SCMO: STOCHASTIC CONTROL FOR OPTIMIZATION OVER PROBABILITY MEASURES ON INFINITE-DIMENSIONAL SPACES

Anonymous authors
Paper under double-blind review

ABSTRACT

We study objective-only optimization of possibly nonconvex and nonsmooth functionals over probability measures on a separable Hilbert space, allowing the optimizer to be intrinsically non-Dirac. We introduce SCMO (Stochastic Control Measure Optimizer), a gradient-free particle method derived from entropy-regularized stochastic control. After finite-particle and Galerkin approximations, a Cole–Hopf transform represents the optimal feedback as a Gibbs-weighted terminal displacement. SCMO approximates this feedback by sampling context clouds, replacing one particle at a time with candidate draws, scoring the resulting empirical measures, and applying exponential reweighting.

SCMO uses separate covariances for candidate proposals and particle updates, termed matched when equal and nonmatched otherwise; our analysis covers both settings. In the matched case, we establish PDE-free qualitative convergence and a projection-first quantitative bound separating particle, Galerkin-projection, and entropy-regularization errors. For the practical multi-context implementation with nonmatched covariance, we prove finite-candidate consistency and show that the signed, curvature-dependent effect of nonmatched covariance can reduce the resulting upper bound on the approximation error. Finally, experiments on function-space, trajectory-law, and contact-rich manipulation problems show that SCMO handles nonsmooth and nonconvex objectives, escapes suboptimal local basins, and recovers prescribed multimodal, non-Dirac law structure.

1 INTRODUCTION

Many problems in machine learning and control seek a probability law over high- or infinite-dimensional solutions rather than a single optimizer, often under nonconvex or nonsmooth objectives; examples include multimodal robot policies (Chi et al., 2023; Ding et al., 2025), function-space generative models (Kerrigan et al., 2024; Li et al., 2026), and nonconvex optimization over probability measures (Luu et al., 2024). Motivated by these settings, we study problems of the form

$$\mathcal{V}_0 := \inf_{\mu \in \mathcal{P}_2(\mathcal{H})} G(\mu), \tag{1}$$

where $\mathcal{H}$ is a separable Hilbert space and $G : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ may be nonconvex, nonsmooth, and interacting. It combines three challenges: the measure-valued decision variable is infinite-dimensional; function-space implementations must control both particle and Galerkin-projection errors; and gradients or Lions derivatives may be unavailable or uninformative for simulated or discontinuous objectives. This motivates an objective-only method that maintains a flexible empirical law and provides particle-wise credit assignment.

This work is inspired by Qiu (2026), which connects optimization over $\mathcal{P}_2(\mathbb{R}^d)$ with entropy-regularized mean-field controls and N-particle approximations. However, that analysis is finite-dimensional and uses the same Brownian covariance in the feedback representation and the particle dynamics. As the corresponding regularization bound scales with the ambient dimension, the estimate is not uniform as $d \to \infty$. The present work nontrivially enlarges this framework to $\mathcal{P}_2(\mathcal{H})$ for a separable Hilbert space $\mathcal{H}$, using trace-class Q-Wiener noise and Galerkin projection. It also introduces an R-context estimator that recovers the unconditional full-cloud feedback as $R \to \infty$ and

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.

separates the proposal covariance used for feedback estimation from the execution covariance used in the particle dynamics. We call it SCMO (*Stochastic Control Measure Optimizer*). Its methodological and theoretical contributions are summarized below.

Our Contributions. Our contributions are fivefold. **(i) Hilbert-space formulation and Galerkin approximation:** we address the optimization over the space of probability measures on infinite-dimensional separable Hilbert spaces, and obtain an objective-only Galerkin implementation for function-valued laws. The resulting method requires neither particle gradients nor Lions derivatives. The infinite-dimensionality refers to the base space on which the measures are supported, rather than to the dimensionality of the Wasserstein space itself. **(ii) Matched-covariance convergence analysis:** we prove a novel PDE-free qualitative convergence for a class of nonsmooth and discontinuous objectives. Under stronger regularity, we derive a projection-first quantitative bound that separates finite-particle, Galerkin-projection, and regularization errors. **(iii) Nonmatched-covariance convergence analysis:** we distinguish the proposal covariance used to estimate the feedback from the execution covariance used to update the particles. We derive a signed, curvature-dependent nonmatched term and show that, under the stated sign condition, it sharpens the matched-covariance upper approximation certificate. **(iv) Multi-context finite-candidate estimation:** we introduce an estimator using R independent complete context clouds and S candidate replacements per context, with joint normalization over the resulting RS scores. We establish consistency and a mean-square error bound, and connect the practical finite-candidate implementation to the exact Euler chain. **(v) Empirical evaluation:** experiments on function-space, trajectory-law, and contact-rich Push-T problems show that SCMO handles nonsmooth, nonconvex, and discontinuous objectives, escapes suboptimal local basins, and recovers prescribed multimodal, non-Dirac law structure.

2 Related Work

Closest precursor. The closest precursor is Qiu (2026). Section 1 states the precise departures developed in the present work.

Entropy-regularized and Gibbs-weighted methods. The soft-constrained Schrödinger bridge of Garg et al. (2024) replaces exact terminal-law matching by a Kullback–Leibler penalty and yields a geometric-mixture characterization of the optimal terminal law; this differs from optimizing a general, potentially interacting law functional. Path-integral methods similarly use Gibbs-weighted rollouts to update controls (Theodorou et al., 2010; Williams et al., 2017b), whereas SCMO uses context-conditioned weights to update an empirical law. Entropy-SGD and related methods also share SCMO's general Gibbs/log-partition structure, exponentially weighting nearby parameter values according to their loss (Chaudhari et al., 2017; Pittorino et al., 2021). They estimate the resulting local-entropy gradient through an auxiliary noisy gradient process and update a single parameter vector. SCMO instead assigns Gibbs weights to context-conditioned one-particle replacements using the objective values of the complete empirical law and executes objective-only particle dynamics.

Black-box, particle, and infinite-dimensional optimization. CEM, CMA-ES, and natural evolution strategies adapt a parameterized search distribution to locate high-quality points (de Boer et al., 2005; Hansen & Ostermeier, 2001; Wierstra et al., 2014); when a complete particle cloud is vectorized, each proposal receives one global score, whereas SCMO assigns one-particle replacement scores and treats the resulting law as the solution. SVGD, particle quantization, mean-field Langevin and stochastic particle descent, entropic fictitious play, and Wasserstein or KL flows instead use gradients, scores, or variational structure (Liu & Wang, 2016; Xu et al., 2022; Chizat & Bach, 2018; Hu et al., 2021; Nitanda & Suzuki, 2017; Chen et al., 2023; Yao et al., 2024; Tankala et al., 2025; Kook et al., 2024; Luu et al., 2024). Consensus-based optimization is derivative-free but designed to concentrate on a point minimizer (Pinnau et al., 2017; Carrillo et al., 2018); recent Hilbert-space extensions and directional pre-basis methods likewise optimize a point or function (Huang & Kouhkouh, 2026; Khatab & Totzeck, 2026; Peixoto et al., 2025). SCMO instead uses only evaluations of a functional on $\mathcal{P}_2(\mathcal{H})$ and may return a multimodal, non-Dirac law.

3 SCMO: STOCHASTIC-CONTROL FORMULATION AND FINITE-CANDIDATE UPDATE

Let $(\mathcal{H}, \langle\cdot,\cdot\rangle, \|\cdot\|)$ be a separable Hilbert space, and let $\mathcal{P}_2(\mathcal{H})$ denote the space of probability measures on $\mathcal{H}$ with finite second moments, equipped with the 2-Wasserstein metric $\mathcal{W}_2$. We study the measure-valued optimization problem $\mathcal{V}_0 := \inf_{\mu\in\mathcal{P}_2(\mathcal{H})} G(\mu)$, where $G : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ is a possibly nonconvex and nonsmooth functional.

3.1 ENTROPY-REGULARIZED MEASURE CONTROL

Let Q be a symmetric, positive-definite, trace-class covariance operator on $\mathcal{H}$. Write $\mathcal{H}_Q := Q^{1/2}(\mathcal{H})$ for the associated Cameron–Martin space, i.e. the space of directions with finite Q-energy, equipped with the norm $\|h\|_Q := \|Q^{-1/2}h\|$, $h \in \mathcal{H}_Q$. Equivalently, if $\{e_j\}_{j\ge1}$ is an orthonormal eigenbasis of Q with eigenvalues $\{\lambda_j\}_{j\ge1}$, then $\mathcal{H}_Q = \left\{h = \sum_{j\ge1} h_j e_j : \sum_{j\ge1} \frac{h_j^2}{\lambda_j} < \infty\right\}$, $\|h\|_Q^2 = \sum_{j\ge1} \frac{h_j^2}{\lambda_j}$. Let $(W_t)_{t\ge0}$ be a Q-Wiener process on $\mathcal{H}$ defined on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$. For $t \in [0,1]$ and $\mu \in \mathcal{P}_2(\mathcal{H})$, let ξ be an $\mathcal{F}_t$-measurable $\mathcal{H}$-valued random variable with law μ, independent of the future increments $(W_s - W_t)_{s\ge t}$. Let Θ_t denote the set of $\mathcal{H}_Q$-valued progressively measurable controls satisfying $\mathbb{E}\int_t^1 \|\theta_s\|_Q^2\, ds < \infty$. For $\theta \in \Theta_t$, define $X_s^{t,\xi;\theta} = \xi + \int_t^s \theta_r\, dr + W_s - W_t$, $s \in [t,1]$. For $\varepsilon \in (0,1)$, define the entropy-regularized value function

$$V_\varepsilon(t,\mu) := \inf_{\theta\in\Theta_t} \left\{ G\big(\mathcal{L}(X_1^{t,\xi;\theta})\big) + \mathbb{E}\left[\frac{\varepsilon}{2}\int_t^1 \|\theta_s\|_Q^2\, ds\right]\right\}, \qquad (t,\mu) \in [0,1]\times\mathcal{P}_2(\mathcal{H}), \tag{2}$$

where the quadratic term $\mathbb{E}\left[\frac{1}{2}\int_t^1 \|\theta_s\|_Q^2\, ds\right]$ by Girsanov's theorem is the relative entropy of the law of $X^{t,\xi;\theta}$ with respect to the law of the uncontrolled process $X^{t,\xi;0}$, providing the rationale for the terminology "entropy-regularized". In particular, for an initial law $\mu_0 = \mathcal{L}(\xi_0)$,

$$V_\varepsilon(0,\mu_0) = \inf_{\theta\in\Theta_0} \left\{ G\big(\mathcal{L}(X_1^{0,\xi_0;\theta})\big) + \mathbb{E}\left[\frac{\varepsilon}{2}\int_0^1 \|\theta_s\|_Q^2\, ds\right]\right\}.$$

Formally applying the dynamic programming principle on $\mathcal{P}_2(\mathcal{H})$ yields the mean-field Hamilton–Jacobi-Bellman (HJB) equation

$$\begin{cases} \partial_t V_\varepsilon(t,\mu) - \dfrac{1}{2\varepsilon}\displaystyle\int_{\mathcal{H}} \|Q^{1/2}\partial_\mu V_\varepsilon(t,\mu)(x)\|^2\, \mu(dx) + \frac{1}{2}\int_{\mathcal{H}} \mathrm{Tr}\left(Q\,\partial_x\partial_\mu V_\varepsilon(t,\mu)(x)\right)\, \mu(dx) = 0, \\ V_\varepsilon(1,\mu) = G(\mu). \end{cases} \tag{3}$$

Here ∂_μ denotes the Lions derivative and $\partial_x\partial_\mu$ denotes its Fréchet derivative in the spatial variable.

3.2 PARTICLE–GALERKIN APPROXIMATION AND COLE–HOPF FEEDBACK

The formal measure-level Equation (3) is computationally intractable: the state variable is a probability measure on an infinite-dimensional space. To tackle this issue, SCMO uses two approximations. First, the measure is represented by an empirical particle law. Second, the Hilbert space is projected to a finite-dimensional Galerkin subspace.

For $\mathbf{x} = (x_1,\dots,x_N) \in \mathcal{H}^N$, define the empirical measure $\mu_{\mathbf{x}}^N := \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$. For $K \ge 1$, set $\mathcal{H}_K := \mathrm{span}\{e_1,\dots,e_K\}$, $Q_K := \Pi_K Q \Pi_K$, where $\Pi_K : \mathcal{H} \to \mathcal{H}_K$ is the orthogonal projection. Let $\iota_K : \mathcal{H}_K \hookrightarrow \mathcal{H}$ be the canonical embedding. For $\nu \in \mathcal{P}_2(\mathcal{H}_K)$, the projected terminal functional is the restriction $G_K(\nu) := G((\iota_K)_\#\nu)$. When no confusion can arise, we write simply $G(\nu)$ for $G_K(\nu)$. The Wasserstein distances used to compare with the original optimizer are computed after embedding laws into $\mathcal{H}$.

For $\mathbf{x} \in (\mathcal{H}_K)^N$, define the projected N-particle value function

$$v_\varepsilon^{K,N}(t,\mathbf{x}) := \inf_{\boldsymbol{\theta}} \mathbb{E}\left[NG(\mu_{\mathbf{X}_1^K}^N) + \frac{\varepsilon}{2}\sum_{i=1}^N \int_t^1 \|\theta_s^i\|_{Q_K}^2\, ds\right], \tag{4}$$

where $\boldsymbol{\theta} = (\theta^1, \dots, \theta^N)$ ranges over progressively measurable $\mathcal{H}_K$-valued controls with finite energy, and $\mathbf{X}_s^K = (X_s^{K,1}, \dots, X_s^{K,N})$ is given by $X_s^{K,i} = x_i + \int_t^s \theta_r^i \, dr + W_s^{K,i} - W_t^{K,i}$, $s \in [t,1]$, $i = 1, \dots, N$. Here $W^{K,1}, \dots, W^{K,N}$ are independent Q_K-Wiener processes on $\mathcal{H}_K$.

When a smooth value function exists, $v_\varepsilon^{K,N}$ formally solves the finite-dimensional HJB equation

$$\begin{cases} \partial_t v_\varepsilon^{K,N} + \frac{1}{2}\sum_{i=1}^N \mathrm{Tr}\big(Q_K D_{x_i}^2 v_\varepsilon^{K,N}\big) - \frac{1}{2\varepsilon}\sum_{i=1}^N \left\| Q_K^{1/2} D_{x_i} v_\varepsilon^{K,N} \right\|^2 = 0, \\ v_\varepsilon^{K,N}(1, \mathbf{x}) = NG(\mu_{\mathbf{x}}^N). \end{cases} \tag{5}$$

This projected particle problem is the point at which the Cole–Hopf transform becomes applicable. Note that the Cole–Hopf transform does not directly linearize the infinite-dimensional Master Equation (3); the obstruction is explained in Appendix A.1. Let $u_\varepsilon^{K,N}(t, \mathbf{x}) := \exp\big(-v_\varepsilon^{K,N}(t, \mathbf{x})/\varepsilon\big)$. Then $u_\varepsilon^{K,N}$ solves the linear heat equation associated with the N independent Q_K-Wiener particles, with terminal condition $\exp(-NG(\mu_{\mathbf{x}}^N)/\varepsilon)$. Equivalently, for $t < 1$ and $\tau := 1 - t$, Feynman–Kac formula gives

$$u_\varepsilon^{K,N}(t, \mathbf{x}) = \mathbb{E}\left[\exp\left(-\frac{N}{\varepsilon} G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}_\tau^K})\right)\right], \tag{6}$$

where

$$\Delta\mathbf{W}_\tau^K = (\Delta W_\tau^{K,1}, \dots, \Delta W_\tau^{K,N}), \qquad \Delta W_\tau^{K,i} \sim \mathcal{N}(0, \tau Q_K)$$

are independent. The optimal feedback drift for particle i is therefore

$$\theta_i^{*,K,N}(t, \mathbf{x}) = -\frac{1}{\varepsilon} Q_K D_{x_i} v_\varepsilon^{K,N}(t, \mathbf{x}) = \frac{1}{\tau} \frac{\mathbb{E}\left[\exp\left(-\frac{N}{\varepsilon} G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}_\tau^K})\right) \Delta W_\tau^{K,i}\right]}{\mathbb{E}\left[\exp\left(-\frac{N}{\varepsilon} G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}_\tau^K})\right)\right]}. \tag{7}$$

Note that this expression is gradient-free: it only requires evaluations of G on empirical measures.

3.3 Finite-candidate SCMO update

Motivated by the projected-particle feedback formula Equation (7), we propose the following algorithm. Full pseudocode is given in Appendix A.3. SCMO evolves a Galerkin coefficient cloud $\mathbf{C} = (C_1, \dots, C_N)^\top \in \mathbb{R}^{N \times K}$, representing $u^{(i)} = \sum_{k=1}^K C_{i,k} e_k$, $i = 1, \dots, N$. For user-chosen scale parameters σ_{prop}, $\sigma_{\mathrm{dyn}} > 0$, let $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_K)$ and define $Q_{\mathrm{prop}} = \sigma_{\mathrm{prop}}^2 \Lambda$ and $Q_{\mathrm{dyn}} = \sigma_{\mathrm{dyn}}^2 \Lambda = \alpha_0 Q_{\mathrm{prop}}$, where $\alpha_0 = \sigma_{\mathrm{dyn}}^2 / \sigma_{\mathrm{prop}}^2$. The proposal covariance Q_{prop} is used to estimate the feedback, whereas the execution covariance Q_{dyn} is used in the Euler update. The cases $\alpha_0 < 1$ and $\alpha_0 > 1$ correspond respectively to colder and hotter execution. A nonmatched choice is motivated empirically (see Appendix A.2) and by the signed curvature contribution derived in Appendix C.1, which can tighten the matched approximation bound.

Initialize $\mathbf{C} \leftarrow \mathbf{C}_{\mathrm{init}}$ and $\Delta t = T/M$. For each of the L outer loops and $m = 0, \dots, M-1$, set $t_m = m\Delta t$ and $\tau_m = T - t_m$. Draw R independent complete context clouds and S candidate replacements per context:

$$\widetilde{C}_p^{(r)} = C_p + \sqrt{\tau_m}\, \sigma_{\mathrm{prop}} Z_p^{(r)} \Lambda^{1/2}, \qquad Y_{i,r,s} = C_i + \sqrt{\tau_m}\, \sigma_{\mathrm{prop}} \xi_{i,r,s} \Lambda^{1/2}, \tag{8}$$

where $r = 1, \dots, R$, $s = 1, \dots, S$ and $Z_p^{(r)}, \xi_{i,r,s} \sim \mathcal{N}(0, I_K)$ are independent. Here $\widetilde{\mathbf{C}}^{(r)} = (\widetilde{C}_1^{(r)}, \dots, \widetilde{C}_N^{(r)})$ is one Gaussian perturbation of the full current cloud. When estimating the feedback of particle i, the r-th context supplies $\{\widetilde{C}_p^{(r)} : p \neq i\}$, while $Y_{i,r,s}$ replaces its i-th component.

For each particle i, define the replacement scores, globally normalized Gibbs weights, and feedback estimate by

$$\begin{aligned} \mathcal{E}_{i,r,s} &= N\, G_K\left(\frac{1}{N}\delta_{Y_{i,r,s}} + \frac{1}{N}\sum_{p \neq i} \delta_{\widetilde{C}_p^{(r)}}\right), \qquad \mathcal{E}_i^{\min} = \min_{r,s} \mathcal{E}_{i,r,s}, \\ w_{i,r,s} &= \frac{\exp\big(-(\mathcal{E}_{i,r,s} - \mathcal{E}_i^{\min})/\varepsilon\big)}{\sum_{q=1}^R \sum_{j=1}^S \exp\big(-(\mathcal{E}_{i,q,j} - \mathcal{E}_i^{\min})/\varepsilon\big)}, \qquad \widehat{\theta}_i = \frac{1}{\tau_m} \sum_{r=1}^R \sum_{s=1}^S w_{i,r,s}(Y_{i,r,s} - C_i), \end{aligned} \tag{9}$$

where we are approximating the ratio $\frac{\mathbb{E}\Big[\mathbb{E}\Big[\exp\Big(-\frac{N}{\varepsilon}G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}^K_\tau})\Big)\Delta W^{K,i}_\tau|\widetilde{\mathbf{C}}^{(r)}_{-i}\Big]\Big]}{\mathbb{E}\Big[\mathbb{E}\Big[\exp\Big(-\frac{N}{\varepsilon}G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}^K_\tau})\Big)|\widetilde{\mathbf{C}}^{(r)}_{-i}\Big]\Big]} = \frac{\mathbb{E}\Big[\exp\Big(-\frac{N}{\varepsilon}G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}^K_\tau})\Big)\Delta W^{K,i}_\tau\Big]}{\mathbb{E}\Big[\exp\Big(-\frac{N}{\varepsilon}G(\mu^N_{\mathbf{x}+\Delta\mathbf{W}^K_\tau})\Big)\Big]}$. The estimator $\widehat{\theta}_i$ is the Gibbs-weighted terminal displacement divided by the remaining time, while subtracting $\mathcal{E}_i^{\min}$ is only a numerical stabilization.

All particles are updated simultaneously. Define $\overline{C}_i = C_i + \Delta t\,\widehat{\theta}_i$ and $\overline{\mathbf{C}} = (\overline{C}_1, \ldots, \overline{C}_N)^\top$, draw a fresh $\boldsymbol{\eta}_m \sim \mathcal{N}(0, I_{N\times K})$, independently of all context and candidate variables, and set

$$\mathbf{C} \leftarrow \overline{\mathbf{C}} + \sqrt{\Delta t}\,\sigma_{\mathrm{dyn}}\boldsymbol{\eta}_m\Lambda^{1/2}. \tag{10}$$

After all outer loops, SCMO returns $\mu^N = N^{-1}\sum_{i=1}^N \delta_{u^{(i)}}$. Finally, note that the frozen-background particle-wise approximation of Qiu (Qiu, 2026, Section 3.3) corresponds to $R = 1$ in our notation. This is exact for additive objectives as $S \to \infty$, because the background contribution cancels from the Gibbs ratio. For general interacting objectives, $R = 1$ gives only a conditional feedback, whereas the unconditional full-cloud feedback is recovered as $R \to \infty$; see Appendix A.4.

4 Convergence Analysis

We first analyze the matched-covariance problem $Q_{\mathrm{prop}} = Q_{\mathrm{dyn}}$, because in this case the Cole–Hopf representation yields a clean approximation theory. We give two guarantees for the idealized projected particle problem: a PDE-free qualitative result under a weak recovery condition and a quantitative projection-first bound separating particle, projection, and regularization errors. We then return to the practical multi-context, finite-candidate implementation with nonmatched covariance. For one outer-loop rollout, we quantify the context–candidate sampling error, the Euler discretization residual, and the signed effect of replacing the proposal covariance by the execution covariance.

4.1 Qualitative convergence

The first result uses only the finite-dimensional risk-sensitive representation. It does not require a Master equation, Lions differentiability, or convexity.

Assumption 4.1 (Robust recovery)**.** *Assume that $G : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ is Borel measurable and bounded from below, and that $-\infty < \mathcal{V}_0 := \inf_{\mu\in\mathcal{P}_2(\mathcal{H})} G(\mu) < \infty$. For every $\eta > 0$ and every $K_0, N_0 \geq 1$, assume that there exist $K \geq K_0$, $N \geq N_0$, a cloud $\mathbf{z} \in (\mathcal{H}_K)^N$, and $r > 0$ such that $\sup_{\mathbf{y}\in B_r(\mathbf{z})} G(\mu^N_{\mathbf{y}}) \leq \mathcal{V}_0 + \eta$.*

Theorem 4.1 (PDE-free convergence under robust recovery)**.** *Assume that Q is trace class with strictly positive eigenvalues and that Assumption 4.1 holds. Let $v_\varepsilon^{K,N}$ be the projected particle value with terminal cost $NG_K(\mu^N_{\mathbf{x}})$. Then there exist sequences $K_m, N_m \to \infty$ and $\varepsilon_m \downarrow 0$ such that*

$$\frac{1}{N_m} v_{\varepsilon_m}^{K_m,N_m}(0, \Pi_{K_m}\mathbf{x}_0) \longrightarrow \mathcal{V}_0.$$

Upper semicontinuous objectives satisfy the recovery condition by density of empirical measures in Galerkin subspaces; this includes indicator-type objectives with closed bad-event sets and open good-event sets. The precise corollaries and proofs are given in Appendix B.1.

4.2 Projection-first quantitative bound

Recall from the Feynman–Kac representation Equation (6) that $v_\varepsilon^{K,N}(t,\mathbf{x}) = -\varepsilon\log\mathbb{E}\left[\exp\left(-\frac{N}{\varepsilon}G_K(\mu^N_{\mathbf{x}+\Delta\mathbf{W}^K_{1-t}})\right)\right]$. This is the idealized projected particle value studied below. For $\nu \in \mathcal{P}_2(\mathcal{H}_K)$, define the projected mean-field regularized value

$$V_\varepsilon^K(t,\nu) := \inf_{\theta\in\Theta_K}\left\{G_K(\mathcal{L}_{\mathcal{H}_K}(X_1^{K,t,\xi;\theta})) + \mathbb{E}\left[\frac{\varepsilon}{2}\int_t^1 \|\theta_s\|^2_{Q_K}\,ds\right]\right\},$$

where $\mathcal{L}_{\mathcal{H}_K}(\xi) = \nu$ and $X_s^{K,t,\xi;\theta} = \xi + \int_t^s \theta_r\,dr + W_s^K - W_t^K$.

The projected Master equation, verification statement, and regularity assumptions for V_ε^K are given in Appendix B.2.

The proof first compares $N^{-1}v_\varepsilon^{K,N}$ with the projected mean-field value V_ε^K, and then compares $V_\varepsilon^K(0,\delta_{\Pi_K x_0})$ with the original optimum $\mathcal{V}_0$, thus obtaining the difference between $N^{-1}v_\varepsilon^{K,N}$ and $\mathcal{V}_0$.

Assume that the infimum $\mathcal{V}_0$ is attained by some $\mu^* \in \mathcal{P}_2(\mathcal{H})$. For $X \sim \mu^\star$, define $\Delta_K^{\mathrm{proj}} := \mathbb{E}_{\mu^\star}\|(I-\Pi_K)X\|^2$, $D_{CM,K}^2(\mu^\star) := \sum_{j=1}^K \frac{\mathbb{E}_{\mu^\star}[(X_j - x_{0,j})^2]}{\lambda_j}$. We assume the projected regularized value V_ε^K is a verified classical solution of the projected Master equation with the derivative and residual bounds stated in Appendix B.2. In particular, the residual constant $L_{\varepsilon,K}$ satisfies

$$\left| \int_{\mathcal{H}_K} \mathrm{Tr}\big(Q_K \partial_{\mu\mu}^2 V_\varepsilon^K(t,\nu)(y)(y)\big)\, \nu(dy) \right| \le L_{\varepsilon,K}.$$

We also assume that G has a local Hölder continuity condition of exponent $\alpha \in (0,1]$ around a minimizer $\mu^\star$, and the projected particle value $v_\varepsilon^{K,N}$ is sufficiently regular for the comparison argument: its spatial gradients are locally Lipschitz on compact time intervals in $[0,1)$ and have at most linear growth. The precise statement is given in Appendix B.2.

Theorem 4.2 (Projection-first convergence of the idealized particle approximation)**.** *Assume $0 < \varepsilon < 1$. Under the assumptions above, let $v_\varepsilon^{K,N}$ be a classical solution of the projected particle HJB. Then*

$$0 \le \frac{1}{N} v_\varepsilon^{K,N}(0,\Pi_K \mathbf{x}_0) - \mathcal{V}_0 \le \frac{L_{\varepsilon,K}}{2N} + C_* \left(\varepsilon^{2/\alpha}\,\mathrm{Tr}(Q) + \Delta_K^{\mathrm{proj}}\right)^{\alpha/2} + \frac{\varepsilon}{2} D_{CM,K}^2(\mu^\star) + \frac{\varepsilon K}{\alpha} \log\frac{1}{\varepsilon}. \tag{11}$$

Here $\Pi_K \mathbf{x}_0 = (\Pi_K x_0, \dots, \Pi_K x_0)$, and C_* is independent of K, N, ε.

The proof decomposes the error into a particle approximation term $L_{\varepsilon,K}/(2N)$ and a projected regularization error; details are given in Appendix B.3. The bound tends to zero along any joint choice $K = K_m$, $N = N_m$, and $\varepsilon = \varepsilon_m \downarrow 0$ such that

$$\frac{L_{\varepsilon_m,K_m}}{N_m} \to 0, \qquad \Delta_{K_m}^{\mathrm{proj}} \to 0, \qquad \varepsilon_m D_{CM,K_m}^2(\mu^\star) \to 0, \qquad \varepsilon_m K_m \log\frac{1}{\varepsilon_m} \to 0.$$

Appendix B.4 gives one sufficient FBSDE-based condition under which $L_{\varepsilon,K}$ is bounded uniformly in (ε, K).

4.3 Finite-candidate implementation with nonmatched covariance

We now connect the practical SCMO update with nonmatched execution covariance to the matched projected particle value. A single outer-loop rollout consists of M Euler updates. Since each outer loop applies the same M-step transition to its current cloud, we analyze one representative rollout ($L = 1$). The results below are stated on the normalized horizon $[0,1]$; a general horizon follows by time rescaling. Let $Q_{\mathrm{prop}} = \sigma_{\mathrm{prop}}^2 Q$, $Q_{\mathrm{dyn}} = \sigma_{\mathrm{dyn}}^2 Q$, $\alpha_0 = \frac{\sigma_{\mathrm{dyn}}^2}{\sigma_{\mathrm{prop}}^2}$, so that $Q_{\mathrm{dyn}} = \alpha_0 Q_{\mathrm{prop}}$. Their projected N-particle extensions are $\mathbf{Q}_{K,\mathrm{prop}} = I_N \otimes \Pi_K Q_{\mathrm{prop}} \Pi_K$, $\mathbf{Q}_{K,\mathrm{dyn}} = \alpha_0 \mathbf{Q}_{K,\mathrm{prop}}$. Throughout this subsection, $v_\varepsilon^{K,N}$ and b denote the matched projected particle value and feedback associated with $\mathbf{Q}_{K,\mathrm{prop}}$.

For fixed $t < 1$ and $\mathbf{x} \in (H_K)^N$, let $\widehat{b}_i^{R,S}(t,\mathbf{x})$ be the globally normalized R-context, S-candidate estimator used by SCMO (see Equation (9)). Under the integrability conditions stated in Appendix C.1, for every fixed $S \ge 1$,

$$\widehat{b}_i^{R,S}(t,\mathbf{x}) \longrightarrow b_i(t,\mathbf{x}) \qquad \text{almost surely as } R \to \infty.$$

If the corresponding fourth moments and inverse-denominator moments are uniformly bounded, then

$$\mathbb{E}\left[\left\|\widehat{b}_i^{R,S}(t,\mathbf{x}) - b_i(t,\mathbf{x})\right\|^2\right] \le C\left(\frac{1}{R} + \frac{1}{RS}\right). \tag{12}$$

A controlled interacting-quadratic diagnostic in Appendix A.4 shows that the empirical feedback mean-square error decays at the predicted rate; see Figure 1.

We next define the exact-feedback nonmatched-covariance chain to which the fully finite implementation converges. Let $h = \frac{1}{M}$, $t_m = mh$, and set

$$\mathbf{X}^h_{m+1} = \mathbf{X}^h_m + h\, b(t_m, \mathbf{X}^h_m) + \sqrt{h}\, \mathbf{Q}^{1/2}_{K,\mathrm{dyn}} Z_{m+1}, \qquad \mathbf{X}^h_0 = \Pi_K \mathbf{x}_0, \tag{13}$$

where $Z_{m+1} \sim \mathcal{N}(0, I_{NK})$ are independent.

