONAL IMAGE GENERATION

**Mechanism and physical deployment.** The official pretrained DLSM score model and noise-conditioned classifier are fixed. All ten classes and five seeds enter the mechanism study. Nested fitting cohorts contain 8, 16, or 32 forward-noised training images at 16 midpoint times; independent test cohorts contain 128 states per class and time. Models, states, and realized Hutchinson probes are paired across methods. The PC-1000 implementation and physical predictor scaling are specified in Appendix H. This independent-state audit is separate from the single-seed terminal experiment below.

Table I.10: **CIFAR-10 mechanism study.** Fifty class–seed cells per fit budget. Risk is held-out Risk/Anchor; harm means risk above Anchor. RCPT improves on VCG in 48/50, 47/50, and 42/50 cells at fits 8, 16, and 32.

| Fit | Controller | Mean Risk/Anchor↓ | Harm↓ |
|---|---|---|---|
| 8 | VCG | 1.1435 | 96% |
| 8 | RCPT | 0.9990 | 32% |
| 16 | VCG | 2.1288 | 92% |
| 16 | RCPT | 1.0659 | 44% |
| 32 | VCG | 3.3404 | 70% |
| 32 | RCPT | 1.7480 | 52% |

Paired RCPT-minus-VCG mean risk differences are $-0.1445$ (95% CI $[-0.2794, -0.0663]$), $-1.063$ ($[-4.051, -0.260]$), and $-1.592$ ($[-5.207, -0.165]$). At fit 8, mean fitting risk falls from approximately 286 for Anchor to $1.7 \times 10^{-5}$ for VCG, while independent risk increases. Protection is incomplete: at fit 32 the RCPT median is 1.003, and its worst class–seed ratio is 23.35. Means and tails must therefore both be reported.

**Metrics and cohort boundaries.** Each generation cohort contains 32 images per class and seed. A separate ResNet-20 supplies independent accuracy and classwise polynomial-kernel KID against 256 test images per class. DLSM accuracy reuses the guidance classifier family. Clean-FID uses Inception-v3 and all 10,000 test images. DINOv2-small uses normalized CLS features: FD is pooled, and classwise RBF MMD uses a target-only median-distance bandwidth. CLIP ViT-B/32 class margin is target similarity minus the largest non-target similarity, averaged over four fixed prompt templates. None enters controller fitting. Pooled FD/FID and means of per-seed FD/FID differ substantially at these sample sizes and are kept apart.

**Image-generation protocol.** The terminal image comparison uses one seed at Anchor scale 1.5. Every method contributes 32 outputs for each of ten classes, for 320 images per method. The protocol uses fit 16/evaluate 128 states at 16 schedule nodes, PC-1000, and SNR 0.16. VCG-4 retains the leading four fitting-covariance eigenvectors per class and time; full VCG and RCPT use ten classifier fields. The target power stays one: an Anchor scale $s$ changes the fitted response to $g_A + (s-1)a_c$, with unchanged target Stein features. The common corrector and predictor Anchor use scale $s$; the fitted correction enters the predictor. This single-seed terminal experiment is reported separately from the repeated-seed mechanism studies. It does not estimate between-fit uncertainty.

Table B.5 reports this comparison.

**Endpoint interpretation.** RCPT improves all seven reported metrics over Anchor, VCG, and both ablated configurations in this experiment. Relative to Anchor, its independent accuracy gain is one correctly classified image, while KID decreases from 0.1707 to 0.1595 and FID from 79.048 to 78.771. Reduced rank repairs part of the full VCG degradation. Direct LOO improves FID and DINO FD over Anchor but has lower independent accuracy and higher KID. The structured rule gives the most consistent endpoint improvement in this run. The single-seed design supports this observed comparison, without establishing seed-level significance or a uniform ordering across tasks. FID and DINO FD compare Gaussian feature summaries; they are not the image-space $W_2$ in Theorem 4.

**Selected qualitative contrasts.** Figure I.1 shows three selected examples from the terminal image experiment. Selection requires RCPT's independent evaluator prediction to match the target, then ranks cases by the number of incorrect alternatives, RCPT margin, and deterministic identifiers. The panel illustrates favorable classification changes and does not measure their frequency or human-perceived visual quality.

