SPIM: SOLVING SCHRÖDINGER BRIDGES VIA ALTERNATING STOCHASTIC OPTIMAL CONTROL

Anonymous authors
Paper under double-blind review

ABSTRACT

Schrödinger bridges (SBs) and stochastic optimal control (SOC) are closely related, but the two-endpoint distribution constraints of SBs make terminal-cost SOC solvers difficult to apply directly. We propose **Schrödinger Potential Iteration Matching (SPIM)**, which solves SBs through alternating SOC that realize the fixed-point iteration of Schrödinger potentials. In each direction, the terminal cost is given by the corresponding relative endpoint energy minus the initial-time value returned by the opposite direction SOC. Under the Cole–Hopf transform, each exact SOC performs one Schrödinger-potential propagation, so a forward–backward pair completes one fixed-point iteration. Under compactness and uniform kernel positivity, exact alternating SOC inherits the Hilbert-metric contraction of the classical potential iteration, yielding geometric convergence of the SOC value functions up to additive constants. The limiting value functions determine the unique optimal generalized Schrödinger bridge. The framework supports energy-to-energy transport from relative endpoint energies without endpoint samples or normalizing constants, extends naturally to generalized Schrödinger bridges with running state costs, and allows different SOC solvers in the two directions. Experiments show that SPIM achieves competitive endpoint and interior transport accuracy across classical and generalized SB benchmarks, scales to higher-dimensional problems, and supports both running-cost and two-endpoint reward fine-tuning.

1 INTRODUCTION

Schrödinger bridges (SBs) (Léonard, 2013) and stochastic optimal control (SOC) (Domingo-Enrich, 2024; Domingo-Enrich et al., 2025) are widely used in generative modeling (Shi et al., 2023; Liu et al., 2022b) and scientific discovery (Liu et al., 2026; Zhang & Chen, 2021). Their theoretical connection is well established (Chen et al., 2021), but using a terminal-cost SOC solver to solve an SB is not straightforward. An SB specifies distributions at both endpoints, whereas a terminal-cost SOC solver takes a prescribed terminal cost as input. This leaves a practical question: how should we choose and update the terminal cost so that the resulting dynamics satisfy both endpoint constraints?

We address this question with **Schrödinger Potential Iteration Matching (SPIM)**, which alternates between two half-bridge SOC problems under the reference diffusion and its time reversal. Each SOC updates the terminal cost of the other: this cost is the corresponding relative endpoint energy minus the value returned by the opposite SOC at its starting time. Figure 1 illustrates this iteration. To see why this alternation works, recall that the Cole–Hopf transform connects Feynman–Kac (Oksendal, 2013) propagation of a Schrödinger potential to SOC value-function propagation. We can therefore carry out the classical fixed-point iteration for generalized Schrödinger potentials (Teter et al., 2025; Chen et al., 2016a) by solving the two SOC problems. With exact solves, each direction performs one half-step, and a forward–backward pair gives one full fixed-point iteration. These updates also correspond to the path-space iterative proportional fitting (IPF) updates (Fortet, 1940; Kullback, 1968; Rüschendorf & Thomsen, 1993). SPIM thus provides an SOC implementation of the classical Schrödinger potential fixed-point iteration.

This connection also gives a convergence guarantee. Under compactness and uniform kernel positivity, the potential fixed-point map contracts in Hilbert's projective metric. Exact SPIM updates therefore yield geometric convergence of the value functions, up to additive constants. After appropriate

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.

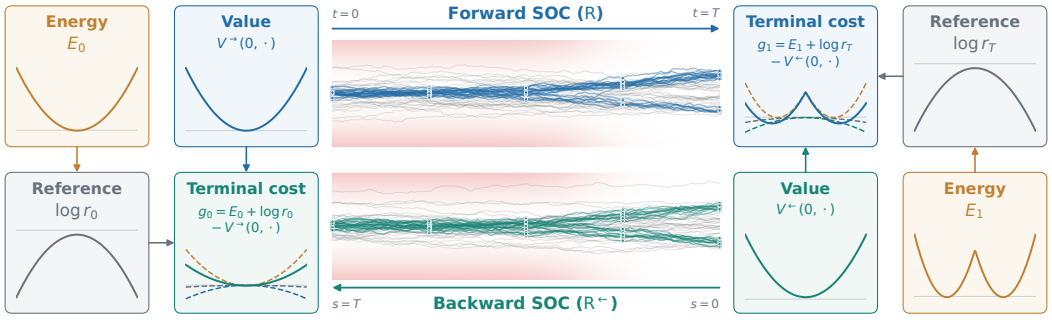

Figure 1: **SPIM for generalized Schrödinger bridges.** Each SOC supplies a boundary value that updates the terminal cost in the opposite direction. Each terminal cost (solid) is the endpoint energy plus the reference log density minus the opposite boundary value (dashed). Gray trajectories show the reference dynamics, and red shading indicates the running state cost.

scaling, the limiting potentials satisfy both endpoint conditions and determine the unique optimal generalized SB.

SPIM has three practical benefits. First, both endpoints can be specified through relative energies rather than samples. Here, a relative energy is the negative log density ratio between an endpoint distribution and the corresponding reference marginal, up to a constant. Second, running state costs enter the two SOC objectives directly, so the same construction covers generalized SBs. Third, the two directions can use different SOC solvers to balance accuracy and computational cost.

Our contributions are threefold:

- We show how to carry out generalized Schrödinger-potential fixed-point iteration through two alternating half-bridge SOC problems.
- We develop SPIM, a framework that supports energy-to-energy transport from relative endpoint energies, includes running state costs, and allows different SOC solvers.
- Under the stated assumptions, we prove geometric convergence of the value functions up to additive constants when the SOC problems are solved exactly, and show that the limit determines the unique optimal generalized SB.

2 Background and Related Work

We first introduce generalized Schrödinger bridges, including the forward and reverse dynamics and the propagation of their potentials. We then review terminal-cost SOC and explain how its value-function propagation relates to these potential propagations.

2.1 Generalized Schrödinger Bridge Problem

Let R be the path law of a reference diffusion with initial law R_0. We consider its forward dynamics and the dynamics of its time reversal $\mathsf{R}^{\leftarrow}$:

$$\begin{aligned} dX_t &= b_t(X_t)\,dt + \sigma_t(X_t)\,dW_t, \qquad X_0 \sim \mathsf{R}_0, \\ d\bar{X}_s &= \bar{b}_s(\bar{X}_s)\,ds + \bar{\sigma}_s(\bar{X}_s)\,d\bar{W}_s, \qquad \bar{X}_0 \sim \mathsf{R}_T, \quad s = T - t. \end{aligned} \tag{1}$$

Here b_t and σ_t are the forward drift and diffusion coefficient, while $\bar{b}_s$ and $\bar{\sigma}_s$ are the corresponding coefficients induced by the time reversal of R (Appendix A.1).

Let Ω be path space and let P_t denote the time-t marginal of a path law P. Given endpoint distributions μ_0, μ_1 and a running state cost c, the generalized Schrödinger bridge minimizes

$$J(\mathsf{P}) = \mathrm{KL}(\mathsf{P}\|\mathsf{R}) + \mathbb{E}_{\mathsf{P}} \int_0^T c_t(X_t)\,dt, \qquad \mathsf{P}_0 = \mu_0, \quad \mathsf{P}_T = \mu_1. \tag{2}$$

The classical Schrödinger bridge is recovered when $c = 0$ (Léonard, 2013; Chen et al., 2021). In backward time, the same running state cost is written as c_{T-s}.

Write $\mathsf{R}_t(dx) = r_t(x)\,dx$ for the reference marginals and $\mu_i(x) = Z_i^{-1}e^{-E_i(x)}$ for the endpoint densities. Define the density ratios and relative endpoint energies by

$$\alpha_0 = \frac{d\mu_0}{d\mathsf{R}_0} = \frac{\mu_0}{r_0}, \qquad \alpha_1 = \frac{d\mu_1}{d\mathsf{R}_T} = \frac{\mu_1}{r_T}, \tag{3}$$

$$\widetilde{E}_0 = E_0 + \log r_0, \qquad \widetilde{E}_1 = E_1 + \log r_T. \tag{4}$$

Then $\alpha_i = Z_i^{-1}e^{-\widetilde{E}_i}$. The generalized bridge has positive Schrödinger potentials $\phi_t, \widehat{\phi}_t$ satisfying (Léonard, 2013; Chen et al., 2021; Liu et al., 2024a)

$$\begin{aligned} \phi_t(x) &= \mathbb{E}_{\mathsf{R}}\left[e^{-\int_t^T c_r(X_r)\,dr}\phi_T(X_T) \mid X_t = x\right], \quad \phi_0\widehat{\phi}_0 = \alpha_0, \\ \widehat{\phi}_t(x) &= \mathbb{E}_{\mathsf{R}}\left[e^{-\int_0^t c_r(X_r)\,dr}\widehat{\phi}_0(X_0) \mid X_t = x\right], \quad \phi_T\widehat{\phi}_T = \alpha_1. \end{aligned} \tag{5}$$

The first expectation uses future paths under R. The second uses past paths, or equivalently future paths under its time reversal $\mathsf{R}^{\leftarrow}$. The bridge marginals and path law are

$$\rho_t^\star(x) = r_t(x)\phi_t(x)\widehat{\phi}_t(x), \tag{6}$$

$$\mathsf{P}^\star(d\omega) = \widehat{\phi}_0(X_0)\phi_T(X_T)e^{-\int_0^T c_t(X_t)\,dt}\mathsf{R}(d\omega). \tag{7}$$

We next relate these potential propagations to SOC value functions.

2.2 Stochastic Optimal Control Problem

Starting from the reference dynamics in equation 1, an admissible control $u \in \mathcal{U}$ modifies the drift along the diffusion directions, yielding

$$dX_t^u = [b_t(X_t^u) + \sigma_t(X_t^u)u_t]\,dt + \sigma_t(X_t^u)dW_t^u. \tag{8}$$

We further write $\mathsf{P}^{u,t,x}$ and $\mathsf{R}^{t,x}$ for the controlled and reference path laws started from $X_t = x$, respectively. Given a running state cost c and a terminal cost g, the terminal-cost SOC problem Domingo-Enrich (2024) is

$$V_g(t,x) := \inf_{u\in\mathcal{U}} \mathbb{E}_{\mathsf{P}^{u,t,x}}\left[\int_t^T \left(\frac{1}{2}\|u_s\|^2 + c_s(X_s)\right) ds + g(X_T)\right]. \tag{SOC-g}$$

Unlike an SB, this problem specifies a terminal cost rather than a terminal distribution. Under standard verification conditions, V_g satisfies the Hamilton–Jacobi–Bellman equation (Bellman, 1966)

$$\partial_t V_g + \mathcal{L}_t V_g + c_t - \frac{1}{2}\left\|\sigma_t^\top \nabla V_g\right\|^2 = 0, \qquad V_g(T,\cdot) = g. \tag{9}$$

Here $\mathcal{L}_t f = b_t \cdot \nabla f + \frac{1}{2}\operatorname{tr}(a_t\nabla^2 f)$ is the reference generator, with $a_t = \sigma_t\sigma_t^\top$. The Cole–Hopf transform (Evans, 2022) linearizes the HJB equation and the Feynman–Kac formula (Oksendal, 2013) yields the path-integral representation (Kappen, 2005):

$$V_g(t,x) = -\log \mathbb{E}_{\mathsf{R}^{t,x}}\left[\exp\left(-\int_t^T c_s(X_s)\,ds - g(X_T)\right)\right]. \tag{PIC-g}$$

Thus, solving a terminal-cost SOC problem propagates the terminal factor e^{-g} and takes its negative logarithm. Comparing with Eq. 5, the choice $g = -\log\phi_T$ gives $V_g(t,x) = -\log\phi_t(x)$. Appendix A.2 gives the conditional path law, optimal control, and invariance to additive constants. SPIM uses this relation to implement potential iteration through SOC solves (Section 3).

2.3 Related Work

Sample-to-sample Schrödinger bridges. The Schrödinger problem has equivalent formulations through path-space entropy, stochastic control, and Schrödinger factors (Dai Pra, 1991; Léonard,

Table 1: **Comparison of different methods.** SPIM combines relative endpoint energies and running state costs through alternating SOC.

Method	No source samples	No target samples	Running state cost support
DSBM (Shi et al., 2023)	×	×	×
GSBM (Liu et al., 2024b)	×	×	✓
ASBS (Liu et al., 2026)	×	✓	×
D2E (Tamogashev & Malkin, 2026)	×	✓	×
D2E (energy-to-energy) (Tamogashev & Malkin, 2026)	✓	✓	×
SPIM (Ours)	✓	✓	✓

2013; Chen et al., 2016b; 2021). Learned methods including diffusion SB, maximum-likelihood SB, bridge matching, and iterative Markovian fitting typically assume samples from both endpoints (De Bortoli et al., 2021; Vargas et al., 2021; Shi et al., 2023). GSBM and twisted bridges add running state costs but train from prescribed endpoint samples (Liu et al., 2024b; Noble et al., 2026; Liu et al., 2022a). SPIM handles relative endpoint energies and a running state cost in one construction.

Energy-specified Schrödinger bridges. PIS, DDS, and their log-variance variants formulate energy-based sampling as a Schrödinger bridge from a prescribed tractable source to an unnormalized target, while Adjoint Sampling considers the corresponding bridge from a Dirac source (Zhang & Chen, 2021; Vargas et al., 2023; Richter & Berner, 2024; Havens et al., 2025). ASBS removes the prescribed-source restriction for sample-to-energy bridges (Liu et al., 2026). D2E presents data-to-energy IPF together with an energy-to-energy extension and a preliminary demonstration (Tamogashev & Malkin, 2026). SPIM realizes the generalized Schrödinger-potential iteration through terminal-cost SOC solves, combining relative endpoint energies, running state costs, and interchangeable SOC solvers.

Stochastic Optimal Control solvers. Recent policy-based SOC solvers directly optimize the control through path space minimization, including Relative Entropy, variance objectives, SOC Matching, and Adjoint Matching (Domingo-Enrich, 2024; Domingo i Enrich et al., 2024; Domingo-Enrich et al., 2025). Value-based methods instead learn the optimal value function and recover the control from it, with recent approaches improving the scalability of classical HJB and path-integral methods (Liao et al., 2026; Jensen et al., 2026; Du et al., 2026). SPIM is solver-agnostic, allowing the two half-bridge stages to use different solvers according to the desired performance–efficiency trade-off.

3 Schrödinger Potential Iteration Matching (SPIM)

SPIM solves two SOC problems in turn, using the value returned by each to update the terminal cost of the other. Section 3.1 identifies the terminal costs that recover the optimal bridge. Section 3.2 turns this relation into an alternating algorithm, and Section 3.3 establishes its convergence and identifies its limit.

3.1 Generalized Schrödinger Bridge as Two SOC Problems

The endpoint conditions in Eq. 5 give $\phi_T = \alpha_1/\widehat{\phi}_T$ and $\widehat{\phi}_0 = \alpha_0/\phi_0$. Taking negative logarithms turns each relation into a terminal cost: the endpoint negative log density ratio minus the opposite SOC value at its starting time. We call this starting-time value the *boundary value*. The following proposition shows that these two SOC problems recover the conditional path laws of the same optimal SB bridge.

Proposition 3.1 (Coupled SOC Self-Consistency)**.** *Under Assumption 1, let $\mathsf{P}^\star$ be the optimal generalized Schrödinger bridge with endpoint distributions μ_0, μ_1. Let $V_g^{\rightarrow}, V_g^{\leftarrow}$ be the forward and reverse SOC values under R and its time reversal $\mathsf{R}^{\leftarrow}$, respectively, with reverse running cost $\bar{c}_s = c_{T-s}$ in time $s = T - t$. There exist functions $V_\star^{\rightarrow}, V_\star^{\leftarrow}$ with associated terminal costs*

$$g_1^\star = -\log \alpha_1 - V_\star^{\leftarrow}(0,\cdot), \qquad g_0^\star = -\log \alpha_0 - V_\star^{\rightarrow}(0,\cdot). \tag{10}$$

The corresponding SOC values satisfy

$$\boxed{V_{g_1^\star}^{\rightarrow} = V_\star^{\rightarrow}, \qquad V_{g_0^\star}^{\leftarrow} = V_\star^{\leftarrow}.} \tag{11}$$

In the original time coordinate, their optimal conditional path laws satisfy

$$\boxed{\begin{aligned}\mathsf{Q}^{\rightarrow}_{\star}(\cdot \mid x) &= \mathsf{P}^{\star}(\cdot \mid X_0 = x),\\ \mathsf{Q}^{\leftarrow}_{\star}(\cdot \mid y) &= \mathsf{P}^{\star}(\cdot \mid X_T = y),\end{aligned}} \tag{12}$$

for μ_0-almost every x and μ_1-almost every y, respectively.