For a deterministic state $\mathbf{x}$, write $\overline{\mathbf{x}}_h = \mathbf{x} + h\, b(t, \mathbf{x})$, and define the matched and execution continuation scores

$$\Phi^{v,\bullet}_h(t, \mathbf{x}) = \mathbb{E}\left[v^{K,N}_\varepsilon \left(t + h, \overline{\mathbf{x}}_h + \sqrt{h}\, \mathbf{Q}^{1/2}_{K,\bullet} Z \right) \right], \qquad \bullet \in \{\mathrm{prop}, \mathrm{dyn}\}.$$

The signed one-step covariance contribution is $c^{K,N,\varepsilon}_h(t, \mathbf{x}) = \Phi^{v,\mathrm{dyn}}_h(t, \mathbf{x}) - \Phi^{v,\mathrm{prop}}_h(t, \mathbf{x})$. Its favorable and unfavorable parts are $g^{K,N,\varepsilon}_h = \left(-c^{K,N,\varepsilon}_h\right)_+$, $p^{K,N,\varepsilon}_h = \left(c^{K,N,\varepsilon}_h\right)_+$. Let $\ell^{K,N}_\varepsilon(t, \mathbf{x}) = \frac{\varepsilon}{2} \|b(t, \mathbf{x})\|^2_{\mathbf{Q}^{-1}_{K,\mathrm{prop}}}$, and define the matched weak Euler residual

$$r^{K,N,\varepsilon}_h(t, \mathbf{x}) = \Phi^{v,\mathrm{prop}}_h(t, \mathbf{x}) + h\ell^{K,N}_\varepsilon(t, \mathbf{x}) - v^{K,N}_\varepsilon(t, \mathbf{x}).$$

Along the chain Equation (13), set

$$\mathcal{E}^{K,N,\varepsilon}_h = \frac{1}{N} \sum_{m=0}^{M-1} \mathbb{E}\left[\left(r^{K,N,\varepsilon}_h(t_m, \mathbf{X}^h_m) \right)_+ \right], \quad \mathcal{G}^{K,N,\varepsilon}_h = \frac{1}{N} \sum_{m=0}^{M-1} \mathbb{E}\left[g^{K,N,\varepsilon}_h(t_m, \mathbf{X}^h_m) \right], \tag{14}$$

$$\mathcal{P}^{K,N,\varepsilon}_h = \frac{1}{N} \sum_{m=0}^{M-1} \mathbb{E}\left[p^{K,N,\varepsilon}_h(t_m, \mathbf{X}^h_m) \right]. \tag{15}$$

Finally, let $\mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon}$ denote the right-hand side of Equation (35), with Q replaced by Q_{prop} and the Cameron–Martin term computed using the eigenvalues of Q_{prop}.

Theorem 4.3 (One-rollout implementation bound)**.** *Assume the hypotheses of Theorem B.7, with Q_{prop} as the covariance of the matched control problem, together with the stability and integrability conditions in Appendix C.1. Then the exact feedback nonmatched-covariance chain satisfies*

$$\mathbb{E} G_K(\mu^N_{\mathbf{X}^h_M}) - \mathcal{V}_0 \le \mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon} + \mathcal{E}^{K,N,\varepsilon}_h - \mathcal{G}^{K,N,\varepsilon}_h + \mathcal{P}^{K,N,\varepsilon}_h. \tag{16}$$

Let $\mathbf{C}^{R,S,h}_m$ be the fully finite SCMO chain using the same execution noises as $\mathbf{X}^h_m$, but replacing b by $\widehat{b}^{R,S}$. Under the quantitative stability assumptions, there is a constant $C_{\mathrm{stab}} < \infty$, independent of R, S, such that

$$\max_{0 \le m \le M} \mathbb{E}\left[\left\| \mathbf{C}^{R,S,h}_m - \mathbf{X}^h_m \right\|^2 \right] \le C_{\mathrm{stab}} \left(\frac{1}{R} + \frac{1}{RS} \right). \tag{17}$$

Consequently,

$$\limsup_{\substack{R \to \infty \\ RS \to \infty}} \left[\mathbb{E} G_K(\mu^N_{\mathbf{C}^{R,S,h}_M}) - \mathcal{V}_0 \right] \le \mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon} + \mathcal{E}^{K,N,\varepsilon}_h - \mathcal{G}^{K,N,\varepsilon}_h + \mathcal{P}^{K,N,\varepsilon}_h. \tag{18}$$

The term $\mathcal{E}^{K,N,\varepsilon}_h$ is the time-discretization contribution. Appendix C.1 gives the identity

$$r^{K,N,\varepsilon}_h(t, \mathbf{x}) = \frac{\varepsilon}{2} \mathbb{E} \int_0^h \left\| b(t, \mathbf{x}) - b(t + s, \mathbf{Y}^{t,\mathbf{x}}_s) \right\|^2_{\mathbf{Q}^{-1}_{K,\mathrm{prop}}} ds,$$

so it is nonnegative and vanishes as $h \downarrow 0$ under suitable Hölder–Lipschitz regularity of the feedback.

The covariance contribution is signed rather than being treated solely as an error. In the smooth case,

$$c^{K,N,\varepsilon}_h(t, \mathbf{x}) = \frac{h}{2} \operatorname{Tr}\left([\mathbf{Q}_{K,\mathrm{dyn}} - \mathbf{Q}_{K,\mathrm{prop}}]\, D^2 v^{K,N}_\varepsilon(t + h, \overline{\mathbf{x}}_h) \right) + o(h).$$

For proportional covariances, define $\kappa_v(t, \mathbf{x}) = \operatorname{Tr}\left(\mathbf{Q}_{K,\mathrm{prop}} D^2 v^{K,N}_\varepsilon(t, \mathbf{x}) \right)$. Then the leading contribution is $\frac{h}{2}(\alpha_0 - 1)\kappa_v$. Thus a colder execution covariance, $\alpha_0 < 1$, is favorable in regions of positive covariance-weighted continuation-value curvature, whereas a hotter covariance, $\alpha_0 > 1$, is favorable in regions of negative weighted curvature. Accordingly, the signed covariance contribution in Equation (16) is favorable whenever $\mathcal{G}^{K,N,\varepsilon}_h > \mathcal{P}^{K,N,\varepsilon}_h$. If, more strongly, $\mathcal{G}^{K,N,\varepsilon}_h > \mathcal{E}^{K,N,\varepsilon}_h + \mathcal{P}^{K,N,\varepsilon}_h$, then the complete one-rollout upper bound is strictly smaller than the matched particle–projection–regularization certificate $\mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon}$. Full proofs, continuous-time analogues, and checkable bounded and polynomial-growth conditions are given in Appendix C.1.

5 EXPERIMENTS

We evaluate SCMO on two structured benchmark suites and one external simulator environment. Across these experiments, we test multimodal-law recovery, escape from suboptimal basins, and optimization under nonsmooth or discontinuous law-level objectives. The main study comprises function-space optimization with PDE energies, open-loop trajectory-law optimization, and a fixed-reset Push-T simulator experiment. Complementary low-dimensional checks isolate these capabilities in controlled settings, including progress from a flat nonsmooth plateau. Complete results and figures for these checks are provided in Appendix D; additional studies examine the effect of R, proposal–execution covariance nonmatch, and latent image generation in Appendices A.4, A.2, and E, respectively.

Protocol and baselines. Within each task, empirical-law methods share the same initial cloud, and all methods use the corresponding common initialization and paired evaluation seeds. Method-specific hyperparameters are selected on separate development runs and frozen before evaluation. Derivative-free methods are matched by maximum primitive objective work: particle-energy evaluations in the PDE suite, trajectory simulations in the passage suite, and simulator calls in Push-T. More details are reported in Appendix D.

We use benchmark-specific subsets of complementary controls: Point Adam and Particle Adam (Kingma & Ba, 2015); a diagonal-Gaussian family optimized using reparameterized Monte Carlo gradients (Kingma & Welling, 2014); whole-cloud CEM (de Boer et al., 2005) and CMA-ES (Hansen & Ostermeier, 2001); greedy whole-cloud random search (Solis & Wets, 1981); and an MPPI-style open-loop optimizer (Williams et al., 2017a). Point Adam, Particle Adam, CEM, and CMA-ES use the standard update rules of their cited optimizers on the stated decision representations. The diagonal-Gaussian, whole-cloud random-search, and MPPI-style empirical-law controls are defined in Appendix D.

Function-space PDE laws. Each empirical law comprises $N = 256$ functions in $L^2([0, 1])$, represented using $K = 32$ Neumann cosine coefficients. The five tasks test smooth balanced two-phase Allen–Cahn recovery (P1); a nonsmooth, nonconvex TV/double-well objective with an absolute-moment penalty (P2); discontinuous interacting three-phase quotas (P3); escape from the shallow phase of a tilted double well to its deeper global phase (P4); and escape from a metastable central phase while balancing mass between two outer global phases (P5).

As shown in Table 1, SCMO attains the lowest mean objective on P1–P3 and is the only method to satisfy the P3 quotas in all 20 runs. SCMO and the diagonal-Gaussian baseline both escape the P4 local phase in every run, with the Gaussian baseline attaining the lower objective. On P5, only SCMO consistently recovers the balanced outer-phase law. Although Particle Adam attains a lower P5 objective, it fails the balanced-law criterion in every run, showing that objective value alone does not certify recovery of the intended law structure.

Trajectory laws. Each empirical law comprises $N = 128$ open-loop, 42-step trajectories. The five tasks test broad single-passage feasibility (T1); narrow local feasibility (T2); continuous two-route mass recovery (T3); discontinuous, interacting route quotas (T4); and escape from a feasible but inferior lower-route basin to a better upper route under the ordinary trajectory objective (T5).

As shown in Table 1, all methods solve T1, whereas SCMO remains effective on the narrow T2 task. SCMO is the only method to satisfy the T3 full-success and two-route criterion in all 20 runs. On T4, SCMO succeeds in $20/20$ runs, compared with $14/20$ for Mixture MPPI and none for the remaining baselines. For T5, SCMO and Mixture MPPI discover the better upper route in every run, whereas random search and CEM remain in the lower-route basin; SCMO attains the lowest mean objective.

Fixed-reset Push-T simulator check. We optimize empirical laws of $N = 8$ open-loop action sequences under the standard contact-rich Push-T dynamics (Chi et al., 2023). The physical task is to control the pushing agent so that its contacts move and rotate the T-shaped block into a prescribed target pose. This is a fixed-reset open-loop optimization study, not policy learning or the canonical state-distribution generalization protocol. We evaluate ten held-out optimization seeds under five

Table 1: Main results. Panels A–B report mean ± standard deviation and structural successes over 20 paired seeds; Panel C reports ten seed-level averages across five Push-T conditions, with pooled counts shown descriptively.

A. Function-space PDE laws: objective [successes/20]

Task	SCMO	Particle Adam	CEM cloud	Gaussian
P1 Smooth Allen–Cahn	$\mathbf{0.128 \pm .003}$	$0.184 \pm .017$	$1.734 \pm .042$	$0.996 \pm .211$
P2 TV–Allen–Cahn	$\mathbf{0.984 \pm .028}$	$1.747 \pm .031$	$1.069 \pm .020$	$1.118 \pm .279$
P3 Three-phase quotas	$\mathbf{-2.541 \pm .635}$ [20]	$2.174 \pm .034$ [0]	$2.835 \pm .083$ [0]	$2.009 \pm .001$ [0]
P4 Tilted local/global	$0.00883 \pm .00008$ [20]	$3.846 \pm .157$ [0]	$5.654 \pm .007$ [0]	$\mathbf{0.00109 \pm .00010}$ [20]
P5 Metastable balanced law	$0.342 \pm .004$ [20]	$\mathbf{0.264 \pm .041}$ [0]	$5.263 \pm .131$ [0]	$5.271 \pm .098$ [0]

B. Trajectory laws: objective [successes/20]

Task	SCMO	Random search	CEM cloud	Mixture MPPI
T1 Easy passage	$0.886 \pm .002$ [20]	$\mathbf{0.846 \pm .000}$ [20]	$0.886 \pm .002$ [20]	$0.848 \pm .000$ [20]
T2 Narrow passage	$\mathbf{1.097 \pm .028}$ [20]	$7.900 \pm .002$ [0]	$7.853 \pm .000$ [0]	$3.188 \pm .416$ [20]
T3 Two-route multimodal	$\mathbf{0.080 \pm .006}$ [20]	$8.005 \pm .252$ [0]	$3.930 \pm .322$ [0]	$1.360 \pm .392$ [0]
T4 Discontinuous quotas	$\mathbf{-0.482 \pm .008}$ [20]	$8.448 \pm .128$ [0]	$6.628 \pm .149$ [0]	$1.073 \pm .777$ [14]
T5 Local/global routes	$\mathbf{2.715 \pm .000}$ [20]	$2.857 \pm .001$ [0]	$2.829 \pm .001$ [0]	$2.840 \pm .004$ [20]

C. Push-T external simulator anchor

Method	Objective ↓	p_s ↑	Quota runs	All-8 runs	Quota seeds (all 5)
SCMO	$\mathbf{0.021 \pm .004}$	$\mathbf{0.965 \pm .027}$	50/50	36/50	10/10
CEM cloud	$0.062 \pm .014$	$0.920 \pm .059$	50/50	27/50	10/10
Random-search cloud	$0.172 \pm .043$	$0.585 \pm .058$	45/50	0/50	6/10
MPPI-style cloud	$0.227 \pm .026$	$0.520 \pm .047$	28/50	0/50	0/10
Full CMA-ES cloud	$0.318 \pm .038$	$0.420 \pm .037$	6/50	0/50	0/10

deterministic scale perturbations $\gamma \in \{.85, .90, 1, 1.05, 1.10\}$ of one frozen reset, using exactly $4{,}408$ simulator calls per method.

The objective combines mean rollout loss with route-balance and minimum-route-mass penalties. Here, p_L and p_R are the fractions of all eight particles that succeed through the left and right routes; failed or unclassified rollouts contribute to neither. The quota $p_L, p_R \geq .25$ therefore requires at least two successful trajectories on each route. As shown in Table 1, SCMO attains the lowest seed-averaged objective, highest success mass, and closest route balance. SCMO and CEM satisfy the quota in all 50 condition–seed cells and under all five conditions for every seed, but SCMO achieves the lower objective and higher success mass; the remaining baselines satisfy the quota less consistently.

6 Conclusion

We introduced SCMO, an objective-only optimizer for empirical probability laws on separable Hilbert spaces. Combining particle–Galerkin approximation with a Cole–Hopf representation, SCMO estimates feedback through context-conditioned one-particle replacement scores. We established matched-covariance convergence, multi-context finite-candidate consistency, and a signed, curvature-dependent analysis of proposal–execution covariance nonmatch. Experiments on function-space PDE energies, open-loop trajectory laws, and fixed-reset Push-T manipulation show that SCMO handles nonsmooth, nonconvex, and discontinuous law-level objectives, escapes suboptimal basins, and recovers prescribed multimodal, non-Dirac law structure.

AI USE STATEMENT

The SCMO methodology, theoretical results, mathematical claims, proof strategies, and proofs presented in this work were developed by the authors and were not generated by generative AI. We used generative AI tools, including OpenAI ChatGPT and Codex, to provide feedback on experimental protocols, suggest experimental parameters, assist with implementing and debugging experimental and plotting code, and help organize and present numerical results. We also used these tools to search for potentially relevant literature, suggest references, restructure and shorten parts of the manuscript, improve the readability of the prose, tables, and captions, and provide critical comments on the exposition of mathematical arguments. Generative AI was not relied upon to formulate the proposed method, originate the paper's theoretical contributions, or provide essential ingredients for the proofs.

The authors reviewed all AI-assisted work. AI-assisted code was inspected and tested by executing the experiments; reported values, tables, and figures were checked against the saved experimental outputs; and suggested citations were verified against their original sources. All mathematical statements, derivations, and proofs were written and independently verified by the authors; AI-assisted comments were treated only as suggestions and were discarded unless substantiated by the authors. Generative AI was not used to generate or alter reported numerical results, translate the manuscript, clean or reformat datasets, or perform qualitative or thematic data analysis. Survey design, interviews, and transcription were not applicable to this work. The authors take responsibility for the final content of this work, including all text, code, claims, proofs, experiments, figures, and other artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work develops a general-purpose optimization method and evaluates it on mathematical benchmarks and simulated control environments. It does not involve human participants, personal or sensitive data, or deployment on physical robotic systems. We are not aware of any direct ethical risks beyond those generally associated with general-purpose optimization software. The authors declare no conflicts of interest.

REPRODUCIBILITY STATEMENT

We provide fully reproducible code in the anonymous supplementary material. It includes environment and dependency specifications, frozen hyperparameter configurations, evaluation seeds and initial conditions, scripts for running all PDE, trajectory-law, and Push-T experiments, and scripts for reproducing the reported tables and figures. Complete objective definitions, benchmark geometries, initialization procedures, evaluation criteria, computational-budget accounting, and task-specific hyperparameters are documented in Appendix D. The assumptions underlying the theoretical results are stated with the corresponding theorems, and complete proofs and auxiliary results are provided in the theoretical appendices. No proprietary datasets are required.

REFERENCES

José A. Carrillo, Young-Pil Choi, Claudia Totzeck, and Oliver Tse. An Analytical Framework for a Consensus-Based Global Optimization Method. *Mathematical Models and Methods in Applied Sciences*, 28(6):1037–1066, 2018. doi: 10.1142/S0218202518500276.

Pratik Chaudhari, Anna Choromanska, Stefano Soatto, Yann LeCun, Carlo Baldassi, Christian Borgs, Jennifer T. Chayes, Levent Sagun, and Riccardo Zecchina. Entropy-SGD: Biasing Gradient Descent Into Wide Valleys. In *International Conference on Learning Representations*, 2017. URL https://openreview.net/forum?id=B1YfAfcgl.

Fan Chen, Zhenjie Ren, and Songbo Wang. Entropic fictitious play for mean field optimization problem. *Journal of Machine Learning Research*, 24(211):1–36, 2023.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin C. M. Burchfiel, and Shuran Song. Diffusion Policy: Visuomotor Policy Learning via Action Diffusion. In *Proceedings*

of Robotics: Science and Systems, Daegu, Republic of Korea, July 2023. doi: 10.15607/RSS.2023.XIX.026.

Lénaïc Chizat and Francis Bach. On the Global Convergence of Gradient Descent for Over-parameterized Models Using Optimal Transport. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Pieter-Tjerk de Boer, Dirk P. Kroese, Shie Mannor, and Reuven Y. Rubinstein. A Tutorial on the Cross-Entropy Method. *Annals of Operations Research*, 134(1):19–67, 2005. doi: 10.1007/s10479-005-5724-z.

Haoran Ding, Noémie Jaquier, Jan Peters, and Leonel Rozo. Fast and Robust Visuomotor Riemannian Flow Matching Policy. *IEEE Transactions on Robotics*, 41:5327–5343, 2025. doi: 10.1109/TRO.2025.3601293.

Jhanvi Garg, Xianyang Zhang, and Quan Zhou. Soft-constrained Schrödinger Bridge: a stochastic control approach. In *International Conference on Artificial Intelligence and Statistics*, pp. 4429–4437. PMLR, 2024.

Nikolaus Hansen and Andreas Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001. doi: 10.1162/106365601750190398.

Kaitong Hu, Zhenjie Ren, David Šiška, and Łukasz Szpruch. Mean-Field Langevin Dynamics and Energy Landscape of Neural Networks. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(4):2043–2065, 2021. doi: 10.1214/20-AIHP1140.

Hui Huang and Hicham Kouhkouh. A Derivative-Free Particle Method for Optimization in Hilbert Spaces. *arXiv preprint arXiv:2605.31565*, 2026.

Gavin Kerrigan, Giosue Migliorini, and Padhraic Smyth. Functional Flow Matching. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, 2024.

Mahmoud Khatab and Claudia Totzeck. A consensus-based optimization algorithm using gaussian processes for global optimization problems in Sobolev spaces. *arXiv preprint arXiv:2603.15337*, 2026.

Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations*, 2015. URL https://arxiv.org/abs/1412.6980.

Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. In *International Conference on Learning Representations*, 2014. URL https://arxiv.org/abs/1312.6114.

Yunbum Kook, Matthew S. Zhang, Sinho Chewi, Murat A. Erdogdu, and Mufan Li. Sampling from the Mean-Field Stationary Distribution. In *Proceedings of the 37th Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 3099–3136. PMLR, 2024.

Zhiqi Li, Yuchen Sun, Greg Turk, and Bo Zhu. Functional Mean Flow in Hilbert Space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1928–1938, June 2026.

Qiang Liu and Dilin Wang. Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Hoang Phuc Hau Luu, Hanlin Yu, Bernardo Williams, Petrus Mikkola, Marcelo Hartmann, Kai Puolamäki, and Arto Klami. Non-geodesically-convex Optimization in the Wasserstein Space. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Atsushi Nitanda and Taiji Suzuki. Stochastic Particle Gradient Descent for Infinite Ensembles. *arXiv preprint arXiv:1712.05438*, 2017.

Caio Peixoto, Daniel Csillag, Bernardo F. P. da Costa, and Yuri F. Saporito. Random Gradient-Free Optimization in Infinite Dimensional Spaces. *arXiv preprint arXiv:2512.20566*, 2025.

René Pinnau, Claudia Totzeck, Oliver Tse, and Stephan Martin. A Consensus-Based Model for Global Optimization and Its Mean-Field Limit. *Mathematical Models and Methods in Applied Sciences*, 27(1):183–204, 2017. doi: 10.1142/S0218202517400061.

Fabrizio Pittorino, Carlo Lucibello, Christoph Feinauer, Gabriele Perugini, Carlo Baldassi, Elizaveta Demyanenko, and Riccardo Zecchina. Entropic Gradient Descent Algorithms and Wide Flat Minima. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=xjXg0bnoDmS.

Jinniao Qiu. Stochastic Control Methods for Optimization. *arXiv preprint arXiv:2601.01248*, 2026.

Francisco J. Solis and Roger J.-B. Wets. Minimization by Random Search Techniques. *Mathematics of Operations Research*, 6(1):19–30, 1981. doi: 10.1287/moor.6.1.19.

Chandan Tankala, Dheeraj M. Nagaraj, and Anant Raj. Beyond Propagation of Chaos: A Stochastic Algorithm for Mean Field Optimization. In *Proceedings of the 38th Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pp. 5410–5440. PMLR, 2025.

Evangelos A. Theodorou, Jonas Buchli, and Stefan Schaal. A Generalized Path Integral Control Approach to Reinforcement Learning. *Journal of Machine Learning Research*, 11(104):3137–3181, 2010.

Daan Wierstra, Tom Schaul, Tobias Glasmachers, Yi Sun, Jan Peters, and Jürgen Schmidhuber. Natural evolution strategies. *Journal of Machine Learning Research*, 15(27):949–980, 2014.

Grady Williams, Andrew Aldrich, and Evangelos A. Theodorou. Model Predictive Path Integral Control: From Theory to Parallel Computation. *Journal of Guidance, Control, and Dynamics*, 40 (2):344–357, 2017a. doi: 10.2514/1.G001921.

Grady Williams, Nolan Wagener, Brian Goldfain, Paul Drews, James M. Rehg, Byron Boots, and Evangelos A. Theodorou. Information Theoretic MPC for Model-Based Reinforcement Learning. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1714–1721, 2017b. doi: 10.1109/ICRA.2017.7989202.

Lantian Xu, Anna Korba, and Dejan Slepčev. Accurate Quantization of Measures via Interacting Particle-based Optimization. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 24576–24595. PMLR, 2022.

Rentian Yao, Linjun Huang, and Yun Yang. Minimizing Convex Functionals over Space of Probability Measures via KL Divergence Gradient Flow. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 2530–2538. PMLR, 2024.

A APPENDIX

A.1 WHY COLE–HOPF IS APPLIED AFTER PARTICLE APPROXIMATION

We explain why the Cole–Hopf transform is applied to the projected N-particle HJB, rather than directly to the mean-field Master equation. The formal infinite-dimensional Master equation has the form

$$\partial_t V_\varepsilon(t,\mu) + \frac{1}{2}\int_{\mathcal{H}} \mathrm{Tr}(Q\partial_x\partial_\mu V_\varepsilon(t,\mu)(x))\,\mu(dx) - \frac{1}{2\varepsilon}\int_{\mathcal{H}} \|Q^{1/2}\partial_\mu V_\varepsilon(t,\mu)(x)\|^2\mu(dx) = 0.$$

In finite-dimensional HJB equations, the Cole–Hopf transform works because the Laplacian of $u = \exp(-V/\varepsilon)$ produces a quadratic gradient term that cancels the Hamiltonian. This cancellation does not occur at the level of the idiosyncratic-noise Master equation.

Indeed, set $u(t,\mu) := \exp(-V_\varepsilon(t,\mu)/\varepsilon)$. Then $\partial_\mu u(t,\mu)(x) = -\frac{1}{\varepsilon}u(t,\mu)\partial_\mu V_\varepsilon(t,\mu)(x)$. Taking the spatial derivative in x gives $\partial_x\partial_\mu u(t,\mu)(x) = -\frac{1}{\varepsilon}u(t,\mu)\partial_x\partial_\mu V_\varepsilon(t,\mu)(x)$. There is no additional quadratic term, because $u(t,\mu)$ depends on the measure variable μ, but not on the microscopic spatial variable x. Consequently,

$$\partial_t u(t,\mu) + \frac{1}{2}\int_{\mathcal{H}} \mathrm{Tr}(Q\partial_x\partial_\mu u(t,\mu)(x))\,\mu(dx) = -\frac{u(t,\mu)}{2\varepsilon^2}\int_{\mathcal{H}} \|Q^{1/2}\partial_\mu V_\varepsilon(t,\mu)(x)\|^2\mu(dx),$$

so the nonlinear control penalty remains. Thus the idiosyncratic Master equation is not linearized by the Cole–Hopf transform.

This is the reason for passing first to the projected N-particle problem. In the particle HJB, the state variable is the finite-dimensional vector $\mathbf{x} = (x_1,\dots,x_N)$, and the second-order operator is the ordinary particle Laplacian $\frac{1}{2}\sum_{i=1}^N \mathrm{Tr}(Q_K D^2_{x_i})$. For $u_\varepsilon^{K,N} = \exp(-v_\varepsilon^{K,N}/\varepsilon)$, the finite-dimensional chain rule produces the missing quadratic gradient terms, which cancel the particle Hamiltonian. This yields a linear heat equation and hence the risk-sensitive Feynman–Kac representation used in Section 3.2.

A.2 A CURVATURE DIAGNOSTIC FOR NONMATCHED EXECUTION COVARIANCE

We give a controlled experiment that isolates the signed covariance term in Theorem C.2. Consider the one-dimensional additive measure objective $G(\mu) = \int_{\mathbb{R}} f(x)\,\mu(dx)$, $f(x) = \frac{1}{4}(x^2-1)^2$. For an empirical measure, the extensive terminal cost is $F_N(\mathbf{x}) = \sum_{i=1}^N f(x_i)$. Consequently, the matched particle value factorizes as

$$v_\varepsilon^N(t,\mathbf{x}) = \sum_{i=1}^N v_\varepsilon(t,x_i),$$

where

$$v_\varepsilon(t,x) = -\varepsilon\log\mathbb{E}\left[\exp\left(-\frac{f(x+\sqrt{(1-t)q_{\mathrm{prop}}}Z)}{\varepsilon}\right)\right], \qquad Z\sim\mathcal{N}(0,1).$$

The exact matched feedback and covariance-weighted curvature are therefore

$$b(t,x) = -\frac{q_{\mathrm{prop}}}{\varepsilon}\partial_x v_\varepsilon(t,x), \qquad \kappa_v(t,x) = q_{\mathrm{prop}}\partial_{xx}v_\varepsilon(t,x).$$

We evaluate these quantities by Gauss–Hermite quadrature. In particular, if $\pi_{t,x}$ denotes the Gibbs-tilted Gaussian proposal, i.e., $\pi_{t,x}(dy) := \dfrac{e^{-f(y)/\varepsilon}\,\mathcal{N}(x,(1-t)q_{\mathrm{prop}})(dy)}{\int_{\mathbb{R}} e^{-f(z)/\varepsilon}\,\mathcal{N}(x,(1-t)q_{\mathrm{prop}})(dz)}$, then $\partial_{xx}v_\varepsilon(t,x) = \mathbb{E}_{\pi_{t,x}}[f''(Y)] - \frac{1}{\varepsilon}\mathrm{Var}_{\pi_{t,x}}(f'(Y))$, so the continuation-value curvature is obtained without finite differences.

For a fixed variance ratio $\alpha_0 > 0$, define the exact one-step effect

$$\begin{aligned} c_h(\alpha_0;t,x) = {} & \mathbb{E}\left[v_\varepsilon\left(t+h,\overline{x}_h+\sqrt{h\alpha_0 q_{\mathrm{prop}}}Z\right)\right] \\ & - \mathbb{E}\left[v_\varepsilon\left(t+h,\overline{x}_h+\sqrt{hq_{\mathrm{prop}}}Z\right)\right], \qquad \overline{x}_h = x + hb(t,x). \end{aligned}$$

The smooth theory in Appendix C predicts

$$c_h(\alpha_0; t, x) = \frac{h}{2}(\alpha_0 - 1)\kappa_v(t + h, \overline{x}_h) + O(h^2).$$

Because

$$f''(x) = 3x^2 - 1,$$

the terminal objective is locally concave near the central barrier and locally convex near the two wells. The continuation value preserves this sign separation over the parameter regime used below. Hence hot execution is predicted to be favorable near the barrier, whereas cold execution is predicted to be favorable near a well.

We use $\varepsilon = 0.2$, $q_{\mathrm{prop}} = 0.25$, $N = 256$, $M = 48$, and compare $\alpha_0 \in \{0.25, 0.5, 1, 1.5, 2\}$. For each value of α_0, we simulate the exact discrete chain

$$X^i_{m+1} = X^i_m + hb(t_m, X^i_m) + \sqrt{h\alpha_0 q_{\mathrm{prop}}}\, Z^i_{m+1}.$$

We consider two initial clouds: a collapsed cloud at the central barrier, $X^i_0 = 0$, and a collapsed cloud at the right well, $X^i_0 = 1$. The experiment uses 50 independent runs and common Gaussian execution noises across the different values of α_0 within each run.

At every step, we evaluate the exact one-step covariance effect and form

$$\mathcal{G}_h = \frac{1}{N}\sum_m \mathbb{E}[(-c_h)_+], \qquad \mathcal{P}_h = \frac{1}{N}\sum_m \mathbb{E}[(c_h)_+],$$

as well as the matched Euler residual $\mathcal{E}_h$ defined in Equation (14). We report the signed covariance contribution $\mathcal{C}_h = -\mathcal{G}_h + \mathcal{P}_h$ and the complete implementation contribution $\mathcal{E}_h - \mathcal{G}_h + \mathcal{P}_h$.

Thus the experiment separately tests whether curvature makes the covariance term favorable and whether this gain is large enough to dominate the Euler residual.

Table 2: Curvature diagnostic for nonmatched execution covariance over 50 independent runs. Negative $\mathcal{C} = -\mathcal{G} + \mathcal{P}$ is favorable; negative $\mathcal{E} + \mathcal{C}$ means that the gain also dominates the Euler residual.