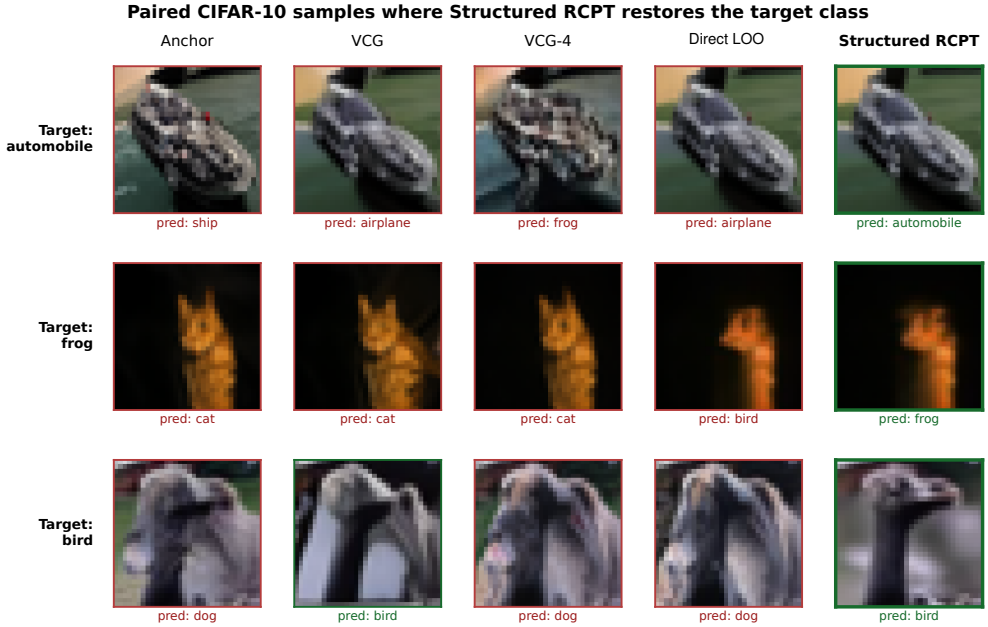


Figure I.1: **Selected CIFAR-10 classification contrasts.** Three matched class–sample pairs from the same seed, with common initial noise as recorded in the source report. Labels are independent ResNet-20 predictions. These examples were selected by the evaluator criterion, not sampled at random. Direct LOO is the calibration ablation; VCG-4 is the capacity ablation. The complete numerical comparison uses all 320 outputs.

**ECG mechanism-only comparison.** An exploratory seed evaluates all classes and 16 nodes at total information budgets 16 and 32. ECG uses potentials $\psi_j = \frac{1}{2}\mathsf{a}(t)\log p_\phi(j \mid x, t)$, whose gradients are exactly the ten classifier fields used by VCG. Full-data rules use the entire budget; split rules allocate half to fitting and half to calibration. Independent holdouts contain 32 or 64 states per node, respectively; ridge is $10^{-4}$. Raw ECG has large risk tails. Direct LOO improves it substantially and at budget 32 has lower mean and worst risk than structured RCPT in this audit. The predeclared mechanism gate stopped the study before image generation; there are no ECG terminal-image metrics. Its smaller holdout and different seed make its ratios distinct from the five-seed mechanism study.

### I.6 Ten-pocket scaffold fixed-rule comparison

We compare eleven prespecified controllers on the same ten scaffold–pocket targets. Every arm uses one $N_{\mathrm{fit}} = 16$ fit (one fitting seed) and a paired, independent $N_{\mathrm{deploy}} = 16$ Proposal-Only deployment per pocket. Anchor, VCG, and RCPT are the immutable matched rows from the full-76-pocket study; the eight fixed-scalar, ridge, EMA, and norm-cap controls reuse the same fitted information and deployment seeds. No decoded molecule or docking result tunes a controller. The ten pockets are the equally weighted experimental units.

Table I.12 confirms that RCPT repairs VCG: validity rises from 0.319 to 0.563, scaffold yield from 0.319 to 0.513, and active-path risk falls from 13.895 to 0.890. It is not uniformly best among simple rules on this small subset. Fixed $\alpha = 0.50$ has higher scaffold, affinity-improved, and unique-useful yields, with paired bootstrap intervals excluding zero, while the norm cap is numerically strong on several endpoints. Conversely, $\alpha = 0.75$, every ridge setting, and EMA deteriorate substantially.