Proof. Take Schrödinger potentials $(\phi^{\star}, \widehat{\phi}^{\star})$ of $\mathsf{P}^{\star}$ and define

$$V^{\rightarrow}_{\star}(t,x) := -\log \phi^{\star}_t(x), \qquad V^{\leftarrow}_{\star}(s,x) := -\log \widehat{\phi}^{\star}_{T-s}(x). \tag{13}$$

By the endpoint products in Eq. 5, we have $g^{\star}_1 = -\log \phi^{\star}_T$ and $g^{\star}_0 = -\log \widehat{\phi}^{\star}_0$. The SOC path integral and Feynman–Kac propagation yield

$$\begin{aligned} V^{\rightarrow}_{g^{\star}_1}(t,x) &= -\log \mathbb{E}_{\mathsf{R}^{t,x}}\left[e^{-\int_t^T c_r(X_r)\,dr}\phi^{\star}_T(X_T)\right]\\ &= -\log \phi^{\star}_t(x) = V^{\rightarrow}_{\star}(t,x).\end{aligned}$$

The same argument under $\mathsf{R}^{\leftarrow}$, using running cost c_{T-s}, gives $V^{\leftarrow}_{g^{\star}_0}(s,x) = -\log \widehat{\phi}^{\star}_{T-s}(x) = V^{\leftarrow}_{\star}(s,x)$. This proves equation 11. The Gibbs variational identity (Appendix A.2) gives the optimal conditional laws, with $C(\omega) = \int_0^T c_t(X_t)\,dt$:

$$\mathsf{Q}^{\rightarrow}_{\star}(d\omega \mid x) = \frac{e^{-C(\omega)}\phi^{\star}_T(X_T)}{\phi^{\star}_0(x)}\mathsf{R}(d\omega \mid X_0 = x), \tag{14}$$

$$\mathsf{Q}^{\leftarrow}_{\star}(d\omega \mid y) = \frac{e^{-C(\omega)}\widehat{\phi}^{\star}_0(X_0)}{\widehat{\phi}^{\star}_T(y)}\mathsf{R}(d\omega \mid X_T = y). \tag{15}$$

These are exactly the conditionals of the bridge factorization in equation 7, proving equation 12. □

Since $-\log \alpha_i = \widetilde{E}_i + \log Z_i$, replacing $-\log \alpha_i$ by $\widetilde{E}_i$ changes only an additive constant in the terminal cost and value function. The optimal control and conditional path law are unchanged. The pointwise SOC updates therefore use relative endpoint energies without requiring the normalizing constants or samples from the prescribed endpoint distributions. This relation describes the solution, but does not yet give an algorithm: the two boundary values are both unknown.

3.2 SPIM Updates via Alternating SOC

SPIM resolves the coupled terminal costs by updating one direction at a time. We keep the value sign convention of Section 3.1: *terminal cost equals relative endpoint energy minus the opposite boundary value*. Let $V^k_1(x)$ be the stored reverse boundary value at the start of round k. The forward solve produces $V^k_0(x) = V^{\rightarrow}_k(0,x)$. The reverse solve then produces $V^{k+1}_1(x) = V^{\leftarrow}_k(0,x)$. The subscripts $0,1$ identify the source and target boundaries. These are evaluations of the directional value functions, not separate objectives. Starting from any finite continuous V^0_1, one round is

$$V^{\rightarrow}_k(t,x) := \inf_{u\in\mathcal{U}} \mathbb{E}_{\mathsf{P}^{u,t,x}}\left[\int_t^T \left(\frac{1}{2}\|u_r\|^2 + c_r(X_r)\right)dr + \widetilde{E}_1(X_T) - V^k_1(X_T)\right], \tag{16a}$$

$$V^k_0 := V^{\rightarrow}_k(0,\cdot).$$

$$V^{\leftarrow}_k(s,y) := \inf_{u\in\mathcal{U}} \mathbb{E}_{\bar{\mathsf{P}}^{u,s,y}}\left[\int_s^T \left(\frac{1}{2}\|u_r\|^2 + c_{T-r}(\bar{X}_r)\right)dr + \widetilde{E}_0(\bar{X}_T) - V^k_0(\bar{X}_T)\right], \tag{16b}$$

$$V^{k+1}_1 := V^{\leftarrow}_k(0,\cdot).$$

Here $\bar{\mathsf{P}}^{u,s,y}$ denotes the reverse controlled law, with $s=0$ at the target and $s=T$ at the source. The reverse solve uses the *new* forward boundary value from the same round. This is sequential boundary exchange. The corresponding positive potentials are $\phi_{0,k} = e^{-V^k_0}$ and $\widehat{\phi}_{T,k} = e^{-V^k_1}$, up to constant factors. Thus each SOC solve propagates one potential, and the pair implements one full potential

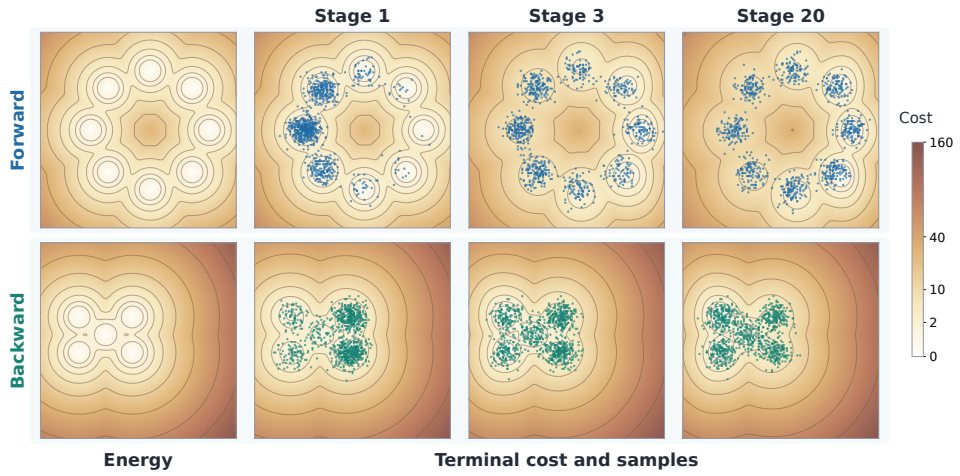

Figure 2: **SPIM from two relative endpoint energies.** Forward (top) and reverse (bottom) rows show the relative endpoint energies (left) and the terminal costs and samples across iterations (right). At each stage, the boundary value obtained from the opposite direction is incorporated into the terminal cost to compensate for the bias induced by the starting energies.

iteration. Algorithm 1 in Appendix A.3 gives the pseudocode, including a constant normalization of the stored values, and derives the equivalent potential updates.

The two directions may use different losses, networks, and SOC solvers, as long as each solver returns the scalar boundary value needed by the next update, in addition to its control. Both solves must use the fixed reference law in their respective time directions, with the running cost evaluated accordingly. A value-based method such as PI-VM supplies these values directly (Liao et al., 2026). A policy-only method such as Adjoint Matching needs value parameterizations (Domingo-Enrich et al., 2025). Appendix C specifies the boundary-value accuracy needed from an approximate solver. The neural experiments use simultaneous (Jacobi) exchange and stabilization described in Appendix E.2.

3.3 Geometric Convergence of Exact SPIM

We now show that exact sequential updates converge to the two SOC values defining the optimal bridge. The result concerns a fixed reference, running cost, and pair of endpoint laws. Its assumptions are compactness, positive continuous endpoint ratios, uniformly positive Feynman–Kac kernels, and exact SOC access. These are stated precisely in Assumption 1 in Appendix A.

Why the iteration contracts. An additive constant in V changes neither its optimal control nor its conditional path law. We therefore measure differences by $\operatorname{osc}(f) = \sup_x f(x) - \inf_x f(x)$, which ignores constants. Let $K_c^{\rightarrow}, K_c^{\leftarrow}$ be the two cost-weighted propagation operators in Appendix A.1. Feynman–Kac turns the two solves into

$$e^{-V_0^k} = K_c^{\rightarrow}\left(e^{-\widetilde{E}_1+V_1^k}\right), \quad e^{-V_1^{k+1}} = K_c^{\leftarrow}\left(e^{-\widetilde{E}_0+V_0^k}\right). \tag{17}$$

Before any constant normalization, these identities are exact. Since $e^{-\widetilde{E}_i} = Z_i\alpha_i$, each update divides an endpoint ratio by the current potential, then propagates the result, up to a constant factor. Propagation contracts Hilbert's metric, while endpoint division preserves it (Chen et al., 2016a). Since $d_H(e^{-V}, e^{-W}) = \operatorname{osc}(V - W)$, a complete round reduces the boundary-value error by $\kappa = \tau_{\rightarrow}\tau_{\leftarrow} < 1$:

$$\operatorname{osc}(V_1^{k+1} - V_1^\star) \le \kappa \operatorname{osc}(V_1^k - V_1^\star). \tag{18}$$

Here $V_1^\star = V_\star^{\leftarrow}(0,\cdot)$, and $\tau_{\rightarrow}, \tau_{\leftarrow}$ are the directional contraction bounds in Appendix B.1.

Theorem 3.2 (Global convergence of SPIM). *Under Assumption 1, exact sequential SPIM converges from any finite continuous V_1^0 to the values in Proposition 3.1, up to additive constants. To fix these constants, choose $x_\dagger \in \mathcal{X}$ and set $\bar{V}_k^d(t,x) = V_k^d(t,x) - V_k^d(0,x_\dagger)$ for $d \in \{\rightarrow, \leftarrow\}$, using each direction's local time. $\bar{V}_\star^d$ is defined similarly. There exist $A < \infty$ and $r \in (0,1)$ such that*

$$\|\bar{V}_k^{\rightarrow} - \bar{V}_\star^{\rightarrow}\|_\infty + \|\bar{V}_k^{\leftarrow} - \bar{V}_\star^{\leftarrow}\|_\infty \le Ar^k, \qquad k \ge 0, \tag{19}$$

Table 2: Classical Schrödinger bridge endpoint errors ($W_2^2 \downarrow$).

Method	Gaussian transport	Ring-mixture transport	Double-well transport	Banana transport
DSBM	0.61436 ± 0.00140	0.08416 ± 0.00335	12.79625 ± 0.04801	0.03297 ± 0.00317
D2E	0.61696 ± 0.00756	0.09363 ± 0.00934	12.06849 ± 0.13837	0.19162 ± 0.01611
ASBS	0.60690 ± 0.00343	1.56359 ± 0.16075	12.49614 ± 0.34460	0.21122 ± 0.15732
SPIM (TR-AM)	0.61100 ± 0.00573	0.43941 ± 0.22154	12.00319 ± 0.02926	0.02780 ± 0.00410
SPIM (PI-VM)	$\mathbf{0.59982 \pm 0.00040}$	$\mathbf{0.02020 \pm 0.00581}$	$\mathbf{11.90751 \pm 0.05322}$	$\mathbf{0.01667 \pm 0.00100}$

Table 3: Generalized Schrödinger bridge interior errors ($W_2^2 \downarrow$).

Method	Sinusoidal tracking	Bimodal splitting	Moving double well	Spherical-meridian transport
GSBM	0.02157 ± 0.00103	0.04917 ± 0.02495	0.02030 ± 0.00143	0.46267 ± 0.39467
DeepGSB	0.00616 ± 0.00040	$\mathbf{0.01333 \pm 0.00044}$	$\mathbf{0.01012 \pm 0.00019}$	4.24018 ± 2.76196
TSBM	0.01849 ± 0.00184	0.07979 ± 0.02642	0.02295 ± 0.00505	0.07913 ± 0.05335
SPIM (TR-AM)	$\mathbf{0.00559 \pm 0.00024}$	0.02458 ± 0.00894	0.01526 ± 0.00267	$\mathbf{0.07020 \pm 0.04411}$
SPIM (TR-PIVM)	0.00771 ± 0.00167	0.01389 ± 0.00187	0.01051 ± 0.00034	1.22095 ± 1.08585

where the norms are over $[0, T] \times \mathcal{X}$. The anchored limits are independent of initialization. After choosing compatible additive constants, their exponential transforms satisfy both endpoint constraints and determine the unique optimal generalized Schrödinger bridge.

What the limit guarantees. Convergence of the boundary values extends to the full value functions because the SOC propagation preserves bounds on terminal-cost differences. At the limit, $V_\star^{\rightarrow}(0, \cdot) + V_\star^{\leftarrow}(T, \cdot) = -\log \alpha_0$ and $V_\star^{\rightarrow}(T, \cdot) + V_\star^{\leftarrow}(0, \cdot) = -\log \alpha_1$ after compatible normalization. Exponentiating these identities gives the required endpoint density ratios. A path-space KL identity then proves that this feasible bridge is globally optimal.

Proofs and further results in the supplement. Appendix B provides the full proof: contraction (B.1), endpoint feasibility (B.2), value and endpoint-error bounds (B.3), and global optimality (B.4). The supplement also establishes three extensions under the corresponding assumptions. For approximate SOC solves, vanishing errors preserve boundary-value convergence, uniformly bounded errors give an explicit asymptotic error bound (Appendix C). Exact Jacobi exchange contracts every two rounds (Appendix C.1). With fixed optimal controls, alternating forward–reverse sampling converges to the prescribed endpoint laws (Appendix D).

4 Experiments

We evaluate SPIM on classical and generalized Schrödinger bridges, dimension scaling, running-cost fine-tuning on real LiDAR terrain, and two-endpoint image reward control. Ablations examine optimization schedules, particle banks, and directional SOC solvers.

4.1 Experimental Settings, Baselines, and Metrics

Settings. SPIM uses endpoint samples with Gaussian smoothing to approximate well-covered endpoint neighborhoods and support stable neural training. Derivations for LiDAR and image control, algorithmic design, and stabilization are given in Appendices E.1 and E.2. Task definitions, architectures, training budgets, homotopy schedules, and evaluation protocols are collected in Appendix E.3.

Metrics. Synthetic experiments report empirical W_2^2 at the endpoints and, where reference marginals are available, along the bridge. Synthetic tables report bidirectional means and cross-seed sample standard deviations. Bold and underlined entries identify the best and second-best reported means. LiDAR reports objective and ground MSE. Image experiments report reward and attribute W_1.

Baselines. Classical-bridge baselines are DSBM (Shi et al., 2023), D2E (Tamogashev & Malkin, 2026), and ASBS (Liu et al., 2026). Generalized-bridge baselines are GSBM (Liu et al., 2024b), DeepGSB (Liu et al., 2022a), and TSBM (Noble et al., 2026). SPIM uses AM (Domingo-Enrich et al.,

Table 4: Gaussian endpoint errors across dimensions ($W_2^2 \downarrow$).

Method	5D	20D	50D
DSBM	0.61436 ± 0.00140	14.27885 ± 0.02800	58.02324 ± 0.02793
D2E	0.61696 ± 0.00756	14.21633 ± 0.09664	59.67658 ± 0.42900
ASBS	$\underline{0.60690 \pm 0.00343}$	14.18891 ± 0.02136	56.65567 ± 0.12854
SPIM (TR-AM)	0.61100 ± 0.00573	$\underline{14.15390 \pm 0.04602}$	$\underline{56.47822 \pm 0.03151}$
SPIM (PI-VM)	$\mathbf{0.59982 \pm 0.00040}$	$\mathbf{13.99780 \pm 0.00777}$	$\mathbf{56.10131 \pm 0.14198}$

Table 5: Particle-bank ablation study.

Source	Target	Endpoint $W_2^2 \downarrow$	Interior $W_2^2 \downarrow$
Fixed	Fixed	0.01340 ± 0.00464	0.00634 ± 0.00097
Fixed	Dynamic	$\mathbf{0.01118 \pm 0.00229}$	$\underline{0.00608 \pm 0.00107}$
Dynamic	Dynamic	$\underline{0.01294 \pm 0.00308}$	$\mathbf{0.00560 \pm 0.00128}$

Table 6: Directional-solver ablation study.

Forward	Backward	Endpoint $W_2^2 \downarrow$	Interior $W_2^2 \downarrow$
PI-VM	PI-VM	0.02280 ± 0.01721	0.01495 ± 0.00660
AM	AM	$\mathbf{0.01340 \pm 0.00464}$	$\mathbf{0.00634 \pm 0.00097}$
PI-VM	AM	0.01688 ± 0.01430	0.00810 ± 0.00284

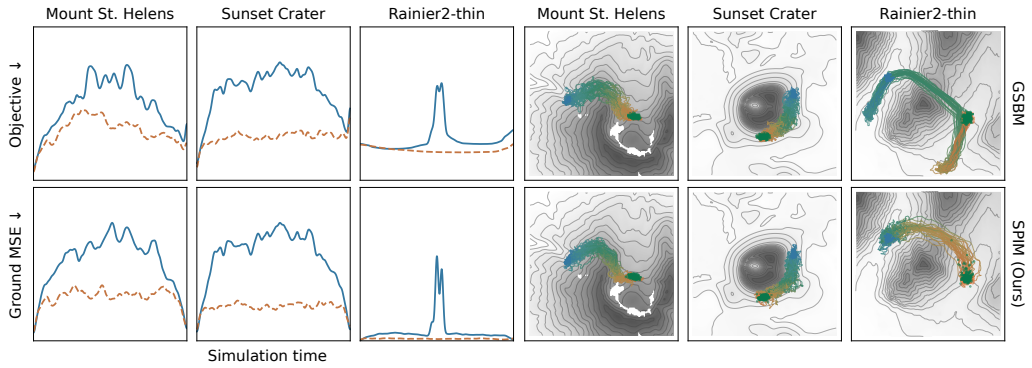

Figure 3: Running-cost fine-tuning on LiDAR terrain. Left: objective and ground MSE (GSBM: blue solid; SPIM: orange dashed). Right: trajectories, with GSBM above and SPIM below.

2025) or PI-VM (Liao et al., 2026) as its SOC solver. TR denotes the trust-region variants (Blessing et al., 2025).

4.2 Classical and Generalized Benchmark Problems

Gaussian, ring-mixture, double-well, and banana transport. These zero-state-cost problems test different endpoint geometries. SPIM (PI-VM) gives the lowest mean endpoint error on all four (Table 2). TR-AM ranks second on double-well and banana transport but has a larger ring-mixture error. Thus PI-VM provides the most consistent endpoint accuracy in this comparison, while multimodal transport remains sensitive to the SOC solver. Gaussian interior errors appear in Table **??**.

Sinusoidal tracking, bimodal splitting, moving double well, and spherical-meridian transport. These problems test transport shaped by running state costs. TR-AM achieves the lowest mean interior error on sinusoidal tracking and spherical-meridian transport (Table 3). TR-PI-VM is second to DeepGSB on bimodal splitting and the moving double well, but performs worse on spherical-meridian transport. Endpoint errors and transport plots are provided in Appendices E.4 and E.7.

High-Dimensional Benchmarks We scale Gaussian transport to $d \in \{5, 20, 50\}$. SPIM (PI-VM) has the lowest mean endpoint error at every dimension, with TR-AM second at 20D and 50D (Table 4).