Initialization	α_0	Final G	$\mathcal{G}$	$\mathcal{P}$	$\mathcal{C}$	$\mathcal{E} + \mathcal{C}$	Sign agr.
barrier	0.25	0.1896 ± 0.003749	0.001428	0.05052	0.04909	0.04967	0.999
	0.50	0.1619 ± 0.004897	0.00355	0.02788	0.02433	0.02485	0.999
	1.00	0.1359 ± 0.005757	0	0	0	0.0005928	–
	1.50	0.1269 ± 0.006398	0.0192	0.01319	-0.006011	-0.005289	0.997
	2.00	0.1267 ± 0.007468	0.03442	0.03374	-0.00068	0.0001851	0.996
well	0.25	0.02055 ± 0.001712	0.08137	3.023e-05	-0.08134	-0.07982	1.000
	0.50	0.03843 ± 0.003048	0.05316	0.0002451	-0.05291	-0.05134	1.000
	1.00	0.0705 ± 0.005405	0	0	0	0.001691	–
	1.50	0.09788 ± 0.008226	0.002241	0.05093	0.04869	0.0505	0.999
	2.00	0.1221 ± 0.01152	0.006136	0.1011	0.09495	0.09687	0.999

Table 2 confirms the predicted curvature-dependent reversal. From the barrier initialization, where continuation-value curvature is predominantly negative, hot execution ($\alpha_0 > 1$) produces a favorable negative covariance contribution; at $\alpha_0 = 1.5$, this gain also dominates the Euler residual. From the well initialization, where curvature is predominantly positive, cold execution ($\alpha_0 < 1$) is favorable and substantially lowers the final objective, whereas hot execution is unfavorable. The sign prediction agrees with the exact one-step effect in at least 99.6% of evaluated cases.

A.3 FULL SCMO PSEUDOCODE

Algorithm 1 Multi-context SCMO with nonmatched covariance

Input: Galerkin dimension K; Galerkin objective G_K; particle number N; horizon T; inner steps M; outer loops L; context clouds R; candidates per context S; temperature $\varepsilon > 0$; proposal scale $\sigma_{\mathrm{prop}} > 0$; execution scale $\sigma_{\mathrm{dyn}} > 0$; and

$$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_K),$$

with

$$Q_{\mathrm{prop}} = \sigma_{\mathrm{prop}}^2 \Lambda, \qquad Q_{\mathrm{dyn}} = \sigma_{\mathrm{dyn}}^2 \Lambda.$$

Initialize

$$\mathbf{C} = (C_1, \dots, C_N)^\top \in \mathbb{R}^{N \times K}, \qquad \Delta t = T/M.$$

for $\ell = 1, \dots, L$ **do**
  for $m = 0, \dots, M-1$ **do**
    Set

$$t_m = m\Delta t, \qquad \tau_m = T - t_m.$$

    Draw

$$\mathbf{Z}^{(r)} \sim \mathcal{N}(0, I_{N \times K}), \qquad r = 1, \dots, R,$$

    and form the complete context clouds

$$\widetilde{\mathbf{C}}^{(r)} = \mathbf{C} + \sqrt{\tau_m}\, \sigma_{\mathrm{prop}} \mathbf{Z}^{(r)} \Lambda^{1/2}.$$

    for $i = 1, \dots, N$ **do**
      Draw

$$\xi_{i,r,s} \sim \mathcal{N}(0, I_K), \qquad r = 1, \dots, R, \quad s = 1, \dots, S,$$

      and set

$$Y_{i,r,s} = C_i + \sqrt{\tau_m}\, \sigma_{\mathrm{prop}} \xi_{i,r,s} \Lambda^{1/2}.$$

      Compute the complete-cloud replacement scores

$$\mathcal{E}_{i,r,s} = N\, G_K \left(\frac{1}{N} \delta_{Y_{i,r,s}} + \frac{1}{N} \sum_{p \neq i} \delta_{\widetilde{C}_p^{(r)}} \right).$$

      Let

$$\mathcal{E}_i^{\min} = \min_{\substack{1 \le r \le R \\ 1 \le s \le S}} \mathcal{E}_{i,r,s},$$

      and define the globally normalized Gibbs weights

$$w_{i,r,s} = \frac{\exp\bigl(-(\mathcal{E}_{i,r,s} - \mathcal{E}_i^{\min})/\varepsilon\bigr)}{\displaystyle\sum_{q=1}^{R} \sum_{j=1}^{S} \exp\Bigl(-(\mathcal{E}_{i,q,j} - \mathcal{E}_i^{\min})/\varepsilon\Bigr)}.$$

      Estimate the feedback drift

$$\widehat{\theta}_i = \frac{1}{\tau_m} \sum_{r=1}^{R} \sum_{s=1}^{S} w_{i,r,s} (Y_{i,r,s} - C_i).$$

    end for
    Form the simultaneous drift update

$$\overline{C}_i = C_i + \Delta t\, \widehat{\theta}_i, \qquad i = 1, \dots, N,$$

    and write

$$\overline{\mathbf{C}} = (\overline{C}_1, \dots, \overline{C}_N)^\top.$$

    Draw an execution noise

$$\boldsymbol{\eta}_m \sim \mathcal{N}(0, I_{N \times K}),$$

    independently of all context and candidate noises, and update

$$\mathbf{C} \leftarrow \overline{\mathbf{C}} + \sqrt{\Delta t}\, \sigma_{\mathrm{dyn}} \boldsymbol{\eta}_m \Lambda^{1/2}.$$

  end for
end for
Output:

$$\mu^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{u^{(i)}}, \qquad u^{(i)} = \sum_{k=1}^{K} C_{i,k} e_k.$$

A.4 A DIAGNOSTIC FOR THE NUMBER OF CONTEXT CLOUDS

We illustrate the error induced by a single complete context cloud using an interacting quadratic measure objective. Let $\kappa > 0$, $G(\mu_{\mathbf{x}}^N) = \frac{\kappa}{2}(\frac{1}{N}\sum_{i=1}^N x_i - a)^2$, and define the scaled terminal cost $F_N(\mathbf{x}) = NG(\mu_{\mathbf{x}}^N) = \frac{\kappa}{2N}(\sum_{i=1}^N x_i - Na)^2$. Unlike an additive empirical average, this objective couples every particle through the empirical mean.

Let $\tau, q > 0$, under the Gaussian reference proposal $Y_i = x_i + \sqrt{\tau q}\, Z_i$, $Z_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, the exact full-cloud Gibbs feedback is

$$b_i(\mathbf{x}) = \frac{1}{\tau}\frac{\mathbb{E}[(Y_i - x_i)e^{-F_N(\mathbf{Y})/\varepsilon}]}{\mathbb{E}[e^{-F_N(\mathbf{Y})/\varepsilon}]} = \frac{\kappa q}{\varepsilon + \kappa q\tau}\,(a - \overline{x})\,,$$

where $\overline{x} = \frac{1}{N}\sum_{i=1}^N x_i$. Consider now the population limit of the one-context estimator. Fix one Gaussian realization $\widetilde{Y}_j = x_j + \sqrt{\tau q}\, Z_j$, $j \neq i$, for the remaining particles and let the number of candidates tend to infinity. The resulting conditional feedback is $b_i^{(1)}\left(\mathbf{x}; \widetilde{\mathbf{Y}}_{-i}\right) = \frac{\kappa q}{N\varepsilon + \kappa q\tau}\left(Na - x_i - \sum_{j\neq i}\widetilde{Y}_j\right)$. Thus, even when the candidate expectation is evaluated exactly, the feedback remains dependent on one random context. Taking expectation over this context gives $\mathbb{E}[b_i^{(1)}] = \frac{N\kappa q}{N\varepsilon + \kappa q\tau}\,(a - \overline{x})$. Consequently, $\frac{\mathbb{E}[b_i^{(1)}]}{b_i} = \frac{N(\varepsilon + \kappa q\tau)}{N\varepsilon + \kappa q\tau}$, which approaches N as $\varepsilon \downarrow 0$. Increasing the number of candidates S therefore does not remove the one-context error; it only approximates the random conditional feedback more accurately.

For R complete contexts and S candidates per context, define $L_{i,r,s} = \exp\left(-\frac{\mathcal{E}_{i,r,s}}{\varepsilon}\right)$, $U_{i,r,s} = Y_{i,r,s} - x_i$, with $\mathcal{E}_{i,r,s}$ as defined in the Equation (9). The globally normalized estimator can be written as $\widehat{b}_i^{R,S} = \frac{1}{\tau}\frac{\sum_{r=1}^R\sum_{s=1}^S L_{i,r,s}U_{i,r,s}}{\sum_{r=1}^R\sum_{s=1}^S L_{i,r,s}}$. For each context, let $A_{i,r}^S = \frac{1}{S}\sum_{s=1}^S L_{i,r,s}U_{i,r,s}$, $B_{i,r}^S = \frac{1}{S}\sum_{s=1}^S L_{i,r,s}$. The pairs $(A_{i,r}^S, B_{i,r}^S)$ are independent and identically distributed over r, with expectations equal to the numerator and denominator of the full-cloud feedback. Consequently, for every fixed $S \geq 1$, $\widehat{b}_i^{R,S} \longrightarrow b_i$, almost surely as $R \to \infty$. This consistency requires the normalization to be performed jointly over all RS context–candidate pairs.

We test this effect using $N = 8$, $\kappa = q = \tau = 1$, $\varepsilon = 0.1$, $a = 0$, with initial empirical mean $\overline{x}_0 = 0.5$. The exact initial feedback is $b_i(\mathbf{x}_0) = -0.454545$, whereas the expected population one-context feedback is $\mathbb{E}[b_i^{(1)}(\mathbf{x}_0)] = -2.222222$. The total score budget is fixed at $RS = 256$, and the experiment is repeated over 50 independent runs for 20 updates. The initial objective is 0.125.

Table 3: Interacting-quadratic diagnostic with fixed context–candidate budget $RS = 256$. Final objectives are mean $\pm$ standard deviation.

R	S	Mean feedback estimate	Feedback RMSE	Final objective	$\Pr(G \leq 10^{-2})$
1	256	-1.7684	1.6139	0.18222 ± 0.17646	0.16
2	128	-1.3169	1.2310	0.10094 ± 0.12844	0.22
4	64	-0.8812	0.7917	0.02242 ± 0.03610	0.60
8	32	-0.6242	0.4493	0.00522 ± 0.00774	0.86
16	16	-0.5148	0.2442	0.00209 ± 0.00295	0.96
32	8	-0.4816	0.1525	0.00141 ± 0.00198	1.00

The one-context method is worse on average than the initial cloud, despite using all 256 candidates in that single context. Reallocating the same score budget over several independent contexts substantially improves both the feedback approximation and the final optimization result. The choices $R = 4$ and $R = 8$ already remove most of the one-context instability while retaining respectively 64 and 32 candidates per context. This provides a computational rationale for using several complete context clouds rather than one context with a very large candidate pool.

Finite-candidate rate diagnostic. To complement the fixed-budget allocation study in Table 3, we directly examine the mean-square rate in Proposition C.1 using the same interacting-quadratic objective, for which the exact full-cloud feedback is available in closed form. We set

$$N = 8, \qquad \kappa = 0.02, \qquad q = \tau = 1, \qquad \varepsilon = 0.1, \qquad a = 0, \qquad \bar{x} = 0.5.$$

Thus,

$$b_i(x) = \frac{\kappa q}{\varepsilon + \kappa q \tau}(a - \bar{x}) = -\frac{1}{12}.$$

For each $S \in \{1, 4, 16, 64\}$, we vary $R \in \{8, 16, 32, 64, 128, 256, 512\}$ and estimate

$$\mathbb{E}\left[\left|\widehat{b}_i^{R,S}(x) - b_i(x)\right|^2\right]$$

over 10,000 independent repetitions. As shown in Figure 1, the empirical MSE decreases approximately proportionally to R^{-1} for every fixed S. The fitted log–log slopes over $R \geq 128$ are -0.991, -0.996, -1.005, and -0.997 for $S = 1, 4, 16, 64$, respectively, consistent with the $O(R^{-1} + (RS)^{-1})$ upper rate.

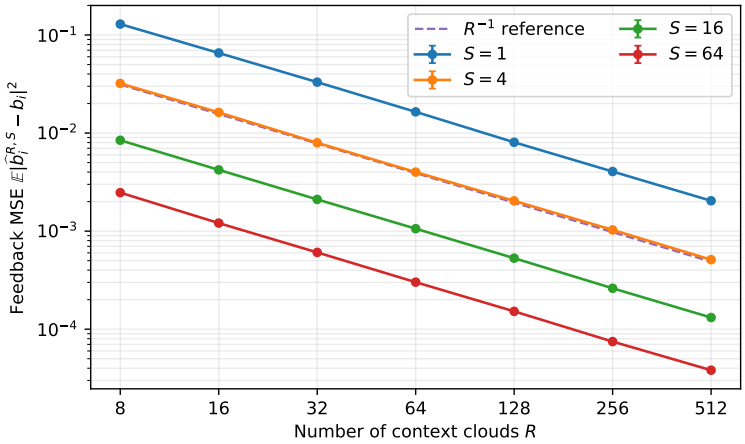


Figure 1: Finite-candidate rate diagnostic for the interacting-quadratic objective. Points show the empirical feedback MSE over 10,000 independent repetitions, with 95% Monte Carlo error bars. For each fixed candidate count S, the MSE follows the predicted R^{-1} decay. Increasing S lowers the within-context candidate contribution, while increasing R controls the context-cloud contribution.

B Convergence Analysis for Matched Covariance Problems

In this section, we give detailed convergence analysis on matched covariance problems, i.e., $Q_{prop} = Q_{dyn}$. Recall the settings in Section 3.2.

Algorithmic approximation. We provide two levels of convergence analysis. The first is a PDE-free qualitative result under weak recovery assumptions. It does not use the Master equation and can cover nonsmooth or discontinuous objectives. The second is a quantitative error decomposition for the idealized continuous-time projected particle value. This sharper result requires a verified classical regularized value function and analytic bounds. Neither result is a convergence theorem for the finite-candidate, time-discretized SCMO implementation.

B.1 Convergence at the Qualitative Level

The qualitative result uses only the finite-dimensional risk-sensitive representation and a robust recovery condition. It does not require a Master equation, Lions differentiability, or convexity.

Assumption B.1 (Robust recovery for possibly discontinuous objectives)**.** *Assume that* $G : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ *is Borel measurable and bounded from below, and that*

$$-\infty < \mathcal{V}_0 := \inf_{\mu \in \mathcal{P}_2(\mathcal{H})} G(\mu) < \infty.$$

For every $\eta > 0$ *and every* $K_0, N_0 \geq 1$*, assume that there exist integers* $K \geq K_0$*,* $N \geq N_0$*, a cloud* $\mathbf{z} \in (\mathcal{H}_K)^N$*, and a radius* $r > 0$ *such that*

$$\sup_{\mathbf{y} \in B_r(\mathbf{z})} G(\mu_{\mathbf{y}}^N) \leq \mathcal{V}_0 + \eta.$$

Theorem B.1 (PDE-free convergence under robust recovery)**.** *Assume that* Q *is trace class with strictly positive eigenvalues and that Assumption B.1 holds. Let* $v_\varepsilon^{K,N}$ *be the projected particle value with terminal cost* $NG_K(\mu_{\mathbf{x}}^N)$*. Then there exist sequences* $K_m, N_m \to \infty$ *and* $\varepsilon_m \downarrow 0$ *such that*

$$\frac{1}{N_m} v_{\varepsilon_m}^{K_m, N_m}(0, \Pi_{K_m} \mathbf{x}_0) \longrightarrow \mathcal{V}_0.$$

Proof. For fixed K, N, let $\gamma_{K,N}$ be the law of $\Pi_K \mathbf{x}_0 + \Delta \mathbf{W}_1^K$ on $(\mathcal{H}_K)^N$, and define

$$f_{K,N}(\mathbf{z}) := G_K(\mu_{\mathbf{z}}^N).$$

The finite-dimensional risk-sensitive representation gives

$$\frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) = -\frac{\varepsilon}{N} \log \int \exp\left(-\frac{N}{\varepsilon} f_{K,N}(\mathbf{z})\right) \gamma_{K,N}(d\mathbf{z}).$$

Because $G \geq \mathcal{V}_0$, the right-hand side is bounded below by $\mathcal{V}_0$.

Fix $\eta > 0$ and choose K_0, N_0. By robust recovery, choose $K \geq K_0$, $N \geq N_0$, $\mathbf{z} \in (\mathcal{H}_K)^N$, and $r > 0$ such that

$$G_K(\mu_{\mathbf{y}}^N) \leq \mathcal{V}_0 + \eta, \qquad \mathbf{y} \in B_r(\mathbf{z}).$$

Since $\gamma_{K,N}$ has full support on $(\mathcal{H}_K)^N$,

$$p := \gamma_{K,N}(B_r(\mathbf{z})) > 0.$$

Therefore

$$\int \exp\left(-\frac{N}{\varepsilon} f_{K,N}(\mathbf{y})\right) \gamma_{K,N}(d\mathbf{y}) \geq p \exp\left(-\frac{N}{\varepsilon}(\mathcal{V}_0 + \eta)\right).$$

It follows that

$$\frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) \leq \mathcal{V}_0 + \eta - \frac{\varepsilon}{N} \log p.$$

Choose $\varepsilon > 0$ small enough so that $-(\varepsilon/N) \log p \leq \eta$. Then

$$\mathcal{V}_0 \leq \frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) \leq \mathcal{V}_0 + 2\eta.$$

Taking $\eta = \eta_m \downarrow 0$ and $K_0 = N_0 = m$ at the m-th step gives the desired diagonal sequence. □

Corollary B.2 (Upper semicontinuous objectives)**.** *Assume that* $G : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ *is bounded from below and upper semicontinuous with respect to* $\mathcal{W}_2$*. Suppose* $-\infty < \mathcal{V}_0 < \infty$ *and that for every* $\eta > 0$ *there exists* $\mu_\eta \in \mathcal{P}_2(\mathcal{H})$ *with*

$$G(\mu_\eta) \leq \mathcal{V}_0 + \eta.$$

Then Assumption B.1 holds, and hence the conclusion of Theorem B.1 follows.

Proof. Fix $\eta > 0$ and $K_0, N_0 \geq 1$. Choose μ_η with $G(\mu_\eta) \leq \mathcal{V}_0 + \eta/3$. Empirical measures with atoms in finite-dimensional Galerkin subspaces are dense in $\mathcal{P}_2(\mathcal{H})$ under $\mathcal{W}_2$. Hence there exist $K \geq K_0$, $N \geq N_0$, and $\mathbf{z} \in (\mathcal{H}_K)^N$ such that $\mathcal{W}_2((\iota_K)_\# \mu_{\mathbf{z}}^N, \mu_\eta)$ is sufficiently small. By upper semicontinuity,

$$G_K(\mu_{\mathbf{z}}^N) = G((\iota_K)_\# \mu_{\mathbf{z}}^N) \leq G(\mu_\eta) + \eta/3 \leq \mathcal{V}_0 + 2\eta/3.$$

The map $\mathbf{y} \mapsto (\iota_K)_\# \mu_{\mathbf{y}}^N$ is continuous from $(\mathcal{H}_K)^N$ to $\mathcal{P}_2(\mathcal{H})$. Applying upper semicontinuity once more, there exists $r > 0$ such that

$$\sup_{\mathbf{y} \in B_r(\mathbf{z})} G_K(\mu_{\mathbf{y}}^N) \leq \mathcal{V}_0 + \eta.$$

This is robust recovery. □

Corollary B.3 (Indicator objectives with robust margins)**.** *Let $H : \mathcal{P}_2(\mathcal{H}) \to \mathbb{R}$ be bounded below and upper semicontinuous. Suppose*

$$G(\mu) = H(\mu) + \sum_{\ell=1}^{L_b} a_\ell \mathbf{1}_{\mathcal{B}_\ell}(\mu) - \sum_{m=1}^{L_g} b_m \mathbf{1}_{\mathcal{A}_m}(\mu), \qquad a_\ell, b_m \geq 0,$$

where each bad-event set $\mathcal{B}_\ell$ is closed and each good-event set $\mathcal{A}_m$ is open in $\mathcal{W}_2$. Then G is upper semicontinuous.

B.2 Convergence at the quantitative level

For any probability measures μ, we write the second moment of μ as $M_2(\mu) := \int \|x\|^2 \mu(dx)$. For $\nu \in \mathcal{P}_2(\mathcal{H}_K)$, define the projected regularized value

$$V_\varepsilon^K(t,\nu) := \inf_{\theta \in \Theta_K} \left\{ G_K(\mathcal{L}_{\mathcal{H}_K}(X_1^{K,t,\xi;\theta})) + \mathbb{E}\left[\frac{\varepsilon}{2} \int_t^1 \|\theta_s\|_{Q_K}^2 \, ds \right] \right\}, \tag{19}$$

where $\mathcal{L}_{\mathcal{H}_K}(\xi) = \nu$, the controls are adapted and $\mathcal{H}_K$-valued, and

$$X_s^{K,t,\xi;\theta} = \xi + \int_t^s \theta_r \, dr + W_s^K - W_t^K.$$

Here $W^K := \Pi_K W$ is a Q_K-Wiener process on $\mathcal{H}_K$, and

$$\|h\|_{Q_K} := \|Q_K^{-1/2} h\|_{\mathcal{H}_K}.$$

Formally, V_ε^K solves the projected Master equation on $\mathcal{P}_2(\mathcal{H}_K)$:

$$\partial_t V_\varepsilon^K(t,\nu) + \frac{1}{2} \int_{\mathcal{H}_K} \mathrm{Tr}\left(Q_K \partial_x \partial_\mu V_\varepsilon^K(t,\nu)(x) \right) \nu(dx) - \frac{1}{2\varepsilon} \int_{\mathcal{H}_K} \|Q_K^{1/2} \partial_\mu V_\varepsilon^K(t,\nu)(x)\|^2 \nu(dx) = 0, \tag{20}$$

with terminal condition $V_\varepsilon^K(1,\nu) = G_K(\nu)$.

For $\mathbf{x} \in (\mathcal{H}_K)^N$, define

$$v_\varepsilon^{K,N}(t,\mathbf{x}) := \inf_{\theta^1,\dots,\theta^N} \mathbb{E}\left[N G_K(\mu_{\mathbf{X}_1^K}^N) + \frac{\varepsilon}{2} \sum_{i=1}^N \int_t^1 \|\theta_s^i\|_{Q_K}^2 \, ds \right], \tag{21}$$

where

$$X_s^{K,i} = x_i + \int_t^s \theta_r^i \, dr + W_s^{K,i} - W_t^{K,i},$$

with independent Q_K-Wiener processes $W^{K,1}, \dots, W^{K,N}$. The projected particle HJB is

$$\partial_t v_\varepsilon^{K,N} + \frac{1}{2} \sum_{i=1}^N \mathrm{Tr}\left(Q_K D_{x_i}^2 v_\varepsilon^{K,N} \right) - \frac{1}{2\varepsilon} \sum_{i=1}^N \|Q_K^{1/2} D_{x_i} v_\varepsilon^{K,N}\|^2 = 0, \tag{22}$$

with terminal condition $v_\varepsilon^{K,N}(1,\mathbf{x}) = N G_K(\mu_{\mathbf{x}}^N)$. The Cole–Hopf transform gives the risk-sensitive representation

$$v_\varepsilon^{K,N}(t,\mathbf{x}) = -\varepsilon \log \mathbb{E}\left[\exp\left(-\frac{N}{\varepsilon} G_K(\mu_{\mathbf{x}+\Delta \mathbf{W}_{1-t}^K}^N) \right) \right].$$

Our proof has two layers. First, $N^{-1} v_\varepsilon^{K,N}$ approximates V_ε^K as $N \to \infty$ for fixed K and ε. Second, $V_\varepsilon^K(0, \delta_{\Pi_K x_0})$ is compared directly with $\mathcal{V}_0$.

Assumption B.2 (Projected classical solution, growth, and particle residual)**.** *For the quantitative results in this section, assume that the terminal objective G is bounded below, continuous and has at most quadratic growth:*

$$|G(\mu)| \leq C_G \left(1 + M_2(\mu)\right), \qquad \mu \in \mathcal{P}_2(\mathcal{H}). \tag{23}$$

For every $\varepsilon > 0$ and $K \geq 1$, we assume that the projected Master equation (20) admits a solution V_ε^K on $[0,1] \times \mathcal{P}_2(\mathcal{H}_K)$, continuous up to $t = 1$ with terminal value G_K. On $[0,1) \times \mathcal{P}_2(\mathcal{H}_K)$, the derivatives

$$\partial_t V_\varepsilon^K, \qquad \partial_\mu V_\varepsilon^K, \qquad \partial_x \partial_\mu V_\varepsilon^K, \qquad \partial^2_{\mu\mu} V_\varepsilon^K$$

exist in the classical Lions sense needed below. On top of that, we assume the Carmona–Delarue chain-rule regularity on $\mathcal{P}_2(\mathcal{H}_K)$: the versions of $\partial_\mu V_\varepsilon^K$ and $\partial_x \partial_\mu V_\varepsilon^K$ are continuous at support points of the measure argument and locally bounded on bounded subsets of $[0,1) \times \mathcal{P}_2(\mathcal{H}_K) \times \mathcal{H}_K$.

We also assume the following growth and local Lipschitz bounds: There exists $C_{\varepsilon,K} < \infty$ such that, for all $t \in [0,1)$,

$$\|\partial_\mu V_\varepsilon^K(t,\nu)(x)\| + \|\partial_x \partial_\mu V_\varepsilon^K(t,\nu)(x)\|_{\mathrm{op}} \leq C_{\varepsilon,K} \left(1 + \|x\| + M_2(\nu)^{1/2}\right). \tag{24}$$

Moreover, we suppose that for every $R < \infty$ and every $\tau < 1$, there is a constant $C_{\varepsilon,K,R,\tau} < \infty$ such that, for all $t \in [0,\tau]$, all $x, x' \in \mathcal{H}_K$ with $\|x\| \vee \|x'\| \leq R$, and all $\nu, \nu' \in \mathcal{P}_2(\mathcal{H}_K)$ with $M_2(\nu) \vee M_2(\nu') \leq R$,

$$\|\partial_\mu V_\varepsilon^K(t,\nu)(x) - \partial_\mu V_\varepsilon^K(t,\nu')(x')\| \leq C_{\varepsilon,K,R,\tau} \left(\|x - x'\| + \mathcal{W}_2(\nu,\nu')\right). \tag{25}$$

Finally, we assume that there exists $L_{\varepsilon,K} < \infty$ such that

$$\left| \int_{\mathcal{H}_K} \mathrm{Tr}\left(Q_K \partial^2_{\mu\mu} V_\varepsilon^K(t,\nu)(y)(y)\right) \nu(dy) \right| \leq L_{\varepsilon,K}, \qquad (t,\nu) \in [0,1) \times \mathcal{P}_2(\mathcal{H}_K). \tag{26}$$

Lemma B.4 (Projected verification)**.** *Under Assumptions B.2, the classical solution V_ε^K coincides with the projected value function defined in Equation (19). In particular, it is unique within this verified classical class.*

Proof. Fix an admissible projected control θ, and set

$$X_s := X_s^{K,t,\xi;\theta}, \qquad \nu_s := \mathcal{L}_{\mathcal{H}_K}(X_s).$$

By Assumptions B.2, the Carmona–Delarue chain rule applies to $s \mapsto V_\varepsilon^K(s,\nu_s)$ in expectation. Using the projected Master equation (20), we obtain, for $r < 1$,

$$V_\varepsilon^K(r,\nu_r) - V_\varepsilon^K(t,\nu) = \int_t^r \left\{ \frac{1}{2\varepsilon} \mathbb{E}\left[\left\| Q_K^{1/2} \partial_\mu V_\varepsilon^K(s,\nu_s)(X_s) \right\|^2 \right] + \mathbb{E}\left[\left\langle \partial_\mu V_\varepsilon^K(s,\nu_s)(X_s), \theta_s \right\rangle \right] \right\} ds.$$

Letting $r \uparrow 1$ and using $V_\varepsilon^K(1,\nu_1) = G_K(\nu_1)$ gives

$$\begin{aligned} &G_K(\nu_1) + \frac{\varepsilon}{2}\mathbb{E}\int_t^1 \|\theta_s\|^2_{Q_K}\, ds - V_\varepsilon^K(t,\nu) \\ &\quad = \frac{\varepsilon}{2}\mathbb{E}\int_t^1 \left\| \theta_s + \frac{1}{\varepsilon} Q_K \partial_\mu V_\varepsilon^K(s,\nu_s)(X_s) \right\|^2_{Q_K} ds \geq 0. \end{aligned}$$

Therefore every admissible control has cost at least $V_\varepsilon^K(t,\nu)$. It remains to check that the minimizing feedback is admissible. By Equation (24) and Equation (25), the finite-dimensional closed-loop McKean–Vlasov equation with drift

$$-\frac{1}{\varepsilon} Q_K \partial_\mu V_\varepsilon^K(s,\nu_s)(X_s^K)$$

is well posed. Moreover, Equation (24) and the standard second-moment estimate for this closed-loop equation imply

$$\mathbb{E}\int_t^1 \left\| -\frac{1}{\varepsilon} Q_K \partial_\mu V_\varepsilon^K(s,\nu_s)(X_s^K) \right\|^2_{Q_K} ds < \infty.$$

For this feedback, the square term vanishes. Hence equality is attained, and V_ε^K is the projected value function. The same verification argument applied to any other function in this class gives uniqueness. □

Assumption B.3. *Assume that the spatial gradients of $v_\varepsilon^{K,N}$ are locally Lipschitz in $\mathbf{x}$ on compact time intervals in $[0,1)$ and have at most linear growth: there is $C_{\varepsilon,K,N} < \infty$ such that*

$$\sum_{i=1}^{N} \|D_{x_i} v_\varepsilon^{K,N}(t,\mathbf{x})\| \le C_{\varepsilon,K,N}(1+\|\mathbf{x}\|), \qquad t \in [0,1). \tag{27}$$

Theorem B.5 (Particle approximation after Galerkin projection)**.** *Let Assumptions B.2 and B.3 hold, and let $v_\varepsilon^{K,N}$ be a classical solution of Equation (22). Then, for every $t \in [0,1]$ and $\mathbf{x} \in (\mathcal{H}_K)^N$,*

$$\left| \frac{1}{N} v_\varepsilon^{K,N}(t,\mathbf{x}) - V_\varepsilon^K(t,\mu_{\mathbf{x}}^N) \right| \le \frac{L_{\varepsilon,K}}{2N}(1-t). \tag{28}$$

Proof. Define $w_\varepsilon^{K,N}(t,\mathbf{x}) := N V_\varepsilon^K(t,\mu_{\mathbf{x}}^N)$. The empirical Lions chain rule gives

$$D_{x_i} w_\varepsilon^{K,N} = \partial_\mu V_\varepsilon^K(t,\mu_{\mathbf{x}}^N)(x_i),$$
$$D_{x_i}^2 w_\varepsilon^{K,N} = \partial_x \partial_\mu V_\varepsilon^K(t,\mu_{\mathbf{x}}^N)(x_i) + \frac{1}{N}\partial_{\mu\mu}^2 V_\varepsilon^K(t,\mu_{\mathbf{x}}^N)(x_i)(x_i).$$

Applying the projected particle HJB operator to $w_\varepsilon^{K,N}$ gives

$$\mathcal{L}^{K,N} w_\varepsilon^{K,N} = N\left[\partial_t V_\varepsilon^K + \frac{1}{2}\int_{\mathcal{H}_K} \mathrm{Tr}(Q_K \partial_x \partial_\mu V_\varepsilon^K) d\mu_{\mathbf{x}}^N - \frac{1}{2\varepsilon}\int_{\mathcal{H}_K} \|Q_K^{1/2}\partial_\mu V_\varepsilon^K\|^2 d\mu_{\mathbf{x}}^N\right]$$
$$+ \frac{1}{2}\int_{\mathcal{H}_K} \mathrm{Tr}\left(Q_K \partial_{\mu\mu}^2 V_\varepsilon^K(t,\mu_{\mathbf{x}}^N)(y)(y)\right) \mu_{\mathbf{x}}^N(dy).$$

The bracketed term vanishes by Equation (20). Thus

$$\mathcal{L}^{K,N} w_\varepsilon^{K,N} = R^{K,N},$$

where

$$R^{K,N}(t,\mathbf{x}) = \frac{1}{2}\int_{\mathcal{H}_K} \mathrm{Tr}\left(Q_K \partial_{\mu\mu}^2 V_\varepsilon^K(t,\mu_{\mathbf{x}}^N)(y)(y)\right) \mu_{\mathbf{x}}^N(dy).$$

By Assumption B.2, $|R^{K,N}| \le L_{\varepsilon,K}/2$.