Table I.11: **ECG image feasibility, one exploratory seed.** Independent Risk/Anchor, with mean, median, and worst class retained. Anchor is one by definition. Both objectives suffer finite-particle estimation error; neither calibrated rule removes all harmful classes.

| Budget | Rule | Mean | Median | Worst | Harm |
|---|---|---|---|---|---|
| 16 | VCG | 734.542 | 16.647 | 5492.323 | 10/10 |
| 16 | ECG | 2799.074 | 187.089 | 20568.155 | 10/10 |
| 16 | ECG + Direct LOO | 9.633 | 0.998 | 86.819 | 4/10 |
| 16 | ECG + split | 14.164 | 1.041 | 131.714 | 9/10 |
| 16 | RCPT | 1.837 | 0.965 | 8.954 | 3/10 |
| 32 | VCG | 4.710 | 1.418 | 25.341 | 9/10 |
| 32 | ECG | 450.485 | 4.500 | 4380.877 | 10/10 |
| 32 | ECG + Direct LOO | 1.111 | 0.978 | 2.273 | 3/10 |
| 32 | ECG + split | 2.234 | 0.999 | 13.056 | 5/10 |
| 32 | RCPT | 1.391 | 0.953 | 2.588 | 4/10 |

Table I.12: **Scaffold simple-baseline comparison.** Equal-pocket means over the matched ten-pocket fit-16/deploy-16 subset. "Affinity" is the fixed-denominator yield that both retains the scaffold and beats the pocket reference; "unique useful" additionally deduplicates molecular identities. Risk is measured on each method's active deployment path because the optional common-state audit did not complete; it is therefore a mechanism diagnostic, not a same-state counterfactual. Bold marks the best value within each of the two controller groups, including ties.

| Controller | Risk/Anchor ↓ | Valid ↑ | Scaffold ↑ | Affinity ↑ | Unique useful ↑ | QED ↑ | SA ↑ | Best QVina ↓ |
|---|---|---|---|---|---|---|---|---|
| Anchor | 1.000 | 0.463 | 0.450 | 0.363 | 0.263 | **0.517** | 0.653 | **−7.95** |
| VCG | 13.895 | 0.319 | 0.319 | 0.294 | 0.200 | 0.409 | 0.584 | −7.66 |
| RCPT | **0.890** | **0.563** | **0.513** | **0.413** | **0.306** | 0.489 | **0.662** | −7.93 |
| Fixed $\alpha = 0.25$ | 1.915 | 0.563 | 0.544 | 0.456 | 0.356 | **0.548** | 0.663 | −7.94 |
| Fixed $\alpha = 0.50$ | 5.116 | **0.588** | **0.569** | **0.506** | **0.406** | 0.529 | 0.629 | **−8.15** |
| Fixed $\alpha = 0.75$ | 11.483 | 0.406 | 0.400 | 0.350 | 0.256 | 0.509 | 0.631 | −8.10 |
| Ridge $\rho = 10^{-3}$ | 18.750 | 0.281 | 0.281 | 0.256 | 0.163 | 0.475 | 0.593 | −7.87 |
| Ridge $\rho = 10^{-2}$ | 18.429 | 0.238 | 0.238 | 0.231 | 0.138 | 0.483 | 0.593 | −7.75 |
| Ridge $\rho = 10^{-1}$ | 20.740 | 0.275 | 0.275 | 0.238 | 0.144 | 0.382 | 0.548 | −7.69 |
| VCG+EMA $\beta = 0.9$ | 2.991 | 0.269 | 0.269 | 0.244 | 0.150 | 0.441 | 0.596 | −7.53 |
| VCG+norm cap 0.25 | **0.994** | 0.544 | 0.538 | 0.456 | 0.356 | 0.507 | **0.664** | −8.02 |

The result supports adaptive negative-transfer protection, not dominance over an oracle-chosen fixed scalar for every subset.

