LEARNING GENERATIVE DYNAMICS WITH SOFT PATH-LAW CONSTRAINTS: A MCKEAN-VLASOV FBSDE APPROACH

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce a unified framework for learning stochastic dynamics from distributional observations and discretely observed trajectories. The key contribution is to combine *soft law constraints*, which regularize fitting to empirical distributions, with *path-dependent supervision*, which captures temporal structure that marginal distributions alone cannot identify. We formulate the problem as a discrete path-dependent McKean–Vlasov control problem, penalizing endpoint, intermediate marginal, and joint sampled-path laws within a single objective. Its optimality system takes the form of a McKean–Vlasov forward-backward stochastic differential equation (FBSDE), in which the marginal terms act continuously in time while the path-law term instead induces conditional adjoint jumps at the observation dates. We derive the reduced FBSDE for quadratic control cost and constant diffusion, establish local well-posedness for a Maximum Mean Discrepancy (MMD) path-law formulation, and develop a sample-based neural solver with causal representations. Experiments on synthetic transport, latent face transport, and human motion validate the framework. In particular, a two-state AMASS experiment under path-law supervision captures both low-to-high and high-to-low transitions, which marginal supervision alone cannot distinguish.

1 INTRODUCTION

Generative modeling is often reduced to terminal sampling: a diffusion model, a deterministic flow, a stochastic interpolant (Song et al., 2020; Ho et al., 2020; Lipman et al., 2023; Albergo and Vanden-Eijnden, 2023) or a Schrödinger bridge (De Bortoli et al., 2021; Shi et al., 2023; Peluchetti, 2023; Gushchin et al., 2024) transports a source law to a target law. This endpoint view is incomplete when the data describe an evolution rather than a single transition. Many stochastic processes share the same initial and terminal laws, and even the same intermediate marginals at every observed time, yet follow very different sample paths. Matching endpoints, or even a whole sequence of marginals, therefore solves the distributional fitting problem without identifying which coupling, and which dynamics, produced the observed transition.

We address this by treating every observed law as a cost inside a single soft law-constrained mean-field control problem. The controlled process starts from a source law and pays a kinetic cost for its drift, while a soft penalty discourages deviations from the target law at the terminal time, from observed marginals at intermediate times, and, when full trajectories are observed, from the joint law of the sampled path at the observation dates. The constraints are soft rather than hard. With few samples, empirical laws are noisy estimates of the true ones, and enforcing them exactly would force the dynamics to fit sampling noise. A soft penalty instead acts as a regularization: the dynamics follow the observed laws only as far as the control cost allows. The selected path is thus the outcome of an explicit tradeoff between control effort and law fidelity, rather than a choice fixed in advance by a reference process. This soft-constraint view is related to recent relaxed Schrödinger bridge formulations (Ma et al., 2025; Jiang et al., 2026), which replace hard terminal constraints with penalty terms or transport relaxations. We extend that viewpoint in two directions at once, to the induced intermediate marginals and to the joint law of the sampled path, and recover the marginal-only formulation exactly when the path-law penalty does not depend on the joint law.

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.

This structure becomes explicit, and interpretable, through the McKean-Vlasov forward-backward stochastic differential equation (FBSDE) that characterizes the optimal control. Endpoint supervision enters as a terminal condition on the adjoint process, intermediate marginal supervision enters as a running drift term in its backward evolution, and joint-path supervision enters through conditional adjoint jumps at the observation dates. The adjoint, the running law forces, and the backward and terminal residuals can therefore all be monitored during training. In the reduced quadratic-control model used throughout, the adjoint directly determines the optimal drift, so marginal and path supervision do not merely reweight a loss, they change the learned dynamics itself. This connects our setting to deep FBSDE solvers (Carmona and Laurière, 2022; Germain et al., 2022; Pham and Warin, 2024), with additional difficulties. The terms induced by the observed laws are not available analytically and must be estimated from samples.

Related work. Trajectory inference and multi-marginal Schrödinger bridges reconstruct dynamics from snapshot marginals (Tong et al., 2020; Huguet et al., 2022; Lavenant et al., 2024; Chen et al., 2023), and generalized Schrödinger bridges add running state costs (Liu et al., 2024); these methods are driven by marginals and therefore cannot select the coupling that joint-path observations determine. Hamdouche et al. (2026) generate time series with a Schrödinger bridge that matches the joint law of the observed path exactly, whereas our path-law term is a soft penalty combined with marginal penalties. Neural SDEs can also be fitted to path laws, adversarially (Kidger et al., 2021) or with signature-kernel scores (Issa et al., 2023), by backpropagating through simulated paths, as in our direct neural SDE baseline; our solver instead regresses on adjoint targets given by the maximum principle.

This paper makes the following contributions: **(i)** A soft law-constrained control problem unifying endpoint, marginal, and joint-path supervision, **(ii)** Its McKean-Vlasov FBSDE optimality system, derived from the stochastic maximum principle, and local well-posedness of this system for the MMD path-law discrepancy (Theorem 1), **(iii)** A sample-based neural residual solver approximating this FBSDE from samples, Algorithm 1, **(iv)** Empirical evidence that path supervision recovers associations marginals cannot: a two-dimensional obstacle experiment and an AMASS motion experiment where only joint-path supervision recovers the correct source-to-route and direction associations.

2 Soft law-constrained generative dynamics

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_t, \mathbb{P})$ a filtered probability space support a d-dimensional Brownian motion W and an independent square-integrable $X_0 \sim \rho_0$, with the augmented filtration they generate. Fix $T > 0$, target marginals $(\rho_t)_{0\le t\le T} \subset \mathcal{P}_2(\mathbb{R}^d)$, and observation times $\Pi = \{0 < s_1 < \cdots < s_K = T\}$. Set

$$\mathbf{X} := (X_{s_1}, \dots, X_{s_K}), \qquad \boldsymbol{\mu}_T := \mathcal{L}(\mathbf{X}), \qquad \mu_t := \mathcal{L}(X_t),$$

When trajectory data are available, let $\boldsymbol{\rho}_T \in \mathcal{P}_2((\mathbb{R}^d)^K)$ denote the target joint sampled-path law, which captures temporal dependence beyond the marginals ρ_{s_k}. The controlled dynamics are

$$\mathrm{d}X_t = \alpha_t \,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \tag{2.1}$$

where $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$. Admissible controls α are A-valued, progressively measurable, and satisfy $\mathbb{E}\big[\int_0^T \|\alpha_t\|^2 \,\mathrm{d}t\big] < \infty$, with A closed and convex. Under the assumptions of Appendix B.2, the SDE is well posed (Hafsi et al., 2026, Proposition 2.4). For penalty weights $\lambda_f, \lambda_g \ge 0$, we minimize

$$\mathcal{J}(\alpha) = \underbrace{\mathbb{E}\left[\int_0^T \frac{1}{2}\|\alpha_t\|^2 \,\mathrm{d}t\right]}_{\text{energy}} + \underbrace{\lambda_f \int_0^T f(\mu_t; \rho_t)\,\mathrm{d}t}_{\text{marginal discrepancy}} + \underbrace{\lambda_g\, g(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T)}_{\text{path-dependent discrepancy}}. \tag{2.2}$$

Here, $f : \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}_+$ and $g : \mathcal{P}_2((\mathbb{R}^d)^K) \times \mathcal{P}_2((\mathbb{R}^d)^K) \to \mathbb{R}_+$ are nonnegative law discrepancies that vanish when their active arguments match the targets. The intuition behind the cost functional (2.2) is as follows : Whenever $\lambda_f, \lambda_g \to \infty$, we expect the resulting $\boldsymbol{\mu}_T$ to converge to $\boldsymbol{\rho}_T$. The targets are fixed, so these terms are deterministic functionals of the induced laws. We characterize solutions using the stochastic maximum principle and the flat derivative $\delta U/\delta m$;

see Appendix B.1. The sampled-path formulation yields adjoint jumps at observation times, in addition to the running marginal-law term (Hafsi et al., 2026). Define the marginal-law gradient $h_t(x, \mu; \rho_t) := \partial_x \frac{\delta f}{\delta m}(\mu; \rho_t)(x)$, and the coordinate path-law gradient

$$r_k(\mathbf{x}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) := \nabla_{x_k} \frac{\delta g}{\delta \boldsymbol{\mu}}(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T)(\mathbf{x}), \qquad k = 1, \dots, K. \tag{2.3}$$

Examples of law-gradients: For Kullback-Leibler divergence (KL), a.k.a. relative entropy, marginal discrepancies, i.e. $f(\mu; \rho_t) = \mathrm{KL}(\mu \,\|\, \rho_t)$,

$$h_t(x, \mu; \rho_t) = \nabla_x \log \frac{\mu(x)}{\rho_t(x)}, \tag{2.4}$$

If $g(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \mathrm{KL}(\boldsymbol{\mu}_T \,\|\, \boldsymbol{\rho}_T)$, we then have

$$r_k(\mathbf{x}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \nabla_{x_k} \log \frac{\boldsymbol{\mu}_T(\mathbf{x})}{\boldsymbol{\rho}_T(\mathbf{x})}. \tag{2.5}$$

This population score identity may be estimated from concatenated paths. Another choice of f and g, which avoids density estimation, is the Maximum Mean Discrepancy (MMD):

$$g(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \tfrac{1}{2} \,\mathrm{MMD}^2_\kappa(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T), \quad r_k(\mathbf{x}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \int \partial_{x_k} \kappa(\mathbf{x}, \mathbf{y}) \, (\boldsymbol{\mu}_T - \boldsymbol{\rho}_T)(\mathrm{d}\mathbf{y}), \tag{2.6}$$

This choice is practical from a theoretical perspective due to the regularity which inherits from the regularity of the kernel map itself, see Sriperumbudur et al. (2010); Gretton et al. (2012).