Let $e_\varepsilon^{K,N} := v_\varepsilon^{K,N} - w_\varepsilon^{K,N}$. Subtracting the HJB equations and using $\|a\|^2 - \|b\|^2 = \langle a+b, a-b\rangle$ gives

$$\partial_t e_\varepsilon^{K,N} + \frac{1}{2}\sum_{i=1}^{N} \mathrm{Tr}(Q_K D_{x_i}^2 e_\varepsilon^{K,N}) + \sum_{i=1}^{N} \langle b_i(t,\mathbf{x}), D_{x_i} e_\varepsilon^{K,N}\rangle = -R^{K,N}(t,\mathbf{x}),$$

with terminal $e_\varepsilon^{K,N}(1,\cdot) = 0$.

Define the comparison drift

$$b_i(t,\mathbf{x}) := -\frac{1}{2\varepsilon} Q_K \left(D_{x_i} v_\varepsilon^{K,N}(t,\mathbf{x}) + D_{x_i} w_\varepsilon^{K,N}(t,\mathbf{x})\right).$$

By the preceding linear-growth and local-Lipschitz Assumptions B.2 and B.3, the comparison diffusion

$$d\Xi_s^i = b_i(s,\boldsymbol{\Xi}_s)\, ds + dW_s^{K,i}, \qquad \boldsymbol{\Xi}_t = \mathbf{x},$$

is nonexplosive on $[t,1]$ and satisfies the standard finite-dimensional moment bound

$$\mathbb{E}\sup_{s\in[t,1]} \|\boldsymbol{\Xi}_s\|^2 < \infty.$$

The common terminal condition for $v_\varepsilon^{K,N}$ and $w_\varepsilon^{K,N}$ gives the local terminal trace $e_\varepsilon^{K,N}(s,\cdot) \to 0$ as $s \uparrow 1$. Together with the linear-gradient bound Equation (27), this gives the quadratic envelope needed to pass the stopped Feynman–Kac terminal term to zero along $\boldsymbol{\Xi}$.

By Assumption B.3, applying Itô's formula to the stopped comparison diffusion and then sending the localization and terminal parameters to their limits yields

$$e_\varepsilon^{K,N}(t,\mathbf{x}) = \mathbb{E}\int_t^1 R^{K,N}(s,\boldsymbol{\Xi}_s)\, ds.$$

Therefore

$$|e_\varepsilon^{K,N}(t,\mathbf{x})| \le \int_t^1 \sup_{\mathbf{z}\in(\mathcal{H}_K)^N} |R^{K,N}(s,\mathbf{z})|\, ds \le \frac{1}{2} L_{\varepsilon,K}(1-t).$$

Dividing by N proves the claim. □

B.3 PROJECTED REGULARIZATION ERROR

Assumption B.4 (Local continuity and comparator)**.** *Assume that the infimum $\mathcal{V}_0$ is attained by some $\mu^* \in \mathcal{P}_2(\mathcal{H})$. Let $X \sim \mu^*$. Assume G satisfies a local Hölder continuity condition with exponent $\alpha \in (0,1]$: there is a constant L_G such that for all $\mu, \nu \in \mathcal{P}_2(\mathcal{H})$,*

$$|G(\mu) - G(\nu)| \le L_G \left(1 + \mathcal{W}_2(\mu,\nu) + M_2(\mu) + M_2(\nu)\right) \mathcal{W}_2(\mu,\nu)^\alpha.$$

For $K \ge 1$, define

$$\Delta_K^{\mathrm{proj}} := \mathbb{E}_{\mu^*} \|(I - \Pi_K)X\|^2,$$
$$D_{CM,K}^2(\mu^*) := \sum_{j=1}^{K} \frac{\mathbb{E}_{\mu^*}[(X_j - x_{0,j})^2]}{\lambda_j}.$$

Theorem B.6 (Projected regularization error)**.** *Let Assumption B.4 hold. Then, for every $K \ge 1$, every $\varepsilon \in (0,1]$, and every $\rho \in (0,1]$,*

$$0 \le V_\varepsilon^K(0, \delta_{\Pi_K x_0}) - \mathcal{V}_0 \le C_* \left(\rho \operatorname{Tr}(Q) + \Delta_K^{\mathrm{proj}}\right)^{\alpha/2} + \varepsilon \left[\frac{1}{2} D_{CM,K}^2(\mu^*) + \frac{K}{2} \log \frac{1}{\rho}\right], \quad (29)$$

where C_* is independent of K, ε, and ρ. In particular, choosing $\rho = \varepsilon^{2/\alpha}$ gives

$$0 \le V_\varepsilon^K(0, \delta_{\Pi_K x_0}) - \mathcal{V}_0 \le C_* \left(\varepsilon^{2/\alpha} \operatorname{Tr}(Q) + \Delta_K^{\mathrm{proj}}\right)^{\alpha/2} + \frac{\varepsilon}{2} D_{CM,K}^2(\mu^*) + \frac{\varepsilon K}{\alpha} \log \frac{1}{\varepsilon}. \quad (30)$$

Proof. Let

$$\gamma_{K,x_0} := \mathcal{L}_{\mathcal{H}_K}(\Pi_K x_0 + W_1^K) = \mathcal{N}(\Pi_K x_0, Q_K)$$

be the uncontrolled projected terminal law. By the finite-dimensional Föllmer drift representation on $\mathcal{H}_K$, for every $\nu \ll \gamma_{K,x_0}$,

$$V_\varepsilon^K(0, \delta_{\Pi_K x_0}) \le G_K(\nu) + \varepsilon \mathcal{H}(\nu \mid \gamma_{K,x_0}), \quad (31)$$

where $\mathcal{H}(\nu \mid \gamma_{K,x_0})$ denotes relative entropy. Fix $\rho \in (0,1]$. Let $X \sim \mu^*$ and let $\zeta_K \sim \mathcal{N}(0, Q_K)$ be independent. Define the $\mathcal{H}_K$-valued random variable

$$Y^{\rho,K} := \Pi_K X + \sqrt{\rho}\, \zeta_K, \qquad \mu_K^\rho := \mathcal{L}_{\mathcal{H}_K}(Y^{\rho,K}).$$

Let

$$\widehat{\mu}_K^\rho := (\iota_K)_\# \mu_K^\rho = \mathcal{L}_{\mathcal{H}}(Y^{\rho,K}).$$

Conditionally on $X = x$,

$$Y^{\rho,K} \mid X = x \sim \mathcal{N}(\Pi_K x, \rho Q_K).$$

Since $\rho > 0$ and Q_K is nondegenerate on $\mathcal{H}_K$, $\mu_K^\rho \ll \gamma_{K,x_0}$.

Define joint measures on $\mathcal{H} \times \mathcal{H}_K$ by

$$\eta(dx, dy) := \mu^*(dx)\, \mathcal{N}(\Pi_K x, \rho Q_K)(dy), \qquad \xi(dx, dy) := \mu^*(dx)\, \mathcal{N}(\Pi_K x_0, Q_K)(dy).$$

The second marginals are μ_K^ρ and γ_{K,x_0}. By data processing for relative entropy,

$$\mathcal{H}(\mu_K^\rho \mid \gamma_{K,x_0}) \le \mathcal{H}(\eta \mid \xi).$$

Since the first marginal is the same, the chain rule for relative entropy gives

$$\mathcal{H}(\eta \mid \xi) = \int_{\mathcal{H}} \mathcal{H}\left(\mathcal{N}(\Pi_K x, \rho Q_K) \mid \mathcal{N}(\Pi_K x_0, Q_K)\right) \mu^*(dx).$$

The Gaussian entropy formula gives

$$\mathcal{H}\left(\mathcal{N}(\Pi_K x, \rho Q_K) \mid \mathcal{N}(\Pi_K x_0, Q_K)\right) = \frac{1}{2} \sum_{j=1}^{K} \frac{(x_j - x_{0,j})^2}{\lambda_j} + \frac{K}{2}(\rho - 1 - \log \rho).$$

Since $\rho \in (0,1]$, $\rho - 1 - \log \rho \le \log(1/\rho)$. Hence

$$\mathcal{H}(\mu_K^\rho \mid \gamma_{K,x_0}) \le \frac{1}{2} D_{CM,K}^2(\mu^*) + \frac{K}{2} \log \frac{1}{\rho}. \quad (32)$$

Couple $\widehat{\mu}_K^\rho$ with μ^* by using $(Y^{\rho,K}, X)$ as $\mathcal{H}$-valued random variables. Then

$$\mathcal{W}_{2,\mathcal{H}}^2(\widehat{\mu}_K^\rho, \mu^*) \le \mathbb{E}\|Y^{\rho,K} - X\|_{\mathcal{H}}^2.$$

Since

$$Y^{\rho,K} - X = -(I - \Pi_K)X + \sqrt{\rho}\,\zeta_K,$$

and the two terms are orthogonal,

$$\mathcal{W}_{2,\mathcal{H}}^2(\widehat{\mu}_K^\rho, \mu^*) \le \Delta_K^{\mathrm{proj}} + \rho\,\mathrm{Tr}(Q_K) \le \Delta_K^{\mathrm{proj}} + \rho\,\mathrm{Tr}(Q). \tag{33}$$

Also,

$$\int_{\mathcal{H}} \|y\|^2 \widehat{\mu}_K^\rho(dy) = \mathbb{E}\|Y^{\rho,K}\|^2 \le 2\mathbb{E}\|X\|^2 + 2\,\mathrm{Tr}(Q).$$

Thus the local Hölder condition and optimality of μ^* imply

$$G(\widehat{\mu}_K^\rho) - \mathcal{V}_0 \le C_* \left(\rho\,\mathrm{Tr}(Q) + \Delta_K^{\mathrm{proj}}\right)^{\alpha/2}. \tag{34}$$

Using Equation (31) with $\nu = \mu_K^\rho$, and recalling that $G_K(\mu_K^\rho) = G(\widehat{\mu}_K^\rho)$, gives Equation (29). The lower bound follows from $G \ge \mathcal{V}_0$ and the nonnegativity of the running cost. Setting $\rho = \varepsilon^{2/\alpha}$ gives Equation (30). □

Theorem B.7 (Projection-first convergence of the idealized particle approximation)**.** *Assume Assumptions B.2, B.3, and B.4 and $0 < \varepsilon \le 1$. Let $v_\varepsilon^{K,N}$ be a classical solution of the projected particle HJB Equation (22). Then*

$$0 \le \frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) - \mathcal{V}_0 \le \frac{L_{\varepsilon,K}}{2N} + C_* \left(\varepsilon^{2/\alpha}\,\mathrm{Tr}(Q) + \Delta_K^{\mathrm{proj}}\right)^{\alpha/2} + \frac{\varepsilon}{2} D_{CM,K}^2(\mu^*) + \frac{\varepsilon K}{\alpha} \log \frac{1}{\varepsilon}. \tag{35}$$

Here $\Pi_K \mathbf{x}_0 = (\Pi_K x_0, \dots, \Pi_K x_0)$.

Proof. Since $\mu_{\Pi_K \mathbf{x}_0}^N = \delta_{\Pi_K x_0}$, the triangle inequality gives

$$\frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) - \mathcal{V}_0 \le \left| \frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) - V_\varepsilon^K(0, \delta_{\Pi_K x_0}) \right| + \left(V_\varepsilon^K(0, \delta_{\Pi_K x_0}) - \mathcal{V}_0\right).$$

The first term is controlled by Theorem B.5; the second term is controlled by Theorem B.6. The lower bound follows from $G \ge \mathcal{V}_0$ and nonnegativity of the running cost. □

Remark B.1 (Asymptotic regime)**.** *The bound in Equation (35) tends to zero along any joint choice $K = K_m$, $N = N_m$, and $\varepsilon = \varepsilon_m$ such that*

$$\frac{L_{\varepsilon_m, K_m}}{N_m} \to 0, \qquad \Delta_{K_m}^{\mathrm{proj}} \to 0,$$

$$\varepsilon_m D_{CM,K_m}^2(\mu^*) \to 0, \qquad \varepsilon_m K_m \log \frac{1}{\varepsilon_m} \to 0, \qquad \varepsilon_m \to 0.$$

Remark B.2. *Theorem B.7 is not a theorem for the fully discrete finite-candidate SCMO implementation. It concerns the idealized continuous-time projected particle value $v_\varepsilon^{K,N}$. Finite-candidate sampling and time-discretization errors must be controlled separately.*

B.4 Optional sufficient condition for uniform particle constants

The main convergence theorem only needs the constants $L_{\varepsilon,K}$ in Assumption B.2. The following assumptions give one sufficient route to make these constants uniform in ε and K.

For $\nu \in \mathcal{P}_2(\mathcal{H}_K)$, write

$$U_\varepsilon^K(t, \nu)(x) := \partial_\mu V_\varepsilon^K(t, \nu)(x).$$

For the projected terminal functional G_K, define

$$A_\nu^K(x) := \partial_x \partial_\mu G_K(\nu)(x), \qquad B_\nu^K(x, y) := \partial_{\mu\mu}^2 G_K(\nu)(x)(y).$$

Assumption B.5 (Projected FBSDE regularity and tangent equations)**.** *Fix $\varepsilon > 0$ and $K \geq 1$. Assume that U_ε^K is continuous on its domain and, for $t < 1$, has continuous spatial derivative $\partial_x U_\varepsilon^K = \partial_x \partial_\mu V_\varepsilon^K$ and continuous Lions derivative $\partial_\mu U_\varepsilon^K = \partial_{\mu\mu}^2 V_\varepsilon^K$. On bounded subsets of $\mathcal{P}_2(\mathcal{H}_K)$ and $\mathcal{H}_K$, these derivatives are locally Lipschitz and have at most polynomial growth.*

Assume further that, for every (t, ν) with $\xi \sim \nu$, the closed-loop SDE

$$dX_s^{t,\xi} = -\frac{1}{\varepsilon} Q_K U_\varepsilon^K(s, \nu_s)(X_s^{t,\xi})\, ds + dW_s^K, \qquad X_t^{t,\xi} = \xi, \qquad \nu_s = \mathcal{L}(X_s^{t,\xi}),$$

has a unique square-integrable solution, and that

$$Y_s^{t,\xi} := U_\varepsilon^K(s, \nu_s)(X_s^{t,\xi}), \quad Z_s^{t,\xi} := \partial_x U_\varepsilon^K(s, \nu_s)(X_s^{t,\xi})$$

satisfies the decoupled FBSDE

$$\begin{cases} dX_s^{t,\xi} = -\dfrac{1}{\varepsilon} Q_K Y_s^{t,\xi}\, ds + dW_s^K, & X_t^{t,\xi} = \xi, \\ dY_s^{t,\xi} = Z_s^{t,\xi}\, dW_s^K, \\ Y_1^{t,\xi} = \partial_\mu G_K(\nu_1)(X_1^{t,\xi}). \end{cases} \tag{36}$$

Also, assume that the following L^2 tangent equations are valid. If $\xi^r = \xi + r\phi(\xi)$ with $\phi \in L^2(\mathcal{H}_K, \nu; \mathcal{H}_K)$, then the first law variation

$$\delta X_s := \frac{d}{dr} X_s^{t,\xi^r}\Big|_{r=0}, \qquad \delta Y_s := \frac{d}{dr} Y_s^{t,\xi^r}\Big|_{r=0}, \qquad \delta Z_s := \frac{d}{dr} Z_s^{t,\xi^r}\Big|_{r=0}$$

exists in L^2 and satisfies

$$\begin{cases} d(\delta X_s) = -\dfrac{1}{\varepsilon} Q_K \delta Y_s\, ds, & \delta X_t = \phi(\xi), \\ d(\delta Y_s) = \delta Z_s\, dW_s^K, \\ \delta Y_1 = A_{\nu_1}^K(X_1^{t,\xi})\delta X_1 + \widetilde{\mathbb{E}}\Big[B_{\nu_1}^K(X_1^{t,\xi}, \widetilde{X}_1^{t,\xi})\delta \widetilde{X}_1\Big]. \end{cases} \tag{37}$$

Finally for fixed deterministic $x \in \mathcal{H}_K$, let $X^{t,x}$ be the decoupled process started from x and driven by the same law flow (ν_s). Its measure-direction variation in the direction ϕ, denoted by $(\delta_\phi X, \delta_\phi Y, \delta_\phi Z)$, exists in L^2 and satisfies

$$\begin{cases} d(\delta_\phi X_s) = -\dfrac{1}{\varepsilon} Q_K \delta_\phi Y_s\, ds, & \delta_\phi X_t = 0, \\ d(\delta_\phi Y_s) = \delta_\phi Z_s\, dW_s^K, \\ \delta_\phi Y_1 = A_{\nu_1}^K(X_1^{t,x})\delta_\phi X_1 + \widetilde{\mathbb{E}}\Big[B_{\nu_1}^K(X_1^{t,x}, \widetilde{X}_1^{t,\xi})\delta \widetilde{X}_1\Big]. \end{cases} \tag{38}$$

All stochastic integrals appearing above are true martingales after localization and passage to the limit. These regularity assumptions are finite-dimensional conditions on $\mathcal{H}_K$; they are stated explicitly here only to justify the FBSDE differentiations used below.

Assumption B.6 (Projected terminal monotonicity and coercivity)**.** *There exist constants $K_G < \infty$ and $\kappa > 0$, independent of (ε, K), such that*

$$\|A_\nu^K(x)\|_{\mathrm{op}} + \|B_\nu^K(x, y)\|_{\mathrm{op}} \leq K_G, \qquad \nu \in \mathcal{P}_2(\mathcal{H}_K),\ x, y \in \mathcal{H}_K, \tag{39}$$

$$\begin{aligned} &\mathbb{E}\left[\left\langle \delta_\phi X_1, A_{\nu_1}^K(X_1^{t,x})\delta_\phi X_1 + \widetilde{\mathbb{E}}\left[B_{\nu_1}^K(X_1^{t,x}, \widetilde{X}_1^{t,\xi})\delta \widetilde{X}_1\right]\right\rangle\right] \\ &\qquad \geq \kappa \mathbb{E}\|\delta_\phi X_1\|^2 - K_G \mathbb{E}\|\delta_\phi X_1\|\, \mathbb{E}\|\delta X_1\|, \end{aligned} \tag{40}$$

and, for every square-integrable pair (X, η) with $\mathcal{L}(X) = \nu$,

$$\mathbb{E}\left[\frac{\mathbf{1}_{\{\|\eta\|>0\}}}{\|\eta\|}\left\langle \eta, Q_K A_\nu^K(X)\eta + Q_K \widetilde{\mathbb{E}}\left[B_\nu^K(X, \widetilde{X})\widetilde{\eta}\right]\right\rangle\right] \geq 0, \tag{41}$$

where $(\widetilde{X}, \widetilde{\eta})$ is an independent copy of (X, η).

Proposition B.8 (Projected second Lions derivative estimate)**.** *Assume Assumptions B.5 and B.6. Then*

$$\|\partial^2_{\mu\mu} V^K_\varepsilon(t,\nu)(x)(y)\|_{\mathrm{op}} \le \frac{K_G^2}{\kappa} + K_G, \qquad \nu\text{-a.e. } y \in \mathcal{H}_K. \tag{42}$$

Consequently,

$$L_{\varepsilon,K} \le \mathrm{Tr}(Q_K)\left(\frac{K_G^2}{\kappa} + K_G\right) \le \mathrm{Tr}(Q)\left(\frac{K_G^2}{\kappa} + K_G\right). \tag{43}$$

The bound is independent of ε and K.

Proof. We prove only the estimate needed for $L_{\varepsilon,K}$.

Step 1: contraction of the first law variation. Let $(\delta X, \delta Y, \delta Z)$ solve Equation (37). Applying Itô's formula to $\langle \delta X_s, Q_K \delta Y_s\rangle$ gives

$$d\langle \delta X_s, Q_K \delta Y_s\rangle = -\frac{1}{\varepsilon}\|Q_K \delta Y_s\|^2\, ds + \langle \delta X_s, Q_K \delta Z_s\, dW_s^K\rangle.$$

Using the regularized norm $(\|z\|^2+\eta^2)^{1/2}$ and then sending $\eta \downarrow 0$, the product rule and Cauchy–Schwarz show that

$$\frac{\mathbf{1}_{\{\|\delta X_s\|>0\}}}{\|\delta X_s\|}\langle \delta X_s, Q_K \delta Y_s\rangle$$

is a supermartingale after localization. The terminal condition and Equation (41) imply that its terminal expectation is nonnegative. Hence

$$\mathbb{E}\left[\frac{\mathbf{1}_{\{\|\delta X_s\|>0\}}}{\|\delta X_s\|}\langle \delta X_s, Q_K \delta Y_s\rangle\right] \ge 0, \qquad s \in [t,1].$$

Since

$$d\|\delta X_s\| = -\frac{\mathbf{1}_{\{\|\delta X_s\|>0\}}}{\varepsilon\|\delta X_s\|}\langle \delta X_s, Q_K \delta Y_s\rangle\, ds,$$

we obtain

$$\mathbb{E}\|\delta X_1\| \le \mathbb{E}\|\phi(\xi)\|. \tag{44}$$

Step 2: measure-direction stability. Let $(\delta_\phi X, \delta_\phi Y, \delta_\phi Z)$ solve Equation (38). Itô's formula for $\langle \delta_\phi X_s, \delta_\phi Y_s\rangle$ gives

$$\mathbb{E}\langle \delta_\phi X_1, \delta_\phi Y_1\rangle = -\frac{1}{\varepsilon}\mathbb{E}\int_t^1 \langle Q_K \delta_\phi Y_s, \delta_\phi Y_s\rangle\, ds \le 0.$$

Using the terminal condition in Equation (38), the mixed coercivity condition Equation (40), and Equation (44),

$$0 \ge \mathbb{E}\langle \delta_\phi X_1, \delta_\phi Y_1\rangle \ge \kappa \mathbb{E}\|\delta_\phi X_1\|^2 - K_G \mathbb{E}\|\delta_\phi X_1\|\,\mathbb{E}\|\phi(\xi)\|.$$

Therefore

$$\mathbb{E}\|\delta_\phi X_1\| \le \frac{K_G}{\kappa}\mathbb{E}\|\phi(\xi)\|. \tag{45}$$

Step 3: second Lions derivative. Differentiating the decoupling relation

$$Y_t^{t,x} = \partial_\mu V^K_\varepsilon(t,\nu)(x)$$

in the measure direction ϕ gives

$$\left\|\mathbb{E}\left[\partial^2_{\mu\mu} V^K_\varepsilon(t,\nu)(x)(\xi)\phi(\xi)\right]\right\| = \|\mathbb{E}[\delta_\phi Y_t]\|.$$

Since $\delta_\phi Y$ is a martingale,

$$\begin{aligned}\|\mathbb{E}[\delta_\phi Y_t]\| &\le K_G \mathbb{E}\|\delta_\phi X_1\| + K_G \mathbb{E}\|\delta X_1\| \\ &\le \left(\frac{K_G^2}{\kappa} + K_G\right)\mathbb{E}\|\phi(\xi)\|.\end{aligned}$$

The standard duality and measurable-selection argument yields Equation (42) for every fixed $x \in H_K$ and ν-a.e. $y \in H_K$. By the continuity of $\partial^2_{\mu\mu} V^K_\varepsilon$ assumed in Assumption B.5, this estimate extends, for every fixed x, to all $y \in \mathrm{supp}\,\nu$. Hence, taking $x = y$, we obtain

$$\left\|\partial^2_{\mu\mu} V^K_\varepsilon(t,\nu)(y)(y)\right\|_{\mathrm{op}} \le \frac{K_G^2}{\kappa} + K_G, \qquad y \in \mathrm{supp}\,\nu.$$

Taking the Q_K-trace and integrating with respect to ν therefore gives Equation (43). □

B.5 EXAMPLES

Example B.9 (Coercive quadratic target)**.** *Let*

$$G_{\mathrm{quad}}(\mu) := \frac{\kappa}{2}\int_{\mathcal{H}} \|x-a\|^2 \mu(dx), \qquad \kappa > 0, \quad a \in \mathcal{H}.$$

Then $\mathcal{V}_0 = 0$ and the unique minimizer is $\mu^* = \delta_a$. The derivatives are

$$\partial_\mu G_{\mathrm{quad}}(\mu)(x) = \kappa(x-a), \qquad \partial_x \partial_\mu G_{\mathrm{quad}}(\mu)(x) = \kappa I, \qquad \partial^2_{\mu\mu} G_{\mathrm{quad}} = 0.$$

Thus the optional uniform-$L_{\varepsilon,K}$ condition holds. In fact, the value function is linear in the measure variable, so the sharper identity $L_{\varepsilon,K} = 0$ holds. For this example,

$$\Delta_K^{\mathrm{proj}} = \|(I-\Pi_K)a\|^2, \qquad D^2_{CM,K}(\delta_a) = \sum_{j=1}^{K} \frac{(a_j - x_{0,j})^2}{\lambda_j}.$$

Therefore Theorem B.7 gives, with $\alpha = 1$ on bounded moment sets,

$$0 \le \frac{1}{N} v_\varepsilon^{K,N}(0, \Pi_K \mathbf{x}_0) \le C_* \left(\varepsilon^2 \operatorname{Tr}(Q) + \|(I-\Pi_K)a\|^2\right)^{1/2} + \frac{\varepsilon}{2}\sum_{j=1}^{K} \frac{(a_j - x_{0,j})^2}{\lambda_j} + \varepsilon K \log\frac{1}{\varepsilon}.$$

C CONVERGENCE ANALYSIS FOR NONMATCHED COVARIANCE PROBLEMS

C.1 FINITE-CANDIDATE CONSISTENCY WITH NONMATCHED EXECUTION COVARIANCE

We now connect the practical SCMO update to the matched projected particle value analyzed in Section B. We analyze one outer-loop rollout, consisting of M updates from its current cloud. When $L > 1$, the same per-rollout analysis applies at each outer iteration; The analysis proceeds in two stages. First, we establish consistency of the R-context, S-candidate Monte Carlo approximation of the Cole–Hopf feedback, thereby relating the fully finite implementation to an exact-feedback nonmatched-covariance chain. Second, we compare this exact-feedback chain, which is driven by the matched feedback but executed with a covariance different from the proposal covariance, with the matched projected particle value. The resulting covariance contribution is retained with its sign and decomposed into a nonnegative favorable gain and a nonnegative unfavorable contribution. In the smooth case, its leading-order one-step effect is governed by $\frac{h}{2}\operatorname{Tr}\left([\mathbf{Q}_{K,\mathrm{dyn}} - \mathbf{Q}_{K,\mathrm{prop}}]\, D^2 v_\varepsilon^{K,N}\right)$. In particular, when $\mathbf{Q}_{K,\mathrm{dyn}} = \alpha_0 \mathbf{Q}_{K,\mathrm{prop}}$, a colder execution covariance, $\alpha_0 < 1$, produces a favorable contribution in regions of positive covariance-weighted continuation-value curvature, whereas a hotter execution covariance, $\alpha_0 > 1$, is favorable in regions of negative weighted curvature. If the favorable curvature regions dominate along the states visited by the chain, the resulting gain exceeds the unfavorable covariance contribution and tightens the matched upper bound. If this gain also dominates the Euler discretization residual, the complete one-outer-loop bound is strictly smaller than the ideal matched particle–projection–regularization bound. Thus, nonmatched execution is not treated solely as an additional error: its interaction with the curvature of the continuation value can produce a genuine improvement in the theoretical certificate.

We recall the notation. For $\mathbf{y} = (y_1, \dots, y_N) \in (\mathcal{H}_K)^N$, write $\mu_{\mathbf{y}}^N = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$, $F_{K,N}(\mathbf{y}) := N\, G_K(\mu_{\mathbf{y}}^N)$. Thus $F_{K,N}$ is the extensive terminal cost appearing in the projected N-particle control problem. Write the full proposal covariance operator on $\mathcal{H}$ by $Q_{\mathrm{prop}} := \sigma^2_{\mathrm{prop}} Q$, and let $Q_{K,\mathrm{prop}} := \Pi_K Q_{\mathrm{prop}} \Pi_K$ be the projected covariance used in the Cole–Hopf control problem and in the context and candidate proposals. Also write $Q_{\mathrm{dyn}} := \sigma^2_{\mathrm{dyn}} Q$ and let $Q_{K,\mathrm{dyn}} := \Pi_K Q_{\mathrm{dyn}} \Pi_K$ be the covariance used in the Euler execution step. In the coefficient implementation, $Q_{K,\mathrm{prop}} = \sigma^2_{\mathrm{prop}}\Lambda$, $Q_{K,\mathrm{dyn}} = \sigma^2_{\mathrm{dyn}}\Lambda$. Since the two covariances have the same spectrum up to scale, define the fixed variance ratio $\alpha_0 := \frac{\sigma^2_{\mathrm{dyn}}}{\sigma^2_{\mathrm{prop}}}$, $Q_{K,\mathrm{dyn}} = \alpha_0 Q_{K,\mathrm{prop}}$, $\alpha_0 \neq 1$. Their N-particle block extensions are $\mathbf{Q}_{K,\mathrm{prop}} = I_N \otimes Q_{K,\mathrm{prop}}$, $\mathbf{Q}_{K,\mathrm{dyn}} = I_N \otimes Q_{K,\mathrm{dyn}} = \alpha_0 \mathbf{Q}_{K,\mathrm{prop}}$.

Throughout this subsection, $v_\varepsilon^{K,N}$ and b denote the matched projected particle value and feedback associated with $Q_{K,\mathrm{prop}}$.

Fix $K, N, \varepsilon > 0$, and $t < 1$. Let $\tau = 1 - t$, $\mathbf{U} = (U_1, \dots, U_N) \sim \mathcal{N}(0, \tau \mathbf{Q}_{K,\text{prop}})$, and define $L_{t,\mathbf{x}}(\mathbf{U}) = \exp\left(-\frac{F_{K,N}(\mathbf{x}+\mathbf{U})}{\varepsilon}\right)$. The exact matched Cole–Hopf feedback is

$$b_i(t, \mathbf{x}) = \frac{1}{\tau} \frac{\mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U}) U_i\right]}{\mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U})\right]}. \tag{46}$$

C.1.1 Multi-context finite-candidate feedback

For each $r = 1, \dots, R$, draw one complete context increment $\mathbf{U}^{(r)} = (U_1^{(r)}, \dots, U_N^{(r)}) \sim \mathcal{N}(0, \tau \mathbf{Q}_{K,\text{prop}})$. Conditionally on each context, draw independent candidate increments $U_i^{(r,s)} \sim \mathcal{N}(0, \tau Q_{K,\text{prop}})$, $s = 1, \dots, S$. All contexts and candidates are independent except for the shared context within a fixed block r. For the feedback of particle i, use $U_{-i}^{(r)} = (U_j^{(r)})_{j \neq i}$, and replace the i-th component by an independent candidate $U_i^{(r,s)}$.Define the complete increment $\mathbf{U}^{(i,r,s)} = \left(U_1^{(r)}, \dots, U_{i-1}^{(r)}, U_i^{(r,s)}, U_{i+1}^{(r)}, \dots, U_N^{(r)}\right) \in (H_K)^N$ and $L_{i,r,s} = L_{t,\mathbf{x}}\big(\mathbf{U}^{(i,r,s)}\big) = \exp\left[-\frac{F_{K,N}\left(\mathbf{x}+\mathbf{U}^{(i,r,s)}\right)}{\varepsilon}\right]$. Equivalently, $F_{K,N}\left(\mathbf{x} + \mathbf{U}^{(i,r,s)}\right)$ is the objective of the complete cloud obtained by perturbing every coordinate $j \neq i$ with the r-th context increment and replacing the i-th coordinate by its (r,s)-candidate increment. Define $\widehat{b}_i^{R,S}(t, \mathbf{x}) = \frac{1}{\tau} \frac{\sum_{r=1}^R \sum_{s=1}^S L_{i,r,s} U_i^{(r,s)}}{\sum_{r=1}^R \sum_{s=1}^S L_{i,r,s}}$.