4.3 Running-cost fine-tuning on LiDAR terrain

We compare GSBM and SPIM (Homotopy TR-AM) on Mount St. Helens, Sunset Crater, and Mount Rainier (Liu et al., 2024b), using point-cloud geometry to define terrain costs (Liu et al., 2024b). SPIM starts from a learned zero-cost bridge and progressively introduces terrain penalties while exchanging boundary values. Figure 3 shows that, compared with GSBM, SPIM finds lower-cost

Table 7: Two-endpoint reward fine-tuning on MNIST–EMNIST : reward r and attribute W_1.

		$\lambda = 1$		$\lambda = 10$		$\lambda = 100$	
Reward task	Method	$r \uparrow$	$W_1 \downarrow$	$r \uparrow$	$W_1 \downarrow$	$r \uparrow$	$W_1 \downarrow$
Area	Base	-1.5931	0.0276	-1.5114	0.0387	-1.5343	0.0435
	AM	-0.9717	0.0145	-0.5592	0.0204	**-0.0760**	**0.0071**
	SPIM (Ours)	**-0.6778**	**0.0097**	**-0.2753**	**0.0121**	-0.1007	0.0117
Aspect ratio	Base	-1.5479	0.1261	-1.4831	0.2054	-1.5304	0.2293
	AM	-1.2826	0.0988	-0.6545	0.1173	-0.2831	0.0820
	SPIM (Ours)	**-0.8906**	**0.0575**	**-0.2967**	**0.0603**	**-0.0670**	**0.0365**
Size	Base	-1.4621	0.0317	-1.4276	0.0437	-1.4651	0.0491
	AM	-1.2599	0.0259	-0.7004	0.0255	-0.1986	0.0131
	SPIM (Ours)	**-0.8919**	**0.0147**	**-0.2587**	**0.0124**	**-0.0572**	**0.0070**

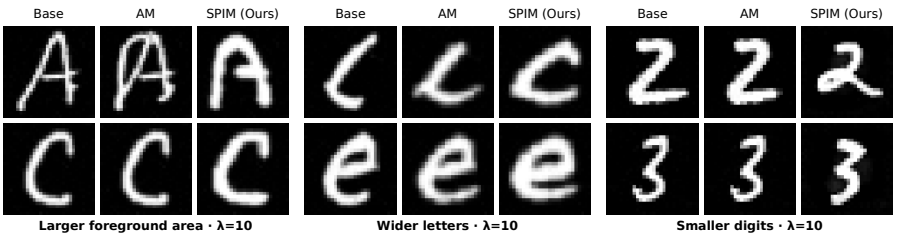

Figure 4: Reward fine-tuning results at $\lambda = 10$: area, aspect ratio, and size, from left to right.

forward paths that follow the terrain more closely. The objective derivation and settings are given in Appendices E.1.2 and E.3.5.

4.4 Two-Endpoint Reward Fine-Tuning on MNIST and EMNIST

We fine-tune a pretrained MNIST (LeCun et al., 1998)–EMNIST (Cohen et al., 2017) bridge using foreground-area, aspect-ratio, and size rewards at both endpoints. AM optimizes each direction independently, while SPIM (Ours) exchanges boundary values to coordinate both endpoint preferences. As shown in Table 7, SPIM achieves higher endpoint-averaged reward and lower attribute W_1 in eight of nine settings, supporting the benefit of two-endpoint coordination. Figure 4 illustrates the corresponding changes in character morphology. Appendix E.1.3 derives the reward-weighted target distributions, data and optimization settings appear in Appendix E.3.6.

4.5 Ablation Studies

We study three components on sinusoidal tracking. These configurations are in Appendix E.3.4. As shown in Table 5, updating both endpoint banks gives the lowest interior error, whereas updating only the target bank gives the lowest endpoint error. Bank refresh helps, but the two metrics favor different refresh patterns. As shown in Table 6, AM in both directions gives the lowest endpoint and interior error. The PIVM solver also offers efficiency trade-off.

5 Conclusion

SPIM implements the generalized Schrödinger-potential iteration through alternating terminal-cost SOC problems. Each solve propagates one potential and provides the boundary value needed to set the next terminal cost. This formulation accommodates relative endpoint energies, running state costs, and different SOC solvers in the two directions. Under the compact positive-kernel assumptions and exact SOC realizability, the anchored values converge geometrically and determine the unique optimal bridge. Across experiments, SPIM provides competitive endpoint and interior transport accuracy across classical and generalized SB benchmarks, remains applicable to higher-dimensional settings, and supports both running-cost and two-endpoint reward fine-tuning.

AI USE STATEMENT

In this work, we used generative AI tools to provide feedback on research methodology and experimental design, assist with method implementation, assist with translation, and support the interpretation of experimental results.

We did not use generative AI tools to generate synthetic datasets, develop theoretical models or conceptual frameworks, formulate or prove mathematical claims, propose or refine hypotheses, clean or reformat datasets, or conduct qualitative or thematic data analysis.

Additionally, we used generative AI tools to assist with code editing, manuscript drafting and revision, literature search and summarization, and improving readability and presentation.

All AI-assisted outputs were reviewed, verified, and revised by the authors. Code and experimental suggestions were validated through implementation and experiments, and suggested references were checked against the original sources. We take responsibility for the final content of this work.

REFERENCES

Richard Bellman. Dynamic programming. *science*, 153(3731):34–37, 1966.

Denis Blessing, Julius Berner, Lorenz Richter, Carles Domingo-Enrich, Yuanqi Du, Arash Vahdat, and Gerhard Neumann. Trust region constrained measure transport in path space for stochastic optimal control and inference. *arXiv preprint arXiv:2508.12511*, 2025.

Yongxin Chen, Tryphon Georgiou, and Michele Pavon. Entropic and displacement interpolation: a computational approach using the hilbert metric. *SIAM Journal on Applied Mathematics*, 76(6): 2375–2396, 2016a.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. On the relation between optimal transport and schrödinger bridges: A stochastic control viewpoint. *Journal of Optimization Theory and Applications*, 169(2):671–691, 2016b.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Stochastic control liaisons: Richard sinkhorn meets gaspard monge on a schrodinger bridge. *Siam Review*, 63(2):249–313, 2021.

Gregory Cohen, Saeed Afshar, Jonathan Tapson, and Andre Van Schaik. Emnist: Extending mnist to handwritten letters. In *2017 international joint conference on neural networks (IJCNN)*, pp. 2921–2926. IEEE, 2017.

Paolo Dai Pra. A stochastic control approach to reciprocal diffusion processes. *Applied mathematics and Optimization*, 23(1):313–329, 1991.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion schrödinger bridge with applications to score-based generative modeling. *Advances in neural information processing systems*, 34:17695–17709, 2021.

Carles Domingo-Enrich. A taxonomy of loss functions for stochastic optimal control. *arXiv preprint arXiv:2410.00345*, 2024.

Carles Domingo-Enrich, Michal Drozdzal, Brian Karrer, and Ricky T. Q. Chen. Adjoint matching: Fine-tuning flow and diffusion generative models with memoryless stochastic optimal control. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=xQBRrtQM8u.

Carles Domingo i Enrich, Jiequn Han, Brandon Amos, Joan Bruna, and Ricky TQ Chen. Stochastic optimal control matching. *Advances in Neural Information Processing Systems*, 37:112459–112504, 2024.

Yuanqi Du, Jiajun He, Dinghuai Zhang, Eric Vanden-Eijnden, and Carles Domingo-Enrich. Rare event analysis via stochastic optimal control. *arXiv preprint arXiv:2604.13213*, 2026.

Lawrence C Evans. *Partial differential equations*, volume 19. American mathematical society, 2022.

Robert Fortet. Résolution d'un système d'équations de m. schrödinger. *Journal de mathématiques pures et appliquées*, 19(1-4):83–105, 1940.

Aaron J Havens, Benjamin Kurt Miller, Bing Yan, Carles Domingo-Enrich, Anuroop Sriram, Daniel S. Levine, Brandon M Wood, Bin Hu, Brandon Amos, Brian Karrer, Xiang Fu, Guan-Horng Liu, and Ricky T. Q. Chen. Adjoint sampling: Highly scalable diffusion samplers via adjoint matching. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=6Eg1OrHmg2.

Cristian Jensen, Luca Schaufelberger, Riccardo De Santi, Kjell Jorner, and Andreas Krause. Value matching: Scalable and gradient-free reward-guided flow adaptation. In *International Conference on Learning Representations*, volume 2026, pp. 144926–144958, 2026.

Hilbert J Kappen. Linear theory for control of nonlinear stochastic systems. *Physical review letters*, 95(20):200201, 2005.

Solomon Kullback. Probability densities with given marginals. *The Annals of Mathematical Statistics*, 39(4):1236–1243, 1968.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Christian Léonard. A survey of the schr\" odinger problem and some of its connections with optimal transport. *arXiv preprint arXiv:1308.0215*, 2013.

Bangyan Liao, Chenglei Yu, Yuchen Yang, Chuanrui Wang, Zhisheng Song, Peidong Liu, and Tailin Wu. Path integral value matching for linear quadratic stochastic optimal control. *arXiv preprint arXiv:2608.10777*, 2026.

Guan-Horng Liu, Tianrong Chen, Oswin So, and Evangelos Theodorou. Deep generalized schrödinger bridge. *Advances in Neural Information Processing Systems*, 35:9374–9388, 2022a.

Guan-Horng Liu, Tianrong Chen, and Evangelos A Theodorou. Deep generalized schr\" odinger bridges: From image generation to solving mean-field games. *arXiv preprint arXiv:2412.20279*, 2024a.

Guan-Horng Liu, Yaron Lipman, Maximilian Nickel, Brian Karrer, Evangelos Theodorou, and Ricky TQ Chen. Generalized schrödinger bridge matching. In *International Conference on Learning Representations*, volume 2024, pp. 14527–14552, 2024b.

Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin K Miller, and Ricky TQ Chen. Adjoint schrödinger bridge sampler. *Advances in Neural Information Processing Systems*, 38:15673–15708, 2026.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022b.

Maxence Noble, Marie Scheid, Yazid Janati, Eric Moulines, and Alain Durmus. Twisted schr\" odinger bridge matching. *arXiv preprint arXiv:2607.16987*, 2026.

Nikolas Nüsken and Lorenz Richter. Solving high-dimensional hamilton–jacobi–bellman pdes using neural networks: perspectives from the theory of controlled diffusions and measures on path space. *Partial differential equations and applications*, 2(4):48, 2021.

Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

Lorenz Richter and Julius Berner. Improved sampling via learned diffusions. In *International Conference on Learning Representations*, volume 2024, pp. 10695–10722, 2024.

Gareth O Roberts and Jeffrey S Rosenthal. General state space markov chains and mcmc algorithms. 2004.

Ludger Rüschendorf and Wolfgang Thomsen. Note on the schrödinger equation and i-projections. *Statistics & probability letters*, 17(5):369–375, 1993.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion schrödinger bridge matching. *Advances in neural information processing systems*, 36:62183–62223, 2023.

Kirill Tamogashev and Nikolay Malkin. Data-to-energy stochastic dynamics. In *International Conference on Learning Representations*, volume 2026, pp. 99944–99964, 2026.

Alexis MH Teter, Wenqing Wang, and Abhishek Halder. Schrödinger bridge with quadratic state cost is exactly solvable. *IEEE Transactions on Automatic Control*, 2025.

Francisco Vargas, Pierre Thodoroff, Austen Lamacraft, and Neil Lawrence. Solving schrödinger bridges via maximum likelihood. *Entropy*, 23(9):1134, 2021.

Francisco Vargas, Will Grathwohl, and Arnaud Doucet. Denoising diffusion samplers. *arXiv preprint arXiv:2302.13834*, 2023.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: a stochastic control approach for sampling. *arXiv preprint arXiv:2111.15141*, 2021.

A Definitions and Assumptions

The proof follows one chain of implications: an exact SOC solve propagates a Schrödinger potential. The resulting potential iteration is a contraction. Its fixed point satisfies both endpoint constraints and minimizes the path-space objective. Appendix C then quantifies errors in the scalar boundary values, and Appendix D studies sampling with the fixed optimal controls. The contraction is the classical Hilbert-metric argument of Chen et al. (2016a). We specialize it to the two reference-weighted operators and transfer its bounds to the SOC values.

A.1 Reference process and running cost

We use a Markov reference path law R with the forward dynamics in equation 1 and marginals $\mathsf{R}_t(dx) = r_t(x)\,dx$. Its time reversal $\mathsf{R}^{\leftarrow} = \Theta_{\#}\mathsf{R}$, where $(\Theta\omega)_s = \omega_{T-s}$, is also assumed to be a Markov diffusion. Both SOC problems use this single reference law, expressed in their respective time directions. This consistency is needed for the two solves to describe the same bridge.

Let $a_t = \sigma_t\sigma_t^{\top}$. Under the usual smoothness and nondegeneracy conditions, the reverse coefficients in equation 1 are

$$\bar{\sigma}_s = \sigma_{T-s}, \qquad \bar{b}_s^i(x) = -b_{T-s}^i(x) + \frac{1}{r_{T-s}(x)}\sum_j \partial_{x_j}\big(a_{T-s}^{ij}(x) r_{T-s}(x)\big). \tag{20}$$

This is the standard diffusion time-reversal formula (Oksendal, 2013). For constant σ, it reduces to $\bar{b}_s = -b_{T-s} + a\nabla \log r_{T-s}$. The forward and reverse drifts are therefore determined by one reference path law. Both reference dynamics are assumed to be available to the SOC solvers.

Absorb the running cost into a reference measure on trajectories:

$$C(\omega) = \int_0^T c_t(X_t)\,dt, \qquad \Gamma(d\omega) = e^{-C(\omega)}\mathsf{R}(d\omega), \qquad J(\mathsf{P}) = \int \log\frac{d\mathsf{P}}{d\Gamma}\,d\mathsf{P}. \tag{21}$$

Here Γ need not have unit mass. The functional J equals equation 2 for $\mathsf{P} \ll \mathsf{R}$, otherwise $J(\mathsf{P}) = +\infty$. Normalizing the finite measure Γ adds a constant to the objective.

For a terminal function, a Feynman–Kac operator takes its conditional expectation weighted by the accumulated cost. In the two directions,

$$\begin{aligned}(K_c^{\rightarrow} f)(x) &= \mathbb{E}_{\mathsf{R}}[e^{-C} f(X_T) \mid X_0 = x] = \int k_c^{\rightarrow}(x,y) f(y)\,dy,\\ (K_c^{\leftarrow} h)(y) &= \mathbb{E}_{\mathsf{R}}[e^{-C} h(X_0) \mid X_T = y] = \int k_c^{\leftarrow}(y,x) h(x)\,dx.\end{aligned} \tag{22}$$

Multiplying either kernel by its corresponding reference marginal gives the same joint endpoint measure $\Gamma_{0T}(dx,dy)$. Therefore,

$$r_0(x) k_c^{\rightarrow}(x,y) = r_T(y) k_c^{\leftarrow}(y,x). \tag{23}$$

Integrating against $h(x)f(y)$ gives the compatibility identity used later to recover both endpoint constraints:

$$\int h\, K_c^{\rightarrow} f\, d\mathsf{R}_0 = \int f\, K_c^{\leftarrow} h\, d\mathsf{R}_T. \tag{24}$$

The density factors r_0, r_T are essential for a nonstationary reference. When $r_0 = r_T = 1$, the two kernels are transposes with respect to volume.

Assumption 1 (Assumptions for convergence)**.** (A1) **Reference and state space.** *$\mathcal{X}$ is compact with a full-support finite volume measure dx, and $T > 0$. The reference R and its time reversal are Markov diffusions. Their marginal densities r_t are strictly positive and continuous. A diffusion on $\mathbb{T}^d$ is one example.*

(A2) **Uniformly positive propagation.** *c is bounded and continuous. Both kernels in equation 22 are continuous, with $0 < \underline{k}_d \le k_c^d \le \overline{k}_d < \infty$ for $d \in \{\rightarrow, \leftarrow\}$.*

(A3) **Endpoints and initialization.** *The endpoint density ratios α_0, α_1 in equation 3 are strictly positive and continuous, and V_1^0 is finite and continuous.*

(A4) **Exact SOC access.** *In both time directions, the reference and admissible controlled martingale problems are well posed. The conditional Girsanov entropy identity equation 25 holds for the admissible finite-energy controls. For every bounded continuous terminal cost used below, Feynman–Kac gives a continuous value function on $[0,T] \times \mathcal{X}$, and SOC verification identifies it with the control infimum. The Gibbs conditional law in equation 27 is realizable for every starting state. Whenever a gradient feedback is used, the value is spatially differentiable for $t < T$ and the feedback is admissible. Each exact solver returns this value function up to one additive constant and realizes its optimal conditional law.*

The path-space control identity in (A4) is standard in SOC (Nüsken & Richter, 2021; Domingo-Enrich et al., 2025). For a fixed starting state it reads

$$\mathrm{KL}(\mathsf{P}^{u,0,x} \| \mathsf{R}^{0,x}) = \frac{1}{2} \mathbb{E}_{\mathsf{P}^{u,0,x}} \int_0^T \|u_t\|^2 \, dt. \tag{25}$$

We use controls adapted to the state filtration and with no component in $\ker \sigma_t(X_t)$. Under the usual Girsanov integrability conditions, this identity holds, for example, for invertible diffusion coefficients. For a rectangular coefficient, the minimum-norm representation $u_t \in \mathrm{range}(\sigma_t(X_t)^\top)$ removes diffusion-null components: they change the control energy but not the state law. Discarding them leaves the SOC optimum unchanged. The reverse problem uses the same convention with $\bar{\sigma}_s$. Assumptions (A1)–(A3) give the potential contraction; (A4) transfers that result to exact SOC solves. The positive lower kernel bound is what makes the contraction uniform over all starting states.