For $\kappa_{\mathrm{add}}(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K w_k \, \kappa_{\mathrm{marg}}(x_k, y_k) + w_{\mathrm{path}} \, \kappa_{\mathrm{path}}(\mathbf{x}, \mathbf{y})$, linearity of the squared MMD in the kernel gives

$$\mathrm{MMD}^2_{\kappa_{\mathrm{add}}}(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T) = \sum_{k=1}^K w_k \,\mathrm{MMD}^2_{\kappa_{\mathrm{marg}}}(\mu_{s_k}, \rho_{s_k}) + w_{\mathrm{path}} \,\mathrm{MMD}^2_{\kappa_{\mathrm{path}}}(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T). \tag{2.7}$$

With these law gradients, the Pontryagin optimality system is written as a forward-backward stochastic differential equation of McKean-Vlasov type:

$$\begin{cases} \mathrm{d}X_t = -Y_t \,\mathrm{d}t + \sigma(t, X_t) \,\mathrm{d}W_t, \qquad X_0 \sim \rho_0, \\ \mathrm{d}Y_t = -\Big(\partial_x \big[\sigma(t, X_t) : Z_t\big] + \lambda_f h_t(X_t, \mu_t; \rho_t)\Big) \,\mathrm{d}t + Z_t \,\mathrm{d}W_t, \quad t \notin \{s_1, \dots, s_{K-1}\}, \\ Y_{s_k} - Y_{s_k^-} = -\lambda_g \mathbb{E}\left[r_k(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) \mid \mathcal{F}_{s_k}\right], \qquad k < K, \\ Y_T = \lambda_g r_K(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T). \end{cases} \tag{2.8}$$

Here $M : N = \mathrm{Tr}(M^\top N)$, and ∂_x differentiates σ with respect to the state argument. The integral form and tower property yield

$$Y_t = \mathbb{E}\bigg[Y_T + \int_t^T \Big(\partial_x \big[\sigma(u, X_u) : Z_u\big] + \lambda_f h_u(X_u, \mu_u; \rho_u)\Big) \,\mathrm{d}u + \lambda_g \sum_{\substack{1 \le k < K \\ t < s_k}} r_k(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) \;\Big|\; \mathcal{F}_t\bigg]. \tag{2.9}$$

Thus, future path sensitivities enter through conditional expectations in the drift of the forward generative dynamics. The following result, which extends Hafsi et al. (2026, Theorem 5.15), establishes well-posedness of the FBSDE for the MMD specification over a short time horizon.

Theorem 1 (Local existence and uniqueness for the MMD specification)**.** *Assume $g(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \frac{1}{2} \,\mathrm{MMD}^2_\kappa(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T)$. Fix the finite number K of observation dates and $\boldsymbol{\rho}_T \in \mathcal{P}_2((\mathbb{R}^d)^K)$. Assume κ is symmetric, positive definite, and C^2, with uniformly bounded $\partial_{\mathbf{x}}\kappa$, $\partial^2_{\mathbf{xx}}\kappa$, and $\partial^2_{\mathbf{ux}}\kappa$. Assume also that $(x, m) \mapsto \partial_x \frac{\delta f}{\delta m}(m; \rho_t)(x)$ is uniformly bounded and Lipschitz in (x, m), uniformly in t. Assume σ is measurable, uniformly Lipschitz in x, has uniform linear growth, and has a bounded, uniformly Lipschitz state derivative. For each column σ^j, assume*

$$\partial_x \sigma^j(t, x) = \Sigma_j(t, x) I_d, \qquad j = 1, \dots, d,$$

for measurable Σ_j. Then there exists $T_0 > 0$, depending on the stated regularity bounds, λ_f, λ_g, and K, such that for $0 < T \le T_0$, (2.8) *with $X_0 = 0$ has a unique solution*

$$(X, Y, Z) \in \mathbb{S}^4([0,T];\mathbb{R}^d) \times \mathbb{S}^\infty([0,T];\mathbb{R}^d) \times \mathbb{H}^2([0,T];\mathbb{R}^{d\times d}), \qquad Z \in \textbf{\textit{BMO}}.$$

Moreover,

$$\sup_{0\le t\le T} \|Y_t\| \le C_T\big(K\lambda_g B_r + \lambda_f T B_f\big) \quad \mathbb{P}\text{-a.s.},$$

where $B_r = \max_k \|r_k\|_\infty$ and $B_f = \sup_{t,x,m} \left\|\partial_x \frac{\delta f}{\delta m}(m;\rho_t)(x)\right\|$, and where the spaces $\mathbb{S}^4$, $\mathbb{S}^\infty$ and $\mathbb{H}^2$ are introduced in Appendix A, and C_T depends only on the diffusion regularity bounds. For constant diffusion, one may take $C_T = 1$, and the same conclusion holds with $X_0 = 0$ replaced by any $X_0 \in L^4(\Omega, \mathcal{F}, \mathbb{P})$ independent of W.

We shall notice that the solution of the forward-backward system given by Theorem 1 is related to the optimal control problem (2.2). We refer to Hafsi et al. (2026) for further details.

3 Deep MKV Algorithm

We now specify the setting used by the solver. The drift is the control itself and the running cost is the quadratic energy of (2.2), and we take

$$\sigma(t,x) \equiv \sigma I_d, \qquad f(\mu;\rho_t) = D_f(\mu;\rho_t), \qquad g(\boldsymbol{\mu}_T;\boldsymbol{\rho}_T) = D_g(\boldsymbol{\mu}_T;\boldsymbol{\rho}_T),$$

where $\sigma > 0$ is a constant and D_f, D_g are MMDs with a Gaussian kernel, evaluated on a feature map of the state and of the sampled path, respectively. Only this feature map changes across experiments (Appendix C.3).

When $\lambda_f = 0$ and g penalizes only the terminal law, letting $\lambda_g \to \infty$ recovers the Schrödinger bridge between ρ_0 and ρ_T, in the sense that the optimal control converges to the Schrödinger bridge control (Ma et al., 2025, Theorem 6.4).

The solver works on the observation grid $0 = t_0 < t_1 < \cdots < t_N = T$, with $t_k = s_k$ for $k \ge 1$ and $N = K$, so that $\mathbf{X} = (X_{t_1}, \dots, X_{t_N})$. For $i \in [\![0, N]\!]$ we write $X_i := X_{t_i}$, $Y_i := Y_{t_i}$, $\mu_i := \mu_{t_i}$, $\rho_i := \rho_{t_i}$, $\mathcal{F}_i := \mathcal{F}_{t_i}$, and for $i \in [\![1, N]\!]$ $\Delta t_i := t_i - t_{i-1}$, $\Delta W_i := W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, \Delta t_i I_d)$. Approximating the time integral in (2.9) by the right-endpoint rule on this grid, and using $\partial_x[\sigma : Z_t] = 0$ since the diffusion is constant, the optimality system (2.8) becomes

$$\begin{cases} X_i = X_{i-1} - Y_{i-1}\,\Delta t_i + \sigma\,\Delta W_i, & i \in [\![1,N]\!], \quad X_0 \sim \rho_0, \\ Y_i = \mathbb{E}\Big[\displaystyle\sum_{k=i+1}^{N} S_k \,\Big|\, \mathcal{F}_i\Big], & i \in [\![0, N-1]\!], \\ S_i = \lambda_f\, h_{t_i}\big(X_i, \mu_i; \rho_i\big)\,\Delta t_i + \lambda_g\, r_i\big(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T\big), & i \in [\![1,N]\!]. \end{cases} \tag{3.1}$$

Starting from the forward–backward optimality system (3.1), we aim to simulate the optimal forward process $X^\star$. This requires the optimal control $\alpha = -Y$, hence the adjoint component Y, which we approximate on this grid by a single neural network $Y_\theta(t_i, \mathcal{H}_i)$ shared across dates.

Here $\mathcal{H}_i$ is the information available at t_i, namely $\mathcal{H}_i = X_i$ for terminal and marginal supervision and the generated prefix $\mathcal{H}_i = (X_0, \dots, X_i)$ for path-law supervision (Appendix C.1), so that the network never sees future states or target trajectories. Let $(\xi_j^b)_{b=1}^B$ be the observable samples of ρ_j, and set $\boldsymbol{\xi}^b = (\xi_1^b, \dots, \xi_N^b)$. Starting from the parameters $\theta^{(0)}$ with a zero final layer and from initial conditions $X_0^{(\ell),b} \sim \rho_0$ for every particle $b \in [\![1, B]\!]$, we iterate the following steps for $\ell = 0, \dots, L-1$, with L the total number of iterations.

1. **Forward simulation.** For every particle $b \in \{1, \dots, B\}$, draw $X_0^{(\ell),b} \sim \rho_0$ and $\Delta W_i^{(\ell),b} \sim \mathcal{N}(0, \Delta t_i I_d)$, and compute

$$X_{i+1}^{(\ell),b} = X_i^{(\ell),b} - Y_{\theta^{(\ell)}}\big(t_i, \mathcal{H}_i^{(\ell),b}\big)\Delta t_{i+1} + \sigma \Delta W_{i+1}^{(\ell),b}, \qquad i = 0, \dots, N-1. \tag{3.2}$$

2. **Backward targets.** We can define the the target empirical laws

$$\widehat{\mu}_j^{(\ell)} = \frac{1}{B}\sum_{b=1}^{B} \delta_{X_j^{(\ell),b}}, \quad \widehat{\boldsymbol{\mu}}^{(\ell)} = \frac{1}{B}\sum_{b=1}^{B} \delta_{\mathbf{X}^{(\ell),b}}, \quad \widehat{\rho}_j = \frac{1}{B}\sum_{b=1}^{B} \delta_{\xi_j^b}, \quad \widehat{\boldsymbol{\rho}} = \frac{1}{B}\sum_{b=1}^{B} \delta_{\boldsymbol{\xi}^b}.$$

Evaluating the law-gradient fields (3.1) at the particles, with all laws replaced by their empirical counterparts, defines the source of particle b at date t_j, for $j = 1, \dots, N$,

$$S_j^{(\ell),b} = \lambda_f \, \Delta t_j \, h_{t_j}\big(X_j^{(\ell),b}, \widehat{\mu}_j^{(\ell)}; \widehat{\rho}_j\big) + \lambda_g \, r_j\big(\mathbf{X}^{(\ell),b}, \widehat{\boldsymbol{\mu}}^{(\ell)}; \widehat{\boldsymbol{\rho}}\big). \tag{3.3}$$

Then we recover the realized target

$$\widetilde{Y}_i^{(\ell),b} = \sum_{j=i+1}^{N} S_j^{(\ell),b}, \qquad i = 0, \dots, N-1. \tag{3.4}$$

3. **Ridge proxy.** The target $\widetilde{Y}_i$ depends on future noise, whereas the feedback quantity required by the maximum principle is $\overline{Y}_i(\mathcal{H}_i) = \mathbb{E}\big[\widetilde{Y}_i \mid \mathcal{H}_i\big]$ which can be approximated with ridge on the features of the past $u_i^{(\ell),b} = \phi_i(\mathcal{H}_i^{(\ell),b})$ where ϕ_i is a fixed (non-learned) map encoding the path history,

$$\beta_i^{(\ell)} = \arg\min_{\beta} \sum_{b=1}^{B} \big\|\widetilde{Y}_i^{(\ell),b} - \beta^\top u_i^{(\ell),b}\big\|^2 + \gamma\|\beta\|^2, \qquad \widehat{\overline{Y}}_i^{(\ell),b} = \big(\beta_i^{(\ell)}\big)^\top u_i^{(\ell),b}. \tag{3.5}$$

4. **Neural projection.** For $i \in [\![0, N-1]\!]$, we take one optimizer step $\theta^{(\ell+1)} = \mathrm{Adam}(\theta^{(\ell)}, \nabla_\theta \mathcal{L}_Y)$ on

$$\mathcal{L}_Y(\theta) = \frac{1}{NB}\sum_{i=0}^{N-1}\sum_{b=1}^{B} \left\| \frac{Y_\theta\big(t_i, \mathcal{H}_i^{(\ell),b}\big) - \widehat{\overline{Y}}_i^{(\ell),b}}{s_i} \right\|^2 \text{ with } s_i = \left(\frac{1}{CB}\sum_{c=1}^{C}\sum_{b=1}^{B} \big\|\widehat{\overline{Y}}_i^{(c),b}\big\|^2 \right)^{1/2} \tag{3.6}$$

The direct neural SDE baseline instead learns the drift by minimizing J directly. Our solver computes the adjoint targets (3.4) explicitly, estimates their conditional expectation by ridge regression (3.5), and fits the network to them.