Proposition C.1 (Consistency of the multi-context feedback estimator)**.** *Assume $\mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U}) \|U_i\|\right] < \infty$, $0 < \mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U})\right] < \infty$. Then, for every fixed $S \geq 1$,*

$$\widehat{b}_i^{R,S}(t, \mathbf{x}) \longrightarrow b_i(t, \mathbf{x}) \qquad \textit{almost surely as } R \to \infty.$$

For a quantitative statement, assume in addition that $\mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U})^4 \left(1 + \|U_i\|^4\right)\right] < \infty$. Writing $B_{i,r}^S := \frac{1}{S} \sum_{s=1}^S L_{i,r,s}$, $\overline{B}_i^{R,S} := \frac{1}{R} \sum_{r=1}^R B_{i,r}^S$, assume that the self-normalized denominator is L^2-stable in the explicit sense that

$$\sup_{R,S \geq 1} \mathbb{E}\left[\left(\overline{B}_i^{R,S}\right)^{-4}\right] < \infty. \tag{47}$$

A stronger but simpler sufficient condition is a deterministic lower bound $\overline{B}_i^{R,S} \geq c > 0$ on the time–state set under consideration. Under these assumptions,

$$\mathbb{E}\left[\left\|\widehat{b}_i^{R,S}(t, \mathbf{x}) - b_i(t, \mathbf{x})\right\|^2\right] \leq C\left(\frac{1}{R} + \frac{1}{RS}\right).$$

The constant may depend on $K, N, \varepsilon, t, \mathbf{x}$, but is independent of R and S. Consequently, under the quantitative assumptions, convergence also holds in L^2, and hence in probability, along every sequence satisfying

$$R_n \to \infty, \qquad R_n S_n \to \infty.$$

Proof. Define the context-block averages $A_{i,r}^S = \frac{1}{S} \sum_{s=1}^S L_{i,r,s} U_i^{(r,s)}$, $B_{i,r}^S = \frac{1}{S} \sum_{s=1}^S L_{i,r,s}$. Then $\widehat{b}_i^{R,S}(t, \mathbf{x}) = \frac{1}{\tau} \frac{\overline{A}_i^{R,S}}{\overline{B}_i^{R,S}}$, $\overline{A}_i^{R,S} := \frac{1}{R} \sum_{r=1}^R A_{i,r}^S$. The pairs $(A_{i,r}^S, B_{i,r}^S)$, $r = 1, \dots, R$, are independent and identically distributed. Since each complete context–candidate pair has the same law as $\mathbf{U}$, $a_i := \mathbb{E}[A_{i,r}^S] = \mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U}) U_i\right]$, $\beta := \mathbb{E}[B_{i,r}^S] = \mathbb{E}\left[L_{t,\mathbf{x}}(\mathbf{U})\right] > 0$. The strong law of large numbers gives $\overline{A}_i^{R,S} \longrightarrow a_i$, $\overline{B}_i^{R,S} \longrightarrow \beta$ almost surely as $R \to \infty$, for every fixed $S \geq 1$. The continuous-mapping theorem therefore yields $\widehat{b}_i^{R,S}(t, \mathbf{x}) \longrightarrow \frac{1}{\tau} \frac{a_i}{\beta} = b_i(t, \mathbf{x})$ almost surely.

We now prove the quantitative estimate. For an $\mathcal{H}_K$-valued square-integrable random variable Y, use the convention $\mathrm{Var}(Y) := \mathbb{E}\|Y - \mathbb{E}Y\|^2$. Let $\mathcal{G}_r := \sigma(U_{-i}^{(r)})$. Conditionally on $\mathcal{G}_r$, the variables $L_{i,r,s} U_i^{(r,s)}$, $s = 1, \dots, S$, are independent and identically distributed. The law of total variance therefore gives

$$\mathrm{Var}(A_{i,r}^S) = \mathrm{Var}\left(\mathbb{E}[L_{t,\mathbf{x}}(\mathbf{U}) U_i \mid U_{-i}]\right) + \frac{1}{S} \mathbb{E}\left[\mathrm{Var}\left(L_{t,\mathbf{x}}(\mathbf{U}) U_i \mid U_{-i}\right)\right],$$

and similarly

$$\mathrm{Var}(B_{i,r}^{S}) = \mathrm{Var}\left(\mathbb{E}[L_{t,\mathbf{x}}(\mathbf{U}) \mid U_{-i}]\right) + \frac{1}{S}\mathbb{E}\left[\mathrm{Var}\left(L_{t,\mathbf{x}}(\mathbf{U}) \mid U_{-i}\right)\right].$$

Since the blocks are independent across r,

$$\mathbb{E}\left\|\overline{A}_i^{R,S} - a_i\right\|^2 \le C_A\left(\frac{1}{R} + \frac{1}{RS}\right),$$

and

$$\mathbb{E}\left|\overline{B}_i^{R,S} - \beta\right|^2 \le C_B\left(\frac{1}{R} + \frac{1}{RS}\right).$$

The fourth-moment assumption, conditional Rosenthal inequalities within each context, and the standard fourth-moment estimate for averages of independent blocks also give

$$\left(\mathbb{E}\left\|\overline{A}_i^{R,S} - a_i\right\|^4\right)^{1/2} \le C'_A\left(\frac{1}{R} + \frac{1}{RS}\right)$$

and

$$\left(\mathbb{E}\left|\overline{B}_i^{R,S} - \beta\right|^4\right)^{1/2} \le C'_B\left(\frac{1}{R} + \frac{1}{RS}\right).$$

Using

$$\frac{\overline{A}_i^{R,S}}{\overline{B}_i^{R,S}} - \frac{a_i}{\beta} = \frac{1}{\overline{B}_i^{R,S}}\left[\left(\overline{A}_i^{R,S} - a_i\right) - \frac{a_i}{\beta}\left(\overline{B}_i^{R,S} - \beta\right)\right],$$

we obtain

$$\mathbb{E}\left\|\widehat{b}_i^{R,S}(t,\mathbf{x}) - b_i(t,\mathbf{x})\right\|^2 \le \frac{2}{\tau^2}\mathbb{E}\left[\frac{\left\|\overline{A}_i^{R,S} - a_i\right\|^2}{\left(\overline{B}_i^{R,S}\right)^2}\right] + \frac{2\|a_i\|^2}{\tau^2\beta^2}\mathbb{E}\left[\frac{\left|\overline{B}_i^{R,S} - \beta\right|^2}{\left(\overline{B}_i^{R,S}\right)^2}\right].$$

By Cauchy–Schwarz and Equation (47),

$$\mathbb{E}\left[\frac{\left\|\overline{A}_i^{R,S} - a_i\right\|^2}{\left(\overline{B}_i^{R,S}\right)^2}\right] \le \left(\mathbb{E}\left\|\overline{A}_i^{R,S} - a_i\right\|^4\right)^{1/2}\left(\mathbb{E}\left[\left(\overline{B}_i^{R,S}\right)^{-4}\right]\right)^{1/2},$$

and the same estimate applies to the denominator fluctuation. Substituting the fourth-moment bounds proves

$$\mathbb{E}\left[\left\|\widehat{b}_i^{R,S}(t,\mathbf{x}) - b_i(t,\mathbf{x})\right\|^2\right] \le C\left(\frac{1}{R} + \frac{1}{RS}\right).$$

□

C.1.2 EXACT-FEEDBACK NONMATCHED-COVARIANCE CHAIN AND ONE OUTER-LOOP ROLLOUT VALUE BOUND

Let

$$h = \frac{1}{M}, \qquad t_m = mh.$$

For a deterministic state $\mathbf{x}$, define the drifted state

$$\overline{\mathbf{x}}_h = \mathbf{x} + h\, b(t,\mathbf{x}).$$

Define the matched and execution continuation scores by

$$\Phi_h^{v,\mathrm{prop}}(t,\mathbf{x}) = \mathbb{E}\left[v_\varepsilon^{K,N}\left(t+h, \overline{\mathbf{x}}_h + \sqrt{h}\,\mathbf{Q}_{K,\mathrm{prop}}^{1/2} Z\right)\right],$$

and

$$\Phi_h^{v,\mathrm{dyn}}(t,\mathbf{x}) = \mathbb{E}\left[v_\varepsilon^{K,N}\left(t+h, \overline{\mathbf{x}}_h + \sqrt{h}\,\mathbf{Q}_{K,\mathrm{dyn}}^{1/2} Z\right)\right], \qquad Z \sim \mathcal{N}(0, I_{NK}).$$

The one-step covariance effect is the signed quantity

$$c_h^{K,N,\varepsilon}(t,\mathbf{x}) = \Phi_h^{v,\mathrm{dyn}}(t,\mathbf{x}) - \Phi_h^{v,\mathrm{prop}}(t,\mathbf{x}). \tag{48}$$

To display explicitly when the nonmatched covariance improves the matched continuation score, define the nonnegative one-step gain and unfavorable contribution

$$\begin{aligned} g_h^{K,N,\varepsilon}(t,\mathbf{x}) &:= \Big(-c_h^{K,N,\varepsilon}(t,\mathbf{x})\Big)_+ = \Big(\Phi_h^{v,\mathrm{prop}}(t,\mathbf{x}) - \Phi_h^{v,\mathrm{dyn}}(t,\mathbf{x})\Big)_+, \\ p_h^{K,N,\varepsilon}(t,\mathbf{x}) &:= \Big(c_h^{K,N,\varepsilon}(t,\mathbf{x})\Big)_+ = \Big(\Phi_h^{v,\mathrm{dyn}}(t,\mathbf{x}) - \Phi_h^{v,\mathrm{prop}}(t,\mathbf{x})\Big)_+. \end{aligned} \tag{49}$$

Then

$$c_h^{K,N,\varepsilon} = p_h^{K,N,\varepsilon} - g_h^{K,N,\varepsilon}. \tag{50}$$

Thus $g_h > 0$ precisely when the execution covariance produces a smaller one-step continuation score than matched execution, while $p_h > 0$ records the opposite event.

Define the matched running cost

$$\ell_\varepsilon^{K,N}(t,\mathbf{x}) = \frac{\varepsilon}{2}\,\|b(t,\mathbf{x})\|^2_{\mathbf{Q}^{-1}_{K,\mathrm{prop}}}. \tag{51}$$

The matched one-step weak Euler residual is

$$r_h^{K,N,\varepsilon}(t,\mathbf{x}) = \Phi_h^{v,\mathrm{prop}}(t,\mathbf{x}) + h\ell_\varepsilon^{K,N}(t,\mathbf{x}) - v_\varepsilon^{K,N}(t,\mathbf{x}). \tag{52}$$

Define the exact-feedback nonmatched-covariance chain by

$$\mathbf{X}_0 = \Pi_K \mathbf{x}_0, \qquad \overline{\mathbf{X}}_{m+1} = \mathbf{X}_m + h\,b(t_m,\mathbf{X}_m),$$

and

$$\mathbf{X}_{m+1} = \overline{\mathbf{X}}_{m+1} + \sqrt{h}\,\mathbf{Q}^{1/2}_{K,\mathrm{dyn}} Z_{m+1}, \qquad Z_{m+1} \sim \mathcal{N}(0, I_{NK}), \tag{53}$$

with independent increments. Set

$$\mathcal{E}_h^{K,N,\varepsilon} = \frac{1}{N}\sum_{m=0}^{M-1} \mathbb{E}\left[\Big(r_h^{K,N,\varepsilon}(t_m,\mathbf{X}_m)\Big)^+\right],$$

$$\mathcal{G}_h^{K,N,\varepsilon} = \frac{1}{N}\sum_{m=0}^{M-1} \mathbb{E}\left[g_h^{K,N,\varepsilon}(t_m,\mathbf{X}_m)\right],$$

and

$$\mathcal{P}_h^{K,N,\varepsilon} = \frac{1}{N}\sum_{m=0}^{M-1} \mathbb{E}\left[p_h^{K,N,\varepsilon}(t_m,\mathbf{X}_m)\right].$$

The cumulative signed covariance effect is therefore

$$\mathcal{C}_h^{K,N,\varepsilon} := \frac{1}{N}\sum_{m=0}^{M-1} \mathbb{E}\left[c_h^{K,N,\varepsilon}(t_m,\mathbf{X}_m)\right] = -\mathcal{G}_h^{K,N,\varepsilon} + \mathcal{P}_h^{K,N,\varepsilon}. \tag{54}$$

For later use, write the existing particle–projection–regularization bound, with the proposal covariance used in the matched problem, as

$$\mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon} := \frac{L_{\varepsilon,K}}{2N} + C^\star\left(\varepsilon^{2/\alpha}\,\mathrm{Tr}(Q_{\mathrm{prop}}) + \Delta_K^{\mathrm{proj}}\right)^{\alpha/2} + \frac{\varepsilon}{2} D^2_{\mathrm{CM},K,\mathrm{prop}}(\mu^\star) + \frac{\varepsilon K}{\alpha}\log\frac{1}{\varepsilon}. \tag{55}$$

Here, if

$$Q_{\mathrm{prop}} e_j = \lambda_{j,\mathrm{prop}} e_j,$$

then

$$D^2_{\mathrm{CM},K,\mathrm{prop}}(\mu^\star) = \sum_{j=1}^{K} \frac{\mathbb{E}_{\mu^\star}\left[(X_j - x_{0,j})^2\right]}{\lambda_{j,\mathrm{prop}}}, \qquad X \sim \mu^\star.$$

Theorem C.2. *Assume the hypotheses of Theorem B.7, with Q_{prop} as the covariance of the matched control problem, and assume that all one-step quantities above are integrable. Then*

$$\mathbb{E}G_K(\mu^N_{\mathbf{X}_M}) - \mathcal{V}_0 \le \mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon} + \mathcal{E}^{K,N,\varepsilon}_h - \mathcal{G}^{K,N,\varepsilon}_h + \mathcal{P}^{K,N,\varepsilon}_h. \tag{56}$$

In particular, the nonmatched covariance gives a negative covariance correction to $\mathcal{R}^{\mathrm{prop}} + \mathcal{E}_h$ whenever

$$\mathcal{G}^{K,N,\varepsilon}_h > \mathcal{P}^{K,N,\varepsilon}_h.$$

If, more strongly,

$$\mathcal{G}^{K,N,\varepsilon}_h > \mathcal{E}^{K,N,\varepsilon}_h + \mathcal{P}^{K,N,\varepsilon}_h,$$

then the complete one outer-loop rollout upper bound is smaller than the ideal matched particle–projection–regularization bound $\mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon}$.

Proof. Conditionally on $\mathbf{X}_m$,

$$\mathbb{E}\left[v^{K,N}_\varepsilon(t_{m+1}, \mathbf{X}_{m+1}) \mid \mathcal{F}_m\right] = \Phi^{v,\mathrm{dyn}}_h(t_m, \mathbf{X}_m).$$

By Equation (48) and Equation (50),

$$\Phi^{v,\mathrm{dyn}}_h(t_m, \mathbf{X}_m) = \Phi^{v,\mathrm{prop}}_h(t_m, \mathbf{X}_m) - g^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m) + p^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m).$$

Using Equation (52),

$$\begin{aligned}
&\mathbb{E}\left[v^{K,N}_\varepsilon(t_{m+1}, \mathbf{X}_{m+1}) \mid \mathcal{F}_m\right] + h\ell^{K,N}_\varepsilon(t_m, \mathbf{X}_m) \\
&\quad \le v^{K,N}_\varepsilon(t_m, \mathbf{X}_m) + \Big(r^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m)\Big)^+ - g^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m) + p^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m).
\end{aligned}$$

Taking expectations and summing over $m = 0, \dots, M-1$ yields

$$\begin{aligned}
&\mathbb{E}v^{K,N}_\varepsilon(1, \mathbf{X}_M) + \sum_{m=0}^{M-1} h\,\mathbb{E}\ell^{K,N}_\varepsilon(t_m, \mathbf{X}_m) \\
&\qquad \le v^{K,N}_\varepsilon\big(0, \Pi_K \mathbf{x}_0\big) + \sum_{m=0}^{M-1} \mathbb{E}\left[\Big(r^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m)\Big)^+\right] \\
&\qquad\quad - \sum_{m=0}^{M-1} \mathbb{E}\left[g^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m)\right] + \sum_{m=0}^{M-1} \mathbb{E}\left[p^{K,N,\varepsilon}_h(t_m, \mathbf{X}_m)\right].
\end{aligned}$$

Since

$$v^{K,N}_\varepsilon(1, \mathbf{X}_M) = F_{K,N}(\mathbf{X}_M) = N\,G_K(\mu^N_{\mathbf{X}_M}),$$

and the running cost is nonnegative, division by N, subtraction of $\mathcal{V}_0$, and Theorem B.7 prove the claim. □

Proposition C.3 (A sufficient condition for a vanishing matched Euler residual)**.** *For $t < 1$ and $\mathbf{x} \in (\mathcal{H}_K)^N$, define the frozen-feedback matched process*

$$\mathbf{Y}^{t,\mathbf{x}}_s = \mathbf{x} + s\,b(t, \mathbf{x}) + \mathbf{Q}^{1/2}_{K,\mathrm{prop}}\mathbf{W}_s, \qquad 0 \le s \le h.$$

Assume that the verification theorem for $v^{K,N}_\varepsilon$ holds. Then

$$r^{K,N,\varepsilon}_h(t, \mathbf{x}) = \frac{\varepsilon}{2}\mathbb{E}\int_0^h \left\|b(t, \mathbf{x}) - b(t+s, \mathbf{Y}^{t,\mathbf{x}}_s)\right\|^2_{\mathbf{Q}^{-1}_{K,\mathrm{prop}}}\, ds. \tag{57}$$

In particular,

$$r^{K,N,\varepsilon}_h(t, \mathbf{x}) \ge 0.$$

Consider a sequence $(K_j, N_j, \varepsilon_j, h_j)$, and write

$$b_j = b^{K_j,N_j}_{\varepsilon_j}, \qquad \mathbf{Q}_{j,\mathrm{prop}} = I_{N_j} \otimes Q_{K_j,\mathrm{prop}}.$$

Assume that there exist $\zeta \in (0,1]$ and constants L_j such that

$$\frac{1}{N_j}\left\|b_j(s,\mathbf{y}) - b_j(t,\mathbf{x})\right\|^2_{\mathbf{Q}^{-1}_{j,\mathrm{prop}}} \le L_j^2\left(|s-t|^{2\zeta} + \frac{1}{N_j}\|\mathbf{y}-\mathbf{x}\|^2\right).$$

Define

$$B_j = \frac{1}{N_j}\sum_{m=0}^{M_j-1} h_j \mathbb{E}\left[\left\|b_j(t_m,\mathbf{X}_m)\right\|^2\right].$$

Then

$$\mathcal{E}_{h_j}^{K_j,N_j,\varepsilon_j} \le \frac{\varepsilon_j L_j^2}{2}\left[\frac{h_j^{2\zeta}}{2\zeta+1} + \frac{h_j}{2}\operatorname{Tr}(Q_{K_j,\mathrm{prop}}) + \frac{h_j^2}{3}B_j\right].$$

Consequently,

$$\varepsilon_j L_j^2\left[h_j^{2\zeta} + h_j \operatorname{Tr}(Q_{K_j,\mathrm{prop}}) + h_j^2 B_j\right] \longrightarrow 0$$

implies

$$\mathcal{E}_{h_j}^{K_j,N_j,\varepsilon_j} \longrightarrow 0.$$

Proof. For brevity, write

$$v = v_\varepsilon^{K,N}, \qquad \mathbf{Q} = \mathbf{Q}_{K,\mathrm{prop}}.$$

The matched particle HJB is

$$\partial_t v + \frac{1}{2}\operatorname{Tr}(\mathbf{Q}D^2 v) - \frac{1}{2\varepsilon}\langle Dv, \mathbf{Q}Dv\rangle = 0,$$

and the matched feedback is

$$b(t,\mathbf{x}) = -\frac{1}{\varepsilon}\mathbf{Q}Dv(t,\mathbf{x}).$$

At a generic point $(s,\mathbf{y})$, for any $\boldsymbol{\theta} \in (\mathcal{H}_K)^N$,

$$\frac{\varepsilon}{2}\|\boldsymbol{\theta} - b(s,\mathbf{y})\|^2_{\mathbf{Q}^{-1}} = \frac{\varepsilon}{2}\|\boldsymbol{\theta}\|^2_{\mathbf{Q}^{-1}} + \langle \boldsymbol{\theta}, Dv(s,\mathbf{y})\rangle + \frac{1}{2\varepsilon}\langle Dv(s,\mathbf{y}), \mathbf{Q}Dv(s,\mathbf{y})\rangle.$$

Using the HJB equation in the final term gives

$$\partial_t v(s,\mathbf{y}) + \langle \boldsymbol{\theta}, Dv(s,\mathbf{y})\rangle + \frac{1}{2}\operatorname{Tr}\left(\mathbf{Q}D^2 v(s,\mathbf{y})\right) + \frac{\varepsilon}{2}\|\boldsymbol{\theta}\|^2_{\mathbf{Q}^{-1}} = \frac{\varepsilon}{2}\|\boldsymbol{\theta} - b(s,\mathbf{y})\|^2_{\mathbf{Q}^{-1}}.$$

Fix $(t,\mathbf{x})$, set

$$\boldsymbol{\theta}_0 = b(t,\mathbf{x}),$$

and define the frozen-feedback process

$$\mathbf{Y}_\rho^{t,\mathbf{x}} = \mathbf{x} + \rho\boldsymbol{\theta}_0 + \mathbf{Q}^{1/2}\mathbf{B}_\rho, \qquad 0 \le \rho \le h,$$

where $\mathbf{B}$ is a standard Brownian motion. Applying Itô's formula to

$$\rho \longmapsto v(t+\rho, \mathbf{Y}_\rho^{t,\mathbf{x}})$$

and using the preceding completion-of-squares identity yields

$$\mathbb{E}v(t+h,\mathbf{Y}_h^{t,\mathbf{x}}) + \frac{\varepsilon h}{2}\|b(t,\mathbf{x})\|^2_{\mathbf{Q}^{-1}} - v(t,\mathbf{x}) = \frac{\varepsilon}{2}\mathbb{E}\int_0^h \left\|b(t,\mathbf{x}) - b(t+\rho,\mathbf{Y}_\rho^{t,\mathbf{x}})\right\|^2_{\mathbf{Q}^{-1}} d\rho.$$

Since

$$\mathbf{Y}_h^{t,\mathbf{x}} \overset{d}{=} \mathbf{x} + h\, b(t,\mathbf{x}) + \sqrt{h}\,\mathbf{Q}^{1/2}Z,$$

the left-hand side is exactly $r_h^{K,N,\varepsilon}(t,\mathbf{x})$. This proves Equation (57).

For the joint sequence, conditionally on $\mathbf{X}_m$, define

$$\mathbf{Y}_{m,\rho}^{(j)} = \mathbf{X}_m + \rho\, b_j(t_m,\mathbf{X}_m) + \mathbf{Q}_{j,\mathrm{prop}}^{1/2}\mathbf{B}_{m,\rho}, \qquad 0 \le \rho \le h_j,$$

where $\mathbf{B}_{m,\cdot}$ is independent of $\mathcal{F}_{t_m}$. Then

$$\mathbf{Y}_{m,\rho}^{(j)} - \mathbf{X}_m = \rho\, b_j(t_m, \mathbf{X}_m) + \mathbf{Q}_{j,\mathrm{prop}}^{1/2} \mathbf{B}_{m,\rho}.$$

The Brownian increment has conditional mean zero, so the cross term vanishes, and

$$\frac{1}{N_j}\mathbb{E}\left[\left\|\mathbf{Y}_{m,\rho}^{(j)} - \mathbf{X}_m\right\|^2 \middle| \mathbf{X}_m\right] = \frac{\rho^2}{N_j}\|b_j(t_m, \mathbf{X}_m)\|^2 + \rho\,\mathrm{Tr}(Q_{K_j,\mathrm{prop}}),$$

where we used

$$\mathrm{Tr}(\mathbf{Q}_{j,\mathrm{prop}}) = N_j\,\mathrm{Tr}(Q_{K_j,\mathrm{prop}}).$$

The assumed Hölder–Lipschitz estimate therefore gives

$$\frac{1}{N_j}\mathbb{E}\left[\left\|b_j(t_m+\rho, \mathbf{Y}_{m,\rho}^{(j)}) - b_j(t_m, \mathbf{X}_m)\right\|^2_{\mathbf{Q}_{j,\mathrm{prop}}^{-1}}\right] \le L_j^2\left[\rho^{2\zeta} + \frac{\rho^2}{N_j}\mathbb{E}\|b_j(t_m, \mathbf{X}_m)\|^2 + \rho\,\mathrm{Tr}(Q_{K_j,\mathrm{prop}})\right].$$

Substituting this estimate into the frozen-feedback identity and integrating over $\rho \in [0, h_j]$ yields

$$\frac{1}{N_j}\mathbb{E}\left[r_{h_j}^{K_j,N_j,\varepsilon_j}(t_m, \mathbf{X}_m)\right] \le \frac{\varepsilon_j L_j^2}{2}\left[\frac{h_j^{2\zeta+1}}{2\zeta+1} + \frac{h_j^3}{3N_j}\mathbb{E}\|b_j(t_m, \mathbf{X}_m)\|^2 + \frac{h_j^2}{2}\,\mathrm{Tr}(Q_{K_j,\mathrm{prop}})\right].$$

Finally, summing over $m = 0, \dots, M_j - 1$, using $M_j h_j = 1$, and recalling

$$B_j = \frac{1}{N_j}\sum_{m=0}^{M_j-1} h_j\,\mathbb{E}\|b_j(t_m, \mathbf{X}_m)\|^2,$$

gives

$$\mathcal{E}_{h_j}^{K_j,N_j,\varepsilon_j} \le \frac{\varepsilon_j L_j^2}{2}\left[\frac{h_j^{2\zeta}}{2\zeta+1} + \frac{h_j}{2}\,\mathrm{Tr}(Q_{K_j,\mathrm{prop}}) + \frac{h_j^2}{3}B_j\right].$$

□

Remark C.1 (Smooth interpretation and magnitude of the covariance effect)**.** *Suppose $v_\varepsilon^{K,N}$ is sufficiently smooth for the following second-order Gaussian expansion to hold uniformly along the states visited by the chain. Define*

$$\chi_v(t, \mathbf{x}) := \mathrm{Tr}\left([\mathbf{Q}_{K,\mathrm{dyn}} - \mathbf{Q}_{K,\mathrm{prop}}]\, D^2 v_\varepsilon^{K,N}(t, \mathbf{x})\right).$$

Then

$$c_h^{K,N,\varepsilon}(t, \mathbf{x}) = \frac{h}{2}\chi_v(t+h, \overline{\mathbf{x}}_h) + o(h).$$

Since the positive-part map is Lipschitz,

$$g_h^{K,N,\varepsilon}(t, \mathbf{x}) = \frac{h}{2}\left(-\chi_v(t+h, \overline{\mathbf{x}}_h)\right)_+ + o(h),$$
$$p_h^{K,N,\varepsilon}(t, \mathbf{x}) = \frac{h}{2}\left(\chi_v(t+h, \overline{\mathbf{x}}_h)\right)_+ + o(h).$$

Consequently,

$$\mathcal{G}_h^{K,N,\varepsilon} = \frac{1}{2N}\sum_{m=0}^{M-1} h\,\mathbb{E}\left[\left(-\chi_v(t_{m+1}, \overline{\mathbf{X}}_{m+1})\right)_+\right] + o(1),$$
$$\mathcal{P}_h^{K,N,\varepsilon} = \frac{1}{2N}\sum_{m=0}^{M-1} h\,\mathbb{E}\left[\left(\chi_v(t_{m+1}, \overline{\mathbf{X}}_{m+1})\right)_+\right] + o(1).$$

If each one-step remainder is $O(h^2)$, then the accumulated remainder is $O(h)$.

For proportional covariances, let

$$\kappa_v(t, \mathbf{x}) = \mathrm{Tr}\left(\mathbf{Q}_{K,\mathrm{prop}} D^2 v_\varepsilon^{K,N}(t, \mathbf{x})\right).$$

Then

$$\chi_v = (\alpha_0 - 1)\kappa_v.$$

If the execution covariance is colder, $\alpha_0 < 1$, then

$$g_h^{K,N,\varepsilon}(t,\mathbf{x}) = \frac{h}{2}(1-\alpha_0)\left(\kappa_v(t+h,\overline{\mathbf{x}}_h)\right)_+ + o(h),$$
$$p_h^{K,N,\varepsilon}(t,\mathbf{x}) = \frac{h}{2}(1-\alpha_0)\left(-\kappa_v(t+h,\overline{\mathbf{x}}_h)\right)_+ + o(h).$$

Thus positive continuation-value curvature contributes to the negative gain term, while negative curvature contributes to the unfavorable term. If the execution covariance is hotter, $\alpha_0 > 1$, the roles of the positive and negative curvature regions are reversed:

$$g_h^{K,N,\varepsilon}(t,\mathbf{x}) = \frac{h}{2}(\alpha_0-1)\left(-\kappa_v(t+h,\overline{\mathbf{x}}_h)\right)_+ + o(h),$$
$$p_h^{K,N,\varepsilon}(t,\mathbf{x}) = \frac{h}{2}(\alpha_0-1)\left(\kappa_v(t+h,\overline{\mathbf{x}}_h)\right)_+ + o(h).$$

Therefore a fixed cold covariance improves the approximation whenever positive weighted curvature dominates negative weighted curvature along the visited path; a fixed hot covariance improves it under the opposite curvature balance.