Algorithm 1 takes the forward and reverse reference dynamics and the relative endpoint energies as inputs. When endpoints are supplied through volume-density energies E_i, equation 4 also requires the reference endpoint log densities $\log r_0, \log r_T$ up to constants. The contraction argument also applies to finite-state Markov references. On noncompact spaces, including $\mathbb{R}^d$, the uniform bounds must be verified separately; the compact-space theorem does not assert convergence there without additional assumptions.

For example, stationary Brownian motion on the unit torus with normalized volume gives $r_t = 1$ and $\widetilde{E}_i = E_i$. A stationary reversible Langevin reference

$$dX_t = \frac{\sigma^2}{2} \nabla \log \pi(X_t) \, dt + \sigma \, dW_t$$

has $r_t = \pi$, the same forward and reverse drift, and endpoint costs $E_i + \log \pi$. Nonstationary and nonreversible references are also covered under Assumption 1.

A.2 SOC VALUE FUNCTIONS

The following logarithmic Feynman–Kac representation is a standard connection between SOC and changes of path measure (Nüsken & Richter, 2021). We retain the conditioning on the starting state because its normalizer is precisely the boundary value needed by the next SPIM update. For a terminal cost g,

$$\phi_g(t,x) = \mathbb{E}_{\mathsf{R}^{t,x}} \left[e^{-\int_t^T c_r(X_r)\, dr - g(X_T)} \right], \qquad V_g(t,x) = -\log \phi_g(t,x). \tag{26}$$

The optimal conditional path law is

$$\mathsf{Q}_g(d\omega \mid x) = \frac{e^{-C(\omega) - g(X_T)}}{\phi_g(0,x)} \mathsf{R}(d\omega \mid X_0 = x). \tag{27}$$

For $\mathsf{P} = \mathsf{P}^{u,0,x}$, taking the logarithm of equation 27 gives

$$\mathrm{KL}(\mathsf{P} \| \mathsf{Q}_g(\cdot \mid x)) = \mathrm{KL}(\mathsf{P} \| \mathsf{R}^{0,x}) + \mathbb{E}_{\mathsf{P}}[C + g(X_T)] - V_g(0,x).$$

Combining this with equation 25 yields

$$\mathbb{E}_{\mathsf{P}^{u,0,x}} \left[\frac{1}{2} \int_0^T \|u_t\|^2 \, dt + C + g(X_T) \right] = V_g(0,x) + \mathrm{KL}(\mathsf{P}^{u,0,x} \| \mathsf{Q}_g(\cdot \mid x)). \tag{28}$$

Algorithm 1 SPIM: alternating Schrödinger-potential updates

Require: Relative endpoint energies $\widetilde{E}_0, \widetilde{E}_1$; fixed reference laws $\mathsf{R}, \mathsf{R}^{\leftarrow}$; running costs $c, \bar{c}_s = c_{T-s}$; SOC solvers $\mathcal{S}_{\rightarrow}, \mathcal{S}_{\leftarrow}$; rounds $M \geq 1$; anchor $x_\dagger$

Ensure: Final boundary functions and directional controls

Choose $V_1^0 \in C(\mathcal{X})$, e.g. $V_1^0 \equiv 0$; set $V_1^0 \leftarrow V_1^0 - V_1^0(x_\dagger)$

for $k = 0, \dots, M-1$ **do**

Forward half-step

$g_1^k \leftarrow \widetilde{E}_1 - V_1^k$

$(V_k^{\rightarrow}, u_k^{\rightarrow}) \leftarrow \mathcal{S}_{\rightarrow}(\mathsf{R}, c, g_1^k)$

$V_0^k \leftarrow V_k^{\rightarrow}(0, \cdot)$

Reverse half-step

$g_0^k \leftarrow \widetilde{E}_0 - V_0^k$

$(V_k^{\leftarrow}, u_k^{\leftarrow}) \leftarrow \mathcal{S}_{\leftarrow}(\mathsf{R}^{\leftarrow}, \bar{c}, g_0^k)$

$V_1^{k+1} \leftarrow V_k^{\leftarrow}(0, \cdot)$

$V_1^{k+1} \leftarrow V_1^{k+1} - V_1^{k+1}(x_\dagger)$

Optional: refresh exploration banks using the current controls.

return $V_0^{M-1}, V_1^M, u_{M-1}^{\rightarrow}, u_{M-1}^{\leftarrow}$

Each directional solver returns a value function and a control for the specified terminal-cost SOC problem. Bank refresh changes the numerical training states; it is not part of the exact potential map.

The nonnegative KL term proves optimality and uniqueness of the conditional path law Q_g. Under the differentiability and verification conditions in (A4), its forward control for $t < T$ is $u_g^\star(t,x) = -\sigma_t(x)^\top \nabla V_g(t,x)$. The same calculation under $\mathsf{R}^{\leftarrow}$ gives the reverse SOC law and control $-\bar{\sigma}_s^\top \nabla V_g^{\leftarrow}$.

Adding a constant a to g gives $V_{g+a} = V_g + a$ and leaves Q_g unchanged. Hence replacing $-\log \alpha_i = \widetilde{E}_i + \log Z_i$ by $\widetilde{E}_i$ only changes additive constants in the value functions. Anchoring the values has the same effect. To simplify the proof, the identities below use $-\log \alpha_i$ and values before anchoring. The convergence bounds remain valid for the relative endpoint energies $\widetilde{E}_i$ and anchoring used in Algorithm 1.

A.3 Schrödinger potential updates

We retain the notation of Section 3.2: $V_0^k = V_k^{\rightarrow}(0,\cdot)$ and $V_1^{k+1} = V_k^{\leftarrow}(0,\cdot)$. The corresponding Schrödinger potentials satisfy

$$\begin{aligned} \phi_{T,k} &= \alpha_1 / \widehat{\phi}_{T,k}, \qquad \phi_{0,k} = K_c^{\rightarrow} \phi_{T,k}, \\ \widehat{\phi}_{0,k} &= \alpha_0 / \phi_{0,k}, \quad \widehat{\phi}_{T,k+1} = K_c^{\leftarrow} \widehat{\phi}_{0,k}, \end{aligned} \tag{29}$$

where $\phi_{0,k} = e^{-V_0^k}$ and $\widehat{\phi}_{T,k} = e^{-V_1^k}$. The reverse SOC uses the newly computed $\phi_{0,k}$. One forward solve followed by one reverse solve therefore gives

$$\widehat{\phi}_{T,k+1} = \mathcal{T}\widehat{\phi}_{T,k}, \qquad \mathcal{T}f = K_c^{\leftarrow}\left(\frac{\alpha_0}{K_c^{\rightarrow}(\alpha_1 / f)}\right). \tag{30}$$

Each division sets an endpoint factor; the following kernel propagates it across the time interval. This is the classical iteration for Schrödinger potentials (Chen et al., 2016a). The initialization $V_1^0 = 0$ gives $\widehat{\phi}_{T,0} = 1$. The optional zero-terminal-cost reverse solve instead gives $\widehat{\phi}_{T,0} = K_c^{\leftarrow}\mathbf{1}$ before anchoring. Both are strictly positive and continuous. The proof permits every positive continuous initial potential, equivalently every $V_1^0 \in C(\mathcal{X})$.

Using $e^{-\widetilde{E}_i} = Z_i \alpha_i$ instead of α_i multiplies the forward output $\phi_{0,k}$ by Z_1 and the map $\mathcal{T}$ by Z_0/Z_1. These constants do not affect normalized potentials or conditional path laws. Thus the updates require neither Z_0 nor Z_1.

These updates also have the classical IPF interpretation. The law $\widehat{\phi}_{0,k}(X_0)\phi_{T,k}(X_T)\Gamma$ has source marginal μ_0. Replacing $\phi_{T,k}$ by $\alpha_1/(K_c^{\leftarrow}\widehat{\phi}_{0,k})$ enforces the target marginal while preserving the conditional paths given X_T. The convergence proof below works directly with the Schrödinger potentials.

B Convergence of SPIM

We first establish convergence modulo the arbitrary potential scale, then use conservation of endpoint mass to identify an actual bridge. Finally, we bound the SOC values and verify global optimality. The iteration in this section uses sequential boundary exchange; the simultaneous variant is discussed in Appendix C.1.

B.1 Potential iteration is a contraction

The positive-kernel argument of Chen, Georgiou, and Pavon (Chen et al., 2016a) applies to the two operators in equation 22. Assumption (A2) supplies a positive lower bound and finite upper bound for each kernel, and (A3) supplies the positive endpoint ratios. The kernels need not integrate to one. Compatibility of the two operators follows from equation 24.

For positive continuous functions, define

$$d_H(f,g) = \operatorname{osc}(\log f - \log g), \qquad \operatorname{osc}(z) = \sup z - \inf z. \tag{31}$$

Thus d_H measures the variation of the log ratio over the state space. It is zero exactly when f and g differ by a positive constant factor, which is the irrelevant scale freedom of the potentials. For $d \in \{\rightarrow, \leftarrow\}$, set

$$\tau_d = \frac{\overline{k}_d - \underline{k}_d}{\overline{k}_d + \underline{k}_d} < 1, \qquad \kappa = \tau_{\rightarrow}\tau_{\leftarrow} < 1. \tag{32}$$

For every positive f, $\underline{k}_d \int f\,dx \le K_c^d f \le \overline{k}_d \int f\,dx$. Consequently its projective diameter $\Delta(K_c^d) := \sup_{f,g>0} d_H(K_c^d f, K_c^d g)$ is at most $2\log(\overline{k}_d/\underline{k}_d)$. The Birkhoff–Bushell contraction theorem gives coefficient $\tanh(\Delta(K_c^d)/4) \le \tau_d$; see Chen et al. (2016a) for this standard result and its use in bridge iteration. In particular,

$$\begin{aligned} d_H(K_c^{\rightarrow} f, K_c^{\rightarrow} g) &\le \tau_{\rightarrow} d_H(f,g), \\ d_H(K_c^{\leftarrow} f, K_c^{\leftarrow} g) &\le \tau_{\leftarrow} d_H(f,g). \end{aligned} \tag{33}$$

Each division $f \mapsto \alpha_i/f$ preserves d_H. Consequently, the full map in equation 30 contracts by κ. To make the fixed-point argument precise, choose $x_\dagger \in \mathcal{X}$ and use the representatives

$$C_\dagger(\mathcal{X}) = \{h \in C(\mathcal{X}) : h(x_\dagger) = 0\}.$$

On this closed subspace, $\|h\|_\infty \le \operatorname{osc}(h) \le 2\|h\|_\infty$, so the oscillation norm is complete. Positive potentials modulo scaling are identified with $C_\dagger(\mathcal{X})$ by the map $f \mapsto -\log f + \log f(x_\dagger)$. Moreover, $\mathcal{T}(af) = a\mathcal{T}(f)$ for $a > 0$, so $h \mapsto -\log \mathcal{T}(e^{-h}) + \log \mathcal{T}(e^{-h})(x_\dagger)$ is a well-defined κ-contraction on this complete space. Banach's theorem therefore gives a unique fixed point of the anchored negative-log-potential map, hence convergence from every positive continuous initialization. Before normalization, this only establishes $\mathcal{T} f^\star = \lambda f^\star$; the identity $\lambda = 1$ is established in Appendix B.2, not assumed here. The familiar τ^2 bound is the case in which both directional bounds equal τ.

In the boundary-value notation of Section 3.2, define the exact maps using normalized endpoint costs:

$$\mathcal{F}(V) = -\log K_c^{\rightarrow}(\alpha_1 e^V), \qquad \mathcal{G}(V) = -\log K_c^{\leftarrow}(\alpha_0 e^V). \tag{34}$$

Then $V_0^k = \mathcal{F}(V_1^k)$ and $V_1^{k+1} = \mathcal{G}(V_0^k)$, and

$$\begin{aligned} \operatorname{osc}(\mathcal{F}(V) - \mathcal{F}(\widetilde{V})) &\le \tau_{\rightarrow} \operatorname{osc}(V - \widetilde{V}), \\ \operatorname{osc}(\mathcal{G}(V) - \mathcal{G}(\widetilde{V})) &\le \tau_{\leftarrow} \operatorname{osc}(V - \widetilde{V}). \end{aligned} \tag{35}$$

The identities $\mathcal{F}(V + a) = \mathcal{F}(V) - a$ and $\mathcal{G}(V + a) = \mathcal{G}(V) - a$ show how constants pass between the two directions. Anchoring after an update therefore preserves every oscillation bound. Replacing the normalized costs $-\log \alpha_i$ by $\widetilde{E}_i$ also changes only constants. Thus the bounds apply directly to the stored SOC boundary values.

B.2 The fixed point satisfies both endpoint constraints

The contraction determines a potential up to scale. The following mass-balance argument shows that its eigenvalue is one, so it solves the normalized Schrödinger system. Choose a representative $\widehat{\phi}_T^\star$ of the limiting potential and define

$$\phi_T^\star = \frac{\alpha_1}{\widehat{\phi}_T^\star}, \qquad \phi_0^\star = K_c^{\rightarrow}\phi_T^\star, \qquad \widehat{\phi}_0^\star = \frac{\alpha_0}{\phi_0^\star}. \tag{36}$$

Convergence up to a positive constant gives $K_c^{\leftarrow}\widehat{\phi}_0^\star = \lambda\widehat{\phi}_T^\star$ for some $\lambda > 0$. Using equation 24 and $\int \alpha_0 \, d\mathsf{R}_0 = \int \alpha_1 \, d\mathsf{R}_T = 1$ gives

$$\begin{aligned} 1 &= \int \widehat{\phi}_0^\star K_c^{\rightarrow}\phi_T^\star \, d\mathsf{R}_0 = \int \phi_T^\star K_c^{\leftarrow}\widehat{\phi}_0^\star \, d\mathsf{R}_T \\ &= \lambda \int \alpha_1 \, d\mathsf{R}_T = \lambda. \end{aligned} \tag{37}$$

Hence the limiting potentials satisfy both propagation equations and both endpoint conditions of the Schrödinger system. In particular,

$$\mathsf{P}^\star(d\omega) = \widehat{\phi}_0^\star(X_0)\phi_T^\star(X_T)\Gamma(d\omega) \tag{38}$$

is a probability law with marginals $\alpha_0\mathsf{R}_0 = \mu_0$ and $\alpha_1\mathsf{R}_T = \mu_1$. At intermediate times, define $\phi_t^\star, \widehat{\phi}_t^\star$ by the two conditional expectations in equation 5. Conditional independence of past and future given X_t under the Markov reference gives $d\mathsf{P}_t^\star / d\mathsf{R}_t = \phi_t^\star\widehat{\phi}_t^\star$. Thus the endpoint construction extends to the full Schrödinger system and the path law in equation 38.

The terminal costs $-\log\phi_T^\star$ and $-\log\widehat{\phi}_0^\star$ give the SOC value functions $V_\star^{\rightarrow}, V_\star^{\leftarrow}$ in Proposition 3.1. Their boundary values satisfy $V_0^\star = -\log\phi_0^\star = \mathcal{F}(V_1^\star)$ and $V_1^\star = -\log\widehat{\phi}_T^\star = \mathcal{G}(V_0^\star)$. Conditioning equation 38 gives the optimal SOC laws in equation 27. We prove global optimality of $\mathsf{P}^\star$ in Appendix B.4 below.

B.3 Transferring convergence to SOC values

We also measure the endpoint constraints directly from the two values. Define the endpoint residuals

$$\begin{aligned} \delta_0^k(x) &= V_k^{\rightarrow}(0,x) + V_k^{\leftarrow}(T,x) - \widetilde{E}_0(x), \\ \delta_1^k(x) &= V_k^{\rightarrow}(T,x) + V_k^{\leftarrow}(0,x) - \widetilde{E}_1(x). \end{aligned} \tag{39}$$

These residuals are zero up to constants when the endpoint products are satisfied. The reverse time argument is T at the source and 0 at the target.

Theorem B.1 (Convergence bounds)**.** *Under Assumption 1, define* $e_k = \operatorname{osc}(V_1^k - V_1^\star)$*. Then*

$$e_{k+1} \le \kappa e_k, \qquad \operatorname{osc}(V_0^k - V_0^\star) \le \tau_{\rightarrow} e_k. \tag{40}$$

Anchor each value function by $\bar{V}(t,x) = V(t,x) - V(0,x_\dagger)$*, as in Theorem 3.2. Over all times and states,*

$$\|\bar{V}_k^{\rightarrow} - \bar{V}_\star^{\rightarrow}\|_\infty + \|\bar{V}_k^{\leftarrow} - \bar{V}_\star^{\leftarrow}\|_\infty \le (1+\tau_{\rightarrow})e_k \le (1+\tau_{\rightarrow})\kappa^k e_0. \tag{41}$$

For the endpoint residuals in equation 39,

$$\operatorname{osc}(\delta_0^k) = 0, \qquad \operatorname{osc}(\delta_1^k) \le (1+\kappa)e_k \le (1+\kappa)\kappa^k e_0. \tag{42}$$

Proof. Boundary values. Apply equation 35 to $V_0^k = \mathcal{F}(V_1^k)$, comparing it with $V_0^\star = \mathcal{F}(V_1^\star)$. Applying the same bound to $\mathcal{G}$ gives equation 40.

Values at all times. If terminal costs satisfy $L \le g - \widetilde{g} \le H$, positivity of the Feynman–Kac expectation gives

$$L \le V_g(t,x) - V_{\widetilde{g}}(t,x) \le H \quad \text{for every } (t,x). \tag{43}$$

The forward terminal costs differ by $-(V_1^k - V_1^\star)$, with oscillation e_k. The reverse terminal costs differ by $-(V_0^k - V_0^\star)$, with oscillation at most $\tau_{\to} e_k$. For the forward problem, take $L = \inf_x[-(V_1^k - V_1^\star)]$ and $H = \sup_x[-(V_1^k - V_1^\star)]$, allowing the same additional constant in both bounds if unnormalized energies are used. The same interval $[L, H]$ bounds the value difference for all times and states, including $(0, x_\dagger)$. Subtracting its value at that anchor therefore gives a uniform bound of $H - L = e_k$, rather than requiring convergence of the additive constants. The identical argument for the reverse problem gives $\tau_{\to} e_k$. Adding the bounds proves equation 41.