Putting these steps together yields Algorithm 1.

4 Experiments

We evaluate three consequences of the soft law-constrained formulation. A controlled two-dimensional obstacle experiment, in which each source cluster must follow a prescribed route, isolates what path supervision adds to marginal supervision. A bidirectional motion experiment on real AMASS data tests the same mechanism, where the model captures the direction of the motion. Finally, terminal-only transport on the high-dimensional AFHQ latent space tests scalability. All experiments use the neural residual solver of Section 3, trained by adjoint regression. The AMASS experiment additionally compares with a path-dependent neural SDE trained by backpropagating the same objective through simulated trajectories, without our adjoint-regression solver. We therefore assess whether each model recovers the correct pairing between initial states and trajectories, not only the per-time distributions.

All law discrepancies are Gaussian-kernel MMDs with several bandwidths, rescaled such that all terms of the objective are on a common scale. Appendix D gives full hyperparameters, dataset construction, and evaluation protocols for every experiment below, and Table 4 reports their compute cost.

4.1 Three-route obstacle transport

We construct a two-dimensional diffusion over $T = 2$ with $N = 20$ time steps and constant diffusion $\sigma = 0.05$. As shown in Figure 1, the source is an equal-weight mixture of three Gaussian components

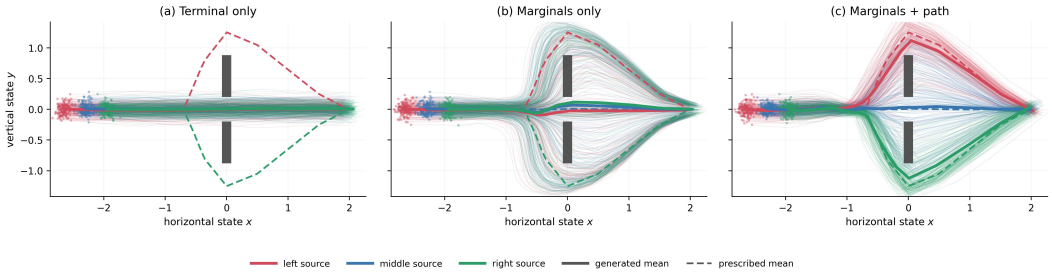


Figure 1: Three-route obstacle transport under terminal-only, marginal-only, and marginal-plus-path supervision. Thin: generated paths; bold: mean path per source cluster; dashed: prescribed routes. Only path supervision assigns each cluster to its route.

Table 1: Three-route obstacle experiment. Discrepancies are calibrated MMDs, normalized so that 0 is the null level of two independent target samples and 1 is the level of the uncontrolled diffusion (Appendix D.1). Lower is better except for route accuracy. Energy is the empirical control cost.

method	terminal MMD ↓	marginal MMD ↓	path MMD ↓	route accuracy ↑	energy ↓
terminal only	**0.000**	0.537	4.675	0.328	**4.65**
marginal only	0.003	**0.023**	1.757	0.216	9.08
marginal + path	**0.000**	0.028	**0.015**	**0.601**	9.25

placed along the horizontal axis, which first merge into a common cloud and must then cross a vertical wall through one of three routes, above, through a central opening, or below, before reaching a shared terminal region. The intermediate target marginals only specify the fraction of particles crossing each route, not which source cluster uses which route. This pairing (left to upper, middle to center, right to lower) is visible only when the initial and crossing positions of the same trajectory are observed jointly, which is what the path law provides. Note that the source cluster of each particle is never given to the model.

We train the same solver under terminal-only, marginal-only, and marginal-plus-path supervision. Route accuracy is the fraction of generated particles that cross through the route assigned to their source cluster. Energy is the average control cost $\frac{1}{B}\sum_{b=1}^{B}\sum_{i=0}^{N-1}\|\alpha_i^b\|^2\Delta t_i$. Figure 1 shows that terminal-only supervision sends most particles through the shortest, central route, marginal-only supervision matches the prescribed fraction on each route but mixes the source clusters across routes, and only marginal-plus-path supervision recovers the prescribed routing.

Table 1 quantifies this. With three routes, an association carrying no information scores $1/3$: terminal-only supervision is at chance (0.328), and marginal-only supervision falls below it (0.216), i.e., the pairing it selects is worse than random. Adding the path term raises route accuracy to 0.601 and lowers the path MMD by two orders of magnitude, at the cost of a slightly higher marginal discrepancy and a higher control energy, since the solver must now steer particles onto specific routes rather than merely onto the correct terminal region. Accuracy stays well below one because once the source clusters merge, particles from different clusters are close, and separating them onto different routes requires more energy while the noise keeps mixing them.

This experiment illustrates the central identifiability claim of the paper. The marginal-only solution is not incorrect, as it matches all the marginal information it receives. It only misses the source–route pairing, which cannot be recovered from the distributions at each time, however many times are observed.

4.2 Bidirectional human motion on AMASS

The three-route experiment is synthetic and its correct routing is known. We now test the same mechanism on real data: whether a single model can generate human motions in the correct direction. Given the first frames of a clip, it must continue either upward or downward. Since both directions produce similar poses, per-frame distributions do not determine the direction, whereas path supervision does.

For that purpose, we use human motion-capture data from AMASS (Mahmood et al., 2019; Loper et al., 2015; Romero et al., 2017) and select two opposite types of motion: *low-to-high* (standing up from a sitting or lying posture) and *high-to-low* (the reverse). Each clip lasts two seconds (61 frames), each frame being a 70-dimensional body pose. The training set contains 522 clips, equally split between the two directions, validation and test use 82 and 66 clips. We compare three models: the marginal-only model, a neural SDE trained by backpropagating the same objective through the simulated trajectories, and the path-dependent model, which adds a path-law term on features describing vertical motion, timing, and limb and foot coordination (Appendix D.4).

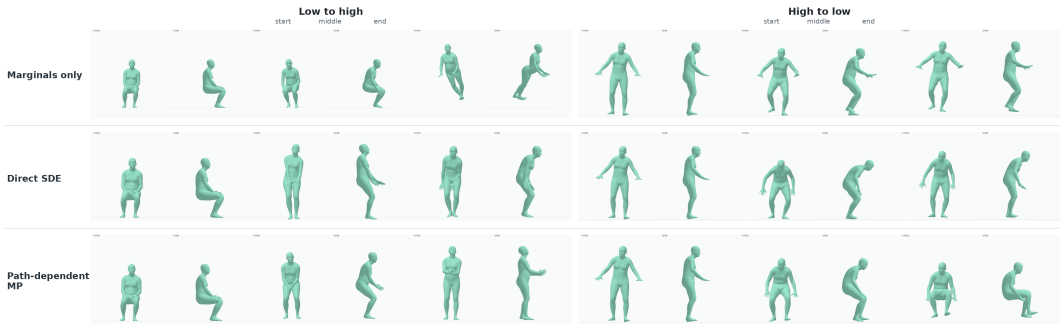


Figure 2: Bidirectional AMASS generation, low-to-high (left) and high-to-low (right). Each strip shows frames from one generated rollout. The marginal-only model does not reliably follow the direction of its initial state, the direct neural SDE weakens one direction, and the path-dependent model generates both.

Figure 2 illustrates these differences qualitatively. The marginal-only model produces realistic individual poses, but they do not form a consistent motion, as the body barely rises or lowers and does not reliably follow the direction of its initial state. The direct neural SDE produces clearer transitions, but one direction remains incomplete. Only the path-dependent model follows the correct direction in both cases, with a smooth and natural movement from the initial to the final posture. Note that a single model learns both directions from unlabeled clips and infers the direction from the initial frames. With marginal supervision only, this would require splitting the data by direction and training one model per direction.

Writing $p_{\uparrow}^{L\to H}$ and $p_{\uparrow}^{H\to L}$ for the fraction of generated rollouts classified as upward motion when starting from a low-to-high or a high-to-low initial condition, we report the direction error

$$E_{\mathrm{dir}} = \tfrac{1}{2}\left(\left|p_{\uparrow}^{L\to H} - 1\right| + \left|p_{\uparrow}^{H\to L}\right|\right),$$

which is zero when every generated motion follows the direction of its initial condition. We also report the motion amplitude $|\Delta h|$, i.e. the change in pelvis height over the generated frames in meters, the normalized MMD between generated and real pelvis heights, and the foot-height error, which measures foot sliding or penetration into the floor. All metrics are computed on held-out clips.

Table 2: Bidirectional AMASS evaluation. Lower is better except for amplitude, whose real-motion reference is shown in the first row.

method	$E_{\mathrm{dir}} \downarrow$	$\lvert\Delta h\rvert$ (m)	pelvis MMD $\downarrow$	foot error (m) $\downarrow$
real motions	0.000	0.440	0.000	0.032
marginal only	0.360	0.138	0.223	0.347
direct neural SDE	0.092	0.384	**0.021**	**0.065**
path-dependent	**0.036**	**0.404**	0.128	0.102

Table 2 shows that the marginal-only model produces motions that are too small ($|\Delta h| = 0.14$ meters versus 0.44 meters) and often in the wrong direction, since per-frame constraints do not link the initial motion to its continuation. The direct neural SDE recovers most of the amplitude and has the lowest foot-height error, but the path-dependent model reduces the direction error by a factor of about three (0.036 versus 0.092). The price is a looser fit of the pelvis-height distribution, the same tradeoff as in Table 1: the path term improves the temporal association at the expense of the marginal fit. Foot errors remain above the real-motion level for all models, as no physical contact model is used.