Proposition C.4 (Continuous-time covariance identity)**.** *Assume that $v_\varepsilon^{K,N}$ is $C^{1,2}$ and that the matched closed-loop feedback*

$$b(t,\mathbf{x}) = -\frac{1}{\varepsilon}\mathbf{Q}_{K,\mathrm{prop}} D v_\varepsilon^{K,N}(t,\mathbf{x})$$

is admissible. Consider

$$d\mathbf{X}_t^{\mathrm{dyn}} = b(t,\mathbf{X}_t^{\mathrm{dyn}})\,dt + \mathbf{Q}_{K,\mathrm{dyn}}^{1/2} d\mathbf{W}_t, \qquad \mathbf{X}_0^{\mathrm{dyn}} = \Pi_K \mathbf{x}_0.$$

Assume that the covariance SDE above admits a nonexplosive strong solution and that

$$\mathbb{E}\int_0^1 \left[\ell_\varepsilon^{K,N}(t,X_t^{\mathrm{dyn}}) + |\chi_v(t,X_t^{\mathrm{dyn}})| + \left\| Q_{K,\mathrm{dyn}}^{1/2} D v_\varepsilon^{K,N}(t,X_t^{\mathrm{dyn}})\right\|^2\right] dt < \infty.$$

Set

$$\chi_t := \mathrm{Tr}\left(\left[\mathbf{Q}_{K,\mathrm{dyn}} - \mathbf{Q}_{K,\mathrm{prop}}\right] D^2 v_\varepsilon^{K,N}(t,\mathbf{X}_t^{\mathrm{dyn}})\right).$$

Then

$$\mathbb{E}F_{K,N}(\mathbf{X}_1^{\mathrm{dyn}}) + \mathbb{E}\int_0^1 \ell_\varepsilon^{K,N}(t,\mathbf{X}_t^{\mathrm{dyn}})\,dt = v_\varepsilon^{K,N}\big(0,\Pi_K\mathbf{x}_0\big) + \frac{1}{2}\mathbb{E}\int_0^1 \chi_t\,dt.$$

Define the nonnegative continuous-time gain and unfavorable contribution

$$\mathfrak{G}^{K,N,\varepsilon} := \frac{1}{2N}\mathbb{E}\int_0^1 (-\chi_t)_+\,dt, \qquad \mathfrak{P}^{K,N,\varepsilon} := \frac{1}{2N}\mathbb{E}\int_0^1 (\chi_t)_+\,dt.$$

Then

$$\mathfrak{R}^{K,N,\varepsilon} := \frac{1}{2N}\mathbb{E}\int_0^1 \chi_t\,dt = -\mathfrak{G}^{K,N,\varepsilon} + \mathfrak{P}^{K,N,\varepsilon}, \tag{58}$$

and consequently

$$\mathbb{E}G_K(\mu^N_{\mathbf{X}_1^{\mathrm{dyn}}}) - \mathcal{V}_0 \le \mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon} - \mathfrak{G}^{K,N,\varepsilon} + \mathfrak{P}^{K,N,\varepsilon}. \tag{59}$$

In the proportional case,

$$\chi_t = (\alpha_0 - 1)\,\mathrm{Tr}\left(\mathbf{Q}_{K,\mathrm{prop}} D^2 v_\varepsilon^{K,N}(t,\mathbf{X}_t^{\mathrm{dyn}})\right).$$

Proof. The matched particle HJB can be written in closed-loop form as

$$\partial_t v_\varepsilon^{K,N} + \left\langle b, D v_\varepsilon^{K,N}\right\rangle + \frac{1}{2}\,\mathrm{Tr}\left(\mathbf{Q}_{K,\mathrm{prop}} D^2 v_\varepsilon^{K,N}\right) + \ell_\varepsilon^{K,N} = 0.$$

Applying Itô's formula to $v_\varepsilon^{K,N}(t, \mathbf{X}_t^{\mathrm{dyn}})$ gives

$$dv_\varepsilon^{K,N}(t, \mathbf{X}_t^{\mathrm{dyn}}) = \left[-\ell_\varepsilon^{K,N}(t, \mathbf{X}_t^{\mathrm{dyn}}) + \frac{1}{2}\chi_t\right] dt + \left\langle Dv_\varepsilon^{K,N}, \mathbf{Q}_{K,\mathrm{dyn}}^{1/2} d\mathbf{W}_t \right\rangle.$$

Integrate, take expectations, and use the terminal condition. Dividing by N, dropping the nonnegative running cost, and applying Theorem B.7 yield Equation (59). The decomposition follows from $\chi_t = (\chi_t)_+ - (-\chi_t)_+$. □

Remark C.2 (Interpretation and vanishing conditions for the covariance terms)**.** *The negative term $-\mathfrak{G}^{K,N,\varepsilon}$ is the part of the covariance nonmatch that lowers the matched continuation-value certificate, while $\mathfrak{P}^{K,N,\varepsilon}$ is the unfavorable part. Hence nonmatched execution improves the matched continuous-time upper bound whenever $\mathfrak{G}^{K,N,\varepsilon} > \mathfrak{P}^{K,N,\varepsilon}$.*

For a fixed cold covariance $\alpha_0 < 1$, writing

$$\kappa_v(t, \mathbf{x}) = \mathrm{Tr}\left(\mathbf{Q}_{K,\mathrm{prop}} D^2 v_\varepsilon^{K,N}(t, \mathbf{x})\right),$$

one has

$$\mathfrak{G}^{K,N,\varepsilon} = \frac{1-\alpha_0}{2N}\mathbb{E}\int_0^1 \left(\kappa_v(t, \mathbf{X}_t^{\mathrm{dyn}})\right)_+ dt,$$
$$\mathfrak{P}^{K,N,\varepsilon} = \frac{1-\alpha_0}{2N}\mathbb{E}\int_0^1 \left(-\kappa_v(t, \mathbf{X}_t^{\mathrm{dyn}})\right)_+ dt.$$

Thus positive weighted curvature contributes directly to the negative gain term. For a hot covariance, the roles of positive and negative curvature are reversed.

Moreover,

$$\mathfrak{G}^{K,N,\varepsilon} + \mathfrak{P}^{K,N,\varepsilon} \le \frac{|\alpha_0 - 1|}{2N}\mathbb{E}\int_0^1 \left|\kappa_v(t, \mathbf{X}_t^{\mathrm{dyn}})\right| dt.$$

Hence both terms vanish along a joint sequence whenever the total weighted particle curvature is $o(N)$, after accounting for any dependence on K and ε. Vanishing is not required for improvement: a strictly positive excess of gain over unfavorable curvature yields a strictly better upper bound.

C.1.3 Consistency of the fully finite one outer-loop rollout implementation

For $\mathbf{x} = (x_1, \ldots, x_N) \in (\mathcal{H}_K)^N$, write $\mathcal{J}_{K,N}(\mathbf{x}) := G_K(\mu_{\mathbf{x}}^N) = \frac{1}{N} F_{K,N}(\mathbf{x})$. For each Euler step m, let $\Xi_m^{R,S}$ denote the complete collection of fresh context and candidate variables used to construct the multi-context estimator, i.e.,

$$\Xi_m^{R,S} := \left(\left(Z_{m,p}^{(r)}\right)_{\substack{1\le r\le R\\ 1\le p\le N}}, \left(\xi_{m,i,r,s}\right)_{\substack{1\le i\le N\\ 1\le r\le R\\ 1\le s\le S}}\right),$$

where all $Z_{m,p}^{(r)},\ \xi_{m,i,r,s} \sim \mathcal{N}(0, I_K)$ are mutually independent.

Write $\widehat{b}_m^{R,S}(\mathbf{x}) = \widehat{b}^{R,S}(t_m, \mathbf{x}; \Xi_m^{R,S}) \in (\mathcal{H}_K)^N$. The variables $\Xi_0^{R,S}, \ldots, \Xi_{M-1}^{R,S}, Z_1^{\mathrm{ex}}, \ldots, Z_M^{\mathrm{ex}}$ are mutually independent, and $Z_m^{\mathrm{ex}} \sim \mathcal{N}(0, I_{NK})$.

Define the finite-candidate chain by

$$\begin{cases} \mathbf{C}_0^{R,S,h} = \Pi_K \mathbf{x}_0, \\ \mathbf{C}_{m+1}^{R,S,h} = \mathbf{C}_m^{R,S,h} + h\widehat{b}_m^{R,S}\left(\mathbf{C}_m^{R,S,h}\right) + \sqrt{h}\,\mathbf{Q}_{K,\mathrm{dyn}}^{1/2} Z_{m+1}^{\mathrm{ex}}. \end{cases} \tag{60}$$

On the same probability space, using the same execution noises Z_{m+1}^{ex}, define the exact-feedback nonmatched-covariance chain by

$$\begin{cases} \mathbf{X}_0^h = \Pi_K \mathbf{x}_0, \\ \mathbf{X}_{m+1}^h = \mathbf{X}_m^h + hb(t_m, \mathbf{X}_m^h) + \sqrt{h}\,\mathbf{Q}_{K,\mathrm{dyn}}^{1/2} Z_{m+1}^{\mathrm{ex}}. \end{cases} \tag{61}$$

Set

$$\delta_{R,S} := \frac{1}{R} + \frac{1}{RS}.$$

Assumption C.1 (Quantitative stability of the finite implementation)**.** *Fix $K, N, \varepsilon > 0$ and $h = 1/M > 0$. Assume the following.*

1. *There exist constants $L_b, B_b < \infty$ such that, for every $m = 0, \dots, M-1$ and every $\mathbf{x}, \mathbf{y} \in (\mathcal{H}_K)^N$,*
$$\|b(t_m, \mathbf{x}) - b(t_m, \mathbf{y})\| \le L_b \|\mathbf{x} - \mathbf{y}\|,$$
and
$$\|b(t_m, \mathbf{x})\|^2 \le B_b \left(1 + \|\mathbf{x}\|^2\right).$$
2. *There exists $C_{\mathrm{MC}} < \infty$, independent of m, R, S, and $\mathbf{x}$, such that*
$$\mathbb{E}_{\Xi_m^{R,S}} \left[\left\| \widehat{b}_m^{R,S}(\mathbf{x}) - b(t_m, \mathbf{x}) \right\|^2 \right] \le C_{\mathrm{MC}} \delta_{R,S} \left(1 + \|\mathbf{x}\|^2\right). \tag{62}$$
Here the expectation is only over the fresh context and candidate variables at step m.
3. *The terminal cloud functional is globally Lipschitz: there exists $L_{\mathcal{J}} < \infty$ such that*
$$|\mathcal{J}_{K,N}(\mathbf{x}) - \mathcal{J}_{K,N}(\mathbf{y})| \le L_{\mathcal{J}} \|\mathbf{x} - \mathbf{y}\|, \qquad \mathbf{x}, \mathbf{y} \in (\mathcal{H}_K)^N. \tag{63}$$

Theorem C.5 (Strong consistency of finite-candidate SCMO)**.** *Under Assumption C.1, there exists a constant*
$$C_{\mathrm{stab}} = C_{\mathrm{stab}}\left(K, N, \varepsilon, h, L_b, B_b, C_{\mathrm{MC}}, \mathbf{Q}_{K,\mathrm{dyn}}, \mathbf{x}_0\right) < \infty$$
which is independent of R and S, such that
$$\max_{0 \le m \le M} \mathbb{E}\left[\left\| \mathbf{C}_m^{R,S,h} - \mathbf{X}_m^h \right\|^2 \right] \le C_{\mathrm{stab}} \left(\frac{1}{R} + \frac{1}{RS} \right). \tag{64}$$
Consequently,
$$\mathbf{C}_m^{R,S,h} \longrightarrow \mathbf{X}_m^h$$
in L^2, and hence in probability, for every $m \le M$, whenever
$$\delta_{R,S} \longrightarrow 0.$$
Moreover,
$$\left| \mathbb{E} G_K\left(\mu^N_{\mathbf{C}_M^{R,S,h}}\right) - \mathbb{E} G_K\left(\mu^N_{\mathbf{X}_M^h}\right) \right| \le L_{\mathcal{J}} \sqrt{C_{\mathrm{stab}} \delta_{R,S}}. \tag{65}$$
It follows that
$$\limsup_{\delta_{R,S} \to 0} \left[\mathbb{E} G_K\left(\mu^N_{\mathbf{C}_M^{R,S,h}}\right) - \mathcal{V}_0 \right] \le \mathcal{R}^{\mathrm{prop}}_{K,N,\varepsilon} + \mathcal{E}_h^{K,N,\varepsilon} - \mathcal{G}_h^{K,N,\varepsilon} + \mathcal{P}_h^{K,N,\varepsilon}.$$

Proof. We first establish a moment bound for the finite chain. By Equation (62) and the linear-growth bound on b,
$$\mathbb{E}_{\Xi_m^{R,S}} \left[\left\| \widehat{b}_m^{R,S}(\mathbf{x}) \right\|^2 \right] \le 2\|b(t_m, \mathbf{x})\|^2 + 2\mathbb{E}_{\Xi_m^{R,S}} \left[\left\| \widehat{b}_m^{R,S}(\mathbf{x}) - b(t_m, \mathbf{x}) \right\|^2 \right] \le B_{\widehat{b}} \left(1 + \|\mathbf{x}\|^2\right),$$
where one may take
$$B_{\widehat{b}} = 2B_b + 4C_{\mathrm{MC}},$$
because
$$\delta_{R,S} = \frac{1}{R} + \frac{1}{RS} \le 2.$$
For arbitrary vectors x, y and $h \in (0, 1]$,
$$\|x + hy\|^2 \le (1 + h)\|x\|^2 + (h + h^2)\|y\|^2.$$
Using the independence and zero mean of the execution noise gives
$$\mathbb{E}\left[\left\| \mathbf{C}_{m+1}^{R,S,h} \right\|^2 \right] \le (1 + h)\mathbb{E}\left[\left\| \mathbf{C}_m^{R,S,h} \right\|^2 \right] + (h + h^2)\mathbb{E}\left[\left\| \widehat{b}_m^{R,S}\left(\mathbf{C}_m^{R,S,h}\right) \right\|^2 \right] + h \operatorname{Tr}\left(\mathbf{Q}_{K,\mathrm{dyn}}\right).$$

Hence there is $c_0 < \infty$, independent of R, S, such that

$$\mathbb{E}\left[\left\|\mathbf{C}_{m+1}^{R,S,h}\right\|^2\right] \le (1 + c_0 h)\mathbb{E}\left[\left\|\mathbf{C}_m^{R,S,h}\right\|^2\right] + c_0 h.$$

The discrete Gronwall inequality yields

$$\sup_{R,S\ge 1} \max_{0\le m\le M} \mathbb{E}\left[\left\|\mathbf{C}_m^{R,S,h}\right\|^2\right] \le K_C \tag{66}$$

for some $K_C < \infty$.

Set

$$D_m^{R,S} = \mathbf{C}_m^{R,S,h} - \mathbf{X}_m^h$$

and

$$e_m^{R,S} = \widehat{b}_m^{R,S}\left(\mathbf{C}_m^{R,S,h}\right) - b\left(t_m, \mathbf{C}_m^{R,S,h}\right).$$

Because the exact and finite chains use the same execution noise, the noise terms cancel and

$$D_{m+1}^{R,S} = D_m^{R,S} + h\left[b\left(t_m, \mathbf{C}_m^{R,S,h}\right) - b(t_m, \mathbf{X}_m^h)\right] + h e_m^{R,S}.$$

Define

$$A_m^{R,S} = D_m^{R,S} + h\left[b\left(t_m, \mathbf{C}_m^{R,S,h}\right) - b(t_m, \mathbf{X}_m^h)\right].$$

The Lipschitz condition gives

$$\|A_m^{R,S}\| \le (1 + hL_b)\|D_m^{R,S}\|.$$

Using

$$\|a + he\|^2 \le (1 + h)\|a\|^2 + (h + h^2)\|e\|^2,$$

we obtain

$$\mathbb{E}\left[\left\|D_{m+1}^{R,S}\right\|^2\right] \le \rho_h \mathbb{E}\left[\left\|D_m^{R,S}\right\|^2\right] + (h + h^2)\mathbb{E}\left[\left\|e_m^{R,S}\right\|^2\right],$$

where

$$\rho_h = (1 + h)(1 + hL_b)^2.$$

Conditioning on $\mathbf{C}_m^{R,S,h}$, using the independence of the fresh context and candidate samples, and applying Equation (62), we get

$$\mathbb{E}\left[\left\|e_m^{R,S}\right\|^2\right] = \mathbb{E}\left[\mathbb{E}\left[\left\|e_m^{R,S}\right\|^2 \middle| \mathbf{C}_m^{R,S,h}\right]\right] \le C_{\mathrm{MC}}\delta_{R,S}\left[1 + \mathbb{E}\left\|\mathbf{C}_m^{R,S,h}\right\|^2\right] \le C_{\mathrm{MC}}(1 + K_C)\delta_{R,S}.$$

Therefore,

$$\mathbb{E}\left[\left\|D_{m+1}^{R,S}\right\|^2\right] \le \rho_h \mathbb{E}\left[\left\|D_m^{R,S}\right\|^2\right] + (h + h^2)C_{\mathrm{MC}}(1 + K_C)\delta_{R,S}.$$

Since $D_0^{R,S} = 0$, iteration gives

$$\mathbb{E}\left[\left\|D_m^{R,S}\right\|^2\right] \le (h + h^2)C_{\mathrm{MC}}(1 + K_C)\left(\sum_{q=0}^{m-1} \rho_h^q\right)\delta_{R,S}.$$

Thus Equation (64) holds with

$$C_{\mathrm{stab}} = (h + h^2)C_{\mathrm{MC}}(1 + K_C)\sum_{q=0}^{M-1} \rho_h^q.$$

Finally, the Lipschitz condition on $\mathcal{J}_{K,N}$ gives

$$\left|\mathbb{E}\mathcal{J}_{K,N}\left(\mathbf{C}_M^{R,S,h}\right) - \mathbb{E}\mathcal{J}_{K,N}\left(\mathbf{X}_M^h\right)\right| \le L_{\mathcal{J}}\mathbb{E}\left[\left\|\mathbf{C}_M^{R,S,h} - \mathbf{X}_M^h\right\|\right] \le L_{\mathcal{J}}\sqrt{C_{\mathrm{stab}}\delta_{R,S}}.$$

Combining this convergence with Theorem C.2 proves the final assertion. □

Remark C.3 (A directly checkable smooth sufficient condition)**.** *Suppose*

$$F_{K,N} \in C_b^2((\mathcal{H}_K)^N),$$

and write

$$M_1 = \sup_{\mathbf{x}} \|DF_{K,N}(\mathbf{x})\|, \qquad M_2 = \sup_{\mathbf{x}} \|D^2 F_{K,N}(\mathbf{x})\|_{\mathrm{op}}.$$

Assume also that

$$F_{\min} \le F_{K,N}(\mathbf{x}) \le F_{\max}, \qquad \mathbf{x} \in (H_K)^N.$$

Let

$$\pi_{t,\mathbf{x}}(d\mathbf{u}) = \frac{\exp(-F_{K,N}(\mathbf{x}+\mathbf{u})/\varepsilon)}{\mathbb{E}[\exp(-F_{K,N}(\mathbf{x}+\mathbf{U})/\varepsilon)]}\, \mathcal{N}\left(0, (1-t)\mathbf{Q}_{K,\mathrm{prop}}\right)(d\mathbf{u}).$$

Gaussian integration by parts gives

$$b(t,\mathbf{x}) = -\frac{1}{\varepsilon}\mathbf{Q}_{K,\mathrm{prop}}\mathbb{E}_{\pi_{t,\mathbf{x}}}\left[DF_{K,N}(\mathbf{x}+\mathbf{U})\right].$$

Consequently,

$$\|b(t,\mathbf{x})\| \le \frac{\|\mathbf{Q}_{K,\mathrm{prop}}\|_{\mathrm{op}} M_1}{\varepsilon},$$

and

$$\|D_{\mathbf{x}} b(t,\mathbf{x})\|_{\mathrm{op}} \le \|\mathbf{Q}_{K,\mathrm{prop}}\|_{\mathrm{op}} \left(\frac{M_2}{\varepsilon} + \frac{M_1^2}{\varepsilon^2}\right).$$

Thus the feedback is globally Lipschitz and bounded.

Moreover,

$$e^{-F_{\max}/\varepsilon} \le L_{t,\mathbf{x}}(\mathbf{U}) \le e^{-F_{\min}/\varepsilon}.$$

Hence every empirical Gibbs denominator has the deterministic lower bound

$$\overline{B}_i^{R,S} \ge e^{-F_{\max}/\varepsilon},$$

and the inverse-denominator condition holds automatically. Since the candidate increments are Gaussian, all required numerator moments are finite, and the uniform estimate

$$\mathbb{E}\left[\left\|\widehat{b}_m^{R,S}(\mathbf{x}) - b(t_m,\mathbf{x})\right\|^2\right] \le C_{\mathrm{MC}}\left(\frac{1}{R} + \frac{1}{RS}\right)$$

holds with a constant independent of $\mathbf{x}, m, R, S$.

Finally,

$$\mathcal{J}_{K,N} = \frac{1}{N} F_{K,N}$$

is globally Lipschitz with

$$L_{\mathcal{J}} \le \frac{M_1}{N}.$$

Therefore Assumption C.1 holds.

Remark C.4 (Localized polynomial-growth conditions and interacting objectives)**.** *The globally bounded assumptions in Remark C.3 are convenient but not necessary. They can be replaced by local regularity, polynomial growth, denominator stability, and moment bounds for the finite and exact chains.*

Fix $K, N, \varepsilon > 0$ *and* $h = 1/M > 0$*, and write*

$$E_{K,N} = (\mathcal{H}_K)^N.$$

For $m = 0, \dots, M-1$*, let*

$$\mathbf{U}_m \sim \mathcal{N}\left(0, \tau_m \mathbf{Q}_{K,\mathrm{prop}}\right), \qquad \tau_m = 1 - t_m.$$

For the finite-candidate estimator, write

$$\overline{B}_{m,i}^{R,S}(\mathbf{x}) = \frac{1}{R}\sum_{r=1}^{R}\left[\frac{1}{S}\sum_{s=1}^{S} L_{m,i,r,s}(\mathbf{x})\right]$$

for its empirical Gibbs denominator, and define the terminal cloud functional

$$\mathcal{J}_{K,N}(\mathbf{x}) = G_K(\mu_{\mathbf{x}}^N).$$

Assume the following conditions.

(i) Structural regularity and polynomial growth. Suppose that

$$F_{K,N} \in C^2(E_{K,N})$$

is bounded from below and satisfies, for some $q \geq 0$,

$$|F_{K,N}(\mathbf{x})| \leq C_F \left(1 + \|\mathbf{x}\|^2\right), \qquad \mathbf{x} \in E_{K,N}, \tag{67}$$

and

$$\|DF_{K,N}(\mathbf{x})\| + \left\|D^2 F_{K,N}(\mathbf{x})\right\|_{\mathrm{op}} \leq C_F \left(1 + \|\mathbf{x}\|^q\right), \qquad \mathbf{x} \in E_{K,N}. \tag{68}$$

The quadratic bound in Equation (67) is the particle-level counterpart of the at-most-quadratic growth assumption on G_K.

(ii) Local inverse-denominator stability. Assume that, for every $R_0 < \infty$,

$$\sup_{\substack{0 \leq m \leq M-1 \\ 1 \leq i \leq N \\ \|\mathbf{x}\| \leq R_0 \\ R,S \geq 1}} \mathbb{E}\left[\left(\overline{B}_{m,i}^{R,S}(\mathbf{x})\right)^{-4}\right] < \infty. \tag{69}$$

(iii) Uniform chain moments. Assume that, for some $p > 0$,

$$\sup_{R,S \geq 1} \max_{0 \leq m \leq M} \mathbb{E}\left[\left\|\mathbf{C}_m^{R,S,h}\right\|^p\right] + \max_{0 \leq m \leq M} \mathbb{E}\left[\left\|\mathbf{X}_m^h\right\|^p\right] < \infty. \tag{70}$$

(iv) Continuity and polynomial growth of the terminal functional. Assume that $\mathcal{J}_{K,N}$ is continuous and that, for some $r < p$,

$$|\mathcal{J}_{K,N}(\mathbf{x})| \leq C_{\mathcal{J}} \left(1 + \|\mathbf{x}\|^r\right), \qquad \mathbf{x} \in E_{K,N}. \tag{71}$$

We now derive the consequences of these assumptions. Define the desirability denominator

$$\beta_m(\mathbf{x}) = \mathbb{E}\left[\exp\left(-\frac{F_{K,N}(\mathbf{x} + \mathbf{U}_m)}{\varepsilon}\right)\right].$$

Since $F_{K,N}$ is bounded from below, the Gibbs factor is bounded from above. Dominated convergence therefore implies that β_m is continuous. Moreover,

$$\beta_m(\mathbf{x}) > 0$$

for every $\mathbf{x}$. Since the set of Euler times is finite, for every localization radius $R_0 < \infty$,

$$\underline{\beta}_{R_0} := \min_{\substack{0 \leq m \leq M-1 \\ \|\mathbf{x}\| \leq R_0}} \beta_m(\mathbf{x}) > 0. \tag{72}$$

Let

$$\pi_{m,\mathbf{x}}(d\mathbf{u}) = \frac{\exp\left(-F_{K,N}(\mathbf{x} + \mathbf{u})/\varepsilon\right)}{\beta_m(\mathbf{x})} \mathcal{N}\left(0, \tau_m \mathbf{Q}_{K,\mathrm{prop}}\right)(d\mathbf{u})$$

be the corresponding Gibbs-tilted proposal law. Gaussian integration by parts gives

$$b(t_m, \mathbf{x}) = -\frac{1}{\varepsilon} \mathbf{Q}_{K,\mathrm{prop}} \mathbb{E}_{\pi_{m,\mathbf{x}}}\left[DF_{K,N}(\mathbf{x} + \mathbf{U}_m)\right]. \tag{73}$$

Differentiating this representation gives

$$D_{\mathbf{x}} b(t_m, \mathbf{x}) = -\frac{1}{\varepsilon} \mathbf{Q}_{K,\mathrm{prop}} \mathbb{E}_{\pi_{m,\mathbf{x}}}\left[D^2 F_{K,N}(\mathbf{x} + \mathbf{U}_m)\right] + \frac{1}{\varepsilon^2} \mathbf{Q}_{K,\mathrm{prop}} \operatorname{Cov}_{\pi_{m,\mathbf{x}}}\left(DF_{K,N}(\mathbf{x} + \mathbf{U}_m)\right). \tag{74}$$

Equation (68) and Equation (72), together with Gaussian polynomial moments, imply that, for every $R_0 < \infty$, there are constants

$$B_{b,R_0}, \qquad L_{b,R_0} < \infty$$

such that

$$\sup_{\substack{0\le m\le M-1\\ \|\mathbf{x}\|\le R_0}} \|b(t_m,\mathbf{x})\| \le B_{b,R_0},$$

and

$$\|b(t_m,\mathbf{x}) - b(t_m,\mathbf{y})\| \le L_{b,R_0}\|\mathbf{x}-\mathbf{y}\|$$

whenever

$$\|\mathbf{x}\| \vee \|\mathbf{y}\| \le R_0.$$

Thus the exact-feedback is locally bounded and locally Lipschitz, although it need not be globally bounded or globally Lipschitz.

Because $F_{K,N}$ is bounded from below, the required local numerator moments are finite automatically:

$$\sup_{\substack{0\le m\le M-1\\ 1\le i\le N\\ \|\mathbf{x}\|\le R_0}} \mathbb{E}\left[L_{t_m,\mathbf{x}}(\mathbf{U}_m)^4\left(1+\|\mathbf{U}_{m,i}\|^4\right)\right] < \infty.$$

Combining these numerator bounds with Equation (69), the block variance argument in Proposition C.1 gives, for every $R_0 < \infty$ and every $R, S \ge 1$,

$$\sup_{\substack{0\le m\le M-1\\ \|\mathbf{x}\|\le R_0}} \mathbb{E}_{\Xi_m^{R,S}}\left[\left\|\widehat{b}_m^{R,S}(\mathbf{x}) - b(t_m,\mathbf{x})\right\|^2\right] \le C_{\mathrm{MC},R_0}\left(\frac{1}{R}+\frac{1}{RS}\right). \tag{75}$$

If only convergence in probability is required, the inverse-moment condition can be replaced by compact-uniform convergence in probability of the numerator and denominator block averages. The inverse-moment condition is used to obtain the explicit L^2 rate.

Define the joint exit time

$$\tau_{R_0} = \inf\left\{m \le M : \left\|\mathbf{C}_m^{R,S,h}\right\| \vee \left\|\mathbf{X}_m^h\right\| > R_0\right\}.$$

Up to τ_{R_0}, the local Lipschitz and local Monte Carlo constants are deterministic. The coupling proof of Theorem C.5 therefore gives

$$\max_{0\le m\le M}\mathbb{E}\left[\left\|\mathbf{C}_{m\wedge\tau_{R_0}}^{R,S,h} - \mathbf{X}_{m\wedge\tau_{R_0}}^h\right\|^2\right] \le C_{\mathrm{stab},R_0}\left(\frac{1}{R}+\frac{1}{RS}\right).$$

Consequently, for every $\eta > 0$,

$$\begin{aligned}\mathbb{P}\left(\max_{0\le m\le M}\left\|\mathbf{C}_m^{R,S,h} - \mathbf{X}_m^h\right\| > \eta\right)\\ \le \frac{(M+1)C_{\mathrm{stab},R_0}}{\eta^2}\left(\frac{1}{R}+\frac{1}{RS}\right) + \mathbb{P}(\tau_{R_0}\le M).\end{aligned} \tag{76}$$

By the union bound, Markov's inequality, and Equation (70),

$$\mathbb{P}(\tau_{R_0}\le M) \le \frac{(M+1)C_p}{R_0^p},$$

where C_p is independent of R, S. First sending

$$R\to\infty, \qquad RS\to\infty,$$

and then $R_0\to\infty$, proves

$$\max_{0\le m\le M}\left\|\mathbf{C}_m^{R,S,h} - \mathbf{X}_m^h\right\| \longrightarrow 0$$

in probability.

Finally, the moment condition Equation (70) and the growth bound Equation (71), with $r < p$, imply uniform integrability of $\{\mathcal{J}_{K,N}(\mathbf{C}_M^{R,S,h})\}_{R,S}$.

Continuity of $\mathcal{J}_{K,N}$, convergence in probability, and uniform integrability therefore yield

$$\mathbb{E}\mathcal{J}_{K,N}\left(\mathbf{C}_M^{R,S,h}\right) \longrightarrow \mathbb{E}\mathcal{J}_{K,N}\left(\mathbf{X}_M^h\right).$$

Thus the finite-candidate consistency result remains valid under polynomial growth, without requiring $F_{K,N}$ or its derivatives to be globally bounded.

Example C.6 (Quadratic-growth interacting measure objectives)**.** *Let $\mathcal{H}_K$ be finite-dimensional, let $a \in \mathcal{H}_K$, and consider*

$$G_K(\mu) = \frac{\kappa_1}{2}\int_{H_K}\|x-a\|^2\,\mu(dx) + \frac{\kappa_2}{2}\left\|\int_{H_K} x\,\mu(dx) - a\right\|^2 + \lambda\,\mathrm{MMD}_k^2(\mu,\nu),$$

where

$$\kappa_1 > 0, \qquad \kappa_2, \lambda \geq 0,$$

$\nu \in \mathcal{P}_2(H_K)$, and k is a bounded symmetric positive-definite C^2 kernel whose first and second derivatives are bounded.

The first term is an additive quadratic confinement, the second term is a nonlinear mean interaction, and the third term is a pairwise distributional interaction. Since

$$\left\|\int x\,\mu(dx) - a\right\|^2 \leq 2M_2(\mu) + 2\|a\|^2$$

and the bounded-kernel MMD term is bounded, there is a constant C such that

$$0 \leq G_K(\mu) \leq C\,(1 + M_2(\mu)).$$

Thus G_K is bounded from below, continuous, and has at most quadratic growth.

For $\mathbf{x} = (x_1, \ldots, x_N)$, $\overline{x} = \frac{1}{N}\sum_{i=1}^N x_i$, the extensive terminal cost is

$$F_{K,N}(\mathbf{x}) = \frac{\kappa_1}{2}\sum_{i=1}^N \|x_i - a\|^2 + \frac{N\kappa_2}{2}\|\overline{x} - a\|^2 + \lambda N\,\mathrm{MMD}_k^2\left(\mu_{\mathbf{x}}^N, \nu\right).$$

Let

$$k_\nu(x) = \int_{H_K} k(x,y)\,\nu(dy).$$

Using symmetry of k, differentiation gives

$$D_{x_i}F_{K,N}(\mathbf{x}) = \kappa_1(x_i - a) + \kappa_2(\overline{x} - a) + \frac{2\lambda}{N}\sum_{j=1}^N D_1 k(x_i, x_j) - 2\lambda D k_\nu(x_i).$$

The bounded derivative assumptions on k imply

$$\|DF_{K,N}(\mathbf{x})\| \leq C_{K,N}\,(1 + \|\mathbf{x}\|),$$

while the block Hessian satisfies

$$\left\|D^2 F_{K,N}(\mathbf{x})\right\|_{\mathrm{op}} \leq C_{K,N}.$$

In particular, Equation (67) and Equation (68) hold with linear gradient growth and bounded Hessian.