### I.7 QM9 STEERING FIXED-RULE COMPARISON

The reward audit uses four seeds, 1,000 integration steps, 16 fitting particles, and eight 16-particle corrected deployment microbatches (128 slots per controller and seed). It uses the normalized three-field class consisting of the two reward-gradient fields and reverse base drift. The first deployment microbatch shares the fitting random stream and the remaining seven are fresh, so this audit supplements rather than replaces the independently seeded eight-seed, 12-field primary result in Table 4.

Table I.13 shows a complementary boundary. No fixed scalar improves mean MAE over VCG, whereas RCPT lowers MAE from 0.8499 to 0.8252 and raises unique success from 33.5 to 39.2 while slightly lowering active-path risk. Fixed ridge at $\rho = 10^{-1}$ leads MAE and unique success, but its effect is nonmonotone: $\rho = 10^{-3}$ is worse than VCG in both metrics. With only four seeds, all paired MAE intervals against VCG include zero; this audit therefore establishes a fixed-rule sensitivity, not statistically decisive reward-domain superiority.

Table I.13: **QM9 steering simple-baseline comparison.** Means over four fit seeds; MAE reports mean ± seed standard deviation. Unique success counts distinct valid molecular identities meeting both targets at the unchanged one-MAD threshold per 128 generation slots, so duplicates and invalid samples consume budget. Risk is the integrated selected/Anchor residual-variance ratio on each controller's own deployment genealogy and is not a common-state comparison. Bold marks the best value within each of the two controller groups, including ties.

| Controller | Risk/Anchor↓ | MAE↓ | Unique success /128↑ |
|---|---|---|---|
| Anchor | 1.000 | $0.9346 \pm 0.0331$ | 20.5 |
| VCG | 0.242 | $0.8499 \pm 0.0205$ | 33.5 |
| RCPT | **0.224** | $\mathbf{0.8252 \pm 0.0462}$ | **39.2** |
| Fixed $\alpha = 0.25$ | 1.215 | $0.9678 \pm 0.0996$ | 20.5 |
| Fixed $\alpha = 0.50$ | 0.330 | $0.8781 \pm 0.0298$ | 29.2 |
| Fixed $\alpha = 0.75$ | 0.086 | $0.8593 \pm 0.0597$ | 35.2 |
| Ridge $\rho = 10^{-3}$ | 0.157 | $0.8892 \pm 0.0341$ | 31.8 |
| Ridge $\rho = 10^{-2}$ | 0.068 | $0.8237 \pm 0.0438$ | 38.5 |
| Ridge $\rho = 10^{-1}$ | **0.041** | $\mathbf{0.8136 \pm 0.0638}$ | **39.8** |
| VCG+EMA $\beta = 0.9$ | 1.746 | $0.8637 \pm 0.0639$ | 32.2 |
| VCG+norm cap 0.25 | 0.045 | $0.8412 \pm 0.0413$ | 33.2 |

### I.8 Bias and penalty-estimator audits

**Coefficient bias.** The expected-risk oracle retains $B = (\lambda^\star)^\top C\, \mathbb{E}[\widehat{\lambda} - \lambda^\star]$; the structured model does not estimate it separately. The Checker audit uses the 12-field basis, $N = 8$, 8,192 reference particles, 10,000 independent fitting ensembles, 80 time points, and the matched ridge convention. The mean bias-aware and no-bias trust are 0.1900 and 0.1934; their mean absolute pointwise difference is 0.0176 (median 0.0077, 95th percentile 0.0661, maximum 0.1071). Ignoring $B$ increases the integrated oracle risk by 0.781%. This supports a practical simplification on Checker, not the general claim $B = 0$ or $B = o(N^{-1})$.

**Penalty approximation.** The large-population Checker audit uses five trajectories, 8,192 population particles per trajectory, 40 fitted times, and 500 repeated fits at $N \in \{16, 32, 64, 128, 256\}$. Across 200 seed–time problems, the structured proxy and $N^{-1}\,\mathrm{tr}(\Omega C^\dagger)$ have Spearman correlation 0.893, but their integrated scale ratio is only 0.167. At $N = 256$, measured penalty is $1.071\times$ the $\Omega$-based prediction while the structured proxy is low by $6.289\times$. At $N = 16$, neither first-order approximation is quantitatively accurate.