4.3 HIGH-DIMENSIONAL TERMINAL TRANSPORT ON AFHQ

The two previous experiments rely on the path term in low or moderate dimension. We finally check that the solver scales to high dimension, in the simpler terminal-only setting ($\lambda_f = 0$) on AFHQ (Choi et al., 2020) without any intermediate observations. The controlled process acts in the 512-dimensional latent space of a variational autoencoder, pre-trained for reconstruction only on cat, dog, and wild-animal images and then frozen. The autoencoder is trained on different animal species, while transport uses only two of them (from cats to dogs). The latent space therefore encodes general animal features rather than those that separate cats from dogs, which makes the transport task less trivial and a fairer test of our model.The source and target images are unpaired.

Matching latent distributions does not guarantee that decoded images look like dogs. The terminal loss therefore combines two normalized MMDs, one on latent codes and one on features of the decoded images from DINOv2 (Oquab et al., 2023), a pre-trained vision network whose features encode object shape and parts rather than color or texture: $g(\mu_T; \rho_T) = \frac{1}{2}\overline{D}_{\text{latent}}(\mu_T, \rho_T) + \frac{1}{2}\overline{D}_{\text{DINO}}(\mu_T, \rho_T)$.

Figure 3 shows a gradual transformation from cat to dog. Features unrelated to the class, such as fur color, pose, and background, are preserved, so each dog remains close to its source cat, although no paired data is used.

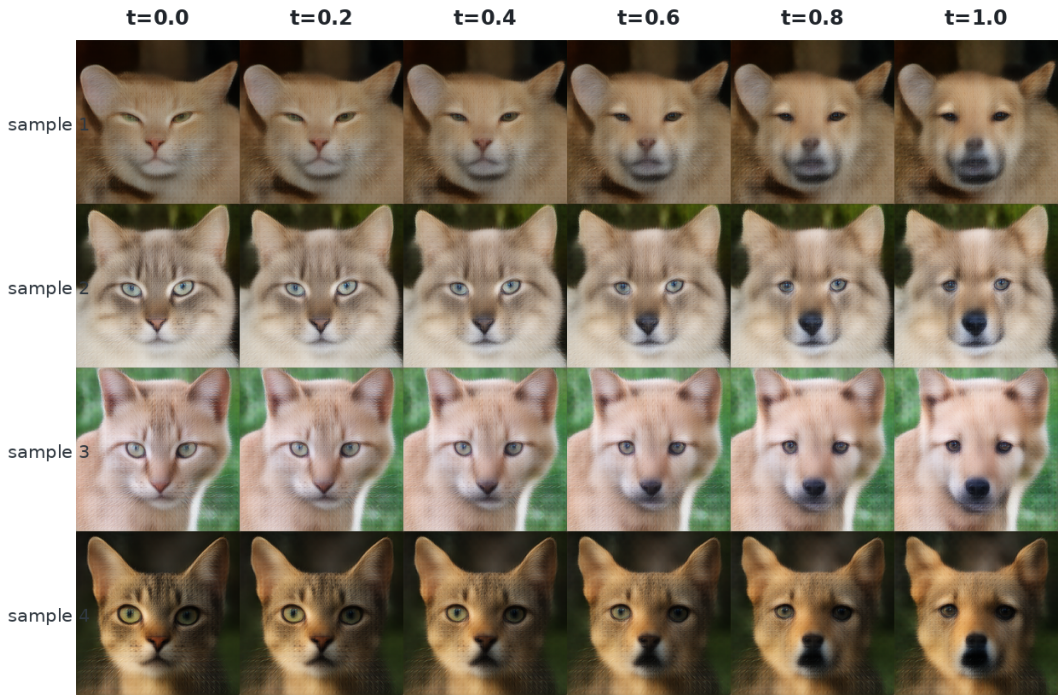


Figure 3: Cat-to-dog transport. Each row is one stochastic rollout decoded at six equally spaced times. Only the terminal law is supervised; intermediate images show the learned dynamics.

We report in Table 3 quantitative metrics. At the final time T, the transported cats reduce the latent sliced Wasserstein distance SW_2^2 (average W_2^2 over random 1D projections) to the dogs from 0.183 to 0.019, and 87% of them are classified as dogs by an independent classifier, not used during training.

This experiment is not meant as an image-translation benchmark. It shows that the same solver scales to a 512-dimensional state space and to a terminal loss combining latent and semantic features.

Summary. Intermediate marginals control how particles spread over time, but not which trajectory each particle follows. When trajectories are observed, the path term recovers this missing structure, such as routes (three-route) or motion directions (AMASS). Terminal supervision alone suffices to transport one population to another, even in high dimension (AFHQ).

Table 3: AFHQ cat-to-dog transport. MMDs and SW_2^2 are computed on held-out validation data, dog rate is the fraction of test cats classified as dogs.

population	latent $\mathrm{MMD}^2 \downarrow$	DINO $\mathrm{MMD}^2 \downarrow$	latent $\mathrm{SW}_2^2 \downarrow$	dog rate $\uparrow$
source cats	1.000	1.000	0.183	0.003
transported cats	**0.025**	**0.079**	**0.019**	**0.872**
target dogs	0.000	0.000	0.000	0.998

5 Conclusion

We proposed a soft law-constrained framework for learning stochastic generative dynamics from endpoint, intermediate marginal, and joint sampled-path observations, which reduces to the marginal-only model when g depends on $\boldsymbol{\mu}_T$ only through its terminal marginal. Marginal snapshots alone do not identify a unique path law. Casting the problem as a mean-field control objective makes the selected dynamics an explicit tradeoff between control effort and law-level fidelity, and the resulting McKean-Vlasov FBSDE makes this structure visible. Terminal observations induce an adjoint boundary condition, intermediate marginals act through running law-gradient forces, and joint-path penalties induce conditional adjoint jumps. The neural residual solver approximates this optimality system from samples, and the three experiments confirm the mechanism it is built on. Terminal supervision suffices for high-dimensional population transport, whereas source-conditional or directional structure requires the joint-path term.

Limitations and future work. The selected path law depends on the control cost, diffusion scale, discrepancy estimator, neural parametrization, and penalty weights, and joint-path supervision needs temporally paired data that unpaired marginal samples cannot provide. Finite observation grids and finite-dimensional memory encoders do not identify the full continuous path law, and sample-based law-gradient estimates can be noisy. The AMASS study also shows that joint-path supervision alone does not settle physical realism, since support and foot-contact errors stay above the real-motion reference. Natural next steps are to add explicit contact constraints, and to study convergence as the grid is refined.

AI use statement

Generative AI tools were used for language editing, code assistance, and debugging. All mathematical derivations, experimental designs, implementation choices, reported results, and final text were checked by the authors, who retain full responsibility for the work.

References

Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *International Conference on Learning Representations*, 2023.

René Carmona and François Delarue. Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics. *Annals of Probability*, 2015.

René Carmona and Mathieu Laurière. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: II—the finite horizon case. *The Annals of Applied Probability*, 32 (6):4065 – 4105, 2022.

Tianrong Chen, Guan-Horng Liu, Molei Tao, and Evangelos A. Theodorou. Deep momentum multi-marginal Schrödinger bridge. In *Advances in Neural Information Processing Systems*, 2023.

Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. StarGAN v2: Diverse image synthesis for multiple domains. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems*, 2021.

Maximilien Germain, Joseph Mikael, and Xavier Warin. Numerical resolution of McKean-Vlasov FBSDEs using neural networks. *Method. Comput. Appl. Prob.*, 24(4):2557–2586, March 2022.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13:723–773, 2012.

Nikita Gushchin, Sergei Kholkin, Evgeny Burnaev, and Alexander Korotin. Light and optimal schrödinger bridge matching. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.

Yadh Hafsi, Samy Mekkaoui, and Huyên Pham. Stochastic maximum principle for McKean–Vlasov control with discrete path dependence. *arXiv preprint arXiv:2609.04588*, 2026.

Mohamed Hamdouche, Pierre Henry-Labordère, and Huyên Pham. Nonparametric generative modeling for time series via Schrödinger bridge. *Journal of Machine Learning Research*, 27(58):1–23, 2026.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, 2020.

Guillaume Huguet, D. S. Magruder, Alexander Tong, Oluwadamilola Fasina, Manik Kuchroo, Guy Wolf, and Smita Krishnaswamy. Manifold interpolating optimal-transport flows for trajectory inference. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Zacharia Issa, Blanka Horvath, Maud Lemercier, and Cristopher Salvi. Non-adversarial training of neural SDEs with signature kernel scores. In *Advances in Neural Information Processing Systems*, 2023.

Yifan Jiang, Renyuan Xu, and Luhao Zhang. Schrödinger bridge with transport relaxation. *arXiv:2602.08118*, 2026.

Patrick Kidger, James Foster, Xuechen Li, and Terry Lyons. Neural SDEs as infinite-dimensional GANs. In *International Conference on Machine Learning*, 2021.

Hugo Lavenant, Stephen Zhang, Young-Heon Kim, and Geoffrey Schiebinger. Toward a mathematical theory of trajectory inference. *The Annals of Applied Probability*, 2024.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

Guan-Horng Liu, Yaron Lipman, Maximilian Nickel, Brian Karrer, Evangelos A. Theodorou, and Ricky T. Q. Chen. Generalized Schrödinger bridge matching. In *International Conference on Learning Representations*, 2024.

Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. SMPL: A skinned multi-person linear model. *ACM Transactions on Graphics*, 2015.

Jin Ma, Ying Tan, and Renyuan Xu. Schrödinger bridge for generative AI: Soft-constrained formulation and convergence analysis. *arXiv preprint arXiv:2510.11829*, 2025.

Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. Amass: Archive of motion capture as surface shapes. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. DINOv2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

Stefano Peluchetti. Diffusion bridge mixture transports, Schrödinger bridge problems, and generative modeling. *Journal of Machine Learning Research*, 2023.

Huyên Pham and Xavier Warin. Mean-field neural networks-based algorithms for mckean-vlasov control problems. *Journal of Machine Learning*, 3(2):176–214, 2024.

Javier Romero, Dimitrios Tzionas, and Michael J. Black. Embodied hands: Modeling and capturing hands and bodies together. *ACM Transactions on Graphics*, 2017.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion Schrödinger bridge matching. In *Advances in Neural Information Processing Systems*, 2023.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Bharath K. Sriperumbudur, Arthur Gretton, Kenji Fukumizu, Bernhard Schölkopf, and Gert R. G. Lanckriet. Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11: 1517–1561, 2010.

Alexander Tong, Jessie Huang, Guy Wolf, David van Dijk, and Smita Krishnaswamy. TrajectoryNet: A dynamic optimal transport network for modeling cellular dynamics. In *International Conference on Machine Learning*, 2020.