The coercive quadratic example is recovered by setting

$$\kappa_2 = \lambda = 0.$$

If $\kappa_2 > 0$, the objective couples all particles through their empirical mean. Indeed,

$$D^2_{x_i x_j}\left[\frac{N\kappa_2}{2}\|\overline{x} - a\|^2\right] = \frac{\kappa_2}{N}I, \qquad 1 \leq i, j \leq N.$$

Thus the objective is no longer an additive sum of independent single-particle costs. If $\lambda > 0$, the MMD term adds a second, kernel-mediated interaction between every pair of particles. Thus this

class contains genuinely interacting empirical-measure objectives while satisfying the structural growth and regularity conditions Equation (67)– Equation (68). Moreover, since the MMD term is bounded and the quadratic part of $F_{K,N}$ has Hessian bounded by $\kappa_1 + \kappa_2$, a sufficient condition for the inverse-denominator condition Equation (69) is

$$4\tau_{\max}(\kappa_1 + \kappa_2)\|Q_{K,\mathrm{prop}}\|_{\mathrm{op}} < \varepsilon, \qquad \tau_{\max} := \max_{0\le m\le M-1} \tau_m.$$

Indeed, under this condition, for every $R_0 < \infty$,

$$\sup_{\substack{0\le m\le M-1\\ \|\mathbf{x}\|\le R_0}} \mathbb{E}\left[\exp\left(\frac{4F_{K,N}(\mathbf{x}+\mathbf{U}_m)}{\varepsilon}\right)\right] < \infty.$$

Since $z \mapsto z^{-4}$ is convex, Jensen's inequality gives

$$\left(\overline{B}^{R,S}_{m,i}(\mathbf{x})\right)^{-4} \le \frac{1}{RS}\sum_{r=1}^{R}\sum_{s=1}^{S} L_{m,i,r,s}(\mathbf{x})^{-4},$$

and hence Equation (69) holds uniformly in R, S.

It remains to verify the chain moment condition Equation (70). Since $\kappa_1 > 0$,

$$F_{K,N}(\mathbf{y}) \ge \frac{\kappa_1}{2}\sum_{i=1}^{N}\|y_i - a\|^2.$$

Moreover, if $E_j \ge 0$ are any finite collection of energies and $w_j \propto e^{-E_j/\varepsilon}$, then for every nondecreasing ϕ,

$$\sum_j w_j\phi(E_j) \le \frac{1}{n}\sum_{j=1}^{n}\phi(E_j),$$

because $\phi(E_j)$ and $e^{-E_j/\varepsilon}$ are oppositely ordered. Taking $\phi(z) = z^{p/2}$, using the preceding coercivity and the quadratic upper bound on $F_{K,N}$, gives, for every $p \ge 2$,

$$\sup_{\substack{R,S\ge 1\\ 0\le m\le M-1}} \mathbb{E}_{\Xi^{R,S}_m}\left[\|\widehat{b}^{R,S}_m(\mathbf{x})\|^p\right] \le C_p(1 + \|\mathbf{x}\|^p).$$

The same argument for the population Gibbs ratio gives

$$\|b(t_m, \mathbf{x})\|^p \le C_p(1 + \|\mathbf{x}\|^p).$$

Since the execution noise is finite-dimensional Gaussian, the exact and finite Euler recursions therefore imply, by induction over the finite number of time steps,

$$\sup_{R,S\ge 1}\max_{0\le m\le M}\mathbb{E}\|C^{R,S,h}_m\|^p + \max_{0\le m\le M}\mathbb{E}\|X^h_m\|^p < \infty$$

whenever the initial cloud has finite p-th moment. In particular, this holds for every $p < \infty$ under the deterministic collapsed initialization used above. Hence, under the displayed inverse-denominator condition, all hypotheses of Remark C.4 are satisfied; one may take any $p > 2$, since the terminal growth exponent is $r = 2$.

D Experimental details and additional figures

D.1 Three controlled toy problems

D.1.1 Four-mode MMD sanity check

Let $\nu^\star = \frac{1}{4}\sum_{s_1,s_2\in\{-1,1\}}\delta_{(s_1,s_2)}$ be the uniform law on the four corners of the square. We optimize $G(\mu) := \mathrm{MMD}^2_k(\mu,\nu^\star)$, where $\mathrm{MMD}^2_k(\mu,\nu) := \iint k(x,x')\,\mu(dx)\mu(dx') - 2\iint k(x,y)\,\mu(dx)\nu(dy) + \iint k(y,y')\,\nu(dy)\nu(dy')$. We use a Gaussian kernel k, which is characteristic, so the unique minimizer is $\mu^\star = \nu^\star$. We compare SCMO with Adam point and Particle

Adam. Adam point optimizes a single state and is evaluated as a degenerate point measure, whereas Particle Adam applies Adam directly to the same $N = 128$ particles using gradients of the empirical MMD objective. SCMO uses $N = 128$, $R = 1$, $S = 256$, $M = 256$, $L = 1$, $T = 1$, $\varepsilon = 10^{-10}$, proposal scale $\sigma_{\text{prop}} = 1.0$, and execution scale $\sigma_{\text{dyn}} = 0.15$. The Gaussian MMD kernel bandwidth is $\ell = 0.7$.

Table 4 reports results over 50 independent runs. Reported metrics are the final MMD^2, the number of occupied modes, the minimum mass assigned to any target mode, and the L^1 error of the four-mode mass vector. Specifically, let $z_1, \dots, z_4$ denote the four target modes and define $B_\rho(z_m) := \{x \in \mathbb{R}^2 : \|x - z_m\| \le \rho\}$, $p_m := \mu^N(B_\rho(z_m))$, $m = 1, \dots, 4$, and $\mu^N := \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ denote the final empirical measure. Thus p_m is the empirical mass assigned to a radius-ρ neighborhood of mode z_m. We count mode m as occupied if $p_m \ge 0.05$, define the minimum mode mass as $\min_m p_m$, and define the L^1 mass error by $\sum_{m=1}^4 \left|p_m - \frac{1}{4}\right|$. Adam point is included as a degenerate point-measure baseline, while Particle Adam serves as a gradient-oracle reference because the MMD objective is smooth. As expected, Particle Adam achieves the lowest MMD value. SCMO uses only objective evaluations and still recovers all four modes in every run, with minimum mode mass 0.229 ± 0.010 and L^1 mass error 0.056 ± 0.027, indicating a nearly balanced four-mode empirical law.

Table 4: Four-mode MMD sanity check over 50 independent runs. Values are mean ± standard deviation.

Method	MMD^2	Modes	Min. mode mass	L^1 mass error
Adam point δ_x	$7.50\times10^{-1} \pm 0.00$	–	–	–
Particle Adam	$1.32\times10^{-3} \pm 8.21\times10^{-4}$	4.0 ± 0.0	0.222 ± 0.012	0.073 ± 0.025
SCMO	$1.49\times10^{-2} \pm 2.49\times10^{-3}$	4.0 ± 0.0	0.229 ± 0.010	0.056 ± 0.027

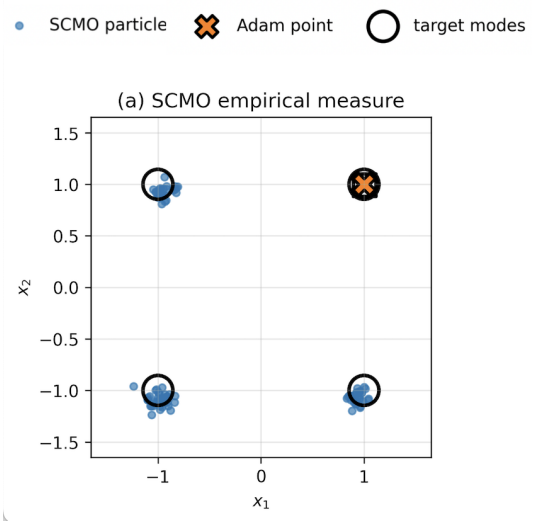


Figure 2: Representative SCMO empirical measure for the four-mode MMD sanity check. The target law is the uniform measure over the four marked modes $(\pm1, \pm1)$. Quantitative results over 50 runs are reported in Table 4.

D.1.2 Asymmetric two-well escape

We consider the two-dimensional asymmetric two-well objective $J(x) = c - \alpha_{\text{near}} \exp\left(-\frac{\|x - x_{\text{near}}\|^2}{2\sigma^2}\right) - \alpha_{\text{far}} \exp\left(-\frac{\|x - x_{\text{far}}\|^2}{2\sigma^2}\right)$, where $x_{\text{near}} = (-1, 0), x_{\text{far}} = (1.8, 0)$, with $c = 0.30$, $\alpha_{\text{near}} = 1.00$, $\alpha_{\text{far}} = 1.35$, and $\sigma = 0.22$. The nearby left well is easier to reach from the origin but is shallower, while the farther right well has strictly lower objective value.

We run 50 independent trials. Adam is initialized from a small Gaussian perturbation around the origin, with initialization scale 0.05, learning rate 10^{-2}, and 2000 steps. SCMO is initialized from a collapsed particle cloud at the origin and uses $N = 128$, $R = 1$, $S = 128$, $M = 128$, $L = 3$, horizon $T = 3$, temperature $\varepsilon = 10^{-10}$, proposal scale $\sigma_{\text{prop}} = 1.0$, and execution scale $\sigma_{\text{dyn}} = 0.15$.

Table 5 shows that Adam consistently converges to the nearby shallow well and never reaches the deeper basin. In contrast, SCMO transports almost all of its empirical mass to the farther deep well: the final deep-well mass is 0.973 ± 0.014, while the shallow-well mass is only 0.027 ± 0.014. The SCMO empirical law attains objective $-1.020 \pm 4.89 \times 10^{-3}$, and the best SCMO atom reaches $-1.050 \pm 1.42 \times 10^{-4}$, close to the bottom of the deeper well. This toy experiment illustrates long-range distributional exploration: SCMO moves the particle law toward the globally better basin, whereas Adam commits to the locally attractive shallow well.

Table 5: Asymmetric two-well escape benchmark over 50 independent runs. Values are mean $\pm$ standard deviation. Lower objective is better.

Output	Objective ↓	Deep-well mass/rate ↑	Shallow-well mass ↓
Adam point	$-0.700 \pm 5.86\times10^{-6}$	0.000 ± 0.000	1.000 ± 0.000
SCMO empirical law	$-1.020 \pm 4.89\times10^{-3}$	0.973 ± 0.014	0.027 ± 0.014

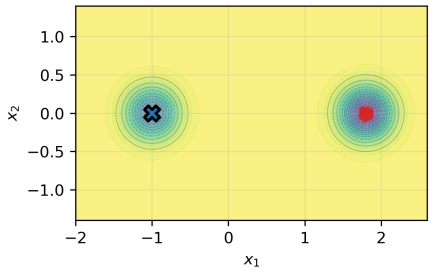


Figure 3: Representative final configuration for the asymmetric two-well escape toy. The background contours show the objective landscape, with darker regions indicating lower objective values. The left basin is closer to the initialization but shallower, while the right basin is farther but deeper. Red points are final SCMO particles, and the blue cross is the final Adam point. Adam remains in the nearby shallow well, whereas SCMO transports most of its empirical measure to the deeper well. Quantitative results over 50 runs are reported in Table 5.

D.1.3 Nonsmooth plateau escape

We use a simple two-dimensional nonsmooth objective to test whether SCMO can escape a flat region where local first-order information is uninformative. The objective is $J(x) = 1_{\{\|x\| \le r\}} + 1_{\{\|x\| > r\}} \min\left\{\frac{\|x-c_1\|^2}{\beta}, \frac{\|x-c_2\|^2}{\beta}\right\}$, with $r = 0.65$, $\beta = 6$, $c_1 = (-1.2, 0.95)$, and $c_2 = (1.2, -0.95)$. We compare the performance of SCMO with Adam. The origin lies inside the flat plateau, so Adam initialized at the origin receives no useful local gradient. SCMO is also initialized from a collapsed cloud at the origin, but uses candidate sampling and exponential reweighting to move particles toward the low-objective basins.

We run 50 independent trials. Adam uses learning rate 10^{-2} and 2000 steps. SCMO uses $N = 128$ particles, $R = 1$, $S = 128$ candidate replacements per particle, $M = 128$ inner steps, $L = 3$ outer iterations, horizon $T = 3$, temperature $\varepsilon = 10^{-10}$, proposal scale $\sigma_{\text{prop}} = 1.0$, and execution scale $\sigma_{\text{dyn}} = 0.15$.

Table 6 shows that Adam remains at the plateau value in every run, while SCMO moves all particle mass outside the plateau and obtains a near-zero empirical objective. The final SCMO cloud also splits approximately evenly between the two low-objective basins.

Table 6: Nonsmooth plateau escape over 50 independent runs. Values are mean ± standard deviation. For SCMO, "mass outside plateau" is the final empirical mass outside $\{x : \|x\| \le r\}$.

Output	Objective	Escape / mass outside plateau	Mass in two wells
Adam point	1.00 ± 0.00	0.00 ± 0.00	–
SCMO empirical law	$2.45\times10^{-4} \pm 2.12\times10^{-5}$	1.00 ± 0.00	0.500 ± 0.050 / 0.500 ± 0.050

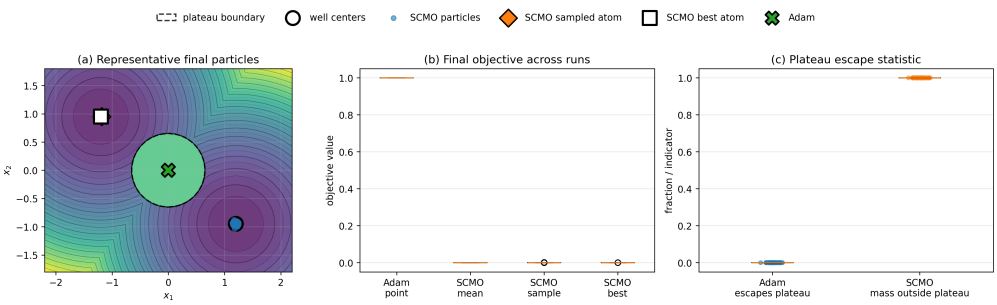


Figure 4: Nonsmooth plateau escape sanity check. (a) Representative final configuration. Adam remains at the origin inside the flat plateau, while SCMO transports its particles into the low-objective basins. (b) Final objective values over 50 runs. (c) Plateau escape statistic: Adam is evaluated by whether the final point leaves the plateau, while SCMO is evaluated by final empirical mass outside the plateau.

D.2 Function-Space Optimization with PDE Energies

PDE problem setting. We next evaluate SCMO on five function-space optimization problems derived from standard phase-field and variational energies. The suite is designed to test recovery of structured function-valued laws under smooth, nonsmooth, discontinuous, and metastable objectives.

Specifically, let $H = L^2([0,1])$, and represent each candidate probability measure on H by an N-particle empirical law. Each particle is expanded using $K = 32$ elements of the orthonormal Neumann cosine basis

$$e_0(x) = 1, \qquad e_k(x) = \sqrt{2}\cos(k\pi x), \quad k = 1, \dots, K-1.$$

Thus, for the i-th particle,

$$u_i(x) = \sum_{k=0}^{K-1} c_{ik} e_k(x) = c_{i0} + \sqrt{2} \sum_{k=1}^{K-1} c_{ik} \cos(k\pi x). \tag{77}$$

This representation automatically satisfies the homogeneous Neumann boundary conditions $u_i'(0) = u_i'(1) = 0$, and its constant coefficient equals the spatial mean, $c_{i0} = \int_0^1 u_i(x)\,\mathrm{d}x$. We evaluate the functions and their energies on the uniform grid $x_g = \frac{g}{127}$, $g = 0, \dots, 127$, using trapezoidal quadrature.

The five problems isolate complementary optimization challenges. P1 is a smooth Allen–Cahn reference problem testing recovery of a balanced two-phase law supported near $u \equiv -1$ and $u \equiv +1$. P2 tests optimization when both the particlewise total-variation energy and the law-level absolute-moment penalty are nonsmooth. P3 tests simultaneous recovery of three separated phases under a discontinuous and interacting occupancy objective with target masses $(0.25, 0.50, 0.25)$. P4 tests escape from the prescribed shallow local phase $u \equiv -1$ toward the unique global phase $u \equiv +1$ in a smooth nonconvex landscape. Finally, P5 tests both escape from the metastable central phase $u \equiv 0$ and distributional splitting between the two outer global phases $u \equiv \pm 1$, with target masses $(0.5, 0, 0.5)$.

We now detail the explicit objective functional for each problem. Recall Equation (77), for a coefficient cloud $c = (c_1, \dots, c_N) \in (\mathbb{R}^K)^N$, define its empirical function-valued law and constant-mode moments by $\mu_c^N = \frac{1}{N}\sum_{i=1}^N \delta_{u_i}$, $m_q(c) = \frac{1}{N}\sum_{i=1}^N c_{i0}^q$. For each problem p, the implemented optimization problem is

$$\min_{c\in(\mathbb{R}^K)^N} G_p(\mu_c^N),$$

where

$$G_p(\mu_c^N) = \frac{1}{N}\sum_{i=1}^N E_p(u_i) + \Phi_p(c).$$

Here, E_p is the particlewise function-space energy and Φ_p is a law-level interaction term.

The five implemented objectives are as follows:

$$E_1(u) = \frac{\eta_1}{2}\int_0^1 |u'(x)|^2\,\mathrm{d}x + \frac{1}{4\eta_1}\int_0^1 \big(u(x)^2-1\big)^2\,\mathrm{d}x,$$
$$\Phi_1(c) = 5m_1(c)^2, \qquad \eta_1 = 0.05;$$

$$E_2(u) = \eta_2\int_0^1 |u'(x)|\,\mathrm{d}x + \frac{1}{4\eta_2}\int_0^1 \big(u(x)^2-1\big)^2\,\mathrm{d}x,$$
$$\Phi_2(c) = 5|m_1(c)|, \qquad \eta_2 = 0.04;$$

$$E_3(u) = \frac{\eta_3}{2}\int_0^1 |u'(x)|^2\,\mathrm{d}x + \frac{4}{\eta_3}\int_0^1 \big(u(x)+1\big)^2 u(x)^2 \big(u(x)-1\big)^2\,\mathrm{d}x,$$
$$\Phi_3(c) \text{ defined below}, \qquad \eta_3 = 0.06;$$

$$E_4(u) = \frac{\eta_4}{2}\int_0^1 |u'(x)|^2\,\mathrm{d}x + \frac{1}{\eta_4}\int_0^1 \left[\frac{1}{4}\big(u(x)^2-1\big)^2 - \beta_4\left(\frac{3}{2}u(x) - \frac{1}{2}u(x)^3\right) + \beta_4\right]\mathrm{d}x,$$
$$\Phi_4(c) = 0, \qquad (\eta_4, \beta_4) = (0.05, 0.12);$$

$$E_5(u) = \frac{\eta_5}{2}\int_0^1 |u'(x)|^2\,\mathrm{d}x + \frac{1}{\eta_5}\int_0^1 \Big[u(x)^2\big(u(x)^2-1\big)^2 + \beta_5\big(1-u(x)^2\big)^2\Big]\,\mathrm{d}x,$$
$$\Phi_5(c) = 5\left\{m_1(c)^2 + \big[m_2(c)-1\big]^2\right\}, \qquad (\eta_5, \beta_5) = (0.05, 0.02).$$

For P3, let

$$z = (-1, 0, 1), \qquad w = (0.25, 0.50, 0.25)$$

denote the three target phases and their prescribed masses. Define the radius-gated empirical mass around phase z_j by

$$q_j(c) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}\{|c_{i0} - z_j| \le 0.28\}, \qquad j = 1, 2, 3.$$

Because the phase centers are separated by one and the gating radius is 0.28, the three neighborhoods are disjoint. The law-level interaction for P3 is

$$\Phi_3(c) = 2\sum_{j=1}^3 |q_j(c) - w_j| - 4\,\mathbf{1}\{q_1(c) \ge 0.15,\ q_2(c) \ge 0.30,\ q_3(c) \ge 0.15\}.$$

Its first term penalizes deviations from the target phase masses, whereas the discontinuous reward tests whether an optimizer can satisfy all three occupancy requirements simultaneously.

Evaluation criteria. The primary performance metric is the implemented objective G_p. We additionally report the structural success for P3, P4, P5. P3 is considered successful when all three occupancy requirements $q_1(c) \ge 0.15$, $q_2(c) \ge 0.30$, $q_3(c) \ge 0.15$ are satisfied. P4 success requires at least 0.90 nearest-phase mass at the global phase $+1$. P5 success requires at least 0.90 combined nearest-phase mass in the two outer phases and an ℓ_1 mass error of at most 0.20 relative to the target $(0.5, 0, 0.5)$.

Baselines. We compare SCMO with three complementary controls. Particle Adam applies Adam (Kingma & Ba, 2015) jointly to all NK coefficients of the empirical cloud, providing a gradient-oracle baseline without restricting the particles to a parametric family. CEM cloud (de Boer et al., 2005) is a derivative-free whole-cloud optimizer: each candidate is a complete N-particle coefficient cloud sampled from a coordinatewise Gaussian proposal, whose mean and standard deviation are updated from the elite candidates. The diagonal-Gaussian baseline restricts the particle law to $c = m + \exp(\ell) \odot \varepsilon$, $\varepsilon \sim \mathcal{N}(0, I_K)$, and optimizes m and ℓ using reparameterized Monte Carlo gradients and Adam (Kingma & Welling, 2014). Fresh samples are used for gradient estimation, while a fixed N-sample common-random-number cloud is used to evaluate and select iterates; this cloud is constructed to reproduce the shared initial empirical law exactly. Particle Adam and the Gaussian family use the gradients supplied by the implemented objective; in particular, the hard occupancy indicators in P3 have zero gradient almost everywhere. These baselines test SCMO against direct empirical-cloud gradient optimization, parametric-law optimization, and derivative-free joint-cloud search.

Hyperparameters and budget accounting. All methods use $K = 32$ cosine coefficients, a 128-point spatial grid, $N = 256$ particles in each reported empirical cloud, and the same initial cloud for each task and random seed. SCMO additionally uses $R = 2$ context clouds, $S = 128$ replacement candidates per particle and context, $M = 128$ inner steps, time horizon $T = 2$, and Gibbs softmin temperature $\varepsilon = 10^{-10}$, giving an inner step size $\Delta t = T/M = 1/64$.

The task-specific settings were selected on development runs for each method's best performance and frozen before evaluation. The resulting settings are:

Task	L	σ_{prop}	σ_{dyn}	Particle Adam lr	CEM $B/\rho/\sigma_0$	Gaussian lr
P1	30	1.0	0.20	0.05	32/.05/.08	.08
P2	500	0.1	0.01	0.10	32/.05/.08	.02
P3	30	1.0	0.20	0.10	256/.20/.35	.08
P4	50	0.4	0.01	0.10	32/.05/.08	.08
P5	30	1.0	0.20	0.10	32/.05/.08	.02

Here, B is the number of complete N-particle clouds sampled in each CEM iteration, ρ is the elite fraction, and σ_0 is the coordinatewise initial proposal standard deviation. CEM uses a minimum standard deviation of 0.005 and smoothing coefficient 0.25, where 0.25 is the weight assigned to the new elite estimate. Particle Adam and the diagonal-Gaussian family use at most $3{,}000$ optimizer updates. Each Gaussian update uses 512 reparameterized Monte Carlo training samples, and its initial standard deviation is floored at 0.1; this floor is not imposed after optimization begins.

We exactly match the configured derivative-free search budgets of SCMO and CEM using particle-wise energy evaluations as the common work unit. For task p, one SCMO inner update evaluates RN context particles and RNS proposed replacements. Its configured search budget is therefore $\mathcal{B}_p^{\text{SCMO}} = L_p M N R(S+1)$. If CEM samples B_p complete N-particle clouds per iteration, we set $I_p^{\text{CEM}} = \left\lceil \frac{L_p M R(S+1)}{B_p} \right\rceil$. For every setting in the table, B_p divides $L_p M R(S+1)$ exactly, and hence $\mathcal{B}_p^{\text{CEM}} = I_p^{\text{CEM}} B_p N = \mathcal{B}_p^{\text{SCMO}}$. The resulting CEM iteration counts for P1–P5 are, respectively, 30,960, 516,000, 3,870, 51,600, 30,960. The corresponding matched configured search-particle budgets are 253,624,320, 4,227,072,000, 253,624,320, 422,707,200, 253,624,320. These are task-specific configured maxima; in particular, P2 receives a substantially larger maximum because it uses $L = 500$. Particle Adam and the diagonal-Gaussian family use gradient oracles and are therefore not equated with the derivative-free particle-evaluation count. Both retain their native cap of $3{,}000$ optimizer updates, and their objective histories had plateaued by their final updates across all tasks.

Plateau stopping can trigger only at or after 60% of each method's configured native budget and only when no meaningful improvement has occurred over 20% of that budget. An improvement is considered meaningful when it is at least $\max\{10^{-4}, 10^{-4}\max(1, |G_{\text{anchor}}|)\}$, where G_{anchor} is the objective at the most recent meaningful improvement.

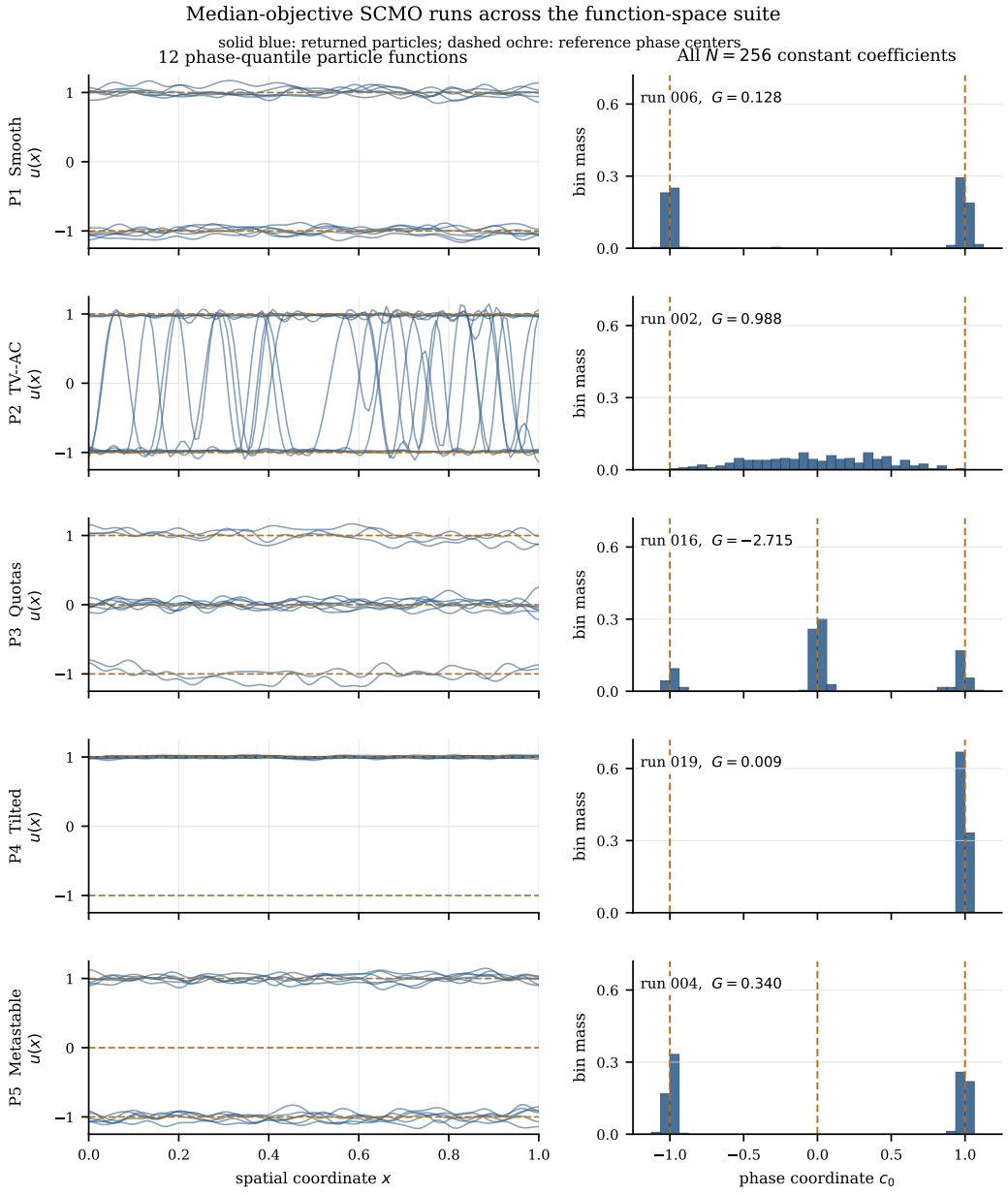
Results. The results are summarized in Table 7. SCMO has the lowest mean objective on P1–P3 and the second-lowest on P4–P5. Its descriptive mean within-seed rank is 1.45, and it is among

the top two methods in all 100 task–seed comparisons. On P3, SCMO encounters and returns a cloud satisfying all three quotas in 20/20 trials, whereas every baseline does so in 0/20. On P4, SCMO and the Gaussian family both reach the global phase in 20/20 trials, with the Gaussian family attaining the lower objective. On P5, SCMO alone satisfies the fixed balanced-global criterion in 20/20 trials, although Particle Adam attains the lower objective. Together, these results demonstrate SCMO's strong performance across smooth, nonsmooth nonconvex, discontinuous interacting, and metastable function-space objectives, particularly when success requires recovering a prescribed multiphase law rather than merely attaining a low scalar objective. Figure 5 shows deterministically selected median-objective SCMO clouds for all five tasks.

Table 7: Five function-space tasks, with 20 paired seeds per task. Panel A reports the objective (mean $\pm$ standard deviation; lower is better); bold and underline mark the best and second-best means. Panel B reports strict structural success.

A. Objective value Task	SCMO	Particle Adam	CEM cloud	Gaussian family
P1 Smooth Allen–Cahn	$\mathbf{0.128 \pm 0.003}$	$\underline{0.184 \pm 0.017}$	1.734 ± 0.042	0.996 ± 0.211
P2 TV–Allen–Cahn	$\mathbf{0.984 \pm 0.028}$	1.747 ± 0.031	$\underline{1.069 \pm 0.020}$	1.118 ± 0.279
P3 Hard three-phase quotas	$\mathbf{-2.541 \pm 0.635}$	2.174 ± 0.034	2.835 ± 0.083	$\underline{2.009 \pm 0.001}$
P4 Tilted local/global	$\underline{0.00883 \pm 0.00008}$	3.846 ± 0.157	5.654 ± 0.007	$\mathbf{0.00109 \pm 0.00010}$
P5 Metastable balanced law	$\underline{0.342 \pm 0.004}$	$\mathbf{0.264 \pm 0.041}$	5.263 ± 0.131	5.271 ± 0.098
Mean within-seed rank $\downarrow$	1.45	2.60	3.57	2.38
B. Strict structural success				
P3 quota success	20/20	0/20	0/20	0/20
P4 global-phase success	20/20	0/20	0/20	20/20
P5 balanced-global success	20/20	0/20	0/20	0/20

Figure 5: Representative runs for function-space optimization with PDE energies



D.3 Trajectory-Law Optimization Through Narrow Passages

Trajectory problem setting. We next consider open-loop trajectory optimization in planar environments with obstacles. A particle is a sequence of $H = 42$ two-dimensional controls, $a_i = (a_{i,0}, \dots, a_{i,H-1}) \in (\mathbb{R}^2)^H$, and induces the discrete trajectory

$$x_{i,0} = x_{\text{start}}, \qquad x_{i,t+1} = x_{i,t} + \Delta t\, a_{i,t}, \qquad \Delta t = 0.15.$$

The optimization variable is the empirical law $\mu_a^N = N^{-1} \sum_{i=1}^N \delta_{a_i}$, represented by $N = 128$ trajectories. Tasks T1–T4 start at $(-4.8, 0)$, whereas T5 starts at $(-4.8, 0.75)$. In every task, the trajectories must pass through one or two openings in a rectangular wall before reaching a circular goal. Table 8 records the task geometry; "strip" denotes an additional rectangular pre-wall obstacle.

Table 8: Geometry of the five trajectory-law tasks. Each interval in the gap column denotes a vertical opening in the wall. The start is $(-4.8, 0)$ for T1–T4 and $(-4.8, 0.75)$ for T5.