Endpoint residuals. The terminal costs and the definitions of V_0^k, V_1^{k+1} give

$$\delta_0^k = 0, \qquad \delta_1^k = V_1^{k+1} - V_1^k. \tag{44}$$

These identities use the relative costs $\widetilde{E}_i$ and values before anchoring. Using $-\log \alpha_i$ or anchoring only adds constants. The first residual therefore has zero oscillation. The second has oscillation at most $e_k + e_{k+1} \le (1+\kappa) e_k$, proving equation 42.

□

The exact identity for δ_0^k follows from using the newly computed V_0^k in the reverse terminal cost. It does not imply that a finite-iteration rollout from an arbitrary initial bank has the desired endpoint marginal. The remaining residual measures the change in the boundary iterate, modulo constants.

Subtracting the endpoint residual at $x_\dagger$ bounds its uniform norm by its oscillation:

$$\max_{i=0,1} \|\delta_i^k - \delta_i^k(x_\dagger)\|_\infty \le (1+\kappa)\kappa^k e_0. \tag{45}$$

Taking any $r \in (\kappa, 1)$ and sufficiently large A in equation 41 proves the value bound in Theorem 3.2. If $\kappa = 0$, the boundary values are exact up to constants after one iteration, and any $r \in (0,1)$ gives the same conclusion.

What the endpoint residual measures. The two directional values define density estimates relative to the reference. Their sum at endpoint i equals $\widetilde{E}_i + \delta_i^k$, so for $i \in \{0,1\}$ define

$$\widehat{\mu}_i^k(dx) = \frac{e^{-\widetilde{E}_i(x) - \delta_i^k(x)} \mathsf{R}_{iT}(dx)}{\int e^{-\widetilde{E}_i - \delta_i^k} \, d\mathsf{R}_{iT}}. \tag{46}$$

Here iT is 0 for the source and T for the target. The relative energy cancels against the target density ratio, giving

$$\log \frac{d\widehat{\mu}_i^k}{d\mu_i} = -\delta_i^k - \log \int e^{-\delta_i^k} \, d\mu_i.$$

Its absolute value is bounded by $\operatorname{osc}(\delta_i^k)$, which tends to zero by equation 42. Since the target densities are continuous on a compact space, their estimated volume densities also converge uniformly. Here r_{iT} converts the estimated density ratio into a volume density.

B.4 The limit minimizes the bridge objective

We use the classical relative-entropy projection argument for Schrödinger bridges (Léonard, 2013). The calculation below includes the running cost through Γ and establishes optimality among all feasible path laws.

The law $\mathsf{P}^\star$ in equation 38 has the required endpoint marginals. Its endpoint potentials are bounded and bounded away from zero, so $J(\mathsf{P}^\star)$ is finite. For any feasible law P with finite $J(\mathsf{P})$, the factorization gives

$$J(\mathsf{P}) = \mathrm{KL}(\mathsf{P} \| \mathsf{P}^\star) + \mathbb{E}_{\mathsf{P}}\left[\log \widehat{\phi}_0^\star(X_0) + \log \phi_T^\star(X_T)\right]. \tag{47}$$

The expectation depends only on the two endpoint marginals. Since P and $\mathsf{P}^\star$ share those marginals, it equals

$$J(\mathsf{P}^\star) = \int \log \widehat{\phi}_0^\star \, d\mu_0 + \int \log \phi_T^\star \, d\mu_1.$$

Consequently,

$$J(\mathsf{P}) - J(\mathsf{P}^\star) = \mathrm{KL}(\mathsf{P} \| \mathsf{P}^\star) \geq 0. \tag{48}$$

Equality holds only for $\mathsf{P} = \mathsf{P}^\star$; laws with infinite objective cannot improve on $\mathsf{P}^\star$. This proves that the limit of SPIM determines the unique optimal generalized Schrödinger bridge and completes the proof of Theorem 3.2.

C Approximate SOC Solvers

The exact proof also explains how inner-solver errors accumulate: each new error is added to earlier errors that have been reduced by the contraction. The relevant error is in the *scalar boundary value over the state space*, modulo a constant.

As in the exact algorithm, $\widetilde{V}_0^k, \widetilde{V}_1^{k+1}$ denote finite continuous approximations to the SOC boundary values, with arbitrary additive normalization. Measure their errors relative to the exact updates by

$$\begin{aligned} \varepsilon_k^{\rightarrow} &= \mathrm{osc}\Big(\widetilde{V}_0^k - \mathcal{F}(\widetilde{V}_1^k)\Big), \\ \varepsilon_k^{\leftarrow} &= \mathrm{osc}\Big(\widetilde{V}_1^{k+1} - \mathcal{G}(\widetilde{V}_0^k)\Big). \end{aligned} \tag{49}$$

These errors are unchanged by additive constants. Define $\widetilde{e}_k = \mathrm{osc}(\widetilde{V}_1^k - V_1^\star)$ and $\eta_k = \tau_{\leftarrow}\varepsilon_k^{\rightarrow} + \varepsilon_k^{\leftarrow}$.

Theorem C.1 (Error bound for SPIM)**.** *Under Assumption 1, except for exactness of the SOC solves, the approximate updates satisfy*

$$\begin{aligned} \widetilde{e}_{k+1} &\leq \kappa \widetilde{e}_k + \eta_k, \\ \widetilde{e}_N &\leq \kappa^N \widetilde{e}_0 + \sum_{k=0}^{N-1} \kappa^{N-1-k} \eta_k. \end{aligned} \tag{50}$$

The forward boundary error also satisfies

$$\mathrm{osc}(\widetilde{V}_0^k - V_0^\star) \leq \tau_{\rightarrow}\widetilde{e}_k + \varepsilon_k^{\rightarrow}. \tag{51}$$

If both directional errors tend to zero, both boundary functions converge up to additive constants. If $\varepsilon_k^d \leq \bar{\varepsilon}_d$, define $\bar{\eta} = \tau_{\leftarrow}\bar{\varepsilon}_{\rightarrow} + \bar{\varepsilon}_{\leftarrow}$. Then

$$\begin{aligned} \limsup_k \widetilde{e}_k &\leq \frac{\bar{\eta}}{1-\kappa}, \\ \limsup_k \mathrm{osc}(\widetilde{V}_0^k - V_0^\star) &\leq \bar{\varepsilon}_{\rightarrow} + \frac{\tau_{\rightarrow}\bar{\eta}}{1-\kappa}. \end{aligned}$$

Proof. The forward error and equation 35 give

$$\mathrm{osc}(\widetilde{V}_0^k - V_0^\star) \leq \varepsilon_k^{\rightarrow} + \tau_{\rightarrow}\widetilde{e}_k.$$

Applying the reverse error bound next gives $\widetilde{e}_{k+1} \leq \varepsilon_k^{\leftarrow} + \tau_{\leftarrow}(\varepsilon_k^{\rightarrow} + \tau_{\rightarrow}\widetilde{e}_k)$. Iterating this inequality proves equation 50. If both errors vanish, split the sum into a finite initial part and a tail: the first decays geometrically, and the second is arbitrarily small because $\eta_k \to 0$. Thus $\widetilde{e}_k \to 0$, and the forward bound gives convergence of $\widetilde{V}_0^k$ as well. A geometric-series bound proves the stated limit for uniformly bounded η_k.

□

Interpreting the error assumption. A uniform boundary-value error of at most ϵ, after subtracting a constant, implies an oscillation error of at most 2ϵ. An empirical training loss alone does not give such a bound. Each solver must estimate the scalar boundary value over the state space; a solver that learns only a control requires an additional value reconstruction step.

For an exploration law ξ with full support, integrating the *excess SOC cost* in equation 28 gives

$$\int \mathrm{KL}\big(\mathsf{P}^{u,0,x} \| \mathsf{Q}_g(\cdot \mid x)\big)\, \xi(dx) \geq 0.$$

Its minimum is zero when the optimal conditional laws are jointly realizable by an admissible control. A minimizer then identifies those laws ξ-almost everywhere, and everywhere if both conditional laws depend weakly continuously on x. This concerns policy optimality; it does not bound the scalar approximation errors in equation 49. A policy-only solver must additionally recover $V_g(0, x)$ up to one constant, for example by estimating its conditional cost-to-go. Finite-sample training also requires adequate coverage of the states where boundary values are exchanged.

These bounds concern boundary values. Convergence of their gradients, or of learned controls, requires additional regularity and approximation bounds. With a changing reference or running cost, the theorem applies within a block where the reference, costs, and endpoint laws are fixed; it gives no rate across different bridge problems.

C.1 Exact Jacobi updates

Under Assumption 1, the same contraction estimates give a separate statement for simultaneous boundary exchange. Take arbitrary $V_0^0, V_1^0 \in C(\mathcal{X})$ and index both functions by their common round: $V_0^{k+1} = \mathcal{F}(V_1^k)$ and $V_1^{k+1} = \mathcal{G}(V_0^k)$. For $e_{i,k} = \mathrm{osc}(V_i^k - V_i^\star)$,

$$e_{0,k+1} \leq \tau_{\rightarrow} e_{1,k}, \qquad e_{1,k+1} \leq \tau_{\leftarrow} e_{0,k}, \qquad e_{i,k+2} \leq \kappa e_{i,k}.$$

Thus each boundary function converges modulo constants over pairs of rounds. This follows by composing the same two inequalities as in the sequential proof. It does not transfer the sequential endpoint identity equation 44 to a Jacobi round, or establish a bound for neural target-network interpolation.

D Endpoint Sampling

We now hold the optimal controls fixed and study repeated forward and reverse sampling. This is a statement about a Markov chain after optimization, separate from convergence of the training iteration.

Invariant endpoint laws. The optimal forward and reverse controls induce the endpoint transition kernels

$$\begin{aligned} K_{01}^\star(x, dy) &= \frac{k_c^{\rightarrow}(x, y)\phi_T^\star(y)}{\phi_0^\star(x)}\, dy, \\ K_{10}^\star(y, dx) &= \frac{k_c^{\leftarrow}(y, x)\widehat{\phi}_0^\star(x)}{\widehat{\phi}_T^\star(y)}\, dx. \end{aligned} \tag{52}$$

They are the two conditional distributions of the endpoint coupling

$$\pi^\star(dx, dy) = \widehat{\phi}_0^\star(x)\phi_T^\star(y) r_0(x) k_c^{\rightarrow}(x, y)\, dx\, dy.$$

Alternating these kernels is therefore a two-block Gibbs sampler (Roberts & Rosenthal, 2004, Section 2.3). The source-to-source transition $M_0 = K_{01}^\star K_{10}^\star$ preserves μ_0, and $\mu_0 K_{01}^\star = \mu_1$.

Convergence from arbitrary starting states. Under (A1)–(A3), the density ratio

$$\frac{d\pi^\star}{d(\mu_0 \otimes \mu_1)}(x, y) = \frac{k_c^{\rightarrow}(x, y)}{r_T(y)\phi_0^\star(x)\widehat{\phi}_T^\star(y)}$$

is continuous and strictly positive on a compact space. It therefore has a minimum $\epsilon > 0$. Denote this ratio by $h(x, y)$. The marginal identities give $\int h(x, y)\mu_1(dy) = \int h(x, y)\mu_0(dx) = 1$ for each fixed value of the other variable, hence $\epsilon \leq 1$. Moreover,

$$K_{01}^\star(x, dy) = h(x, y)\mu_1(dy), \qquad K_{10}^\star(y, dx) = h(x, y)\mu_0(dx).$$

In particular, $K_{10}^\star(y, A) \geq \epsilon\mu_0(A)$ for every measurable A. Integrating against $K_{01}^\star(x, dy)$ proves $M_0(x, A) \geq \epsilon\mu_0(A)$. This is a Doeblin minorization, so the standard uniform-ergodicity theorem applies (Roberts & Rosenthal, 2004, Theorem 8). Here its bound follows directly: for $\epsilon < 1$, write $M_0 = \epsilon\mu_0 + (1 - \epsilon)N$, where N is a Markov kernel preserving μ_0. Using $\|\nu - \mu\|_{\mathrm{TV}} = \sup_A |\nu(A) - \mu(A)|$ and the nonexpansiveness of a Markov kernel, any initial source law ν obeys

$$\|\nu M_0^n - \mu_0\|_{\mathrm{TV}} \leq (1 - \epsilon)^n \|\nu - \mu_0\|_{\mathrm{TV}}.$$

Applying $K_{01}^\star$ gives convergence of the target samples to μ_1. If $\epsilon = 1$, one full cycle is sufficient.

Scope. This sampling result assumes exact simulation of the fixed optimal kernels. Errors in learned controls or time discretization require a separate analysis. Refreshing an exploration bank with the changing controls $u_k^{\rightarrow}, u_k^{\leftarrow}$ during Algorithm 1 generally uses different kernels at different iterations. The stationarity and minorization argument above applies to the limiting optimal kernels and supplies no convergence claim for that adaptive bank-refresh procedure.

E SUPPLEMENTARY EXPERIMENTS

We first derive the LiDAR and image-control formulations without endpoint energy models, then describe the algorithmic design, experimental settings, additional metrics, and transport visualizations.

E.1 APPLICATIONS WITHOUT ENDPOINT ENERGY MODELS

The SOC boundary conditions depend on endpoint density *ratios* relative to the reference. A reference bridge learned from samples can therefore remove the need to evaluate endpoint densities. Below, the exact statements assume a Markov reference with mutually consistent forward and reverse dynamics and the bridge existence and SOC-verification conditions of Appendix A.

E.1.1 KNOWN REWEIGHTING OF REFERENCE MARGINALS

Let $\nu_0 = \mathsf{R}_0$, $\nu_1 = \mathsf{R}_T$, and suppose the desired endpoints are known reweightings

$$\mu_i(dx) = Z_i^{-1} w_i(x)\nu_i(dx), \qquad Z_i = \int w_i \, d\nu_i \in (0,\infty), \quad i \in \{0,1\}, \tag{53}$$

with positive weights w_i. Although the densities of ν_i may be unknown, $-\log(d\mu_i/d\nu_i) = -\log w_i + \log Z_i$ is known up to a constant. Proposition 3.1 thus gives the optimal coupled terminal costs, up to additive constants,

$$\begin{aligned} g_1^\star &= -\log w_1 - V_\star^{\leftarrow}(0,\cdot), \\ g_0^\star &= -\log w_0 - V_\star^{\rightarrow}(0,\cdot). \end{aligned} \tag{54}$$

Neither endpoint density nor its normalizing constant is needed by these SOC problems.

To see why the resulting law is optimal, write $C(\omega) = \int_0^T c_t(X_t)\,dt$ and let the endpoint factors enforce $\mathsf{P}_0^\star = \mu_0$, $\mathsf{P}_T^\star = \mu_1$ in

$$\frac{d\mathsf{P}^\star}{d\mathsf{R}} = f_0(X_0) f_1(X_T) e^{-C}.$$

For every feasible finite-entropy path law P, the endpoint terms have the same expectations under P and $\mathsf{P}^\star$. Taking logarithms therefore gives

$$\begin{aligned} &\mathrm{KL}(\mathsf{P}\|\mathsf{R}) + \mathbb{E}_{\mathsf{P}} C - \mathrm{KL}(\mathsf{P}^\star\|\mathsf{R}) - \mathbb{E}_{\mathsf{P}^\star} C \\ &\qquad = \mathrm{KL}(\mathsf{P}\|\mathsf{P}^\star) \geq 0. \end{aligned} \tag{55}$$

Equality holds only at $\mathsf{P} = \mathsf{P}^\star$. This proves that the coupled SOC solution has the prescribed endpoints and uniquely minimizes the generalized bridge objective.

E.1.2 LIDAR: TERRAIN CONTROL WITH FIXED ENDPOINTS

Let B be the Brownian reference and let $\mathsf{R}^{(0)}$ be its zero-cost bridge between the desired terrain endpoint laws μ_0, μ_1. This reference can be learned from endpoint samples. Because its endpoint marginals already equal μ_i, the weights in equation 53 are $w_0 = w_1 = 1$. Consequently,

$$g_1^\star = -V_\star^{\leftarrow}(0,\cdot), \qquad g_0^\star = -V_\star^{\rightarrow}(0,\cdot).$$

The geometry-based cost in equation 61 then penalizes terrain departures while the boundary corrections enforce the endpoint laws, without evaluating their energies.

The reference change also preserves the desired Brownian-relative optimum. More generally, suppose the preceding homotopy level is the exact optimum for strength $\lambda_{\ell-1}$, so that

$$\frac{d\mathsf{R}^{(\ell-1)}}{d\mathsf{B}} = a_{0,\ell-1}(X_0) a_{1,\ell-1}(X_T) e^{-\lambda_{\ell-1} C}.$$

For any P with the same prescribed endpoints,

$$\begin{aligned} &\mathrm{KL}(\mathsf{P}\|\mathsf{R}^{(\ell-1)}) + (\lambda_\ell - \lambda_{\ell-1})\mathbb{E}_{\mathsf{P}} C \\ &\qquad = \mathrm{KL}(\mathsf{P}\|\mathsf{B}) + \lambda_\ell \mathbb{E}_{\mathsf{P}} C - \kappa_{\ell-1}, \end{aligned} \tag{56}$$

where $\kappa_{\ell-1} = \int \log a_{0,\ell-1}\, d\mu_0 + \int \log a_{1,\ell-1}\, d\mu_1$ is independent of P. Starting at $\lambda_0 = 0$, exact incremental solves and reference updates therefore recover the full terrain-cost optimum at $\lambda_L = 1$. This justifies the incremental cost and zero relative endpoint-energy terms used in the LiDAR homotopy design.