A NOTATIONS

For a given random variable χ over $(\Omega, \mathcal{F}, \mathbb{P})$, we denote by $\mathbb{P}_\chi$ its law under $\mathbb{P}$. The expectation under $\mathbb{P}$ is denoted by $\mathbb{E}$. We also denote the following spaces

$$\begin{cases} \mathbb{S}^p([0,T];\mathbb{R}^d) & := \Big\{ Y = (Y_t)_{0\le t\le T} : Y \text{ is } \mathbb{F}\text{-prog. measurable, càdlàg and} \\ & \qquad \mathbb{E}\Big[\sup_{0\le t\le T} |Y_t|^p\Big] < \infty \Big\}, \quad p \ge 1, \\ \mathbb{S}^\infty([0,T];\mathbb{R}^d) & := \Big\{ Y = (Y_t)_{0\le t\le T} : Y \text{ is } \mathbb{F}\text{-prog. measurable, càdlàg and} \\ & \qquad \Big\| \sup_{0\le t\le T} |Y_t| \Big\|_{L^\infty} < \infty \Big\}, \\ \mathbb{H}^2([0,T];\mathbb{R}^{d\times n}) & := \Big\{ Z = (Z_t)_{0\le t\le T} : Z \text{ is } \mathbb{F}\text{-prog. measurable and} \\ & \qquad \mathbb{E}\Big[\int_0^T |Z_t|^2 \,\mathrm{d}t\Big] < \infty \Big\}, \\ \mathrm{BMO}([0,T];\mathbb{R}^{d\times n}) & := \Big\{ Z = (Z_t)_{0\le t\le T} : Z \text{ is } \mathbb{F}\text{-prog. measurable and} \\ & \qquad \sup_\tau \Big\| \mathbb{E}\Big[\int_\tau^T |Z_t|^2 \,\mathrm{d}t \,\Big|\, \mathbb{F}_\tau \Big] \Big\|_{L^\infty} < \infty \Big\}, \end{cases}$$

where the supremum in the definition of BMO is taken over all $\mathbb{F}$-stopping times τ valued in $[0,T]$ and where $\|\chi\|_{L^\infty} := \inf\big\{C \ge 0 : |\chi| \le C, \quad \mathbb{P}-\text{a.s}\big\}$ for any random variable χ over $(\Omega, \mathcal{F}, \mathbb{P})$.

For a Polish space E endowed with a metric d_E, we denote by $\mathcal{P}_2(E)$ the set of Borel probability measures μ on E with finite second moment, i.e., $\int_E d_E(x,x_0)^2 \mu(\mathrm{d}x) < \infty$ for every $x_0 \in E$. For $\mu, \nu \in \mathcal{P}_2(E)$, we write $\Pi(\mu,\nu)$ for the set of probability measures on $E \times E$ with first marginal μ and second marginal ν. The associated 2-Wasserstein distance is

$$\mathcal{W}_2(\mu,\nu) := \left(\inf_{\pi\in\Pi(\mu,\nu)} \int_{E\times E} d_E(x,y)^2 \,\pi(\mathrm{d}x,\mathrm{d}y) \right)^{1/2}.$$

When $E = \mathbb{R}^d$, the metric d_E is the Euclidean distance. When $E = \mathcal{C}([0,T];\mathbb{R}^d)$, we also denote by

$$\mathcal{W}_{2,T}(\mu,\nu) := \left(\inf_{\pi\in\Pi(\mu,\nu)} \int \sup_{0\le t\le T} |\omega_t - \tilde\omega_t|^2 \pi(\mathrm{d}\omega,\mathrm{d}\tilde\omega) \right)^{1/2}.$$

B ADDITIONAL BACKGROUND

B.1 FUNCTIONAL DERIVATIVES ON $\mathcal{P}_2(\mathbb{R}^d)$

We recall the derivative convention used in Section 2. For a functional $U : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$, the linear functional derivative $\frac{\delta U}{\delta m}(\mu)(\cdot) : \mathbb{R}^d \to \mathbb{R}$ is characterized, up to an additive constant, by

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} U\big(\mu + \varepsilon(\mu'-\mu)\big) = \int \frac{\delta U}{\delta m}(\mu)(v)\,(\mu'-\mu)(dv), \qquad \mu' \in \mathcal{P}_2(\mathbb{R}^d),$$

whenever the derivative exists. A common normalization is

$$\int \frac{\delta U}{\delta m}(\mu)(v)\,\mu(dv) = 0.$$

When the map $v \mapsto \frac{\delta U}{\delta m}(\mu)(v)$ is differentiable, the spatial derivative

$$\partial_v \frac{\delta U}{\delta m}(\mu)(v)$$

is the law-gradient field that appears in the adjoint equation. For a function $F(x,\mu)$ depending on both a state and a law, the same construction is applied to $\mu \mapsto F(x,\mu)$ with x held fixed. For the joint-law argument of $g(\boldsymbol{\mu}_T;\boldsymbol{\rho}_T)$, the same definition on the product space gives $\delta g/\delta \boldsymbol{m}$, whose gradient has K vector blocks $\partial_{x_k}(\delta g/\delta \boldsymbol{m})$; these are the separate sensitivities entering the adjoint at each observation date.

B.2 MAXIMUM-PRINCIPLE DERIVATION

We use the discrete-path maximum principle of (Hafsi et al., 2026), which reduces to the classical McKean–Vlasov principle of (Carmona and Delarue, 2015) when g is independent of its joint-law argument. This section specializes that maximum principle to the generative objective (2.2).

For the model (2.1), assume $\sigma(t,x)$ is uniformly Lipschitz in x with linear growth, differentiable in x with the regularity required below, and that the derivatives of f and g are continuous and square-integrable along the controlled process. These conditions are a specialization of Assumptions 2.3, 2.5, and 4.1 in (Hafsi et al., 2026); path dependence is confined to the law functional g, not to the state coefficients. Keeping the pure law penalties outside the Hamiltonian,

$$H(t,x,y,z,a) = a\cdot y + \sigma(t,x) : z + \tfrac{1}{2}\|a\|^2,$$

where $M : N = \mathrm{Tr}(M^\top N)$, and with h_t, r_k as in (2.3), the adjoint specialized to (2.1)–(2.2) is

$$\begin{cases} \mathrm{d}X_t = \alpha_t\,\mathrm{d}t + \sigma(t,X_t)\,\mathrm{d}W_t, \\ \mathrm{d}Y_t = -\Big(\partial_x[\sigma(t,X_t) : Z_t] + \lambda_f h_t(X_t,\mu_t;\rho_t)\Big)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \qquad t\notin\{s_1,\dots,s_{K-1}\}, \\ -\Delta Y_{s_k} = \lambda_g \mathbb{E}\left[r_k(\mathbf{X},\boldsymbol{\mu}_T;\boldsymbol{\rho}_T)\mid \mathcal{F}_{s_k}\right], \qquad k<K, \\ Y_T = \lambda_g r_K(\mathbf{X},\boldsymbol{\mu}_T;\boldsymbol{\rho}_T). \end{cases}$$

For an optimal control, the variational necessary condition is $\partial_a H(t,X_t,Y_t,Z_t,\alpha_t)\cdot(a-\alpha_t)\ge 0$ for every $a\in A$, almost everywhere. If $a\mapsto H$ is strictly convex, pointwise minimization gives $\alpha_t = \mathrm{Proj}_A(-Y_t)$; in particular, if $A=\mathbb{R}^d$, then $\alpha_t = -Y_t$.

Proposition 1 (Reduced first variation and optimality)**.** *Assume the reduced dynamics and the regularity and integrability conditions above. If (X,Y,Z) is the state and adjoint associated with an admissible control α, then, for any admissible perturbation v,*

$$DJ(\alpha)[v] = \mathbb{E}\int_0^T (\alpha_t + Y_t)\cdot v_t\,\mathrm{d}t. \tag{B.1}$$

Consequently, every optimum over $A=\mathbb{R}^d$ satisfies $\alpha=-Y$. If, in addition, the L^2 lifts of the active running marginal functional and of the terminal functional g are convex and $x\mapsto\sigma(t,x) : Z_t$ is convex almost surely for almost every t, every solution of the reduced optimality system is a global minimizer.

Proof. For the additive controlled drift, the state variation satisfies

$$\mathrm{d}V_t = v_t\,\mathrm{d}t + \partial_x\sigma(t,X_t)V_t\,\mathrm{d}W_t, \qquad V_0=0,$$

and reduces to $V_t=\int_0^t v_u\,\mathrm{d}u$ when σ is state-independent. Differentiating the objective gives

$$\begin{aligned} DJ(\alpha)[v] = \mathbb{E}\int_0^T [\alpha_t\cdot v_t + \lambda_f h_t(X_t,\mu_t;\rho_t)\cdot V_t]\,\mathrm{d}t \\ + \lambda_g\sum_{k=1}^K \mathbb{E}[r_k(\mathbf{X},\boldsymbol{\mu}_T;\boldsymbol{\rho}_T)\cdot V_{s_k}]. \end{aligned}$$

Since V is continuous and V_{s_k} is $\mathcal{F}_{s_k}$-measurable, the terminal term and the path terms equal $\mathbb{E}[Y_T\cdot V_T] - \sum_{k<K}\mathbb{E}[\Delta Y_{s_k}\cdot V_{s_k}]$. Itô integration by parts on the intervals between observation dates, using the linearized diffusion term in V to cancel the $\partial_x[\sigma : Z]$ contribution in the adjoint, therefore gives

$$\mathbb{E}[Y_T\cdot V_T] - \sum_{k<K}\mathbb{E}[\Delta Y_{s_k}\cdot V_{s_k}] = \mathbb{E}\int_0^T [-\lambda_f h_t\cdot V_t + Y_t\cdot v_t]\,\mathrm{d}t,$$

which proves (B.1). The sufficient claim follows from convexity of the lifted law costs together with convexity of the Hamiltonian in (x,a); here the latter reduces to convexity of $x\mapsto\sigma(t,x) : Z_t$, since $a\mapsto a\cdot Y_t + \frac{1}{2}\|a\|^2$ is convex. These are the reduced counterparts of the necessary and sufficient conditions in (Hafsi et al., 2026, Theorems 4.6–4.7). □

For the MMD specialization $g = G$, convexity of G under mixtures of measures is not the same as convexity of its L^2 lift; in particular, a characteristic MMD kernel alone does not establish the sufficient convexity assumption in Proposition 1.

The adjoint of the reduced system also admits an equivalent conditional-expectation representation. Its integral form (Hafsi et al., 2026, Remark 3.5) and the tower property give, for $0 \le t \le T$,

$$Y_t = \mathbb{E}\left[Y_T + \int_t^T \Big(\partial_x [\sigma(u, X_u) : Z_u] + \lambda_f h_u(X_u, \mu_u; \rho_u) \Big) \, \mathrm{d}u + \lambda_g \sum_{\substack{1 \le k < K \\ t < s_k}} r_k(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) \,\middle|\, \mathcal{F}_t \right], \tag{B.2}$$

where the terminal value is the one in (2.8), so its final-coordinate path contribution is counted once. The strict inequality $t < s_k$ selects the post-jump value and preserves the terminal condition at $t = T$; for $t < s_k$, (B.2) follows from $\mathbb{E}[\mathbb{E}[r_k \mid \mathcal{F}_{s_k}] \mid \mathcal{F}_t] = \mathbb{E}[r_k \mid \mathcal{F}_t]$ once the stochastic integral has zero conditional mean. This representation makes explicit that full-path sensitivities enter the adjoint through conditional expectations, never as anticipative inputs to the control, and it is what the neural solver in Section 3 approximates by regression.