**Realized calibration is a separate empirical question.** Neither preceding expected-penalty audit estimates the realized $\kappa$ required by Theorem 3. The corresponding target-law quantities are $q = \mathrm{Var}_p(a^\top \widehat{\lambda})$, $\kappa = \mathrm{Cov}_p(g^0 + a^\top \widehat{\lambda}, a^\top \widehat{\lambda})$, and $\alpha_{\mathrm{ray}} = \Pi_{[0,1]}(1 - \kappa/q)$ when $q > 0$. The relevant errors are $S - q$ and $\widehat{P} - \kappa$, together with selected-amplitude error, realized ray regret, and negative-transfer frequency. When $\kappa < 0$, the ray oracle retains the full update, and a nonnegative penalty has error at least $-\kappa$. Appendix I.2 measures these quantities on an independent corrected-SMC reference. It finds lower selected-ray regret despite substantial proxy error. The reference is approximate and dependent within each particle system, so the audit does not certify the target-law bound for Algorithm 1.

## J Additional Mechanism Diagnostics

The diagnostics below supplement the main figures with fitting-budget, paired-risk, and genealogy measurements.

### J.1 Mechanism diagnostics

The fitting-population sweep contains ten paired Checker SMC runs at each of $N_{\mathrm{fit}} \in \{4, 8, 16, 32, 64, 128, 256\}$, with 10,000 evaluation particles from a fresh deployment population.

Table I.14: **Covariance-penalty ablation.** Ten paired Checker seeds. Replacing the structured effective-rank proxy with a full $\Omega$ estimate gives no consistent endpoint gain. This changes the penalty estimator, separately from the Direct LOO amplitude ablation.

| Mode | Fit | Estimator | $W_2 \downarrow$ | MMD$\downarrow$ |
|---|---|---|---|---|
| SMC | 8 | Full $\Omega$ | 2.010 | **2.917** |
| SMC | 8 | RCPT | **1.784** | 3.181 |
| SMC | 16 | Full $\Omega$ | 2.170 | 2.213 |
| SMC | 16 | RCPT | **2.030** | **2.069** |
| Proposal | 8 | Full $\Omega$ | 1.769 | 2.519 |
| Proposal | 8 | RCPT | **1.768** | 2.519 |
| Proposal | 16 | Full $\Omega$ | 1.469 | 1.275 |
| Proposal | 16 | RCPT | **1.443** | **1.269** |

All 70 runs per method completed. Figure 1 compares the integrated fitting and independent residual ratios, each with its own matched Anchor denominator. Low-particle VCG nearly interpolates its fitting states, although its independent risk can be orders of magnitude larger. RCPT gives up some fitting improvement to reduce independent risk. At fit 8 the full sweep has median risks 105.22 and 0.966, and mean risks 524.17 and 1.084 for VCG and RCPT; the mean reflects a remaining RCPT tail. This fitting-population sweep is separate from the fixed-schedule endpoint comparison in Table 2.

Figure J.1 displays the same sweep against particle count on both panels. The median support ratios $N_{\text{fit}}/r_{\text{eff}}$ for RCPT are 1.33, 1.15, 1.48, 2.80, 5.50, 10.88, and 21.75; the effective rank itself changes at small $N$. Mean amplitude rises from 0.143 at $N = 4$ to 0.921 at $N = 256$, while VCG and RCPT independent risks converge. The method therefore retains more of the fitted update as the independent risks converge; this observation does not identify the population-optimal headroom.