E.1.3 MNIST–EMNIST: OPTIMAL REWARD-WEIGHTED ENDPOINTS

Let the pretrained reference have digit and letter marginals ν_0, ν_1, and let r_i be the corresponding reward. The desired endpoint $\mu_i(dx) = Z_i^{-1} e^{\lambda r_i(x)} \nu_i(dx)$ is itself the unique KL-regularized reward optimum:

$$\mu_i = \arg\max_{\pi \ll \nu_i} \{\lambda \mathbb{E}_\pi r_i - \mathrm{KL}(\pi \| \nu_i)\}. \tag{57}$$

Indeed, $\mathrm{KL}(\pi\|\mu_i) = \mathrm{KL}(\pi\|\nu_i) - \lambda \mathbb{E}_\pi r_i + \log Z_i$, and its minimum is zero. Thus the target rewards desirable attributes while controlling departure from the data distribution.

Taking $w_i = e^{\lambda r_i}$ and $c = 0$ in equation 54 yields

$$\begin{aligned} g_1^\star &= -\lambda r_1 - V_\star^{\leftarrow}(0,\cdot), \\ g_0^\star &= -\lambda r_0 - V_\star^{\rightarrow}(0,\cdot), \\ \mathsf{P}^\star &= \arg\min_{\mathsf{P}_0=\mu_0,\, \mathsf{P}_T=\mu_1} \mathrm{KL}(\mathsf{P}\|\mathsf{R}). \end{aligned} \tag{58}$$

The required terminal gradients contain only reward gradients and opposite-direction value gradients. The correction terms account for the coupling between the two endpoints. They separate terminal-reward optimizations generally do not impose both reward-weighted marginals. Endpoint samples, reward evaluations, and the pretrained reference therefore suffice without an image-density or endpoint-energy model.

Connection to neural training. These identities characterize the exact reference and coupled SOC fixed point. Learned reference marginals and reverse dynamics, finite SOC optimization, and inter-level reference fitting are approximations. Exact endpoint identities also require sampling the conditional controls from the prescribed starting marginal (μ_i for image control). The Gaussian augmentation and mixed starting banks described below provide training coverage. They do not by themselves establish exact target-marginal sampling.

E.2 ALGORITHM DESIGN AND STABILIZATION

SPIM couples two neural SOC solvers through their boundary values. The following implementation choices address changing boundary targets, large inner-solver updates, and difficult task initialization. Core task settings are summarized in Appendix E.3.

E.2.1 JACOBI BOUNDARY EXCHANGE

At the start of outer cycle k, we freeze copies $V^{\rightarrow}_{\mathrm{targ},k}$ and $V^{\leftarrow}_{\mathrm{targ},k}$ of the two value networks. The subscript targ denotes a frozen training target, distinct from the anchored values $\bar{V}$ in Section 3.3. Both solves use these snapshots:

$$g_{1,k} = \widetilde{E}_1 - V^{\leftarrow}_{\mathrm{targ},k}(0,\cdot), \qquad g_{0,k} = \widetilde{E}_0 - V^{\rightarrow}_{\mathrm{targ},k}(0,\cdot).$$

The snapshots are replaced only after both solves finish. This gives each inner solve a fixed terminal objective and is intended to limit feedback from partially fitted opposite-direction values. Jacobi exchange is the schedule used for the coupled neural experiments. The idealized iteration in Section 3.2 uses sequential boundary exchange. For image fine-tuning, slowly updated boundary target networks further smooth the exchange. Their interpolation rate is specified in Appendix E.3.6.

E.2.2 TRUST-REGION INNER SOLVERS

TR-AM and TR-PI-VM regularize the directional SOC update around a frozen local policy, following the trust-region formulation of Blessing et al. (2025). Tempering the update toward this anchor is intended to limit abrupt policy changes when the boundary objective changes. TR-AM selects its tempering strength from sampled conditional path costs using an empirical KL threshold ε, then fits weighted adjoint targets. TR-PI-VM fits tempered path-integral value targets with an anchor correction. The anchor remains fixed between scheduled refreshes. Replay is collected for that anchor and training stage. This local anchor is distinct from the Brownian or pretrained reference defining the task. Core trust-region parameters and task budgets are summarized below. The sinusoidal-tracking ablations disable TR to isolate the schedule, particle-bank, and solver choices.

E.2.3 ENDPOINT COVERAGE, PARTICLE BANKS, AND REPLAY

For neural training, SPIM uses endpoint samples with Gaussian smoothing as a practical approximation to well-covered endpoint neighborhoods. Adding Gaussian perturbations in state space broadens local sample coverage and is intended to stabilize value and control fitting. Endpoint particle banks supply the samples used to construct this training coverage. Replay stores trajectory-derived samples for the inner SOC solves. Evaluation uses the prescribed endpoint targets. The dynamic-bank variants replace endpoint particles with transported samples across outer cycles. Their initial normalization/gauge bank remains fixed, so refreshing the sampling bank does not also change the normalization reference. Scalar gauge offsets leave the induced controls unchanged. Parameter-gradient clipping limits large gradients passed to the optimizer, and control-norm caps bound learned controls during synthetic and LiDAR rollouts. Core network and optimization settings are summarized below.

E.2.4 HOMOTOPY FOR DIFFICULT TASKS

Homotopy initializes a difficult task through a sequence of simpler problems. Spherical-meridian transport gradually introduces its geometry. LiDAR gradually introduces the terrain cost from a pretrained zero-cost bridge. The preceding level supplies the initialization or updated reference for the next level. This continuation is intended to ease optimization as the geometry or cost becomes more demanding. For LiDAR, the reference is updated between levels. The experimental settings report the initialization and refinement budgets. The remaining synthetic experiments, including Gaussian dimension scaling and ablations, use direct training. Image reward strengths are separate fixed-strength tasks.

E.3 EXPERIMENTAL SETTINGS

E.3.1 SHARED PROTOCOL

All tasks use $T = 1$ and both transport directions. Synthetic tasks use Brownian references. LiDAR and image fine-tuning use pretrained bridges. SPIM uses endpoint samples for training coverage and prescribed energies for synthetic terminal costs. DSBM, GSBM, and TSBM use endpoint samples. D2E uses its energy-to-energy variant, ASBS uses source samples and target energy gradients, and DeepGSB also uses endpoint log densities.

Except for images, SPIM uses scalar 3×128 networks initialized at zero, Adam with learning rate 10^{-4}, batch 3,200, and time step 0.01. Batch sizes and auxiliary computation differ across methods. Table 8 lists optimizer updates separately from initialization.

Table 8: Optimizer-update budgets, including Gaussian dimension scaling.

Experiment	Method	Main	Auxiliary	Initialization
Zero-cost bridges	SPIM (PI-VM / TR-AM)	180,000	0	0
Zero-cost bridges	DSBM / D2E	90,000	0	0
Zero-cost bridges	ASBS	90,000	90,000	0
Sinusoidal tracking, bimodal splitting, moving double well	SPIM (TR-AM / TR-PIVM)	180,000	0	0
Spherical-meridian transport	SPIM (Homotopy TR-AM)	200,000	0	6,000
Spherical-meridian transport	SPIM (Homotopy TR-PIVM)	240,000	0	6,000
State-cost bridges	GSBM	90,000	600,000	0
State-cost bridges	DeepGSB	90,000	0	0
State-cost bridges	TSBM	90,000	109,500	0
Sinusoidal-tracking ablations	All solver / bank configurations	120,000	0	0
LiDAR	SPIM (Homotopy TR-AM)	180,000	0	400,000
LiDAR	GSBM	400,000	20,000	0
MNIST–EMNIST	AM	12,000	0	290,000
MNIST–EMNIST	SPIM (Ours)	60,000	0	290,000

Synthetic evaluation. We evaluate 2,048 samples per direction using unregularized empirical W_2^2 in the full state space, without dimension normalization. With generated samples $\widehat{\rho}_t^q$ and independent

reference samples $\widehat{\rho}_t^\star$ at physical time t,

$$D_{\mathrm{end}}^{\mathrm{f}} = W_2^2(\widehat{\rho}_1^{\mathrm{f}}, \widehat{\mu}_1), \qquad D_{\mathrm{end}}^{\mathrm{b}} = W_2^2(\widehat{\rho}_0^{\mathrm{b}}, \widehat{\mu}_0), \tag{59}$$

$$D_{\mathrm{int}}^{q} = \frac{1}{19} \sum_{t \in \{0.05, 0.10, \ldots, 0.95\}} W_2^2(\widehat{\rho}_t^{q}, \widehat{\rho}_t^\star). \tag{60}$$

Ring-mixture, double-well, and banana transport report endpoint errors only. Directions are averaged within each seed, followed by cross-seed means and sample standard deviations. TSBM retains its native starting draws.

E.3.2 Zero-cost bridges and Gaussian dimension scaling

All zero-cost tasks use $\sigma = \sqrt{2}$ and $c = 0$. **Gaussian transport**: $\mathcal{N}(-0.1\mathbf{1}_d, I_d)$ to $\mathcal{N}(0.1\mathbf{1}_d, I_d)$, with $d = 5$ in the main comparison and $d \in \{5, 20, 50\}$ in dimension scaling. The 5D run is shared; analytical Gaussian bridges supply the interior references. **Ring-mixture transport**: equal-weight mixtures of 8 and 6 Gaussians on the unit circle, each with covariance $0.01 I_2$. **Double-well transport**: $\mathcal{N}(0, I_{16})$ to a separable double well (center 0, radius 1.5, scale 0.35, tilts $(0.15, -0.1)$ repeated eight times). **Banana transport**: $\mathcal{N}(0, I_2)$ to $(Z_1, 0.5(Z_1^2 - 1) + 0.25 Z_2)$, with independent standard Gaussian Z_i.

Training. SPIM uses $(K, U) = (30, 3000)$ at every dimension, with PI-VM or TR-AM ($\varepsilon = 0.1$). DSBM, D2E, and ASBS use batches 512, 128, and 512, with 20, 20, and 200 integration intervals, respectively.

E.3.3 Bridges with state costs

All state-cost tasks use $\sigma = 1$. **Sinusoidal tracking**: $\mathcal{N}((-2, 0), 0.25 I_2)$ to $\mathcal{N}((2, 0), 0.25 I_2)$, with $c(t, x, y) = 32[y - 2\sin(2\pi t)]^2$ and an analytical Gaussian/Mehler reference. **Bimodal splitting**: $c(x, y) = 2y^2$, Mehler parameters $(0, 4)$, and $\rho_t \propto (K_t f_0)(K_{1-t} g_1)$. Here f_0 is centered at $(-2.25, 0)$ and g_1 is an equal mixture at $(2.25, \pm 2.25)$, all with precision $(16, 16)$; ρ_0, ρ_1 define the endpoints. **Moving double well**: writing $z = x - (-2, 0) - (4, 0)t$,

$$U(z) = z_x^2 + \frac{(z_y^2 - 1.5^2)^2}{4}, \quad \rho_t(x) \propto e^{-2U(z)}, \quad c(t, x) = \tfrac{1}{2}\|\nabla U(z)\|^2 - \tfrac{1}{2}\Delta U(z) + C,$$

where C makes the minimum cost zero. **Spherical-meridian transport**: Gaussian endpoints at $(0, 0, \pm 2.3)$ with covariance $0.01 I_3$, and a radius-2 spherical-meridian reference with uniform longitude and nonzero Euclidean thickness. Its signed cost is constructed from the reference drift through the HJB equation.

Training. The first three tasks use SPIM with $(K, U) = (30, 3000)$; TR-AM and TR-PI-VM use $\varepsilon = 0.1$ and 0.3. Spherical-meridian transport uses a 6,000-update zero-cost initialization, geometry homotopy, and batch 8,192; evaluation uses the final task. GSBM, DeepGSB, and TSBM train directly on the final task with batches 128, 64, and 128 and 100 integration intervals. TSBM uses its native unidirectional variant. Only SPIM uses homotopy in this comparison.

E.3.4 Ablation configurations

Sinusoidal-tracking ablations disable TR and homotopy, use AM/PI-VM batches of 3,200/12,800, and allocate 60,000 updates per direction. Final checkpoints are evaluated. The schedule comparison uses $(K, U) \in \{(15, 4000), (20, 3000), (30, 2000), (40, 1500), (60, 1000)\}$ for AM and the first, third, and fifth pairs for PI-VM.

At $(K, U) = (30, 2000)$, bank variants use fixed endpoint samples, a fixed source/dynamic target, or two dynamic banks. Dynamic banks start from Gaussian particles and insert 1,024 transported endpoints per cycle. One-/two-sided refresh adds 30,720/61,440 rollouts. Thus initialization and sampling cost also differ. The same update budget compares PI-VM/PI-VM, AM/AM, and forward PI-VM/backward AM, retaining each solver's batch size.

E.3.5 LiDAR terrain experiments

Tasks. Helens, Sunset, and Rainier use horizontal domain $[-5,5]^2$ and vertical scales $6, 4, 1.5$, respectively, with $\sigma = 1$. Source/target horizontal centers are $(-2.275, 0.2)/(2.275, -0.9)$, $(2, 0.7)/(-0.1, -2.3)$, and $(-2.7, 1.8)/(2.7, -1)$. Gaussian endpoints with pre-projection covariances $0.02I_3/0.03I_3$ are projected onto local planes fitted to the 20 nearest terrain points. Rainier uses smoothed terrain and Sunset localized-peak terrain. With local projection Π and horizontal boundary penalty B,

$$c(x) = w_{\text{close}}\|x - \Pi(x)\|^2 + w_{\text{height}} e^{\Pi(x)_3} + w_{\text{boundary}} B(x). \tag{61}$$

Closeness/boundary weights are 5,000. The height weight is 0 for Helens/Sunset and 5,000 for Rainier.

Training. SPIM refines a 400,000-update zero-cost reference using TR-AM ($\varepsilon = 0.1$) over eight cost levels from $\lambda = 0$ to 1. Each level uses incremental cost $(\lambda_\ell - \lambda_{\ell-1})c$ and zero relative endpoint-energy terms. Refinement comprises 96,000 SOC updates and 84,000 reference-fitting updates between levels. GSBM uses 400,000 matching and 20,000 auxiliary updates, with batch 64 and 100 integration intervals.

Evaluation. For 256 trajectories per direction and 1,000 integration steps, Figure 3 reports

$$J(t) = \mathbb{E}[\tfrac{1}{2}\|b_{\text{total}}(t, X_t)\|^2 + c(X_t)], \qquad \text{GroundMSE}(t) = \mathbb{E}\|X_t - \Pi(X_t)\|^2,$$

where b_{total} is the Brownian-relative total drift. Here $J(t)$ is an instantaneous cost. Its time integral is the dynamic objective. Profiles use Gaussian smoothing with standard deviation eight time bins.

E.3.6 MNIST–EMNIST reward experiments

Reference and fine-tuning. We use 60,000 MNIST digits and 24,000 case-merged EMNIST A–E images, with matched orientation, 28×28 resolution, and values in $[-1, 1]$. A bidirectional U-Net reference is pretrained for 290,000 optimizer updates using batch 128 and Adam 10^{-4}, $\sigma^2 = 1.02$ and rollouts use 30 steps with a deterministic final step.

Fine-tuning freezes this reference and learns scalar U-Nets whose gradients define relative controls initialized at zero, using Adam 10^{-5} and batch 64. AM omits opposite-direction boundary corrections. SPIM (Ours) includes Jacobi corrections with target interpolation 0.25. Their fine-tuning budgets are 12,000 and 60,000 updates, respectively, so the comparison does not isolate coupling at equal budget. Training/test banks contain 2,048/768 images per endpoint.

Rewards and evaluation. For foreground mask $m_p = \text{sigmoid}(10[(x_p + 1)/2 - 0.5])$, let c_x, v_x, v_y, c_{xy} be its horizontal centroid, variances, and covariance on $[-1, 1]^2$, with $\epsilon = 10^{-6}$. The attributes are

$$h_{\text{area}} = \text{mean}_p\, m_p, \qquad h_{\text{size}} = \sqrt{v_x + v_y + \epsilon},$$
$$h_{\text{aspect}} = \tfrac{1}{2}\log\frac{\max(v_x - c_{xy}^2/(v_y + \epsilon), 0) + \epsilon}{v_y + \epsilon}, \quad h_{\text{slant}} = \frac{c_{xy}}{v_y + \epsilon}, \qquad h_{\text{position}} = c_x.$$

Shape rewards are $r_i = -\frac{1}{2}[(h - h_i^\star)/\delta_i]^2$. Digit/letter targets are training-bank 25th/75th percentiles, and δ_i is half the corresponding interquartile range. Class-set reward is frozen-classifier log probability of $\{0, 2, 4, 6, 8\}$ or $\{A, E\}$ at temperature 1. Targets are

$$\mu_i(dx) = Z_i^{-1} e^{\lambda r_i(x)} \nu_i(dx), \qquad i \in \{0, 1\}, \tag{62}$$

with unweighted data law ν_i. All six tasks use $\lambda = 1$. Area, aspect, and size additionally use $\lambda = 10, 100$, as does the SPIM class-set extension. Each strength is trained separately. Starts follow μ_i at $\lambda = 1$ and $(\mu_i + \nu_i)/2$ otherwise. Evaluation targets remain μ_i. Mixed starts do not imply exact sampling of the target endpoint law, and attribute W_1 measures only the selected attribute distribution. Table 7 averages endpoint metrics from 256 samples per direction: unscaled reward and attribute W_1 against μ_i.

E.4 Additional Endpoint and Interior Metrics

Table 2 supplements the endpoint comparison. Table 9 reports state-cost bridge endpoint errors, supplementing the interior comparison in Table 3. Table 10 reports the fixed-budget allocation ablation on sinusoidal tracking.

Table 9: State-cost bridge endpoint errors ($W_2^2 \downarrow$).