B.3 KL SCORE IDENTITY

For the KL law penalty, define the cost functional $U : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ by

$$U(\mu) = \mathrm{KL}(\mu \,\|\, \rho) = \int \log \frac{\mu(x)}{\rho(x)} \, \mu(dx), \quad \text{whenever } \mu \ll \rho, \text{ and } +\infty \text{ otherwise,}$$

where densities are denoted by the same symbols as the corresponding measures. For $\mu_\varepsilon = \mu + \varepsilon(\mu' - \mu)$, direct differentiation gives

$$\frac{d}{d\varepsilon} U(\mu_\varepsilon) = \int \log \frac{\mu_\varepsilon}{\rho} \, d(\mu' - \mu) + \int d(\mu' - \mu).$$

The second term is zero because both measures have unit mass. Setting $\varepsilon = 0$ identifies

$$\frac{\delta U}{\delta m}(\mu)(x) = \log \mu(x) - \log \rho(x)$$

up to an additive constant. Taking the spatial derivative gives

$$\partial_x \frac{\delta}{\delta m} \mathrm{KL}(\mu \,\|\, \rho)(x) = \nabla_x \log \mu(x) - \nabla_x \log \rho(x) = \nabla_x \log \frac{\mu(x)}{\rho(x)}.$$

This is the score-difference identity used in (2.4). The same calculation applies at intermediate times by replacing ρ with the observed marginal ρ_t for any t. On the product space, it also gives $r_k = \nabla_{x_k} \log(\boldsymbol{\mu}_T / \boldsymbol{\rho}_T)$ whenever the joint densities and their derivatives exist; this identity by itself does not provide the bounded Lipschitz regularity used in Theorem 1.

The same argument extends to the path-space term G. Under the density assumptions of (2.5), and for perturbations for which differentiation under the integral is justified, the same calculation on $(\mathbb{R}^d)^K$ yields

$$\frac{\delta G}{\delta \boldsymbol{m}}(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T)(\mathbf{x}) = \log \frac{\boldsymbol{\mu}_T(\mathbf{x})}{\boldsymbol{\rho}_T(\mathbf{x})} + c(\boldsymbol{\mu}_T),$$

where the additive constant is immaterial. Differentiation in x_k gives (2.5). If *both* laws factorize, $\boldsymbol{\mu}_T = \bigotimes_k \mu_{s_k}$ and $\boldsymbol{\rho}_T = \bigotimes_k \rho_{s_k}$, then

$$\mathrm{KL}(\boldsymbol{\mu}_T \,\|\, \boldsymbol{\rho}_T) = \sum_{k=1}^K \mathrm{KL}(\mu_{s_k} \,\|\, \rho_{s_k}), \qquad r_k = \nabla_{x_k} \log \frac{\mu_{s_k}(x_k)}{\rho_{s_k}(x_k)}.$$

A product-form target alone does not imply this reduction for a dependent generated law. Unlike this factorization, the additive-kernel identity (2.7) requires no product-law assumption.

B.4 MMD PATH-LAW FIELD

We first record the structure of the kernel used in the experiments, before turning to well-posedness of the full path-law formulation. Let $\mu_{s_k} = (\mathrm{pr}_k)_{\#}\boldsymbol{\mu}_T$ and $\rho_{s_k} = (\mathrm{pr}_k)_{\#}\boldsymbol{\rho}_T$ denote the coordinate marginals of the joint laws. For nonnegative weights $w_1, \dots, w_K, w_{\mathrm{path}}$, a positive-definite kernel κ_{marg} on $\mathbb{R}^d$, and a positive-definite kernel κ_{path} on $(\mathbb{R}^d)^K$, both with finite kernel expectations, define

$$\kappa_{\mathrm{add}}(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{K} w_k \, \kappa_{\mathrm{marg}}(x_k, y_k) + w_{\mathrm{path}} \, \kappa_{\mathrm{path}}(\mathbf{x}, \mathbf{y}),$$

$$\mathrm{MMD}^2_{\kappa_{\mathrm{add}}}(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T) = \sum_{k=1}^{K} w_k \, \mathrm{MMD}^2_{\kappa_{\mathrm{marg}}}(\mu_{s_k}, \rho_{s_k}) + w_{\mathrm{path}} \, \mathrm{MMD}^2_{\kappa_{\mathrm{path}}}(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T).$$

If $w_{\mathrm{path}} = 0$, a zero discrepancy only implies equality of the marginals $\mu_{s_k} = \rho_{s_k}$, not of the joint laws once $K > 1$, whereas if $w_{\mathrm{path}} > 0$ and κ_{path} is characteristic, it implies $\boldsymbol{\mu}_T = \boldsymbol{\rho}_T$.

Substituting κ_{add} into (2.6), the force splits accordingly:

$$r_k^{\mathrm{add}}(\mathbf{x}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \underbrace{w_k \int_{\mathbb{R}^d} \nabla_1 \kappa_{\mathrm{marg}}(x_k, u) \, (\mu_{s_k} - \rho_{s_k})(\mathrm{d}u)}_{\text{marginal part}} + \underbrace{w_{\mathrm{path}} \int_{(\mathbb{R}^d)^K} \partial_{x_k} \kappa_{\mathrm{path}}(\mathbf{x}, \mathbf{y}) \, (\boldsymbol{\mu}_T - \boldsymbol{\rho}_T)(\mathrm{d}\mathbf{y})}_{\text{path part}}.$$

The marginal part depends on the path only through x_k and on the laws only through their marginals at s_k. Evaluated at $\mathbf{X}$, it is therefore $\mathcal{F}_{s_k}$-measurable and passes unchanged through the conditional expectation in the jump of (2.8). The path part depends on the whole sampled path, including the future states X_{s_j}, $j > k$, and on the coupling across observation dates. It is the only term for which the conditional expectation in (2.8) is non-trivial, and the only one that can distinguish two processes with the same marginals.

Before proving local well-posedness, we record the regularity that the argument needs. For a symmetric kernel κ, the convention in (2.6) means

$$G(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = \tfrac{1}{2} \iint \kappa(\mathbf{x}, \mathbf{y})(\boldsymbol{\mu}_T - \boldsymbol{\rho}_T)(\mathrm{d}\mathbf{x})(\boldsymbol{\mu}_T - \boldsymbol{\rho}_T)(\mathrm{d}\mathbf{y}).$$

Differentiating in the measure and then in x_k gives the field r_k in (2.6), without a factor of two because G is *half* the squared MMD. In the estimates below the target law $\boldsymbol{\rho}_T$ is fixed, so we suppress it and write $r_k(\mathbf{x}, \boldsymbol{\mu}_T)$. If $\partial_{x_k}\kappa$ is bounded by B_κ and is jointly Lipschitz, then

$$\|r_k\|_\infty \le 2B_\kappa, \qquad \|r_k(\mathbf{x}, \boldsymbol{\mu}_T) - r_k(\mathbf{x}', \boldsymbol{\mu}'_T)\| \le L_\kappa\big(\|\mathbf{x} - \mathbf{x}'\| + \mathcal{W}_2(\boldsymbol{\mu}_T, \boldsymbol{\mu}'_T)\big). \tag{B.3}$$

The law term follows by integrating against a coupling of $\boldsymbol{\mu}_T, \boldsymbol{\mu}'_T$; the two spatial terms follow by the Lipschitz property in the first argument. These are the estimates of (Hafsi et al., 2026, Remark 5.12).

Proof of local MMD well-posedness. We use the notations introduced in Appendix A. The forward process X is continuous; Y may jump only at $s_1, \dots, s_{K-1}$. The assertion $Z \in \mathrm{BMO}$ means

$$\sup_{\tau \le T} \left\| \mathbb{E}\left[\int_\tau^T \|Z_t\|^2 \, \mathrm{d}t \,\middle|\, \mathcal{F}_\tau \right] \right\|_\infty < \infty, \tag{B.4}$$

where the supremum is over stopping times. By (B.3), $B_r \le 2B_\kappa$.

Proof of Theorem 1. For $X_0 = 0$, this is (Hafsi et al., 2026, Theorem 5.15), with (N, λ, K) in that paper replaced by (K, λ_g, κ) here. The kernel assumptions give the boundedness and Lipschitz estimates (B.3); the diffusion assumptions are precisely the relevant specializations of its Assumptions 2.3, 4.1, and 5.13. The bounded Lipschitz running field h_t is included below exactly as in the FBSDE (2.8). We recall the construction to explain the local-time restriction.

Given $\overline{Y} \in \mathcal{S}^4$, solve

$$\mathrm{d}X_t = -\overline{Y}_t \, \mathrm{d}t + \sigma(t, X_t) \, \mathrm{d}W_t, \qquad X_0 = 0,$$

and freeze $\mathbf{X}$ and $\boldsymbol{\mu}_T = \mathcal{L}(\mathbf{X})$ in the backward equation. Solving the resulting linear BSDE successively backward between observation dates, with the prescribed conditional jumps, defines

$\Psi(\overline{Y}) = Y$. Since $\partial_x[\sigma(t, X_t) : Z_t] = Z_t\Sigma(t, X_t)$ with bounded $\Sigma = (\Sigma_1, \dots, \Sigma_d)^\top$, the Girsanov argument of the source removes the linear Z-term on each open observation interval. Under the resulting equivalent measure, the BSDE has bounded running driver $\lambda_f h_t$, terminal value and jumps bounded by $\lambda_g B_r$, and therefore

$$\sup_{0\le t\le T} \|Y_t\| \le C_T\big(K\lambda_g B_r + \lambda_f T B_f\big) \quad \text{almost surely,}$$

where C_T depends only on the diffusion regularity bounds. Applying Itô's formula to $\|Y\|^2$ on these intervals and summing the finitely many bounded jump contributions gives (B.4). The forward and backward stability estimates in the proof of (Hafsi et al., 2026, Theorem 5.15, Equations (5.22)–(5.23)), with the additional bounded Lipschitz running field $\lambda_f h_t$, give

$$\|\Psi(\overline{Y}) - \Psi(\overline{Y}')\|_{\mathcal{S}^4}^4 \le CT^4 \|\overline{Y} - \overline{Y}'\|_{\mathcal{S}^4}^4, \qquad T \le 1,$$

with C controlled by the stated constants, including λ_f and the Lipschitz bound of h. Choosing $T_0 \le 1$ such that $CT_0^4 < 1$ yields the unique fixed point and the claimed solution spaces.