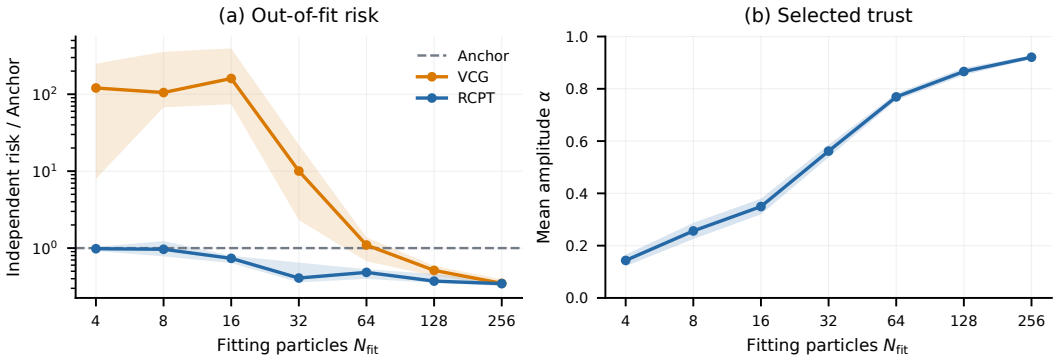


Figure J.1: **Increasing fitting information closes the gap to VCG.** Risk curves show medians and interquartile ranges; amplitude shows mean and standard error over ten seeds at each particle count.

Figure J.2 compares independent risk across the paired methods. RCPT improves 66/70 Checker runs and repairs 34/47 harmful VCG runs to Anchor or below. The separate Scaffold common-state audit contains 75 pockets; pocket 45 was reserved for the engineering smoke test before this audit. RCPT improves every pocket and repairs 68/73 harmful VCG schedules. This diagnostic cohort is distinct from the full 76-pocket deployment table. The ratios are not pooled across domains.

### J.2 Selected amplitude across tasks

Figure 4 summarizes time-resolved amplitudes. Thick lines are means across units, bands are interquartile ranges, and thin lines are individual schedules. The summary averages the saved time nodes without claiming a common continuous-time integral. Native time runs in opposite generation directions in some domains. Amplitude is relative to each fitted coefficient; it is not a comparable drift norm or an independent measurement of population headroom.

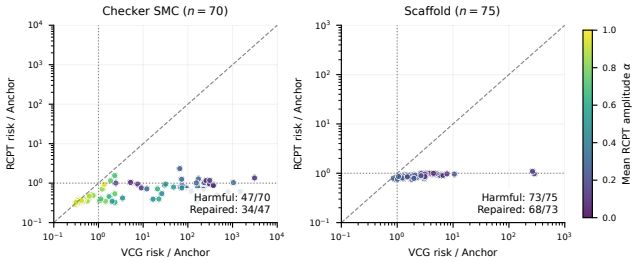


Figure J.2: **Risk repair and retained adaptation.** Points below the diagonal favor RCPT. Points to the right of the vertical Anchor line and below the horizontal one are harmful VCG fits repaired to Anchor or below. Color denotes mean selected amplitude, not terminal quality.

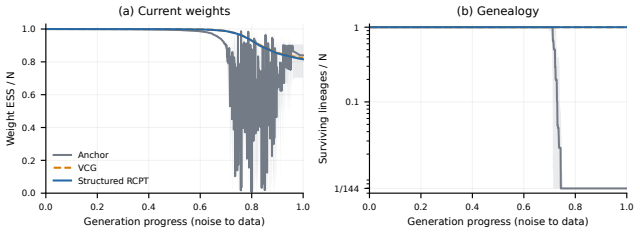


Figure J.3: **Weight ESS and genealogy in online QM9 steering.** Median and interquartile range across eight $N = 144$ seeds, ordered in generation time. Anchor resamples 85.8 times on average and ends with one ancestor in every seed; VCG and RCPT never resample and retain all 144. Resampling can restore weight ESS without restoring lost ancestry. These are observed path-dependent outcomes, not a universal consequence of local residual ordering.

The cohorts are: 76 Scaffold pockets from the common-state audit; eight online QM9 steering seeds with 144 live particles; ten Checker fit-16 schedules from the optimizer-comparison audit; and 20 CIFAR seed–class cells from an earlier two-seed scale-1.5 calibration audit. Those CIFAR traces are not from the single-seed endpoint experiment. The Checker alpha cohort is distinct from both the primary fixed-schedule endpoint table and the single-seed replay. Mean structured amplitudes are 0.190, 0.945, 0.825, and 0.111, respectively. Structured RCPT is more conservative than Direct LOO on Scaffold and CIFAR, while retaining more of the VCG fit in online QM9 steering. This is an observed response of the rule, not a proof that it identifies the terminal-optimal amplitude.