Method	Sinusoidal tracking	Bimodal splitting	Moving double well	Spherical-meridian transport
GSBM	0.01225 ± 0.00450	0.24188 ± 0.18477	0.02210 ± 0.00878	$\underline{0.00332 \pm 0.00059}$
DeepGSB	$\mathbf{0.00950 \pm 0.00059}$	$\mathbf{0.01539 \pm 0.00881}$	$\mathbf{0.01607 \pm 0.00090}$	21.66403 ± 36.33950
TSBM	0.02057 ± 0.00549	0.12198 ± 0.08407	0.03246 ± 0.02657	0.00514 ± 0.00076
SPIM (TR-AM)	$\underline{0.01099 \pm 0.00426}$	0.13856 ± 0.11212	0.02425 ± 0.01164	$\mathbf{0.00131 \pm 0.00004}$
SPIM (TR-PIVM)	0.01441 ± 0.00289	$\underline{0.05209 \pm 0.02282}$	$\underline{0.01757 \pm 0.00051}$	0.03326 ± 0.01673

Table 10: Budget allocation for SPIM (AM) on sinusoidal tracking.

Cycles	Updates/direction/cycle	Endpoint $W_2^2 \downarrow$	Interior $W_2^2 \downarrow$
15	4000	0.01978 ± 0.00784	$\mathbf{0.00534 \pm 0.00042}$
20	3000	0.01645 ± 0.00636	0.00641 ± 0.00142
30	2000	$\mathbf{0.01340 \pm 0.00464}$	0.00634 ± 0.00097
40	1500	$\underline{0.01494 \pm 0.01046}$	0.00602 ± 0.00059
60	1000	0.01701 ± 0.00889	$\underline{0.00575 \pm 0.00134}$

E.5 High-Dimensional Gaussian Interior Results

Table 11 reports Gaussian interior errors in $d \in \{5, 20, 50\}$, complementing the endpoint results in Table 4 in the main text. Both use the evaluation protocol in Appendix E.3.

Table 11: Gaussian interior errors across dimensions ($W_2^2 \downarrow$).

Method	5D	20D	50D
DSBM	0.67948 ± 0.00138	16.05857 ± 0.01197	$\underline{62.77338 \pm 0.14383}$
D2E	$\mathbf{0.67131 \pm 0.00642}$	$\mathbf{15.17721 \pm 0.05359}$	$\mathbf{62.29092 \pm 0.31739}$
ASBS	0.68467 ± 0.00201	15.96491 ± 0.01682	63.75178 ± 0.06124
SPIM (TR-AM)	0.68253 ± 0.00310	15.96296 ± 0.03400	63.63845 ± 0.04530
SPIM (PI-VM)	$\underline{0.67819 \pm 0.00023}$	$\underline{15.91111 \pm 0.00402}$	63.63964 ± 0.04364

E.6 Additional MNIST–EMNIST Samples

Figures 5–8 provide additional selected examples for the reward experiments in Appendix E.3.6. Base denotes the frozen reference; AM and SPIM (Ours) denote independent and coupled fine-tuning. These examples illustrate attribute changes; quantitative comparisons use the metrics and budgets reported above.

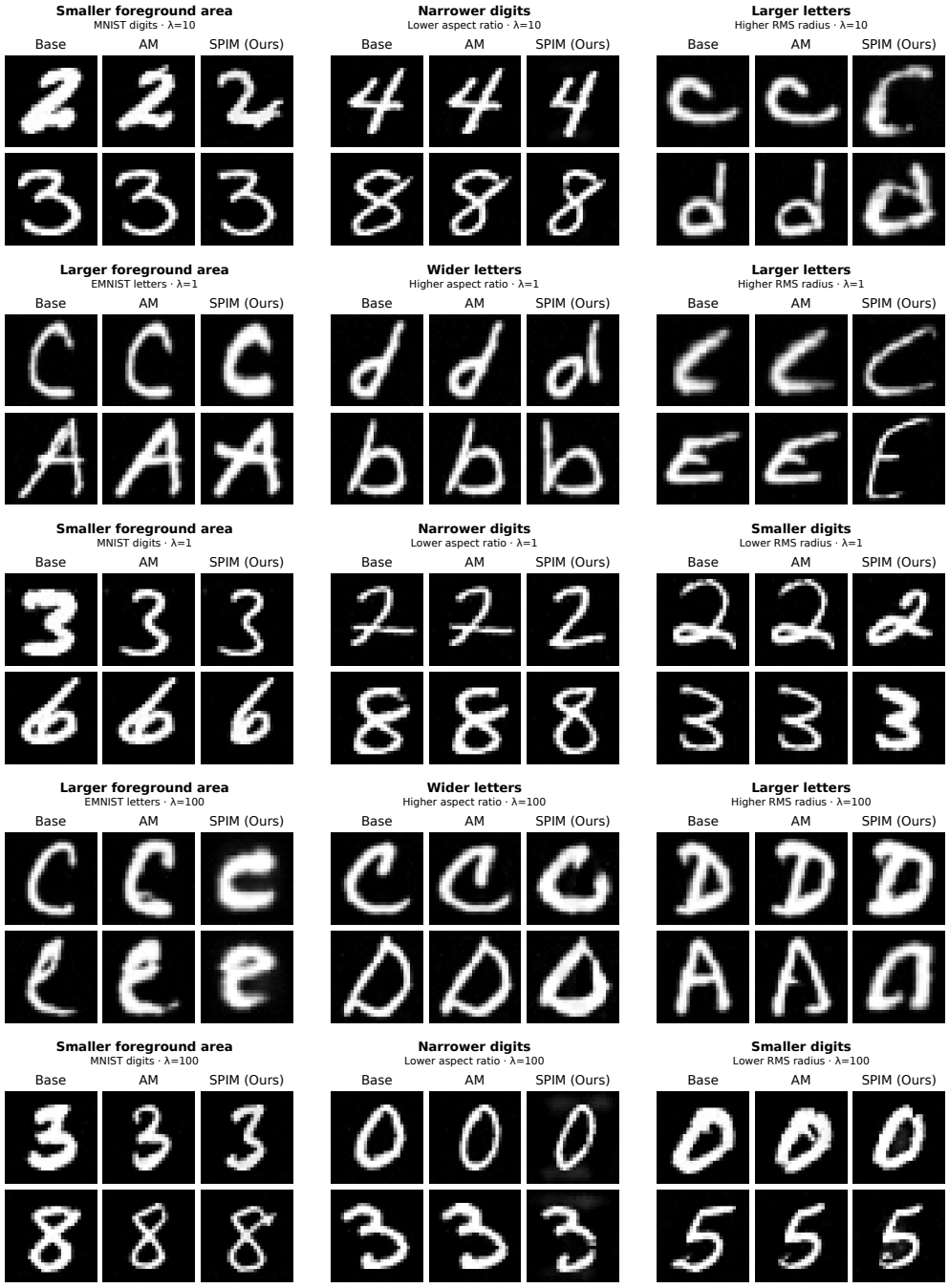

Figure 5: Additional selected outputs for area, aspect-ratio, and size rewards. Each panel identifies the endpoint domain and reward strength and compares Base, AM, and SPIM (Ours).

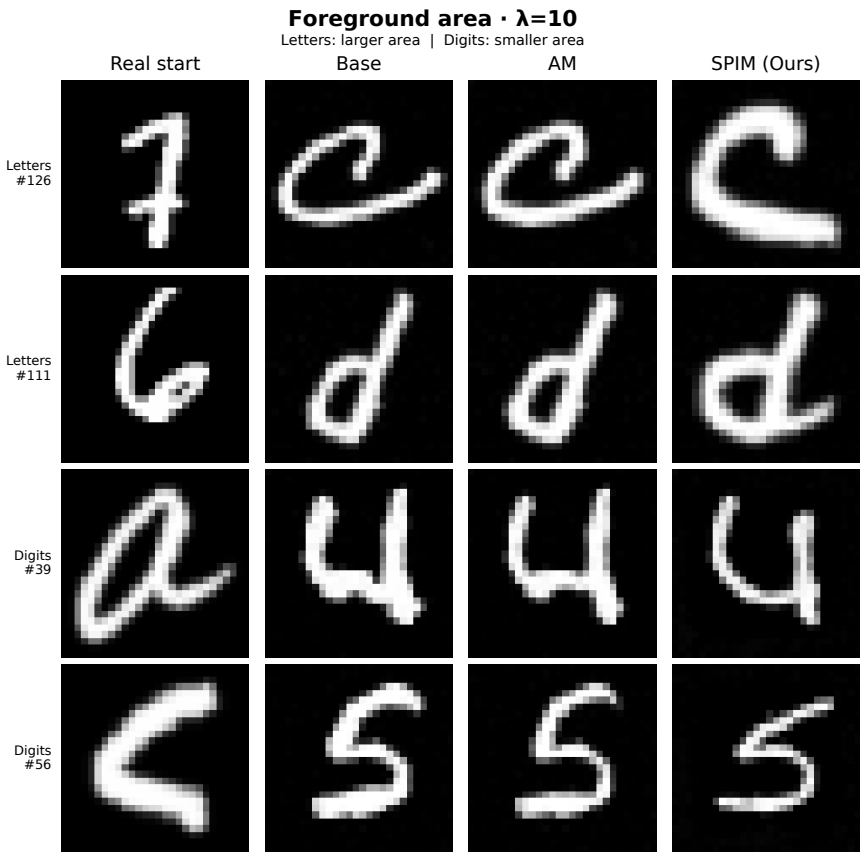

Figure 6: Area reward at $\lambda = 10$. Each row shows a starting image and the outputs of Base, AM, and SPIM (Ours). Both transport directions are included. Row labels identify the starting domain and sample.

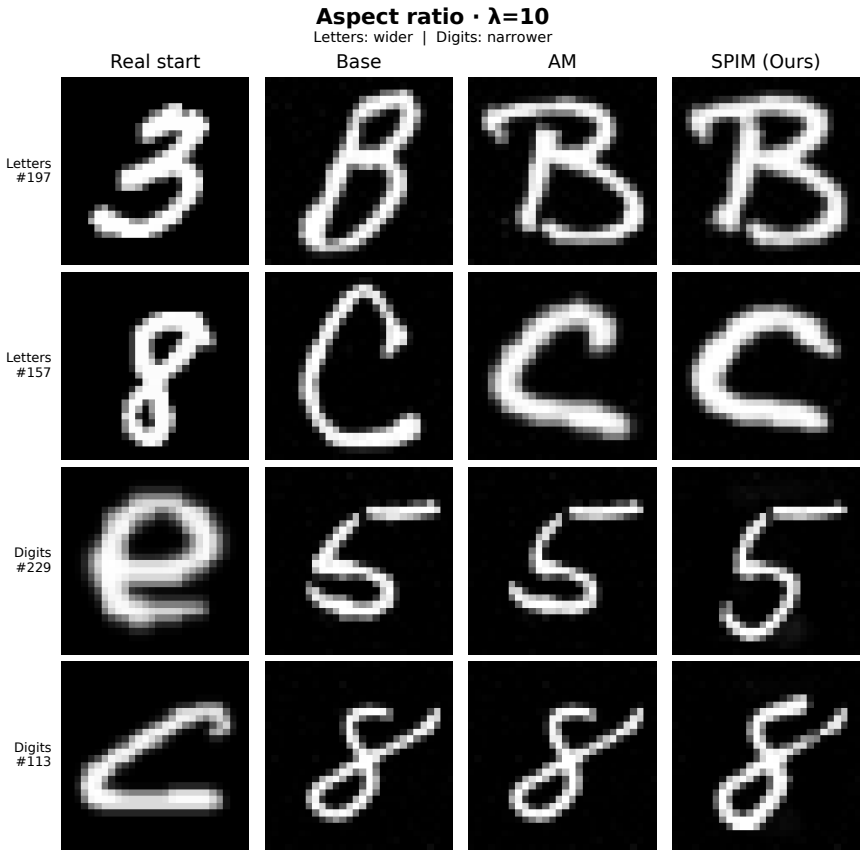

Figure 7: Aspect-ratio reward at $\lambda = 10$, with the same layout as Figure 6.

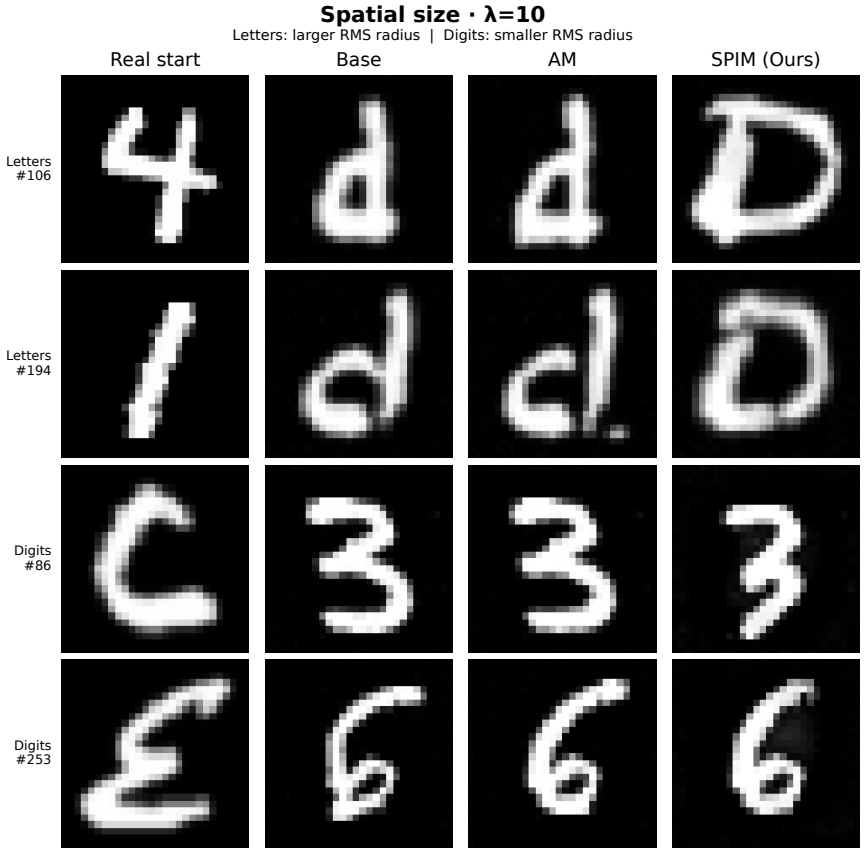

Figure 8: Size reward at $\lambda = 10$, with the same layout as Figure 6.

E.7 Full Transport Visualizations

The supplied transport panels are collected below. Zero-cost panels show endpoint samples projected onto the first two coordinates. State-cost panels overlay samples at physical times $t \in \{0, 0.25, 0.5, 0.75, 1\}$, colored from blue to orange. For spherical-meridian transport, the displayed three-dimensional window is restricted, so the quantitative tables remain necessary to assess mass outside the displayed region. Method headers identify each SPIM instantiation by its SOC solver.

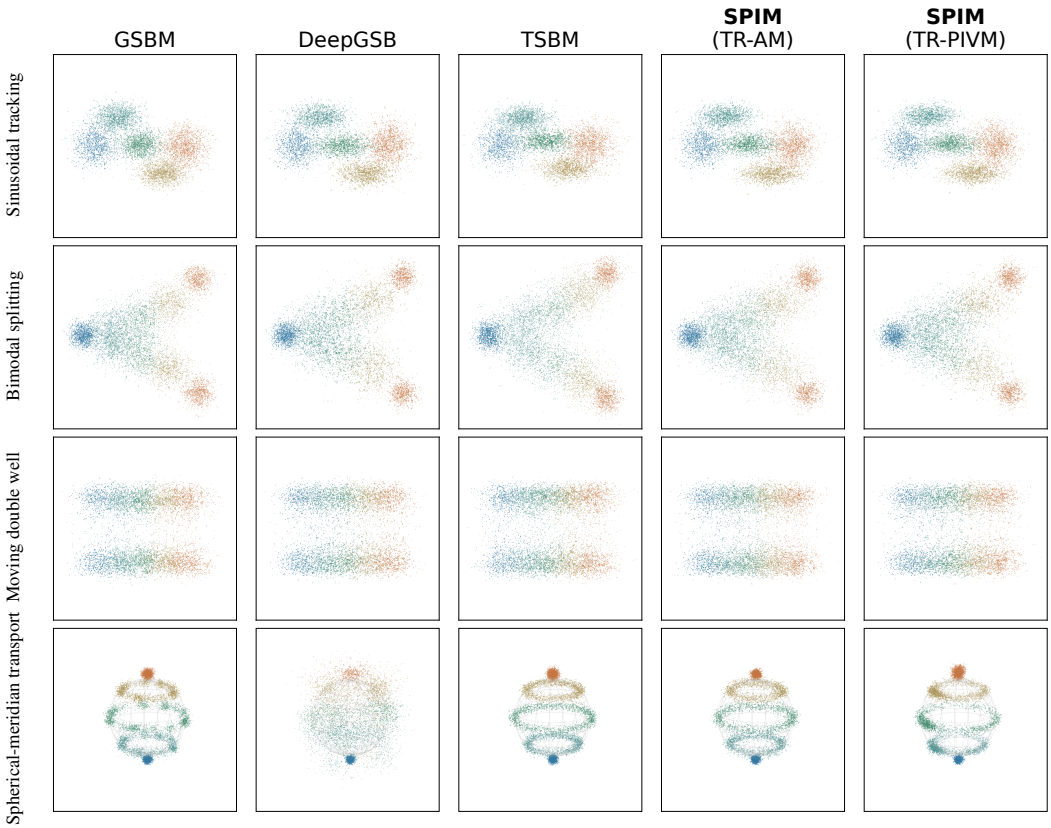

Figure 9: Forward transport with state costs. Rows are tasks and columns are methods.