For completeness, the extension to a randomized L^4 source under *constant* diffusion follows directly, without changing the argument's contraction constant. For an input $\overline{Y} \in \mathcal{S}^4$, set $X_t = X_0 - \int_0^t \overline{Y}_u \,\mathrm{d}u + \sigma W_t$ and $\boldsymbol{\mu}_T = \mathcal{L}(\mathbf{X})$. The backward solution is explicitly

$$Y_t = \mathbb{E}\left[\lambda_g r_K(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) + \lambda_g \sum_{k<K:\, t<s_k} r_k(\mathbf{X}, \boldsymbol{\mu}_T; \boldsymbol{\rho}_T) + \lambda_f \int_t^T h_u(X_u, \mu_u; \rho_u)\,\mathrm{d}u \,\middle|\, \mathcal{F}_t\right]. \tag{B.5}$$

Brownian martingale representation in the filtration augmented by the independent initial state gives Z on each interval. The strict inequality $t < s_k$ in (B.5) gives the right-continuous convention and exactly the negative forward-time jumps in (2.8). The terminal value is included once, and the uniform bound $\sup_t \|Y_t\| \le K\lambda_g B_r + \lambda_f T B_f$ follows immediately.

For two inputs driven by the same X_0, W,

$$\|X - X'\|_{\mathcal{S}^4} \le T\|\overline{Y} - \overline{Y}'\|_{\mathcal{S}^4}, \qquad \mathcal{W}_2(\boldsymbol{\mu}_T, \boldsymbol{\mu}'_T) \le \sqrt{K}\|X - X'\|_{\mathcal{S}^4}.$$

By (B.3), the Lipschitz bound of h, and Doob's inequality applied to the conditional expectations in (B.5),

$$\|\Psi(\overline{Y}) - \Psi(\overline{Y}')\|_{\mathcal{S}^4} \le C_K\big(\lambda_g L_\kappa T + \lambda_f L_f T^2\big)\|\overline{Y} - \overline{Y}'\|_{\mathcal{S}^4},$$

where L_f is a uniform Lipschitz constant for h_t. The map is again a contraction for sufficiently small T, while the explicit forward equation preserves the fourth moment of X_0. The same conditional Itô estimate gives the BMO bound and completes this constant-diffusion extension. □

C Implementation details

C.1 Neural parametrization

A single network $Y_\theta(t_i, \mathcal{H}_i) \in \mathbb{R}^d$, shared across all dates, approximates the adjoint. Its final layer is initialized at zero, so that training starts from the uncontrolled diffusion. The three supervision regimes differ only in the objective and in the information $\mathcal{H}_i$ given to the network.

Terminal only. We take $\lambda_f = 0$ and $g(\boldsymbol{\mu}_T; \boldsymbol{\rho}_T) = D_g(\mu_T; \rho_T)$. Only the terminal law μ_T is penalized, while the intermediate laws μ_t are selected by the control energy $\frac{1}{2}\mathbb{E}\int_0^T \|\alpha_t\|^2\,\mathrm{d}t$ alone, which yields the Schrödinger bridge interpolation in the limit $\lambda_g \to \infty$. Then $r_k = 0$ for $k < K$, the adjoint has no jumps, and $\mathcal{H}_i = X_i$.

Marginal. We take $\lambda_f > 0$, $f(\mu_t; \rho_t) = D_f(\mu_t; \rho_t)$, and g as in the terminal case. Each marginal μ_t is constrained against ρ_t separately, so the coupling between dates is left unconstrained and $\mathcal{H}_i = X_i$.

Path. We take $g(\boldsymbol{\mu}_T;\boldsymbol{\rho}_T) = D_g(\boldsymbol{\mu}_T;\boldsymbol{\rho}_T)$ on the joint law of $\mathbf{X} = (X_{s_1},\dots,X_{s_K})$ with $\lambda_f > 0$. The whole path is constrained jointly, which also constrains the coupling between dates. Each field r_k depends on the whole path, so the adjoint jumps at every s_k and depends on the past, and $\mathcal{H}_i = (X_0,\dots,X_i)$. In practice, the network receives the zero-padded prefix in the three-route experiment and a compressed causal memory in AMASS (Appendix D.4).

In all three cases, the backward targets, the ridge projection, and the regression loss are identical (Algorithm 1); only the source terms S_j in (3.3) change.

Algorithm 1 Shared-adjoint solver for soft marginal and path-law constraints

Require: source law ρ_0; target samples $\boldsymbol{\xi}^b = (\xi_1^b,\dots,\xi_N^b)$, $b \in [\![1,B]\!]$; grid $0 = t_0 < \cdots < t_N = T$ with $t_j = s_j$, $N = K$; weights λ_f, λ_g; diffusion σ; batch size B; iterations L; ridge penalty γ; feature maps $(\phi_i)_{i<N}$; scales $(s_i)_{i<N}$ from (3.6)

Ensure: trained parameters $\theta^{(L)}$

initialize $\theta^{(0)}$ with a zero final layer
for $\ell = 0,\dots,L-1$ **do**
  draw $X_0^{(\ell),b} \sim \rho_0$, $b \in [\![1,B]\!]$
  for $i = 0,\dots,N-1$ **do**
    $\mathcal{H}_i^{(\ell),b} \leftarrow X_i^{(\ell),b}$ (terminal/marginal) or $(X_0^{(\ell),b},\dots,X_i^{(\ell),b})$ (path)
    draw $\Delta W_{i+1}^{(\ell),b} \sim \mathcal{N}(0,\Delta t_{i+1} I_d)$
    $X_{i+1}^{(\ell),b} \leftarrow X_i^{(\ell),b} - Y_{\theta^{(\ell)}}\big(t_i,\mathcal{H}_i^{(\ell),b}\big)\Delta t_{i+1} + \sigma\Delta W_{i+1}^{(\ell),b}$ ▷ forward simulation (3.2)
  end for
  detach $\mathbf{X}^{(\ell),b} = (X_1^{(\ell),b},\dots,X_N^{(\ell),b})$; form $\widehat{\mu}_j^{(\ell)}, \widehat{\boldsymbol{\mu}}^{(\ell)}, \widehat{\rho}_j, \widehat{\boldsymbol{\rho}}$
  $S_j^{(\ell),b} \leftarrow$ source (3.3), $j = 1,\dots,N$ ▷ backward targets
  $\widetilde{Y}_{N-1}^{(\ell),b} \leftarrow S_N^{(\ell),b}$; $\widetilde{Y}_i^{(\ell),b} \leftarrow S_{i+1}^{(\ell),b} + \widetilde{Y}_{i+1}^{(\ell),b}$, $i = N-2,\dots,0$ ▷ (3.4)
  for $i = 0,\dots,N-1$ **do**
    $u_i^{(\ell),b} \leftarrow \phi_i\big(\mathcal{H}_i^{(\ell),b}\big)$
    $\beta_i^{(\ell)} \leftarrow \arg\min_\beta \sum_{b'=1}^B \big\|\widetilde{Y}_i^{(\ell),b'} - \beta^\top u_i^{(\ell),b'}\big\|^2 + \gamma\|\beta\|^2$
    $\widehat{\widetilde{Y}}_i^{(\ell),b} \leftarrow \big(\beta_i^{(\ell)}\big)^\top u_i^{(\ell),b}$ ▷ ridge proxy (3.5)
  end for
  $\mathcal{L}_Y(\theta) \leftarrow \frac{1}{NB}\sum_{i=0}^{N-1}\sum_{b=1}^B \big\|\big(Y_\theta(t_i,\mathcal{H}_i^{(\ell),b}) - \mathrm{sg}(\widehat{\widetilde{Y}}_i^{(\ell),b})\big)/s_i\big\|^2$ ▷ (3.6); sg: stop-gradient
  $\theta^{(\ell+1)} \leftarrow \mathrm{Adam}\big(\theta^{(\ell)}, \nabla_\theta \mathcal{L}_Y\big)$
end for
return $\theta^{(L)}$

C.2 Numerical stabilization

The rollout can use antithetic Brownian increments during training: for an even batch, half the increments are sampled as $\Delta W \sim \mathcal{N}(0,\Delta t\, I_d)$ and the other half are set to $-\Delta W$. This reduces variance without changing the discretized equation. The forward update may use a norm-clipped adjoint

$$Y_{\mathrm{eff}} = Y \min\left(1, \frac{c}{\|Y\|}\right),$$

where c is a drift scale. This is a numerical safeguard for finite-step particle rollouts and is not part of the continuous FBSDE. We therefore log the fraction of updates affected by clipping. Solver-gradient clipping is optional; each experimental protocol states whether it is active and gives its threshold.

C.3 MMD discrepancies and feature maps

For a feature map φ, we use

$$\mathrm{MMD}_\varphi^2(\mu,\rho) = \iint \kappa\big(\varphi(x),\varphi(y)\big)\,(\mu-\rho)(\mathrm{d}x)\,(\mu-\rho)(\mathrm{d}y), \quad \kappa(z,z') = \frac{1}{M}\sum_{m=1}^M \exp\Big(-\frac{\|z-z'\|^2}{2p\,\varepsilon_m^2}\Big),$$

where p is the dimension of the feature space and $\varepsilon_1,\dots,\varepsilon_M$ are bandwidths calibrated on training data. We set $D_f(\mu;\rho_t) = \frac{1}{2}\mathrm{MMD}_\psi^2(\mu,\rho_t)$ for a marginal feature map ψ and $D_g(\boldsymbol{\mu}_T;\boldsymbol{\rho}_T) =$

$\frac{1}{2}\,\mathrm{MMD}^2_\Psi(\boldsymbol{\mu}_T, \boldsymbol{\rho}_T)$ for a path feature map Ψ.Features within a block are concatenated, and blocks are combined as a weighted sum of calibrated MMDs, i.e., an additive kernel as in (2.7). Throughout the experiments, the kernel family is fixed (multiscale Gaussian, with bandwidths calibrated on training data) and only the feature map changes. For a path $\mathbf{x} = (x_1, \dots, x_K)$ with initial state x_0, the feature maps used in the experiments are:

- *three-route*: the full sampled path $(x_1, \dots, x_K)$, the terminal state x_K, and the pair formed by the initial state x_0 and the crossing height;
- *AMASS*: vertical progression, root-height evolution, limb timing and coordination, and support-foot evolution along the path, together with framewise pose and sole-height features (Appendix D.4);
- *AFHQ*: the standardized latent code and the DINOv2 CLS feature of the decoded image (Appendix D.3).

.