Task	Goal (tolerance)	Wall x-range	Wall y-range	Gap interval(s)	Pre-wall strip
T1	$(4.6, 0)$ (0.50)	$[-0.55, 0.55]$	$[-3.2, 3.2]$	$[-0.80, 0.80]$	none
T2	$(4.6, 0.55)$ (0.34)	$[-0.90, 0.90]$	$[-3.2, 3.2]$	$[1.16, 1.39]$	$[-2.55, -1.00] \times [-0.40, 0.40]$
T3	$(4.6, 0)$ (0.38)	$[-0.75, 0.75]$	$[-3.2, 3.2]$	$[1.05, 1.45]$, $[-1.45, -1.05]$	none
T4	$(4.6, 0)$ (0.38)	$[0.12, 1.48]$	$[-3.5, 3.5]$	$[1.38, 1.70]$, $[-1.70, -1.38]$	none
T5	$(4.6, 0.75)$ (0.50)	$[-0.40, 0.40]$	$[-3.4, 3.4]$	$[0.25, 1.10]$, $[-0.30, 0.15]$	none

The five tasks isolate distinct optimization regimes. T1 is a wide single-passage reference problem with a feasible initialization. T2 narrows the opening and tests precise local feasibility. T3 has reflection-symmetric upper and lower passages together with continuous-in-mass law-level terms that reward successful mass, progress along both routes, and route balance; it therefore tests multimodal-law recovery without a discontinuous route-coverage bonus. T4 retains the symmetric two-route geometry but introduces discontinuous success and route-quota rewards, thereby testing an interacting, discontinuous objective. T5 tests escape from a feasible but inferior lower-route basin.

Trajectory objective. For task p, let $[x_p^-, x_p^+] \times [y_p^-, y_p^+]$ be the rectangular wall and let $\mathcal{I}_p$ be the union of its enabled vertical gap intervals, as specified in Table 8. The solid part of the wall is

$$\mathcal{W}_p = \left\{(x, y) : x \in [x_p^-, x_p^+],\ y \in [y_p^-, y_p^+],\ y \notin \mathcal{I}_p\right\}.$$

Let $\mathcal{S}_p$ denote the additional pre-wall rectangular strip on T2 and set $\mathcal{S}_p = \varnothing$ on T1 and T3–T5. The complete obstacle set is

$$\mathcal{O}_p = \mathcal{W}_p \cup \mathcal{S}_p.$$

Let g_p and r_p denote the goal center and tolerance in task p. For particle i, define

$$\begin{aligned}
c_i &= \mathbf{1}\{x_{i,t} \in \mathcal{O}_p \text{ for some } t \in \{1, \dots, H\}\}, \\
d_i &= \|x_{i,H} - g_p\|_2^2, \\
e_i &= \sum_{t=0}^{H-1} \|a_{i,t}\|_2^2, \\
v_i &= \sum_{t=0}^{H-2} \|a_{i,t+1} - a_{i,t}\|_2^2, \\
s_i &= \mathbf{1}\left\{c_i = 0,\ \|x_{i,H} - g_p\|_2 \le r_p\right\}.
\end{aligned}$$

For any scalar particle statistic z_i, write $\bar{z} = N^{-1} \sum_i z_i$. On the two-route tasks, each successful trajectory is assigned an upper or lower label according to its interpolated height at the wall midpoint. The unconditional successful-route masses are

$$\begin{aligned}
p_u &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{s_i = 1,\ i \text{ uses the upper gap}\}, \\
p_l &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{s_i = 1,\ i \text{ uses the lower gap}\}.
\end{aligned}$$

Thus, a failed or unclassified trajectory contributes to neither route mass. For T3–T4, h_i is the mean squared vertical distance from the trajectory to the interior of its nearest enabled passage, evaluated at five equally spaced cross-sections of the wall. Each side of a passage is inset by 20% of its width before this distance is computed. This shaping term is invariant under vertical reflection and contains no preferred route sign.

The implemented task objective is

$$\begin{aligned} G_p(\mu_a^N) = \lambda_{\mathrm{col},p}\bar{c} + \lambda_{\mathrm{g},p}\bar{d} + \lambda_{\mathrm{e},p}\bar{e} + 0.012\bar{v} + w_{h,p}\bar{h} - w_{s,p}\bar{s} \\ - w_{0,p}\mathbf{1}\{\bar{s} \geq c_{0,p}\} \\ - w_{c,p}\mathbf{1}\{p_u \geq q_{u,p},\ p_l \geq q_{l,p}\} \\ - w_{j,p}\min(p_u, q_{u,p}) - w_{j,p}\min(p_l, q_{l,p}) + w_{b,p}(p_u - p_l)^2, \end{aligned} \tag{78}$$

where terms with zero weights are omitted. The task-specific values are

Task	λ_{col}	$(\lambda_{\mathrm{g}}, \lambda_{\mathrm{e}})$	c_0	w_0	w_s	w_c	w_j	w_b	w_h	(q_u, q_l)
T1	6	$(1.1, .02)$	.20	1	0	0	0	0	0	$(0, 0)$
T2	6	$(1.1, .02)$	.25	1	0	0	0	0	0	$(0, 0)$
T3	6	$(1.1, .02)$	–	0	1.5	0	1	.75	1	$(.20, .20)$
T4	6	$(1.1, .02)$	.45	1	0	1.25	1	.35	1	$(.15, .15)$
T5	7	$(5.0, .04)$	.30	1	0	0	0	0	0	$(0, 0)$

In particular, T5 increases the terminal-error and control-effort weights to 5.0 and 0.040, respectively, while setting every route-specific objective weight to zero. Its objective therefore does not explicitly favor the upper or lower route. T5 initializes every method around the same feasible but inferior lower-route template, whose wall-crossing height is $y = -0.075$, while the start and goal at $y = 0.75$ are aligned with the more direct upper passage; it therefore tests escape from the initial route basin under route-neutral proposals.

Evaluation criteria. The primary metric is the implemented objective $G_p(\mu_a^N)$, for which lower is better. We additionally report successful mass $\bar{s}$, all-particle success, and the following task-specific structural criteria: $\bar{s} \geq .20$ on T1; $\bar{s} \geq .25$ on T2; $\bar{s} = 1$ and $p_u, p_l \geq .20$ on T3; $\bar{s} \geq .45$ and $p_u, p_l \geq .15$ on T4; and $\bar{s} \geq .30$ and $p_u \geq .90$ on T5. The T3 requirement deliberately combines full success with two-route mass, so merely placing particles near both gaps is insufficient. The T5 upper-route requirement is an evaluation criterion rather than an optimized route reward: G_5 contains no term depending on p_u or p_l. It therefore measures whether the optimizer escapes the initial lower-route basin and discovers the geometrically preferred upper passage without being directly rewarded for its route label.

Baselines. We compare SCMO with three derivative-free controls, all initialized from the same empirical cloud within each task and seed and all using route-neutral proposals. Greedy whole-cloud random search (Solis & Wets, 1981) samples batches of complete N-trajectory perturbation clouds around the incumbent and accepts the lowest-objective proposal only when it improves G_p. Whole-cloud CEM (de Boer et al., 2005) maintains a coordinatewise Gaussian proposal over the complete empirical cloud, evaluates G_p for every sampled cloud, and updates its mean and standard deviation from the elite clouds. The two-component MPPI-style optimizer (Williams et al., 2017a) maintains two Gaussian distributions over open-loop control sequences. Each component is updated using Gibbs weights computed from point-trajectory costs; equal numbers of trajectories are then sampled from the two components to form an N-particle cloud, which is evaluated and selected using G_p. This method is an open-loop MPPI-style law optimizer rather than a canonical receding-horizon controller. These controls test SCMO against greedy whole-cloud search, elite-based joint-cloud adaptation, and Gibbs-weighted parametric trajectory sampling.

Hyperparameters and budget accounting. All SCMO runs use $N = 128$, $R = 4$, $S = 96$, $M = 90$, $L = 13$, algorithmic horizon $T_{\mathrm{alg}} = 1$, $\varepsilon = 10^{-4}$, and covariance-coordinate decay 0.035. The task-specific proposal and diffusion scales are $(1, .20)$ on T1–T2, $(.60, .035)$ on T3–T4, and $(.30, .01)$ on T5. The task-specific settings were selected on development runs for each method's best performance and frozen before evaluation. The resulting settings are reported in Table 9.

Table 9: Task-specific trajectory optimizer parameters. CEM entries give initial (x, y) standard deviations, elite fraction, and smoothing. MPPI entries give (x, y) noise standard deviations, temperature, and mean smoothing.

Task	SCMO $(\sigma_{\text{prop}}, \sigma_{\text{dyn}})$	RS batch	RS scale	CEM (σ_x, σ_y)/elite/smooth	MPPI (σ_x, σ_y)/temp/smooth
T1	$(1, .20)$	12	.0025	$(.10, .10)$/.20/.70	$(.03, .03)$/.01/.25
T2	$(1, .20)$	12	.0100	$(.10, .10)$/.20/.50	$(.15, .80)$/.08/.25
T3	$(.60, .035)$	12	.0250	$(.05, .80)$/.10/.50	$(.10, .20)$/.02/.25
T4	$(.60, .035)$	12	.0250	$(.05, .80)$/.10/.50	$(.10, .40)$/.04/.25
T5	$(.30, .010)$	53	.0050	$(.03, .12)$/.10/.50	$(.03, .15)$/.02/.25

Whole-cloud CEM uses population 192, re-inflates every 25 iterations by 1.25, and has a minimum standard deviation of .03 on T1–T2 and .02 on T3–T5. The two-component MPPI-style optimizer uses 384 sampled trajectories per component, standard-deviation smoothing .15 on T1–T4 and .10 on T5, re-inflation every 20 iterations by 1.15, and minimum (x, y) standard deviations $(.03, .05)$ on T1–T2, $(.02, .02)$ on T3–T4, and $(.02, .03)$ on T5. Random search uses 37,928 updates on T1–T4 and 8,588 on T5.

We match the maximum number of simulated trajectories rather than nominal iterations. SCMO's exact configured cap is $B_{\text{traj}} = N + LMN\{R(S+1)+1\} = 58{,}256{,}768$ trajectory rollouts, and the baseline iteration ceilings above are chosen so that every method has this same cap; the last atomic batch is shortened when necessary. Budget-aware plateau stopping is shared across methods. It can activate only after 40% of the configured cap and requires no significant incumbent improvement for 15% of the cap, with improvement threshold $\max\{10^{-4}, 10^{-4}\max(1, |G_{\text{anchor}}|)\}$.

Results. The results are summarized in Table 10. SCMO attains the lowest mean objective on T2–T5, including the explicitly law-valued two-route tasks T3 and T4. It satisfies the strict T3 and T4 structural criteria in all 20 runs; Mixture MPPI satisfies the T4 criterion in $14/20$ runs, and no baseline satisfies the full T3 criterion. On the narrow T2 task, SCMO reaches mean successful mass $.996 \pm .006$, compared with $.516 \pm .053$ for Mixture MPPI and zero for Random Search and CEM. On T5, SCMO attains the lowest mean objective, 2.7152 ± 0.0001; SCMO and Mixture MPPI satisfy the strict upper-route criterion in all 20 runs, whereas Random Search and CEM satisfy it in none. The descriptive mean within-seed ranks are 1.50 for SCMO, 2.20 for Mixture MPPI, 2.90 for CEM, and 3.40 for Random Search. Together, these results show that SCMO is reliable on the easy reference problem and performs strongly on narrow-feasibility, multimodal-law, and discontinuous interacting objectives, while also discovering the geometrically better route from an inferior initialization under route-neutral proposals.

Table 10: Five trajectory-law tasks with 20 paired seeds per task. Panel A reports the objective (mean $\pm$ sample standard deviation; lower is better); bold and underline mark the best and second-best means. Panel B reports the strict task-specific structural criterion.

A. Objective value Task	SCMO	Random search	CEM cloud	Mixture MPPI
T1 Easy passage	0.886 ± 0.002	$\mathbf{0.846 \pm 0.000}$	0.886 ± 0.002	$\underline{0.848 \pm 0.000}$
T2 Narrow passage	$\mathbf{1.097 \pm 0.028}$	7.900 ± 0.002	7.853 ± 0.000	$\underline{3.188 \pm 0.416}$
T3 Two-route multimodal	$\mathbf{0.080 \pm 0.006}$	8.005 ± 0.252	3.930 ± 0.322	$\underline{1.360 \pm 0.392}$
T4 Hard route balance	$\mathbf{-0.482 \pm 0.008}$	8.448 ± 0.128	6.628 ± 0.149	$\underline{1.073 \pm 0.777}$
T5 Local/global routes	$\mathbf{2.7152 \pm 0.0001}$	2.8569 ± 0.0010	$\underline{2.8293 \pm 0.0008}$	2.8403 ± 0.0038
Mean within-seed rank ↓	1.50	3.40	2.90	2.20
B. Strict structural success				
T1: $\bar{s} \geq .20$	20/20	20/20	20/20	20/20
T2: $\bar{s} \geq .25$	20/20	0/20	0/20	20/20
T3: $\bar{s} = 1, p_u, p_l \geq .20$	20/20	0/20	0/20	0/20
T4: $\bar{s} \geq .45, p_u, p_l \geq .15$	20/20	0/20	0/20	14/20
T5: $\bar{s} \geq .30, p_u \geq .90$	20/20	0/20	0/20	20/20

Figure 6 displays the SCMO cloud whose objective is closest to the median SCMO objective for each task, with ties resolved deterministically. Each panel plots 40 trajectories selected at evenly

spaced wall-crossing order statistics from the returned 128-particle cloud. In particular, the T3 and T4 panels show both routes without any endpoint-preserving route proposal.

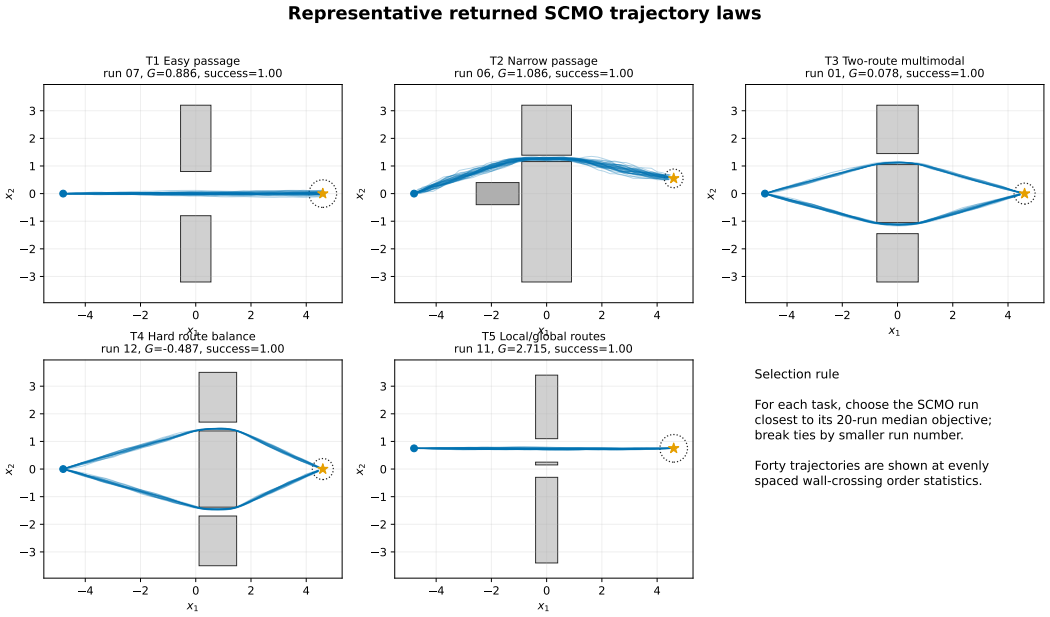


Figure 6: Deterministically selected median-objective SCMO clouds for T1–T5. Each panel shows 40 trajectories selected at evenly spaced wall-crossing order statistics from the returned $N = 128$ cloud, together with the wall, enabled gaps, and goal.

D.4 Push-T as a Contact-Rich External Benchmark

Push-T problem setting. To complement the synthetic diagnostics, we evaluate on the contact-rich Push-T manipulation environment used in Diffusion Policy (Chi et al., 2023). In Push-T, a circular agent must push a T-shaped block into a fixed T-shaped target region centered at $(256, 256)$; a rollout succeeds when the block covers more than 95% of the target region. We study fixed-state, open-loop optimization of a law over complete pushing plans, rather than the canonical policy-learning protocol or a learned-policy evaluation. For particle i, reshape the optimization variable $z_i \in \mathbb{R}^{32}$ into 16 two-dimensional vectors $z_{i,0}, \dots, z_{i,15} \in \mathbb{R}^2$. At the equally spaced times $\tau_j = \frac{299j}{15}$, $j = 0, \dots, 15$, these vectors define bounded target-position knots

$$q_{i,j} = 5 + 502\ \mathrm{sigmoid}(z_{i,j}) \in [5, 507]^2, \qquad \mathrm{sigmoid}(r) = \frac{1}{1 + e^{-r}},$$

where the sigmoid is applied coordinatewise. A shape-preserving piecewise cubic Hermite interpolant is then fitted separately to the x- and y-coordinates of the 16 knots and evaluated at $t = 0, \dots, 299$. This produces the open-loop action sequence $a_i = (a_{i,0}, \dots, a_{i,299})$, $a_{i,t} \in [5, 507]^2$, where $a_{i,t}$ is the absolute target position sent to the circular pushing agent at step t. The interpolation does not overshoot the range of adjacent knots, so every action remains within the environment bounds without subsequent clipping. This mapping is fixed and contains no learned parameters. The optimized empirical law is

$$\mu_z^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{z_i}, \qquad N = 8,$$

and therefore represents a distribution over eight open-loop pushing plans.

The simulator state consists of the center position $p_{\mathrm{push}} = (x_{\mathrm{push}}, y_{\mathrm{push}})$ of the circular pusher and the planar pose (p_T, θ_T) of the movable T-shaped block, where $p_T = (x_T, y_T)$ is the block center and θ_T is its orientation in radians. Coordinates are measured in pixels in the 512×512 workspace. We use the exactly serialized reset state

$$\big(x_{\mathrm{push}}, y_{\mathrm{push}}, x_T, y_T, \theta_T\big) = (450, 62, 340, 172, \pi/4).$$

Thus, the pusher initially has center $(450, 62)$, while the T-shaped block has center $(340, 172)$ and orientation $\pi/4$, and goal center being $(256, 256)$. Five deterministic conditions are obtained by scaling the agent and block positions about the goal by $\gamma \in \{.85, .90, 1.00, 1.05, 1.10\}$; the initial action center is scaled in the same way while a frozen residual cloud is preserved. These are controlled symmetric perturbations of one state, not independent random task draws. Within each fixed condition, every method receives the same serialized reset state and the same serialized $N = 8$ initial latent cloud.

Push-T objective. For a decoded action sequence $a = (a_0, \ldots, a_{299})$, let $r_{\max}(a)$ be the maximum Push-T reward attained during its rollout and define

$$\mathcal{L}(a) = \sum_{t=1}^{299} \|a_t - a_{t-1}\|_2,$$
$$\mathcal{A}(a) = \sum_{t=2}^{299} \|a_t - 2a_{t-1} + a_{t-2}\|_2.$$

Let $I(a)$ indicate an invalid rollout summary. The implemented particle loss is

$$\ell(a) = 1 - r_{\max}(a) + 0.02\frac{\mathcal{L}(a)}{512} + 0.01\frac{\mathcal{A}(a)}{512} + 10I(a). \tag{79}$$

Here $I(a) = 1$ if the rollout summary contains a non-finite reward or coverage, a non-finite or negative $\mathcal{L}(a)$ or $\mathcal{A}(a)$, or an unparseable success flag, and $I(a) = 0$ otherwise; invalid numerical fields are set to zero when computing $\ell(a)$. We use the official `gym-pusht==0.1.6` success criterion. Goal coverage is the fraction of the fixed T-shaped goal region overlapped by the movable T-shaped block, and a rollout succeeds when this coverage strictly exceeds 0.95. Each rollout ends upon success or after the 300-step horizon. To classify the pushing route, we consider the first pusher–block contact located more than 10 pixels to either side of the directed line from the initial block center to the goal center. Contacts within 10 pixels of this line are treated as ambiguous, and classification continues until a later contact leaves this central band. A qualifying contact on the positive side is labelled left, whereas one on the negative side is labelled right. Let $s_i \in \{0, 1\}$ indicate rollout success and let ρ_i denote its left, right, or unclassified route label. Define the unconditional masses by

$$p_s = \frac{1}{N}\sum_{i=1}^{N} s_i,$$
$$p_L = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{s_i = 1, \rho_i = L\},$$
$$p_R = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{s_i = 1, \rho_i = R\}.$$

Failures contribute to neither route mass. Successful rollouts whose contact remains inside the dead zone count toward p_s but toward neither p_L nor p_R. The empirical-law objective is

$$G_{\mathrm{PushT}}(\mu_z^N) = \frac{1}{N}\sum_{i=1}^{N} \ell(a_i) + (p_L - .5)^2 + (p_R - .5)^2 + 2\big[(.25 - p_L)_+^2 + (.25 - p_R)_+^2\big]. \tag{80}$$

Evaluation criteria. The primary metric is G_{PushT}, and lower is better. We additionally report successful mass, the both-quota event $p_L, p_R \geq .25$, the both-route event $p_L, p_R \geq .125$, and all-particle success $p_s = 1$. Because $N = 8$, the both-route threshold requires at least one successful trajectory on each classified route. Route balance and quotas are explicit terms of Equation (80); they are therefore target-aligned diagnostics rather than independent downstream outcomes.

Baselines. We compare SCMO with four derivative-free whole-cloud baselines. For every baseline, one candidate is a complete $N \times 32$ latent cloud and is scored using the same empirical-law objective G_{PushT}. Greedy whole-cloud random search (Solis & Wets, 1981) samples isotropic Gaussian perturbations around the incumbent cloud and retains the best proposal only when it improves the objective. Whole-cloud CEM (de Boer et al., 2005) maintains a coordinatewise Gaussian proposal over complete clouds and updates its mean and standard deviation using the elite candidates. Our MPPI-style baseline (Williams et al., 2017a) samples complete-cloud perturbations and updates the nominal cloud using their Gibbs-weighted average; it is an open-loop empirical-law optimizer rather than a canonical receding-horizon controller. Full CMA-ES (Hansen & Ostermeier, 2001) treats the flattened cloud as one 256-dimensional individual and applies positive-weight full-covariance and cumulative step-size adaptation. Its population is fixed at ten for exact rollout-budget matching rather than using the dimension-dependent default. Within every condition and seed, all methods start from the same serialized cloud and return the lowest-objective cloud visited.

Hyperparameters and budget accounting. The task-specific settings were selected for each method's best performance. Table 11 gives the frozen method parameters.

Table 11: Frozen Push-T optimizer parameters. Every method performs 50 updates and receives exactly 4,408 optimization simulator calls.

Method	Population/update structure	Selected parameters
SCMO	$R = 2$, $S = 4$, $M = 5$, $L = 10$	proposal .075, execution .01, temperature .3, $T_{\mathrm{alg}} = 1$
CEM cloud	10 complete clouds	elite fraction .3, init. std. .075, floor .01, smoothing .5
Random-search cloud	11 complete clouds	proposal scale .04
MPPI-style cloud	10 complete clouds	proposal .075, step .6, temperature .1
Full CMA-ES cloud	10 complete clouds	parent fraction .3, initial scale .075, eigenvalue floor 10^{-12}

The budget is exact at the simulator-call level. The initial cloud costs eight rollouts. Each of SCMO's $LM = 50$ updates costs $N\{R + RS + 1\} = 8(2 + 8 + 1) = 88$ rollouts, so its total is $8 + 50 \times 88 = 4{,}408$. CEM, MPPI, and full-covariance CMA-ES each evaluate ten complete clouds plus one monitor cloud per update, also costing $8(10 + 1) = 88$ rollouts. Random search evaluates 11 complete clouds, again costing $11 \times 8 = 88$. No early stopping is used, and every method returns its best visited eligible incumbent.

Results. The results are summarized in Table 12. SCMO has the lowest selected incumbent objective and the highest successful mass. Its seed-macro objective is 0.0207 ± 0.0036, compared with 0.0619 ± 0.0141 for the strongest baseline, CEM; the SCMO-minus-CEM difference is -0.0412 with a 95% paired-t interval $[-0.0512, -0.0312]$. SCMO also has the lowest mean selected objective in each of the five fixed conditions. Both SCMO and CEM meet both route quotas in all 50 state–seed runs and in all five conditions for every one of the ten seeds. SCMO makes all eight particles successful in 36/50 runs, compared with 27/50 for CEM, while Random Search, MPPI, and full CMA-ES do so in none. Figure 7 replays all eight selected particles for every method at the unit-scale condition.

Table 12: Push-T final evaluation. Continuous entries are mean $\pm$ sample standard deviation over ten seed-level macro-averages across the five fixed conditions. Event counts pool the 50 state–seed runs. Lower objective and higher successful mass are better; bold and underline mark the best and second-best continuous means.

Method	Objective ↓	Success mass ↑	Both routes	Both quotas	All 8 succeed
SCMO	$\mathbf{0.021 \pm 0.004}$	$\mathbf{0.965 \pm 0.027}$	50/50	50/50	36/50
CEM cloud	$\underline{0.062 \pm 0.014}$	$\underline{0.920 \pm 0.059}$	50/50	50/50	27/50
Random-search cloud	0.172 ± 0.043	0.585 ± 0.058	50/50	45/50	0/50
MPPI-style cloud	0.227 ± 0.026	0.520 ± 0.047	50/50	28/50	0/50
Full CMA-ES cloud	0.318 ± 0.038	0.420 ± 0.037	49/50	6/50	0/50

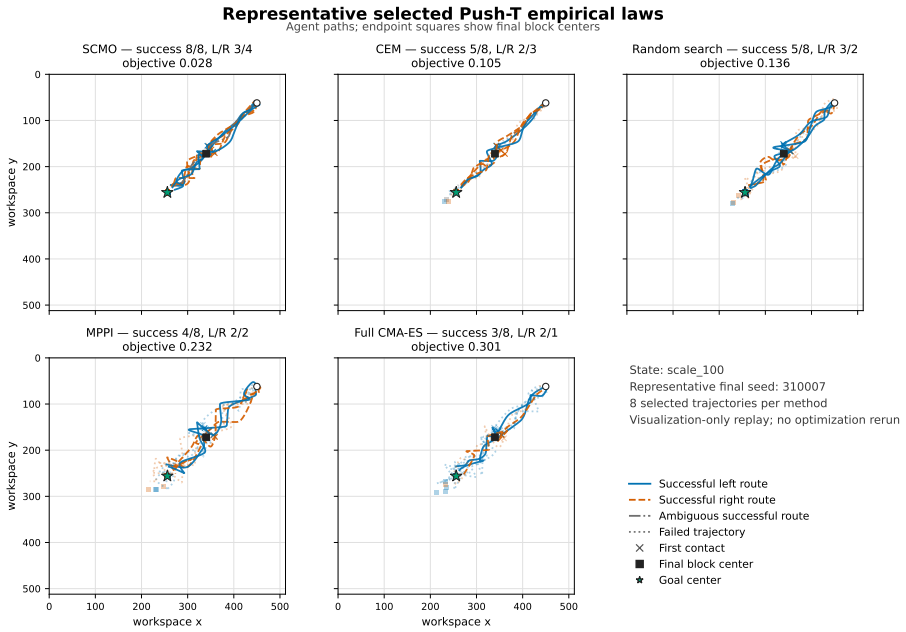


Figure 7: Visualization-only replays of the eight selected particles for each method under the deterministically selected representative seed 310007 at scale 1.00.

E ADDITIONAL ILLUSTRATION: LATENT DISTRIBUTION MATCHING

We include an additional learned-representation experiment to illustrate that SCMO can optimize an empirical law in a latent space. This is not intended as a generative-modeling benchmark. We use a pretrained convolutional VAE on MNIST and keep both the encoder and decoder fixed. Training images are encoded into $d = 16$-dimensional latent vectors, and the resulting empirical latent law is used as the target ν_{MNIST}. SCMO evolves $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{z_t^i}$ by minimizing $G(\mu_t^N) = \mathrm{MMD}^2(\mu_t^N, \nu_{\mathrm{MNIST}})$ with a Gaussian kernel. The decoder is used only for visualization after the latent particles have been optimized.

For the direct SCMO run, we use 12000 target latents obtained by encoding MNIST training images, kernel bandwidth 1.0, $N = 64$, $R = 1$, $S = 8$, $M = 64$, $L = 8$, horizon $T = 1$, and $\varepsilon = 10^{-8}$. The latent covariance is diagonal with decay parameter 0.12, proposal scale $\sigma_{\mathrm{prop}} = 1.0$, and execution scale $\sigma_{\mathrm{dyn}} = 0.08$.

We also test whether the expensive candidate-evaluation step can be amortized. We collect 32,000 cloud transitions from 100 independent teacher SCMO rollouts, each using $N = 128$, $R = 1$, $S = 128$, $M = 64$, and $L = 5$. For each step, we store the time feature $\phi_k = \left(\frac{\tau_k}{T}, \frac{\ell}{L-1}\right)$, the current cloud $Z_k = (z_k^1, \ldots, z_k^N)$, and the SCMO displacement $D_k = (d_k^1, \ldots, d_k^N)$, $d_k^i = \Delta t\, \widehat{\theta}_k^i$. The amortized model is a DeepSets-type permutation-equivariant network. It computes a set summary $h(Z) = \frac{1}{N}\sum_{j=1}^N \psi(z^j)$, and predicts the displacement of particle i by $\widehat{d}_\eta^i = \rho_\eta\big(\phi_k, z^i, h(Z), \overline{z}, \mathrm{Var}(Z)\big)$. The model is trained to imitate the law-level SCMO transition using

$$\mathcal{L}(\eta) = \lambda_{\mathrm{law}}\ \mathrm{MMD}^2\left(\frac{1}{N}\sum_{i=1}^N \delta_{z_k^i + \widehat{d}_\eta^i}, \frac{1}{N}\sum_{i=1}^N \delta_{z_k^i + d_k^i}\right) + \lambda_{\mathrm{disp}}\ \frac{1}{N}\sum_{i=1}^N \|\widehat{d}_\eta^i - d_k^i\|^2,$$

with $\lambda_{\mathrm{law}} = 1.0$, $\lambda_{\mathrm{disp}} = 0.5$, and MMD bandwidth 1.0. The network is trained for 150 epochs with Adam, learning rate 10^{-3}, and batch size 128.

At inference time, the learned network replaces the SCMO candidate pool and replacement-score evaluation. Starting from a fresh Gaussian latent cloud, we iterate the learned measure-flow map and

then decode the final latent particles. Figure 8 shows decoded samples from the true SCMO teacher and the amortized network. Both rows contain recognizable samples across all ten digit classes, suggesting that the learned permutation-equivariant map captures the law-level SCMO evolution well enough to preserve broad latent-mode coverage.

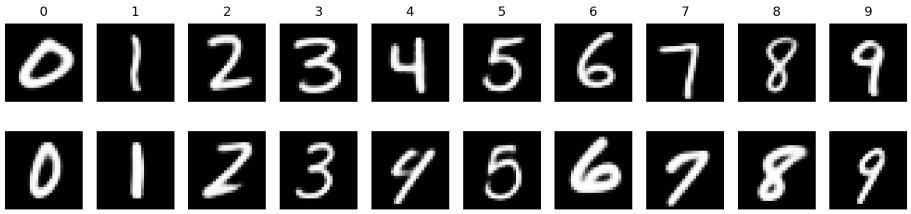


Figure 8: Additional latent distribution-matching illustration on MNIST. A pretrained VAE is kept fixed, and SCMO optimizes an empirical law over 16-dimensional latent vectors toward the encoded MNIST latent distribution using an MMD objective. The first row shows decoded samples from a true SCMO rollout. The second row shows decoded samples from the learned DeepSets measure-flow network, which imitates SCMO cloud transitions and avoids candidate replacement scoring at inference time. Columns are grouped by classifier-predicted digit class $0, \ldots, 9$; the classifier is used only for visualization.