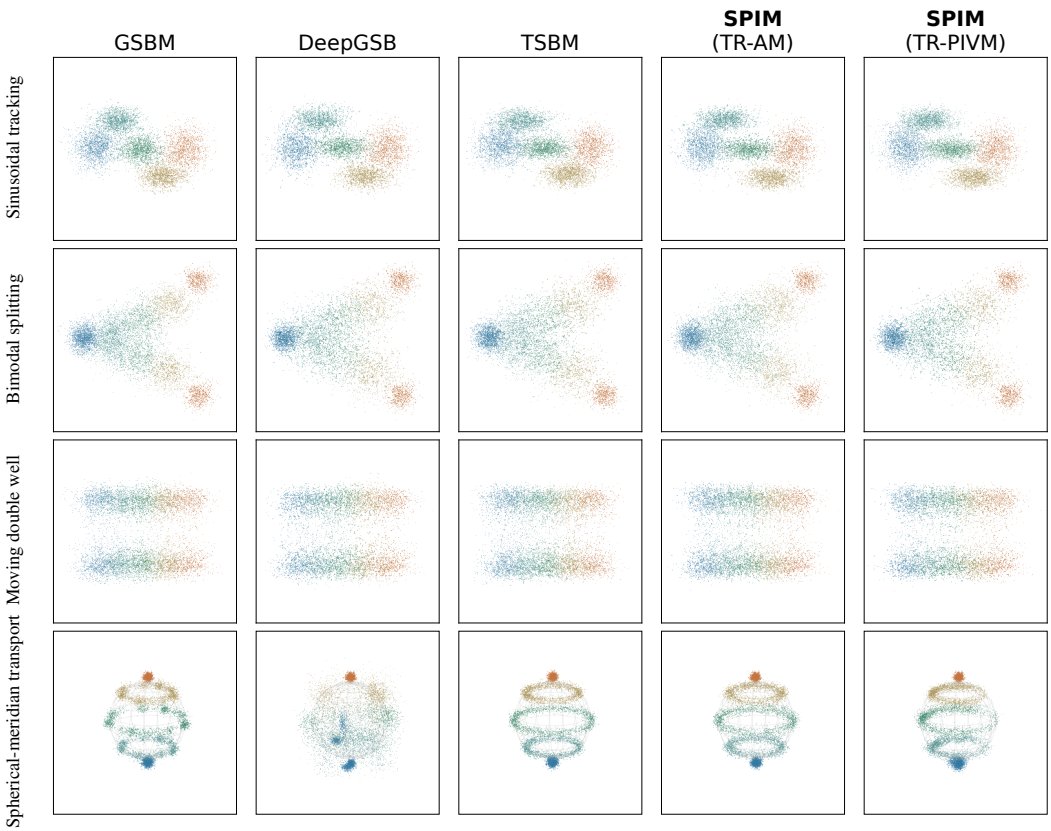

Figure 10: Backward transport with state costs. Rows are tasks and columns are methods.

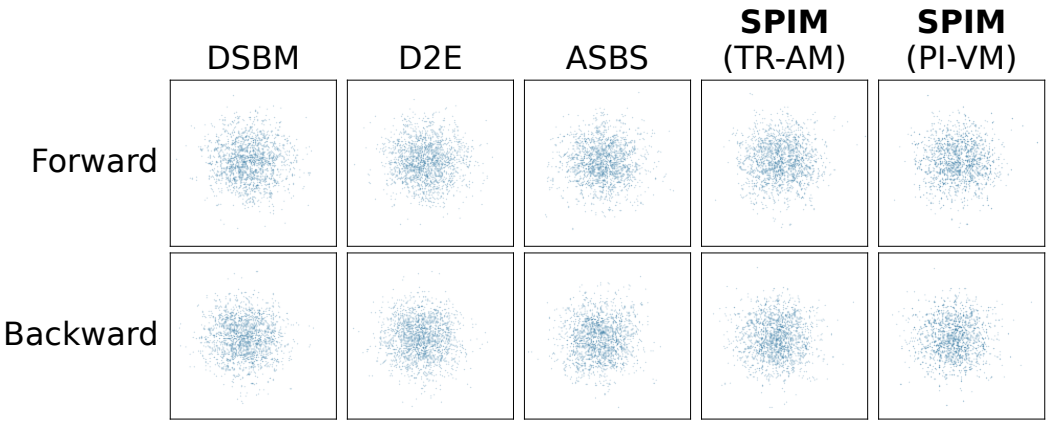

Figure 11: Endpoint samples for gaussian transport (5D): forward above, backward below.

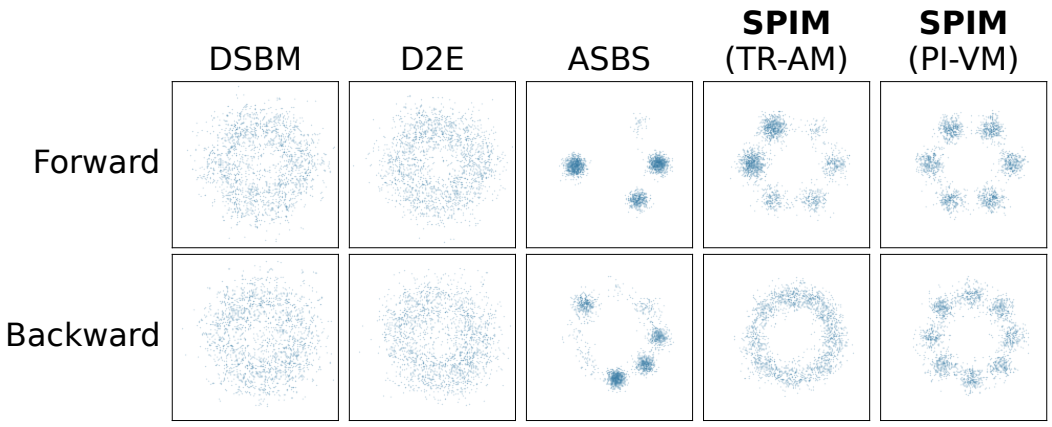

Figure 12: Endpoint samples for ring-mixture transport (2D): forward above, backward below.

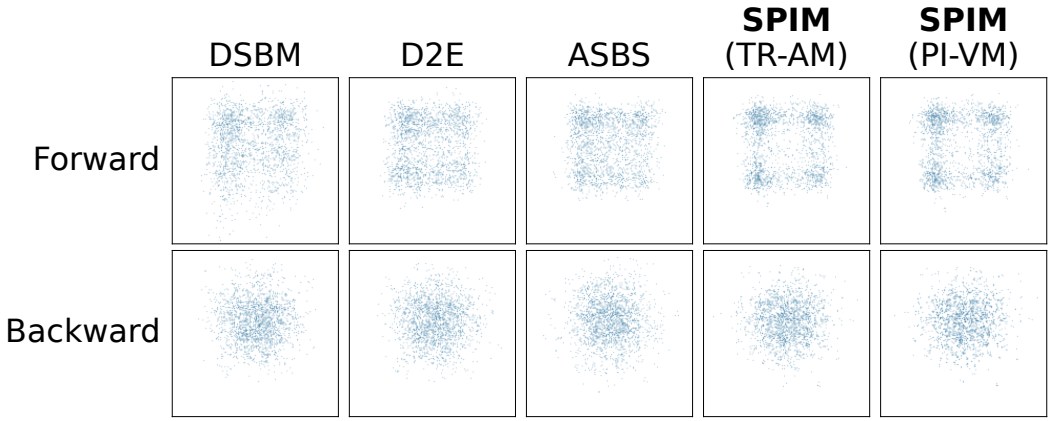

Figure 13: Endpoint samples for double-well transport (16D): forward above, backward below.

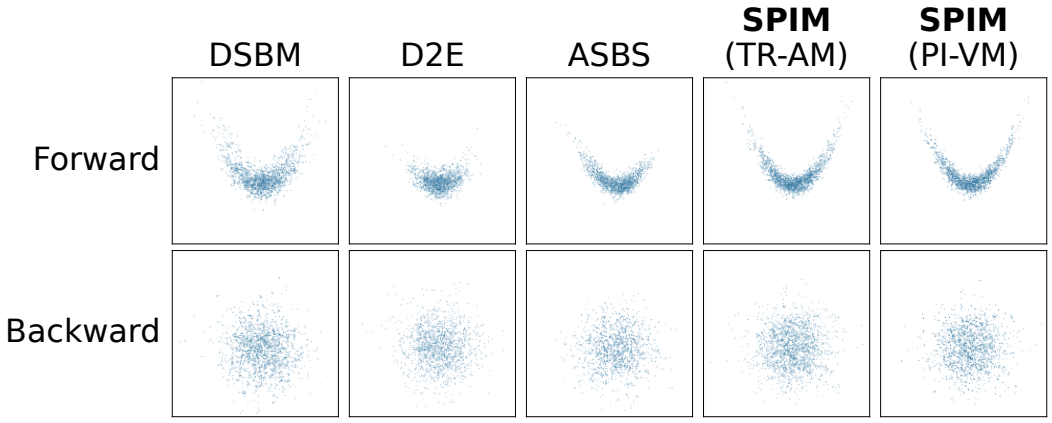

Figure 14: Endpoint samples for banana transport (2D): forward above, backward below.

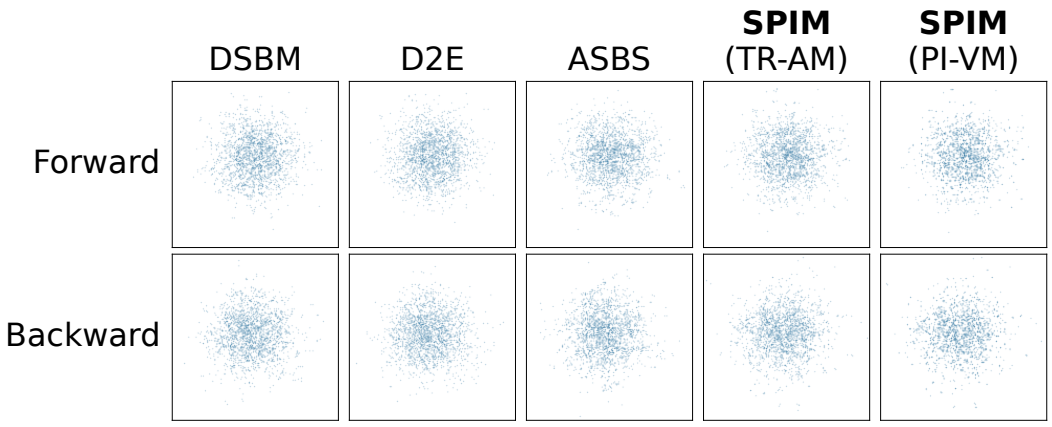

Figure 15: Endpoint samples for gaussian transport (20D): forward above, backward below.

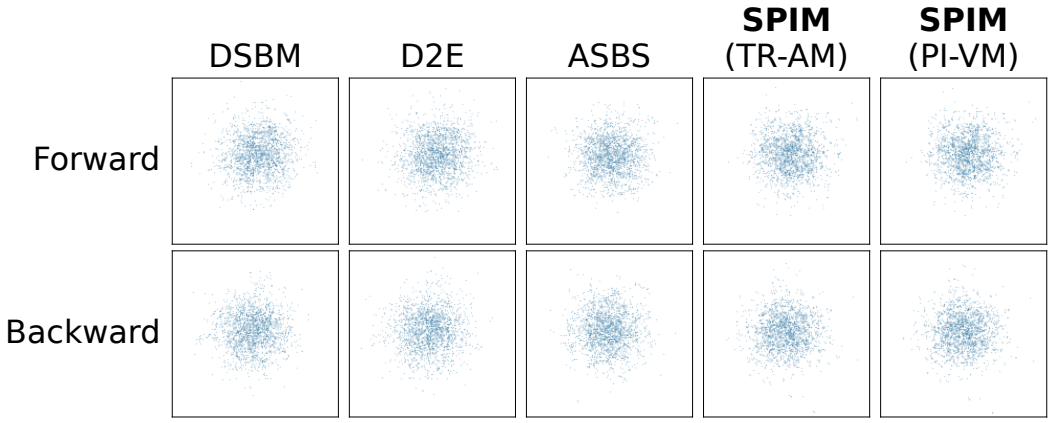

Figure 16: Endpoint samples for gaussian transport (50D): forward above, backward below.

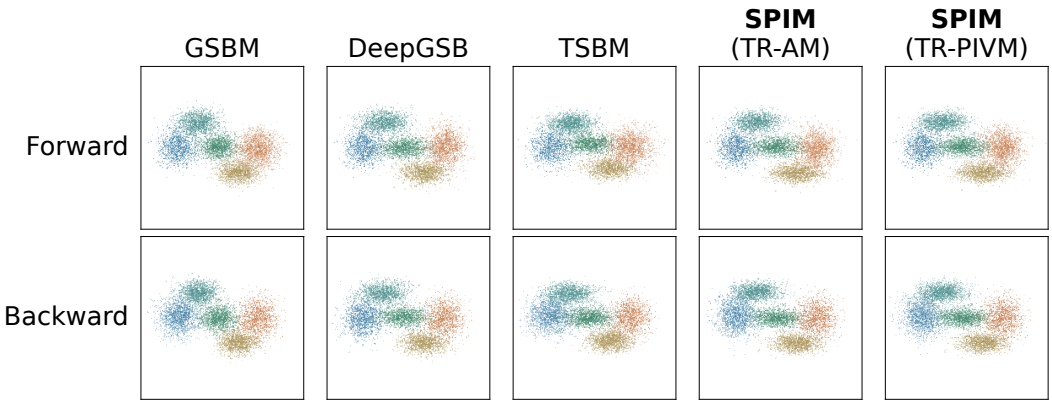

Figure 17: Sinusoidal tracking samples: forward above, backward below.

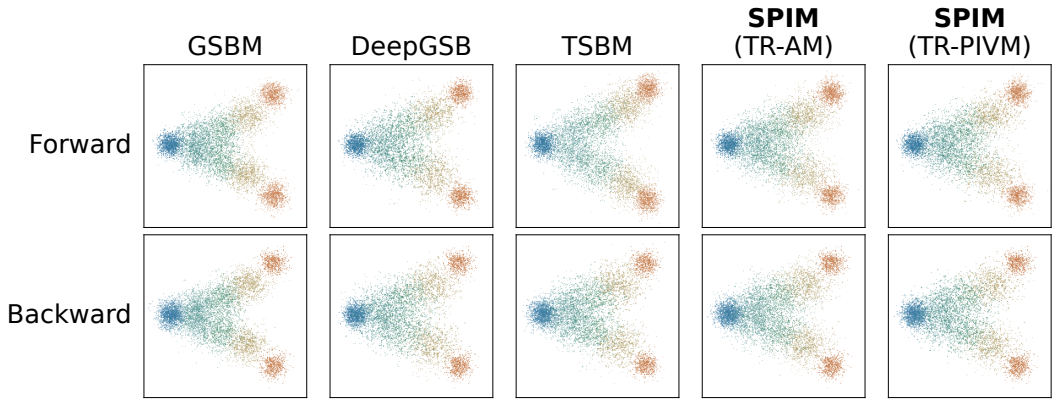

Figure 18: Bimodal splitting samples: forward above, backward below.

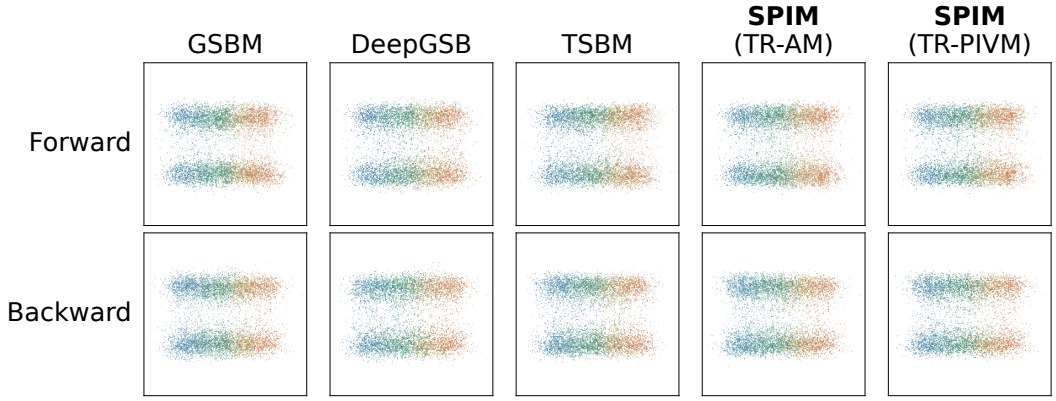

Figure 19: Moving double well samples: forward above, backward below.

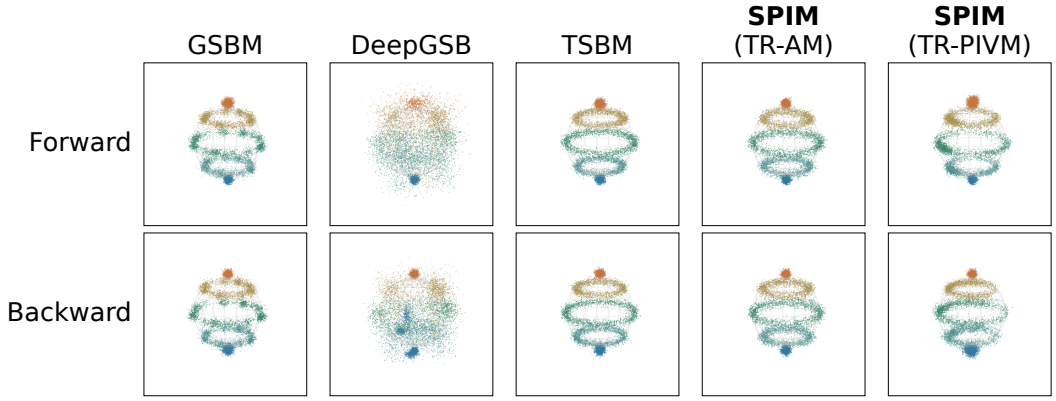

Figure 20: Spherical-meridian transport samples: forward above, backward below.