D Experimental protocol details

The weights λ_g and λ_f control the tradeoff between endpoint fidelity and intermediate or path fidelity, and are treated as numerical penalty parameters rather than intrinsic properties of the data. In each setting, λ_g is first chosen so that a terminal-only solver reaches a stable endpoint match without persistent instability in the estimated law-gradient fields. When intermediate marginals or a path term are added, λ_g is kept fixed and the remaining weight is increased until the corresponding metrics improve without destroying the terminal fit or producing unstable rollouts. Larger weights enforce the corresponding law constraint more strongly, while finite values let the solver balance noisy empirical estimates against control effort and numerical stability.

D.1 Joint-path evaluation metrics

Let $\mathbf{x}^1, \dots, \mathbf{x}^B$ and $\mathbf{u}^1, \dots, \mathbf{u}^M$ be generated and held-out target paths on the same observation grid Π, with empirical laws $\widehat{\boldsymbol{\mu}}_T = B^{-1}\sum_p \delta_{\mathbf{x}^p}$ and $\widehat{\boldsymbol{\rho}}_T = M^{-1}\sum_q \delta_{\mathbf{u}^q}$. The joint-path Wasserstein distance is

$$\mathcal{W}_{2,\Pi}(\widehat{\boldsymbol{\mu}}_T, \widehat{\boldsymbol{\rho}}_T) = \left[\min_{\gamma \in \Gamma_{B,M}} \sum_{p=1}^{B}\sum_{q=1}^{M} \gamma_{pq} \sum_{k=1}^{K} \|x_k^p - u_k^q\|^2\right]^{1/2},$$

where $\Gamma_{B,M}$ contains nonnegative matrices with row sums $1/B$ and column sums $1/M$, so a single coupling matches complete paths rather than integrating separately coupled marginal distances. We use exact empirical transport, optimal matching when $B = M$.

The joint-path MMD uses a characteristic full-path Gaussian evaluation kernel $\kappa_{\mathrm{eval}}(\mathbf{x}, \mathbf{u}) = \exp(-\|\mathbf{x} - \mathbf{u}\|^2/(2\varepsilon^2_{\mathrm{eval}}))$, with bandwidth selected from training or validation data and frozen across arms. Its empirical squared discrepancy is the V-statistic

$$\widehat{\mathrm{MMD}}^2_{\kappa_{\mathrm{eval}}} = \frac{1}{B^2}\sum_{p,p'=1}^{B} \kappa_{\mathrm{eval}}(\mathbf{x}^p, \mathbf{x}^{p'}) + \frac{1}{M^2}\sum_{q,q'=1}^{M} \kappa_{\mathrm{eval}}(\mathbf{u}^q, \mathbf{u}^{q'}) - \frac{2}{BM}\sum_{p=1}^{B}\sum_{q=1}^{M} \kappa_{\mathrm{eval}}(\mathbf{x}^p, \mathbf{u}^q),$$

including diagonal pairs.including diagonal pairs. Table 1 reports this discrepancy after calibration, so that 0 is the null level of two independent target samples and 1 is the level of the uncontrolled diffusion. A full-path kernel is sensitive to the coupling across observation dates, unlike an additive kernel evaluated one marginal at a time.

For each pair of observation dates $j < k$, let $\widehat{C}^{\boldsymbol{\mu}}_{jk}$ and $\widehat{C}^{\boldsymbol{\rho}}_{jk}$ be the $d \times d$ matrices of empirical coordinatewise Pearson correlations between observations at s_j and s_k in the generated and target populations. The lagged-correlation error is

$$e_{\mathrm{corr}} = \max_{(j,k,a,b)} \left|\widehat{C}^{\boldsymbol{\mu}}_{jk}(a,b) - \widehat{C}^{\boldsymbol{\rho}}_{jk}(a,b)\right|,$$

maximized over all date pairs and coordinate pairs with nonzero training variance. This is a second-order diagnostic: a small correlation error does not by itself establish equality of the joint laws.

Let $h_{i,k}$ denote the root-height coordinate of clip i at frame k and w_i its held-out mixture weight. Mean displacement is $|\Delta h| = \sum_i w_i |h_{i,60} - h_{i,4}|$. The support error E_{sup} pools the four heel and toe heights reconstructed from the SMPL-H mesh, relative to a fixed initial floor, into a single root-mean-square statistic over frames 5 to 60. The root-height MMD uses a three-scale Gaussian kernel on the standardized height, averaged over frames and normalized by a training-calibration constant. Bandwidths, standardization, and normalizers are fitted on the training split and frozen for evaluation.

D.2 THREE-ROUTE OBSTACLE TRANSPORT

We use $T = 2$, $N = 20$, $\Delta t = 0.1$, and constant diffusion $\sigma = 0.05$, with 20 retained observation dates spanning the crossing region. The source is an equal-weight three-component mixture, and the prescribed upper, opening, and lower crossing heights are evenly spaced. Training and evaluation use independent banks of $4{,}095$ and $1{,}536$ paths; component identities are used only to construct and evaluate the reference coupling and are never given to the controller.

All three arms share one adjoint network with two hidden SiLU layers of width 64 and a zero-initialized output layer. The terminal-only and marginal-only arms take the Markov input (t_i, X_i). The path-dependent arm instead takes normalized time concatenated with the zero-padded two-dimensional prefix over all 21 dates. Terminal pressure is fixed across arms; the marginal pressure is shared by the marginal-only and path-dependent arms, and the path arm adds a whole-path pressure and a direction pressure, built from the joint law of the initial component and the crossing height rather than from an explicit class label.

Discrepancies are calibrated multiscale Gaussian-kernel MMDs, with bandwidths and null and separation floors fitted once on the training bank and frozen. We use batches of 512, Adam with learning rate 10^{-3}, and $3{,}000$ updates, with target preconditioning and gradient clipping disabled. Metrics are evaluated periodically on the fixed evaluation bank, and each arm reports the checkpoint with the smallest logged empirical objective.

D.3 HIGH-DIMENSIONAL TERMINAL TRANSPORT ON AFHQ

The AFHQv2 cache contains $5{,}065$ cat, $4{,}678$ dog, and $4{,}593$ wild-animal training images, together with 493, 491, and 483 official-test images, respectively. A deterministic $90/10$ split with seed 42 gives $4{,}559$ cat and $4{,}210$ dog samples for transport training and 506 cat and 468 dog samples for internal validation. Wild-animal images are used to train the representation but never appear in the transport source or target laws.

The frozen representation is a convolutional residual VAE operating at 512×512 pixels, with base width 64 and latent dimension 512. It is trained without domain labels, paired images, translation loss, or adversarial loss. Its objective combines pixel reconstruction, LPIPS with weight 0.5, a 10^{-4} KL penalty with a 10-epoch warm-up, and a local latent-cycle penalty of weight 0.2. We use AdamW with learning rate 2×10^{-4}, weight decay 10^{-4}, effective batch size 120, cosine decay to 10^{-5}, and gradient clipping at 1. The frozen epoch-51 checkpoint is selected by the internal reconstruction-validation criterion. DINO features are the normalized CLS representation of `facebook/dinov2-small`, with inputs resized to 256 and center-cropped to 224 pixels.

Latent coordinates are standardized using the dog-training mean and standard deviation. The controlled process has $T = 1$, $N = 40$, and constant diffusion $\sigma = 0.05$. The single shared adjoint network maps $(t_k, X_k) \in \mathbb{R}^{513}$ through two width-512 SiLU layers to $Y_k \in \mathbb{R}^{512}$; it has $788{,}480$ parameters, and its output layer is initialized to zero. Conditional expectations are estimated separately at each decision date by linear ridge regression with penalty 100. The terminal pressure is $\lambda_g = 640$, and the terminal penalty g have equal weight. Both use biased multiscale Gaussian-kernel MMD^2. The latent bandwidths are $(0.493, 0.697, 0.985, 1.394, 1.971)$, with null floor 0.002416 and source–target gap 0.046307; the DINO bandwidths are $(0.0227, 0.0321, 0.0454, 0.0642, 0.0907)$, with null floor 0.002349 and gap 0.221676. These calibration constants are fitted on the training split and then frozen.

We train with batches of 512 source states and 512 target states, two target draws per update, AdamW at learning rate 2×10^{-4} with zero weight decay, timewise target preconditioning calibrated from eight batches, gradient-norm threshold 100, and seed 42. The run is scheduled for $3{,}000$ updates and saves every 250 updates. Table 3 and Figure 3 use checkpoint 750, selected using the internal-validation law metrics before the official test classifier is evaluated. The latent SW_2^2 uses 256 fixed Gaussian projections and the 468-sample internal validation cohort; dog rate uses the 491 official-test cats, rollout seed 2026, and the frozen `openai/clip-vit-base-patch32` classifier.

D.4 Bidirectional human motion on AMASS

Both directions are curated without action labels, from SMPL-H body joints resampled to 30 frames per second. We search for two-second windows whose support-relative lower-back height changes monotonically by a prescribed amount, keep windows with stable endpoint postures and bounded yaw, vertical speed, and root translation, and reject ambiguous candidates by visual inspection. Accepted clips are canonicalized and resampled to 61 frames, with performer-disjoint splits of 522, 82, and 66 clips. The first five states are observed and $X_5, \dots, X_{60}$ are generated with $\Delta t = 1/30$ and $\sigma = 0.05$.

The path-dependent adjoint network is one time-shared, two-hidden-layer SiLU MLP of width 128, taking normalized time and a causal memory built from training-only prefix features, the current body state, and sole-height coordinates. Each outer iteration rolls a batch of paths, forms detached backward adjoint targets from the calibrated marginal and joint-path MMD derivatives, estimates the conditional expectations by ridge regression, and takes one Adam update on the timewise-normalized regression loss. The joint-path objective combines calibrated features for vertical progression, root-height evolution, limb timing and coordination, and support-foot evolution, together with framewise pose and sole-height terms that preserve local plausibility. Kernel bandwidths and normalizers are fitted on the training split and frozen. The direct neural SDE baseline uses a causal GRU encoder with an MLP drift head and optimizes the full rollout objective directly. Table 2 reports the checkpoint selected on validation metrics only.

E Compute resources

Table 4 reports wall-clock fitting time for a single run of each main experiment on one NVIDIA A100 80 GB GPU. The three three-route arms are trained separately for $3{,}000$ updates each; the AFHQ time is taken from the checkpoint-750 run log, and the AMASS time from the reported shared-MLP run.

Table 4: Compute for the main reported neural experiments, on one NVIDIA A100 80 GB GPU.

experiment	updates	wall-clock time
three-route (terminal / marginal / path)	$3 \times 3{,}000$	5.9 / 7.3 / 8.3 min
AFHQ cat-to-dog transport	750	1 h 54 min
bidirectional AMASS motion	$2{,}025$	≈ 3 h