### J.3 Residual risk as a terminal surrogate

Across 83 finite Proposal-Only profiles, integrated held-out residual has Spearman correlation $0.775$ with empirical TV (Pearson $0.881$ after log scaling its many-order-of-magnitude range), while the theorem-aligned $B_1 = \frac{1}{2}\int_0^1 \sqrt{V_t}\,\mathrm{d}t$ has Spearman correlation $0.730$. Retaining the time-resolved profile matters: $B_1$ is a median $47\%$ tighter than the additional time–Cauchy relaxation $B_2 = \frac{1}{2}\sqrt{\int_0^1 V_t\,\mathrm{d}t}$. Nevertheless, every evaluated $B_1$ exceeds one and its median is $5.99\times$ empirical TV. We therefore use residual risk as a correlated comparative diagnostic, not an exact terminal error estimate.

### J.4 Cross-domain portability and fixed controls

The CIFAR-10 study in Appendix I.5 extends the finite-particle mechanism to a pretrained image model. Its independent-state risk audits and separate terminal cohorts answer different questions: residual generalization and downstream behavior. The current terminal comparison is the single-seed scale-1.5 experiment.

Appendix Tables I.12 and I.13 compare fixed shrinkage, ridge, and engineering safeguards. The best fixed heuristic does not transfer: a fixed scalar can be strong on selected scaffold yields, while

no fixed scalar improves QM9 steering MAE over VCG; ridge behavior is similarly task-dependent and nonmonotone. RCPT is the common arm that estimates trust from the available fit rather than selecting a task-specific regularization value.

## K Computational Cost

After expert fields and divergences are available, VCG costs $O(Nm^2 + m^3)$ per fitted step. Leave-one-out refitting reuses those features and adds small dense linear algebra, but no expert, denoiser, or evaluator call. Table B.6 reports matched costs with hardware; all times exclude downstream docking and offline evaluation.

Scaffold timing means cover all 76 pockets on an RTX A6000. QM9 times use four A100-SXM4 runs; runs on other hardware are excluded. The 1,000-step QM9 runs predated memory instrumentation, so peak memory comes from a matched 10-step probe. Small VCG–RCPT timing reversals reflect run variability. RCPT adds 3.13 seconds over VCG in approximately 3,880 seconds (0.081%) with equal probed memory; the recorded Anchor time is 3877.460 seconds.

The CIFAR terminal-evaluation record gives 211.41 seconds per class for VCG and 214.26 for RCPT, a 1.34% difference. These records do not separate controller work from run variability or identify hardware, so they are a paired wall-time observation rather than a controller-only overhead estimate.

**Why the small-particle regime matters computationally.** RCPT's main systems benefit is not that leave-one-out refitting makes one trajectory cheaper; it is that a controller can be fit on the small ensemble the user actually wants to generate. Table K.1 measures this scaling in our implementations.

Table K.1: Particle scaling for the 12-field scaffold RCPT fit on NVIDIA RTX A6000 48GB GPUs, averaged over the complete 76-pocket benchmark. The *Solve* column is aggregate proposal-controller solve time; the remainder is dominated by trajectory and field/model computation.

| $N_{\mathrm{fit}}$ | s/pocket | Peak GB | Solve s |
|---|---|---|---|
| 16 | 197.445 | 4.740 | 5.282 |
| 32 | 268.852 | 9.365 | 5.473 |
| 64 | 419.193 | 18.611 | 5.478 |

From 16 to 64 scaffold particles, allocated VRAM rises by $3.93\times$ and wall time by $2.12\times$, whereas aggregate controller-solve time changes by only 0.20 seconds. On the same A100, matched QM9 steering probes require 4.569 GB at $N = 16$ and 40.204 GB at $N = 144$ ($8.80\times$ memory for $9\times$ as many live particles). These are direct measurements of particle scaling, not an extrapolation to thousands of particles.

These matched measurements support calibration at low particle budgets: 12-field controllers fitted from 8–16 particles add negligible computation over VCG, while larger fitting populations increase model cost and memory. They do not measure a speedup over DriftLite's published high-particle experiments (Ren et al., 2026).

A frozen schedule also amortizes its fitting time over independent outputs. Deployment remains microbatchable: it retains the fields needed to apply the combination while removing online covariance, LOO, and particle-synchronization work.